



INSA

N°d'ordre NNT : 2019LYSEI118

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
L'INSA Lyon

Ecole Doctorale 160
Electronique, Electrotechnique et Automatique

Spécialité/ discipline de doctorat :
Traitement du signal et de l'image

Soutenue publiquement le 03/12/2019, par :
Matthieu Martin

**Reconstruction 3D de données
échographiques du cerveau du prématuré et
segmentation des ventricules cérébraux et
thalami par apprentissage supervisé**

Devant le jury composé de :

Delachartre, Philippe	PU	INSA Lyon	Directeur de thèse
Mateus, Diana	PU	Ecole Centrale de Nantes	Rapporteuse
PetitJean, Caroline	MC-HDR	Université de Rouen	Rapporteuse
Sciolla, Bruno	Docteur	Eco-SI	Examineur
Thiran, Jean-Philippe	PU	EPFL	Examineur
Quétin, Philippe	PH	CH d'Avignon	Examineur

Département FEDORA – INSA Lyon - Ecoles Doctorales – Quinquennal 2016-2020

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	CHIMIE DE LYON http://www.edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr INSA : R. GOURDON	M. Stéphane DANIELE Institut de recherches sur la catalyse et l'environnement de Lyon IRCELYON-UMR 5256 Équipe CDFA 2 Avenue Albert EINSTEIN 69 626 Villeurbanne CEDEX directeur@edchimie-lyon.fr
E.E.A.	ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE http://edeea.ec-lyon.fr Sec. : M.C. HAVGOUDOUKIAN ecole-doctorale.eea@ec-lyon.fr	M. Gérard SCORLETTI École Centrale de Lyon 36 Avenue Guy DE COLLONGUE 69 134 Écully Tél : 04.72.18.60.97 Fax 04.78.43.37.17 gerard.scorletti@ec-lyon.fr
E2M2	ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION http://e2m2.universite-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : H. CHARLES secretariat.e2m2@univ-lyon1.fr	M. Philippe NORMAND UMR 5557 Lab. d'Ecologie Microbienne Université Claude Bernard Lyon 1 Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69 622 Villeurbanne CEDEX philippe.normand@univ-lyon1.fr
EDISS	INTERDISCIPLINAIRE SCIENCES-SANTÉ http://www.ediss-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : M. LAGARDE secretariat.ediss@univ-lyon1.fr	Mme Emmanuelle CANET-SOULAS INSERM U1060, CarMeN lab, Univ. Lyon 1 Bâtiment IMBL 11 Avenue Jean CAPELLE INSA de Lyon 69 621 Villeurbanne Tél : 04.72.68.49.09 Fax : 04.72.68.49.16 emmanuelle.canet@univ-lyon1.fr
INFOMATHS	INFORMATIQUE ET MATHÉMATIQUES http://edinfomaths.universite-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	M. Luca ZAMBONI Bât. Braconnier 43 Boulevard du 11 novembre 1918 69 622 Villeurbanne CEDEX Tél : 04.26.23.45.52 zamboni@maths.univ-lyon1.fr
Matériaux	MATÉRIAUX DE LYON http://ed34.universite-lyon.fr Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bât. Direction ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIÈRE INSA de Lyon MATEIS - Bât. Saint-Exupéry 7 Avenue Jean CAPELLE 69 621 Villeurbanne CEDEX Tél : 04.72.43.71.70 Fax : 04.72.43.85.28 jean-yves.buffiere@insa-lyon.fr
MEGA	MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE http://edmega.universite-lyon.fr Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bât. Direction mega@insa-lyon.fr	M. Jocelyn BONJOUR INSA de Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69 621 Villeurbanne CEDEX jocelyn.bonjour@insa-lyon.fr
ScSo	ScSo* http://ed483.univ-lyon2.fr Sec. : Véronique GUICHARD INSA : J.Y. TOUSSAINT Tél : 04.78.69.72.76 veronique.cervantes@univ-lyon2.fr	M. Christian MONTES Université Lyon 2 86 Rue Pasteur 69 365 Lyon CEDEX 07 christian.montes@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

REMERCIEMENTS

Je tiens dans cette partie à remercier toutes les personnes qui m'ont permis directement ou indirectement d'avoir la chance de travailler sur cette thèse et d'aller jusqu'au bout de mon doctorat.

Pour commencer, je veux remercier les personnes avec qui j'ai interagi professionnellement au sein du laboratoire CREATIS. Tout d'abord mes encadrants, qui m'ont transmis leurs connaissances et guidé pendant ces 3 ans :

Mon directeur de thèse, Philippe DELACHARTRE, pour m'avoir offert l'opportunité de travailler sur cette thèse, suivi régulièrement pendant ces trois ans et pour avoir mis à ma disposition tous les moyens techniques nécessaires à ce travail.

Philippe QUETIN, pour son énorme implication dans la mise à disposition et l'annotation des données cliniques ainsi que pour sa pédagogie et le temps qu'il a consacré à m'expliquer les aspects médicaux du projet.

Bruno SCIOLLA, pour m'avoir fait partager ses connaissances techniques et scientifiques lors de ma première année de thèse et pour avoir continué à s'impliquer dans cette dernière après la fin de son post-doctorat.

Michaël SDIKA, pour avoir accepté de faire partie de mon encadrement pendant ma dernière année de doctorat et avoir apporté une expertise technique et scientifique qui ont contribué à la qualité de ce travail.

J'aimerais aussi remercier les membres de l'équipe ULTIM, dont j'ai fait partie pendant ces 3 ans, pour leur accueil et les échanges scientifiques que l'on a eu. Merci en particulier à :

Adeline BERNARD pour son investissement sur la partie reconstruction de ma thèse.

Hervé LIEBGOTT pour sa bonne humeur et son accompagnement personnel.

Damien GARCIA et François VARRAY pour leurs retours pertinents sur mes présentations.

Merci également à tous les membres du CREATIS avec qui j'ai pu avoir des discussions scientifiques enrichissantes, en particulier autour de mon sujet de thèse : Juan ABASCAL, Olivier BERNARD, Denis FRIBOULET, Pierre-Antoine GANAYE, Carole LARTIZIEN, Sarah LECLERC, Yufei LI, Fabien MILLOZ, ...

Enfin, je veux remercier le personnel technique et administratif pour leur aide et leur disponibilité au quotidien. En particulier, un immense merci à Pierre FERRIER pour m'avoir installé un bureau debout improvisé lors de mes derniers mois de thèse.

Ensuite je tiens à remercier les personnes du CREATIS que j'ai côtoyées dans un cadre moins formel :

Valentin BARON pour nos échanges sportifs et pour m'avoir fait découvrir le triathlon.

Maxime DI FOLCO pour nos échanges de compliments dans les couloirs du laboratoire.

Damien GARCIA pour les moments sportifs passés à l'escalade, à la piscine (en particulier pour ses conseils sur les battements en papillon) et en course à pied.

Pierre FERRIER, Tom HOWHEILLER, Goulven LEMOIGNE et Maxime POLICHETTI pour tous les bons moments passés en pause sur l'escalier de secours.

Mes co-bureaux Tom HOWHEILLER, Hao XU et Ruifen ZHANG pour nos échanges scientifiques et culturels. Merci en particulier à Ruifen pour sa bonne humeur et ses blagues ainsi qu'à Tom pour les moments de craquages mutuels.

Maxime POLICHETTI pour son soutien moral, le crabe sur les quais de Rhône et les rencontres inopinées au parc de la tête d'or.

Enfin, merci également aux autres personnes avec qui j'ai eu plaisir à échanger : Suzanne BUSSOD, Eloïse CHASSAING, Noëlie DEBS, Audrey DURAN, Ane ETXEBESTE, Pierre-Antoine GANAYE, Nina GHIGO, Victor LACROSAZ, Sarah LECLERC, Antonio LORENTE, Vincent PERROT, Emilie PINATEL, Emeline TURQUIN, Florian VIXEGE, ...

Mon environnement social personnel a également joué un rôle important dans la réussite de mon doctorat par le soutien, l'ouverture et les moments de décompression qu'il m'a apporté. Je souhaite dire un grand merci à Vincent BIHLER, Gwenaëlle BORN, Willy BREVET, Sophie CERF, Margaux CORONINI, Jibril COULIBALI, Elise DANDACHE, Pauline DEMOUSTIER, Christelle DESPLAT, Anne-Lise DUROY, Benoit FACHE, Gaëtan GRENOT, Nicolas GUILLET, Philippe JOOS, Matthieu Nault, Guillaume OUAR, Pierre PINEAU, Emilie RENOUX, Zoé RENOUX, Valentin SEREY, Paul SOULLIE, Hugo VIDAL, Stanislas VOISARD, ... Des remerciements particulier à :

Sophie CERF pour sa présence, nos échanges enrichissants et tous les moments ressourçants passés à Grenoble.

Margaux CORONINI pour tous ses encouragements.

Jibril COULIBALI pour cette soirée improbable et mémorable au bal des pompiers lors de ma période de rédaction.

Elise DANDACHE pour les randonnées, sorties sportives et moments d'échanges passés ensemble.

Gaëtan GRENOT pour avoir initié une cagnotte pour mon pot de thèse qui m'a permis d'acheter un vélo couché, son soutien et toutes les séances de sport passées ensemble.

Gaëtan GRENOT et Guillaume OUAR pour les randonnées partagées et pour m'avoir fait découvrir l'escalade.

Philippe JOOS pour ses limoncellos et ses spritzs recommandés jusqu'en Italie.

Willy BREVET, Lucie LAMBRETH et Hugo VIDAL pour les fous rires, les week-ends et les vacances passées ensemble.

Nicolas GUILLET et Pierre PINEAU pour toutes les soirées passées ensemble aux BDJ, à la pinte douce et à l'ET.

Zoé RENOUX pour m'avoir accueilli quelques jours à Londres lorsque j'ai perdu mon passeport à IUS.

Valentin SEREY et Stanislas VOISARD pour avoir continué à me faire découvrir les bars et restaurants Nantais.

Paul SOULLIE pour sa présence, tous nos échanges personnels et pour m'avoir tenu à jour musicalement pendant 3 ans.

Hugo VIDAL pour la qualité de ses relectures de mes articles rédigés en anglais.

Enfin, merci à ma famille pour leur soutien lors de mes années d'études et en particulier à mes parents pour m'avoir soutenu financièrement jusqu'à mon doctorat.

RÉSUMÉ

Environ 15 millions d'enfants naissent prématurément chaque année dans le monde. Ces patients sont susceptibles de présenter des anomalies du développement cérébral qui peuvent causer des troubles du neuro-développement : paralysie cérébrale, surdité, cécité, retard du développement intellectuel, ... Le volume des structures cérébrales est un paramètre clinique qui permet de pronostiquer les troubles du neuro-développement afin d'orienter les patients dans des parcours de soins adaptés. Dans le cas du système ventriculaire cérébral (SVC), le volume est également utilisé pour déterminer si une opération chirurgicale doit être réalisée. Aujourd'hui, ces mesures quantitatives peuvent être obtenues uniquement en analysant des données acquises par IRM, qui est un examen réalisé chez seulement 15 % des prématurés. En routine clinique, c'est l'échographie transfontanellaire (ETF) 2D qui est pratiquée sur tous les prématurés. Cet examen permet de diagnostiquer les dilations ventriculaires mais pas de quantifier précisément le volume ventriculaire ou celui des autres structures cérébrales car l'information 3D est indisponible. L'objectif de cette thèse est de montrer que, sous réserve d'avoir une qualité d'image suffisante, l'ETF 3D permettrait d'acquérir des données dans lesquelles le volume des structures cérébrales pourrait être quantifié chez 100 % des prématurés. Les problématiques principales associées à cet objectif sont l'obtention de données échographiques 3D de haute qualité et l'étiquetage des millions de voxels qu'elles contiennent dans un temps clinique (quelques secondes). Ce mémoire se focalise sur la segmentation du SVC et des thalami.

Les quatre contributions principales apportées par ce travail sont : le développement d'un algorithme qui permet la reconstruction 3D de haute qualité d'ETF 2D (1), la création de bases de données 3D annotées du SVC et des thalami (2), la segmentation du SVC et des thalami dans un temps clinique par des réseaux de neurones convolutifs (CNN) (3) et enfin le début de la création d'une architecture de CNN dédiée à ce problème de segmentation qui apprend la position anatomique du SVC (4). Notre algorithme de reconstruction a permis de reconstruire 25 volumes échographiques de haute qualité. Il a été validé in-vivo avec une précision de 0.69 ± 0.14 mm mesurée sur le corps calleux. Des premiers essais de validation sur fantôme ont aussi été réalisés. Au total, 25 segmentations 3D de référence ont été obtenues dans le cas du SVC et 16 dans le cas des thalami. Les meilleurs résultats de segmentation ont été obtenus avec le CNN 3D V-net qui a segmenté le SVC et les thalami avec des Dice respectifs de 0.828 ± 0.044 et 0.891 ± 0.016 . Ces segmentations ont été réalisées en quelques secondes dans des volumes de taille $320 \times 320 \times 320$ voxels. L'apprentissage de la position anatomique du SVC a été réalisé en intégrant un CPPN (Compositional Pattern Producing Network) aux CNNs avec lesquels nous avons travaillé. Il a permis d'améliorer significativement la précision des CNNs lorsqu'ils avaient peu de couches. Par exemple, dans le cas du réseau V-net à 7 couches, le Dice a progressé de 0.524 ± 0.076 à 0.724 ± 0.107 .

Cette thèse montre qu'il est possible de segmenter automatiquement, avec précision et dans un temps clinique, des structures cérébrales de l'enfant prématuré dans des données échographiques 3D. Cela montre qu'une échographie 3D de haute qualité pourrait être utilisée en routine clinique pour quantifier le volume des structures cérébrales et ouvrir la voie aux études d'évaluation de son bénéfice pour les patients.

Mots-clés : Données échographique 3D, reconstruction 3D, enfants prématurés, segmentation, thalami, système ventriculaire cérébral, apprentissage supervisé, CNN, CPPN

ABSTRACT

About 15 million children are born prematurely each year worldwide. These patients are likely to suffer from brain abnormalities that can cause neurodevelopmental disorders : cerebral palsy, deafness, blindness, intellectual development delay, ... The volume of the brain structures is a clinical parameter that can be used to predict these disorders and to guide patients into appropriate health care pathways. In the case of the cerebral ventricular system (CVS), the volume is also used to determine when surgery should be performed. Today, these quantitative measurements can only be obtained by analyzing MRI data, which is an examination performed in only 15 % of premature infants. In clinical routine, 2D transfontanellar ultrasound (ETF) is performed on all premature infants. This examination is used to diagnose ventricular dilation but not to quantify precisely the ventricular volume or that of other brain structures because the 3D information is unavailable. The aim of this thesis is to show that, provided that the image quality is sufficient, 3D ETF would make it possible to acquire data in which the volume of brain structures could be quantified in 100 % of premature infants. The main issues associated with this objective are to obtain high quality 3D ultrasound data and to label the millions of voxels they contain in a clinical time (a few seconds). This thesis focuses on the segmentation of CVS and thalami.

The four main contributions of this work are : the development of an algorithm that enables the high-quality 3D reconstruction of 2D ETFs (1), the creation of annotated 3D databases of the CVS and thalami (2), the segmentation of the CVS and the thalami in a clinical time by convolutional neural networks (CNN) (3) and finally the beginning of the creation of a CNN architecture dedicated to this segmentation problem that learns the anatomical position of CVS (4). Our reconstruction algorithm was used to reconstruct 25 high-quality ultrasound volumes. It was validated in-vivo where an accuracy of 0.69 ± 0.14 mm was obtained on the corpus callosum. First validation attempts were also performed on a neonate brain phantom. In total, 25 reference 3D segmentations were obtained for the CVS and 16 for the thalami. The best segmentation results were obtained with the V-net, a 3D CNN, which segmented the CVS and the thalami with respective Dice of 0.828 ± 0.044 et 0.891 ± 0.016 . These segmentations were performed in a few seconds in volumes of size $320 \times 320 \times 320$ voxels. Learning the anatomical position of the CVS was achieved by integrating a CPPN (Compositional Pattern Producing Network) into the CNNs. It significantly improved the accuracy of CNNs when they had few layers. For example, in the case of the 7-layer V-net network, the Dice has increased from 0.524 ± 0.076 to 0.724 ± 0.107 .

This thesis shows that it is possible to automatically segment brain structures of the premature infant into 3D ultrasound data with precision and in a clinical time. This proves that high quality 3D ultrasound could be used in clinical routine to quantify the volume of brain structures and paves the way for studies to evaluate its benefit to patients.

Keywords : 3D ultrasound data, 3D reconstruction, premature infants, segmentation, thalami, cerebral ventricular system, supervised learning, CNN, CPPN

SOMMAIRE

SOMMAIRE	xiii
LISTE DES FIGURES	xvii
LISTE DES TABLES	xxi
ABRÉVIATIONS & ACRONYMES	xxiii
NOTATIONS & SYMBOLES	xxv
I INTRODUCTION	1
I.1 Prématurité, complications neurologiques et diagnostic par l'imagerie cérébrale	3
I.1.1 Prématurité et complications neurologiques associées	3
I.1.2 Anatomie cérébrale	4
I.1.3 Imagerie cérébrale chez le prématuré	7
I.1.4 Anomalies cérébrales chez le prématuré	10
I.1.5 Limites et pistes d'amélioration du diagnostic des anomalies cérébrales par l'imagerie chez l'enfant prématuré	13
I.1.6 Conclusion	15
I.2 Problématique scientifique associée	16
I.2.1 Reconstruction 3D d'une séquence d'images 2D	16
I.2.2 Segmentation du SVC et des thalami dans des volumes reconstruits en 3D	20
I.3 Conclusion	25
II ÉTAT DE L'ART : RECONSTRUCTION 3D	27
II.1 Estimation de la position relative de deux images	30
II.1.1 Acquisition US main-libre réalisée sans capteur	31
II.1.2 Enregistrement du mouvement d'acquisition	34
II.1.3 Conclusion	36
II.2 Interpolation dans la grille cartésienne de reconstruction	38
II.2.1 Méthodes basées sur les pixels	38
II.2.2 Méthodes basées sur les voxels	39
II.2.3 Méthodes basées sur les fonctions	41
II.2.4 Comparaison des méthodes et conclusion	42
II.3 Repositionnement dans le repère anatomique	42
III ETAT DE L'ART : SEGMENTATION	45
III.1 Segmentation des structures cérébrales chez le nouveau né en IRM	47
III.1.1 Segmentation multi-atlas	47
III.1.2 Méthode paramétrique utilisant des atlas probabilistes	48
III.1.3 Segmentation du SVC : ventriculomégalie et hydrocéphalie	50
III.1.4 Classification	50

III.2 Segmentation en échographie	53
III.2.1 Segmentation du SVC par contours actifs	53
III.2.2 Segmentation du SVC par deep learning	57
III.3 Conclusion	58
IV RECONSTRUCTION 3D D'ÉCHOGRAPHIE TRANSFONTANELLAIRE 2D	59
IV.1 Introduction	63
IV.1.1 Données cliniques utilisées pour créer les volumes échographiques	63
IV.1.2 Hypothèses sur le mouvement d'acquisition	64
IV.1.3 Travail réalisé sur les fichiers DICOM	66
IV.2 Recalage deux à deux des images contenues dans les séquences dynamiques	68
IV.2.1 Méthode	68
IV.2.2 Expériences et résultats	71
IV.2.3 Conclusion	72
IV.3 Positionnement des images dans la grille de reconstruction	73
IV.3.1 Grille de reconstruction	73
IV.3.2 Modèle d'acquisition paramétrique	75
IV.3.3 Optimisation des paramètres des modèles paramétriques	78
IV.3.4 Validation des modèles de reconstruction	87
IV.3.5 conclusion	94
IV.4 Interpolation des volumes reconstruits	95
IV.4.1 Méthodes	95
IV.4.2 Temps d'interpolation et résultats qualitatifs	98
IV.4.3 Conclusion	103
IV.5 Positionnement dans le repère anatomique	104
IV.5.1 Méthode	104
IV.5.2 Résultats qualitatifs	105
IV.6 Débruitage	108
IV.7 Créations de volume échographiques à partir d'ETFs réalisées avec d'autres sondes	110
IV.8 Conclusion et perspectives	112
V CRÉATION DE SEGMENTATION 3D À PARTIR D'ANNOTATIONS MANUELLES 2D	115
V.1 Création des segmentations 3D du SVC	118
V.1.1 Segmentations 2D manuelles des séquences dynamiques	118
V.1.2 Recalage des images segmentées manuellement et interpolation des annotations dans les images non annotées	120
V.1.3 Positionnement dans la grille de reconstruction et création d'une segmentation 3D	124
V.1.4 Positionnement dans le repère commun	126
V.1.5 Validation globale de l'algorithme de création des segmentations 3D	126
V.1.6 Conclusion	129
V.2 Création des segmentations 3D des thalami	130
V.2.1 Segmentations 2D manuelles	130
V.2.2 Interpolation des annotations sur les images non annotées et positionnement dans le repère commun	131
V.2.3 Validation de la méthode d'interpolation	132
V.2.4 Conclusion	134
V.3 Conclusion	135

VI SEGMENTATION AUTOMATIQUE DU SYSTÈME VENTRICULAIRE CÉRÉBRAL ET DES THALAMI PAR APPRENTISSAGE SUPERVISÉ	137
VI.1 Bases de données	140
VI.1.1 Pré-traitement des données	140
VI.1.2 Base de données du SVC	142
VI.1.3 Base de données des thalami	144
VI.2 Réseaux de neurones convolutifs : définition, optimisation et comparaison	145
VI.2.1 Apprentissage profond et réseaux de neurones convolutifs	145
VI.2.2 Optimisation, précision et comparaison de réseaux de neurones convolutifs	150
VI.3 Fonction de coût et taille des patches d'entraînement	153
VI.3.1 Définition des fonctions de coût	153
VI.3.2 Optimisation par rapport à l'entropie croisée	154
VI.3.3 Optimisation par rapport à l'entropie croisée puis le softDice	154
VI.3.4 Comparaison de différentes fonctions de coût	156
VI.3.5 Définition de la taille des patches d'entraînement	156
VI.3.6 Conclusion	158
VI.4 Optimisation du nombre de couches de U-net et V-net	159
VI.4.1 Méthode	159
VI.4.2 Expérience et résultats	160
VI.4.3 Volumes mesurés après segmentation du SVC par V-net.	163
VI.4.4 Conclusion	163
VI.5 Apprentissage de la localisation du SVC	164
VI.5.1 Motivations	164
VI.5.2 Architecture du CPPN et intégration à U-net et V-net	164
VI.5.3 Expériences et résultats	166
VI.5.4 Discussion	169
VI.5.5 Conclusion	170
VI.6 Taille adéquate du jeu d'entraînement et augmentation artificielle de données	171
VI.6.1 Taille adéquate de l'ensemble d'entraînement	171
VI.6.2 Augmentation artificielle de données	173
VI.7 Segmentation de données atypiques	177
VI.7.1 Présentation des données atypiques	177
VI.7.2 Expériences et résultats	178
VI.7.3 Conclusion	181
VI.8 Segmentation des thalami	182
VI.8.1 Expérience et résultats	182
VI.8.2 Volume thalamique mesuré après segmentation par V-net.	184
VI.8.3 Conclusion	185
VI.9 Conclusion	186
VI.9.1 Résumé du travail effectué	186
VI.9.2 Résumé des contributions	186
VI.9.3 Perspectives	187
CONCLUSION	189
BIBLIOGRAPHIE	193

LISTE DES FIGURES

I.1	Plans anatomiques.	5
I.2	Images échographiques du cerveau en coupe coronale, sagittale et axiale.. . . .	5
I.3	Anatomie simplifiée de l'encéphale.	6
I.4	Anatomie d'un neurone.	7
I.5	Échographe et position de la fontanelle antérieure.	8
I.6	Image échographique 2D en coupe coronale provenant d'une ETF.	8
I.7	Appareil d'IRM.	9
I.8	Images IRM (pondération T2) en coupe coronale, sagittale et axiale.. . . .	10
I.9	Ventricules cérébraux normaux, dilatés et très dilatés.	11
I.10	Comparaison d'images échographiques acquises avec une sonde 9L4 et 10 V4.	14
I.11	PWMLs visibles, en coupe coronale, dans des volumes 3D US et IRM en pondération T2.	14
I.12	Balayage angulaire et séquence dynamique.	17
I.13	Comparaison d'une image acquise en coupe sagittale et de la séquence dynamique.	17
I.14	Association d'un repère à deux images successives.	18
I.15	Paramètres permettant d'estimer la position relative entre deux images successives.. . . .	18
I.16	Grille cartésienne de reconstruction et images replacées dans celle-ci.	19
I.17	Remplacement d'un volume 3D par rapport aux plans anatomiques de référence.	20
I.18	Définition de la segmentation.	20
I.19	Anatomie du système ventriculaire cérébral dans les trois vues anatomiques.	21
I.20	Aspect du SVC en coupe coronale dans des données US 3D reconstruites.	22
I.21	Anatomie des thalami dans les trois vues anatomiques.	23
I.22	Aspect des thalami en coupe coronale dans des données US 3D reconstruites.	23
I.23	Histogramme d'intensité pour l'arrière-plan, les thalami et le SVC.	24
I.24	Image US 2D et magnitude du gradient calculé sur cette image.	24
II.1	Les mouvements d'acquisition en imagerie US.. . . .	30
II.2	Décorrélation du speckle.	31
II.3	Estimation du mouvement hors du plan de l'image.	31
II.4	Définition du mouvement dans le plan de l'image.	32
II.5	Distribution des valeurs de chaque pixel au voxel le plus proche.	38
II.6	Interpolation de la grille à partir des voxels non vides dans un certain voisinage.	39
II.7	Distribution de la valeur d'un pixel à plusieurs voxels de son voisinage.	39
II.8	Attribution à chaque voxel de la valeur du pixel le plus proche.	40
II.9	Attribution de la valeur d'un voxel à partir des pixels contenus dans un certain voisinage.	40
II.10	Attribution de la valeur d'un voxel à partir de deux images adjacentes.	41
II.11	Interpolation de la grille de reconstruction à partir de fonctions.	41
III.1	Segmentation par recalage multi-atlas.	48
III.2	Segmentation à partir d'un modèle paramétrique.	49
III.3	Méthode classique pour segmenter une image avec un classifieur.	51
III.4	Approche de representation learning pour segmenter une image avec un classifieur.	51

III.5	CNN multi-échelle proposé par Moeskops <i>et al.</i> (2016).	53
III.6	Segmentation du SVC en imagerie US 3D : méthode proposée par Qiu <i>et al.</i> (2017).	55
III.7	Segmentation du SVC en imagerie US 3D : méthode proposée par Boucher <i>et al.</i> (2018b).	56
III.8	Segmentation du SVC en imagerie US 2D : méthode proposée par Wang <i>et al.</i> (2018).	57
IV.1	Algorithme permettant de créer un volume échographique à partir d'une ETF 2D.	61
IV.2	Balayage angulaire et séquence dynamique.	63
IV.3	Définition d'un repère orthonormé R_1 dans chaque SD.	64
IV.4	Paramètres permettant d'estimer la position relative entre deux images successives.	64
IV.5	Remplacement d'un volume 3D par rapport aux plans anatomiques de référence.	64
IV.6	Mouvement d'acquisition respectant l'hypothèse 1.	65
IV.7	Mouvement d'acquisition respectant l'hypothèse 2.	66
IV.8	Paramètres caractérisant le mouvement dans le plan de l'image.	68
IV.9	Définition de l'ensemble E .	69
IV.10	Recalage des images d'une séquence dynamique.	71
IV.11	Optimisation du facteur s .	71
IV.12	Image acquise en coupe sagittale et SD en coupe sagittale.	73
IV.13	Grille de reconstruction dans le repère R_2 a) et séquence dynamique dans le repère R_1 .	74
IV.14	Représentation graphique des paramètres du modèle de rotation d'axe mobile.	77
IV.15	Définition du plan de coupe de la SD correspondant à l'image de référence.	79
IV.16	Visualisation du plan de coupe de la SD correspondant à l'image de référence.	81
IV.17	Points utilisés pour optimiser les paramètres du modèle d'acquisition	81
IV.18	Points de repères anatomiques en coupe sagittale avant et après reconstruction.	83
IV.19	Coupe sagittale des volumes échographiques obtenus avec les modèles d'acquisition.	83
IV.20	Représentation graphique des paramètres du modèle f_1 .	85
IV.21	Représentation graphique des paramètres du modèle f_1 .	86
IV.22	Segmentations manuelles du corps calleux.	88
IV.23	Corps calleux tronqués ou difficilement visibles.	88
IV.24	Segmentations manuelles et reconstruites du corps calleux.	89
IV.25	Corps calleux avec le moins bon Dice après reconstruction.	91
IV.26	Corps calleux avec le meilleur Dice après reconstruction.	91
IV.27	Structures cérébrales du fantôme utilisé pour valider les modèles d'acquisition.	92
IV.28	Comparaison de l'anatomie du corps calleux in-vivo et sur fantôme.	93
IV.29	Images acquises en coupe coronale in-vivo et sur fantôme.	93
IV.30	Interpolation par la méthode PNN.	96
IV.31	Interpolation par la méthode VNN.	96
IV.32	Interpolation à partir des valeurs des plus proches voisins.	97
IV.33	Interpolation à partir des valeurs des pixels des images adjacentes.	98
IV.34	Résultats de l'interpolation avec la méthode PNN.	100
IV.35	Résultats de l'interpolation avec la méthode VNN.	100
IV.36	Résultats de l'interpolation dans un voisinage sphérique de rayon 3.	101
IV.37	Résultats de l'interpolation dans un voisinage sphérique de rayon 5.	101
IV.38	Résultats de l'interpolation à partir d'un pixel sur les images adjacentes.	102
IV.39	Résultats de l'interpolation à partir de neuf pixels sur les images adjacentes.	102
IV.40	Comparaison des méthodes d'interpolation.	103
IV.41	Emplacement du splénium du corps calleux en coupe sagittale et coronale.	104
IV.42	Orientation d'un volume 3D en coupe sagittale.	105
IV.43	Données échographiques 3D dans le repère commun.	106
IV.44	Comparaison d'un IRM et d'un volume échographique reconstruit.	107
IV.45	Volumes échographiques avec et sans débruitage.	108
IV.46	Volumes échographiques créés à partir d'ETFs acquises avec différentes sondes.	110

V.1	Segmentations manuelles du SVC en coupe coronale, visibles en coupe sagittale.	118
V.2	Parties du SVC qui ont systématiquement été segmentées manuellement dans les SDs. .	118
V.3	Segmentation manuelle des thalami sur tablette avec le logiciel MITK.	119
V.4	Étapes ayant permis de créer les segmentations 3D du SVC.. . . .	119
V.5	Segmentation du SVC et carte de distance aux bord du SVC.	120
V.6	Interpolation des segmentations du SVC dans les images non annotées.. . . .	121
V.7	Segmentations manuelles du SVC et segmentation obtenue après interpolation.	122
V.8	Obtention d'une segmentation de référence à partir d'un modèle CAO du SVC.	122
V.9	Evaluation de la précision de l'interpolation des contours du SVC dans les SDs.	123
V.10	Précision de l'interpolation du SVC en fonction de l'espacement entre les annotations. .	123
V.11	Comparaison d'une segmentation de référence à des segmentations interpolées.	124
V.12	Vue sagittale d'une segmentation 3D interpolées du SVC.	125
V.13	Obtention d'une segmentation 3D de référence.	125
V.14	Estimation de l'erreur commise lors de la création d'une segmentation 3D.	126
V.15	Estimation de l'erreur globale commise lors de la création d'une segmentation 3D. . . .	127
V.16	Précision de la création d'une segmentation 3D.	128
V.17	Comparaison d'une segmentation 3D et de segmentations interpolées.	128
V.18	Segmentation manuelle des thalami.	130
V.19	Étapes ayant permis de créer les segmentations 3D des thalami.	130
V.20	Segmentations des thalami après interpolation sur toutes les images.	132
V.21	Étapes ayant permis de créer les segmentations 3D de référence des thalami.	132
V.22	Erreur commise lors de l'interpolation d'une segmentation 3D des thalami.	133
V.23	Précision des segmentations 3D des thalami obtenues.	133
VI.1	Données échographiques 3D et segmentations associées du SVC et des thalami.	140
VI.2	Données (X, Y) dans le cas de la segmentation du SVC.	141
VI.3	Données (X, Y) dans le cas de la segmentation du SVC et des thalami.. . . .	141
VI.4	SVC dilatés contenus dans le jeu de test.	142
VI.5	SVC normaux contenus dans le jeu de test.	143
VI.6	Thalami contenus dans le jeu de test.. . . .	145
VI.7	Classes des méthodes d'apprentissage automatique.. . . .	146
VI.8	Représentation d'une couche convolutive.	147
VI.9	Réseau de neurones convolutifs U-net.	148
VI.10	Réseau de neurones convolutifs V-net.	149
VI.11	Histogramme du Dice moyen obtenu après avoir optimisé 30 U-net.	152
VI.12	Courbes d'apprentissage de U-net par rapport à l'entropie croisée.	154
VI.13	Courbes d'apprentissage de U-net par rapport à l'entropie croisée puis le softDice. . . .	155
VI.14	Courbes d'apprentissage de U-net avec et sans changement de fonction de coût.	155
VI.15	Augmentation du nombre de couches de l'architecture U-net.	159
VI.16	Augmentation du nombre de couches de l'architecture V-net.. . . .	160
VI.17	Résultats qualitatifs obtenus pour la segmentation du SVC par U-net et V-net.. . . .	162
VI.18	Données d'entrée et architecture du CPPN.. . . .	165
VI.19	Intégration du CPPN à U-net et V-net.	166
VI.20	Images des formes produites par le CPPN.	167
VI.21	Évolution du Dice et de la MAD avec et sans usage du CPPN.	168
VI.22	Évolution du Dice et de la MAD avec et sans usage du CPPN.	169
VI.23	Segmentations du SVC pour par un V-net 7 couches avec et sans CPPN.	170
VI.24	Influence du nombre de volumes d'entraînement sur le Dice.	171
VI.25	Augmentation artificielle de données.	174
VI.26	Données échographique et segmentations 3D d'un SVC extrêmement dilaté.	177
VI.27	Sonde Philips Purewave : données échographique et segmentations 3D du SVC associé..	178

VI.28	Segmentations 3D du SVC produites par U-net et V-net : cas d'hydrocéphalie..	179
VI.29	Segmentations 3D du SVC produites par U-net et V-net : sonde Philips PureWave.	179
VI.30	Faux-positif proche d'un champ DICOM éliminé..	180
VI.31	Segmentation des thalami produites par U-net et V-net.	183
VI.32	Segmentation du SVC et des thalami produites par V-net..	184

LISTE DES TABLES

I.1	Classification de la prématurité.	3
I.2	Prévalence des handicaps moteurs et des déficiences intellectuelles chez le prématuré.	4
IV.1	Erreur quadratique moyenne calculée sur les SDs et SDs recalées avec $s \in \{0.5, 1\}$	72
IV.2	Sens physique des paramètres du modèle défini par l'Équation.IV.10	76
IV.3	Valeurs initiales et bornes des paramètres à optimiser dans le modèle f_1	82
IV.4	Distance moyenne entre les points de repère anatomique après optimisation.. . . .	84
IV.5	Valeur des paramètres du modèle f_1 après optimisation.	85
IV.6	Précision du placement des images dans la grille de reconstruction..	90
IV.7	Temps moyen d'interpolation de la valeur des voxels de la grille de reconstruction.	99
IV.8	Valeurs des angles de rotation en coupe sagittale et coronale	105
IV.9	Distance moyenne entre les points de repère anatomique après reconstruction.. . . .	111
V.1	Précision des segmentations 3D obtenues pour les espacements utilisés.	135
VI.1	Base de données du SVC..	142
VI.2	Base de données des thalami.	144
VI.3	Taille et nombre de patchs donnés en entrée des réseaux pendant l'entraînement.	150
VI.4	Dice moyens obtenus pour différentes fonction de coût.	156
VI.5	Dice moyens obtenus pour différentes fonctions de coût et tailles de patch.	157
VI.6	p-valeurs obtenues pour différentes fonctions de coût et tailles de patch.	157
VI.7	Résultats quantitatifs obtenus par U-net et V-net avec différents nombres de couches.. . . .	161
VI.8	p-valeurs obtenues en comparant U-net et V-net..	161
VI.9	Erreurs volumiques commise par V-net dans le cas du SVC..	163
VI.10	Résultats quantitatifs obtenus en utilisant un CPPN avec U-net et V-net pour différents nombres de couches.	167
VI.11	p-valeurs comparant les performances de chaque réseau avec et sans CPPN.	168
VI.12	Dice moyen obtenu en fonction du nombre de volume d'entraînement..	172
VI.13	Entraînement de U-net avec augmentation de données.	175
VI.14	Entraînement de V-net avec augmentation de données..	176
VI.15	Dice obtenus par U-net et V-net dans le cas de la segmentation de données atypiques.	178
VI.16	Résultats quantitatifs obtenus par U-net et V-net : SVC et thalami.	182
VI.17	Erreur volumique commise par V-net dans le cas des thalami..	184
VI.18	Précision des segmentations 3D obtenues pour les espacements utilisés.	190

ABRÉVIATIONS & ACRONYMES

Abréviation ou acronyme	Signification
A	Antérieur
ASD	Absolute surface difference
CC	Corps calleux
CEC	Couche entièrement connectée
CH	Centre hospitalier
CNN	Réseau de neurones convolutionnels
CPPN	Réseau générateur de formes par compositions
CSP	Cavum septi pellucidi
D	Droite
DS	Distribution step
EC	Entropie croisée
ECP	Entropie croisée pondérée
EM	Espérance-maximisation
ETF	Échographie transfontanellaire
EQM	Erreur quadratique moyenne
FBM	Function based method
G	Gauche
HFS	Hole-filling method
IVH	Hémorragie intraventriculaire
IRM	Imagerie par résonance magnétique
LCS	Liquide cérébro-spinal
MAD	Mean absolute distance
MV	Meilleure validation
P	Postérieur
PBM	Pixel based method
PNN	Pixel-nearest neighbor
PVL	Leucomalacie périventriculaire
PWML	Lésions punctiformes de la substance blanche
SA	Semaine d'aménorrhée
SB	Substance blanche
SD	Séquence dynamique
SVC	Système ventriculaire cérébral

Abréviation ou acronyme	Signification
US	Ultrasonore
VBM	Voxel based method
VL	Ventricules latéraux
VM	Ventriculomégalie
VNN	Voxel-nearest neighbor
VV	Volume ventriculaire

NOTATIONS & SYMBOLES

Symboles	Signification
A^{-1}	Inverse la matrice A
A^t	Transposée de la matrice A
$\operatorname{argmin}_x f(x)$	Élément x qui minimise la fonction f
$\min_x f(x)$	Minimum de f par rapport à la variable x
\cos	Fonction cosinus
\sin	Fonction sinus
\exp	Fonction exponentielle
\ln	Fonction logarithme népérien
$\forall x$	Quel que soit x
$E \cap F$	Intersection des ensembles E et F
$E \cup F$	Union des ensembles E et F
$ E $	Nombre de 1 dans l'ensemble E
$\ \cdot\ _1$	Norme 1
$[[a, b]]$	Ensemble des nombres entiers compris entre a et b
$[a, b]$	Ensemble des nombres réels compris entre a et b
$\{a, b\}$	Ensemble contenant les éléments a et b
$\widehat{(i, j)}$	Angle entre les vecteurs i et j
\overrightarrow{AB}	Vecteur reliant les points A et B
\mathbb{R}	Ensemble des nombres réels
\mathbb{R}^N	N produit cartésien de \mathbb{R} avec lui même
$\sum_{i=1}^N$	Somme pour i allant de 1 à N
\wedge	Produit vectoriel

Notations	Signification
A_t	Matrice permettant de recalculer I_{t+1} sur I_t
C	Nombre de classes associées au problème de segmentation
C_t	Position du centre de rotation de la sonde au temps t
d	Distance euclidienne
δ_{ref}	Taille des pixels de I_{ref}
δ_{SD}	Taille des pixels des images d'une SD
f_1	Modèle de rotation d'axe mobile
f_2	Modèle de rotation d'axe fixe
f_3	Modèle de translation linéaire
O_t	Centre de l'image t d'une SD
t	Index d'une image dans une SD
T	Nombre d'images dans une ETF
I_t	Image t d'une SD
I_{3D}	Coupe sagittale d'une SD après reconstruction
I_{ref}	Acquisition de référence en coupe sagittale
I_{sag}	Coupe sagittale choisie dans une SD
\mathcal{X}	Ensemble des données échographiques 3D
\mathcal{Y}	Ensemble des segmentations de références réalisées dans \mathcal{X}
$(d_{x_t}, d_{y_t}, d_{z_t})$	Vecteur reliant O_t à O_{t+1}
$\alpha_t, \beta_t, \gamma_t$	Angles entre les images I_t et I_{t+1}
φ, ϕ, ψ	Angle entre un volume 3D interpolé et le repère anatomique de référence
$(O_2, u_{x_2}, u_{y_2}, u_{z_2})$	Repère orthonormé associé aux volumes 3D interpolés
$(O_1, u_{x_1}, u_{y_1}, u_t)$	Repère orthonormé associé aux séquences dynamiques
(O_{CC}, u_x, u_y, u_z)	Repère orthonormé associé aux volumes dans le repère commun
(C_t, i_t, j_t, k_t)	Référentiel associée à l'image I_t
(O_t, i_t, j_t, k_t)	Repère orthonormée associée à l'image I_t

CHAPITRE I

INTRODUCTION

Contents

I.1	Prématurité, complications neurologiques et diagnostic par l'imagerie cérébrale .	3
I.1.1	Prématurité et complications neurologiques associées	3
	Prématurité	3
	Complications neurologiques associées à la prématurité	4
I.1.2	Anatomie cérébrale	4
	Plans anatomiques de référence	4
	Anatomie de l'encéphale	6
	Anatomie du neurone	6
I.1.3	Imagerie cérébrale chez le prématuré	7
	Échographie transfontanellaire	7
	Imagerie par résonance magnétique	9
I.1.4	Anomalies cérébrales chez le prématuré	10
	Hémorragie intraventriculaire et ventriculomégalie	10
	Anomalies de la substance blanche et retard de croissance cérébrale	12
I.1.5	Limites et pistes d'amélioration du diagnostic des anomalies cérébrales par l'imagerie chez l'enfant prématuré	13
	Limite du diagnostic des anomalies cérébrales par l'imagerie	13
	Pistes d'améliorations du diagnostic des anomalies cérébrales par l'imagerie	13
I.1.6	Conclusion	15
I.2	Problématique scientifique associée	16
I.2.1	Reconstruction 3D d'une séquence d'images 2D	16
	Acquisition d'une séquence d'images 2D	16
	Repositionnement des images les unes par rapport aux autres	18
	Interpolation des volumes reconstruits	19
	Orientation du volume reconstruit dans les trois plans anatomiques	19
I.2.2	Segmentation du SVC et des thalami dans des volumes reconstruits en 3D	20
	Définition	20
	Anatomie du SVC et des thalami et aspect dans des données US.	21
	Création d'une base de données annotées	25
I.3	Conclusion	25

CHAPITRE I

Chaque année, on estime que 15 millions d'enfants naissent prématurément dans le monde (Blencowe *et al.* (2012)). Ce chiffre, en augmentation, correspond à plus d'une naissance sur dix. Ces patients sont extrêmement fragiles et donc susceptibles de subir de nombreuses complications qui sont la première cause de décès en période néonatale : 1 millions d'enfants par an (Liu *et al.* (2016a)). Ce premier chapitre regroupe des informations sur la prématurité et les complications qui lui sont associées, en particulier au niveau du développement du système nerveux central. Les méthodes cliniques de diagnostic des anomalies cérébrales et de pronostic des troubles du neuro-développement seront présentées avec leurs limites. Les problématiques scientifiques et techniques que ces dernières soulèvent seront ensuite définies.

I.1 Prématurité, complications neurologiques et diagnostic par l'imagerie cérébrale

I.1.1 Prématurité et complications neurologiques associées

Prématurité

La durée d'une grossesse, aussi appelée âge gestationnel, se mesure en semaine d'aménorrhée (SA). Ce chiffre correspond au nombre de semaines qui se sont écoulées depuis les dernières règles. A terme, il se situe entre 40 et 41 SA pour une grossesse normale. On parle de prématurité lorsqu'un enfant naît avant 37 SA, on en distingue plusieurs niveaux qui sont définis dans la Table.I.1 et chiffrés à partir des données recueillies dans Blencowe *et al.* (2012).

TABLE I.1 – Classification de la prématurité.

	SA	Prévalence
Très grande prématurité	< 28	5.2 %
Grande prématurité	28 à 32	10.4 %
Prématurité moyenne	32 à 37	84.4 %

Le pourcentage mondial de naissances prématurées a augmenté en moyenne de 14.7% de 1990 à 2010. En 2010, il était en moyenne de 11.1% au niveau mondial, avec un minimum de 7.2% en Asie de l'Est, un maximum de 13.6% en Asie du Sud-Est et une moyenne de 8.6% dans les pays développés. La prématurité concerne donc un grand nombre d'enfants qui nécessitent une prise en charge médicale et un suivi particuliers.

Complications neurologiques associées à la prématurité

Lorsqu'un enfant naît prématurément, il ne bénéficie plus de l'environnement utérin adéquat dans lequel il se développait. Ses organes, qui n'avaient pas atteint un stade de développement leur permettant d'assurer complètement leurs fonctions, doivent alors terminer leur croissance et jouer leurs rôles de manière autonome. Cela peut entraîner des complications, notamment au niveau du système nerveux central, du système pulmonaire, du système cardiovasculaire et du système digestif.

Cette thèse se focalise sur les anomalies qui touchent le système nerveux central et plus particulièrement l'encéphale. Ce dernier est très fragile chez les prématurés car il termine son développement complexe autour de 35 SA. Des lésions ou des retards de croissance de celui-ci peuvent entraîner des handicaps moteurs tels que la paralysie cérébrale, la surdité ou la cécité ainsi que des déficiences intellectuelles au niveau social, verbal ou analytique. Plusieurs travaux de recherche ont étudié la prévalence de ces complications (Serenius *et al.* (2013), Arnaud *et al.* (2007), Delobel-Ayoub *et al.* (2006)), et d'autres sont en cours pour les corroborer et les affiner. En particulier, l'étude EPI-PAGE 2 (Ancel *et Goffinet* (2014)) a pour objectif de suivre le neuro-développement à 2, 5, 8 et 12 ans de plusieurs milliers d'enfants nés prématurés (entre 22 à 34 SA) en France en 2011. Des premiers résultats, à 2 ans (Pierrat *et al.* (2017)), sont donnés dans la Table.I.2.

TABLE I.2 – Prévalence des handicaps moteurs et des déficiences intellectuelles chez le prématuré.

SA	Handicap moteur	Déficiência intellectuelle
<32	10 %	15 %
32-33	4 %	10 %
34-36	0.5 %	5 %

Ces études, en particulier cette dernière, confirment que des complications neurodéveloppementales touchent un grand nombre d'enfants et que la prévalence de celles-ci augmentent avec le niveau de prématurité. La prise en charge clinique des nouveaux nés, au niveau neurologique, a pour objectifs d'estimer et de réduire les risques de ces complications. Cela permet de mettre en place un suivi et une prise en charge adaptés pendant les premières années du développement de l'enfant et ainsi de réduire leurs impacts (Roberts *et al.* (2008)).

I.1.2 Anatomie cérébrale

Cette partie décrit l'anatomie cérébrale de manière simplifiée. Les notions qui y sont présentées seront utiles pour comprendre les causes et les conséquences des anomalies cérébrales susceptibles de toucher les enfants prématurés.

Plans anatomiques de référence

La position anatomique de référence est celle d'un corps humain, debout, dont la tête est orientée de manière à ce que le regard soit horizontal. Pour cela, le plan de Francfort (Figure.I.1.a), qui passe par le bord inférieur de l'orbite et le bord supérieur du méat acoustique externe, est orienté

de sorte à être parallèle au sol. L'anatomie est décrite dans trois plans orthogonaux : le plan axial, le plan coronal et le plan sagittal. Ils sont représentés au niveau de la tête sur la Figure.I.1.b.

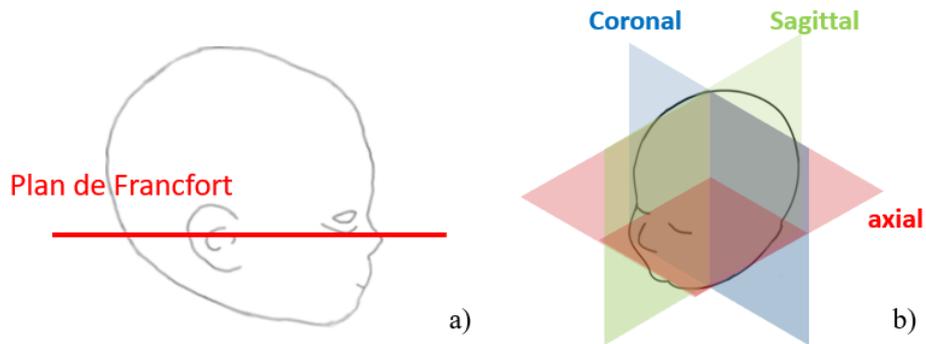
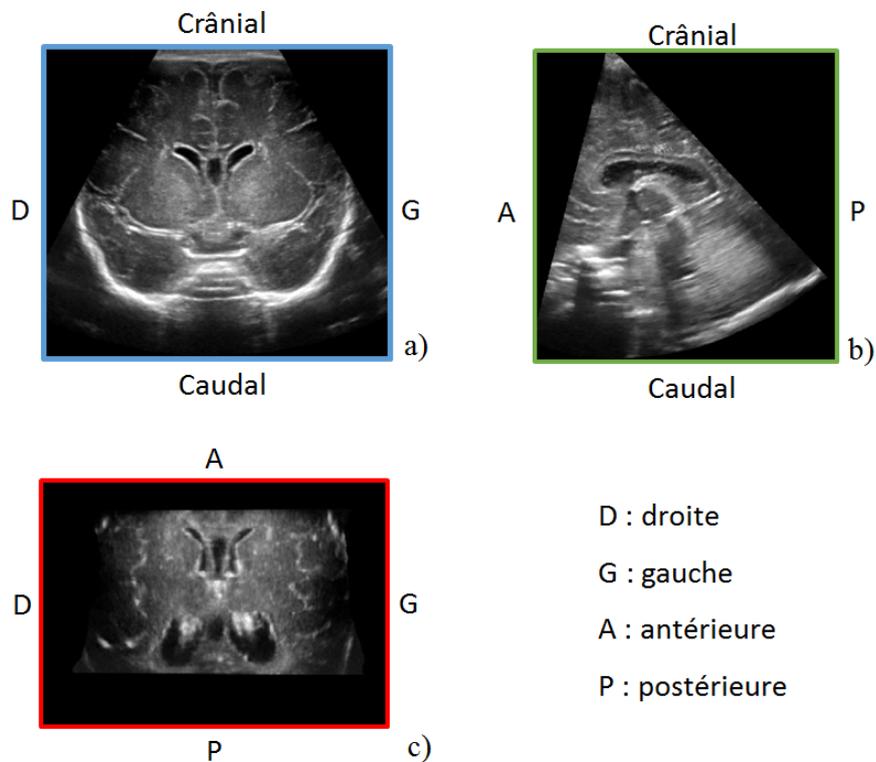


FIGURE I.1 – Plan de Francfort (a) et plans anatomiques de référence (b).

Le plan axial est horizontal, il sépare le corps en une partie haute (crâniale) et une partie basse (caudale), ce plan (Figure.I.2.c) est orienté horizontalement selon l'axe droite-gauche et verticalement selon l'axe antéro-postérieur; le plan coronal est vertical, il divise le corps en une partie avant (antérieure) et une partie arrière (postérieure), (Figure.I.2.a), il est orienté horizontalement selon l'axe droite-gauche et verticalement selon l'axe crânio-caudal; le plan sagittal est vertical, il sépare le corps en une partie droite et une partie gauche, celui-ci (Figure.I.2.b) est orienté horizontalement selon l'axe antéro-postérieur et verticalement selon l'axe crânio-caudal.



D : droite
 G : gauche
 A : antérieure
 P : postérieure

FIGURE I.2 – Images échographiques du cerveau en coupe coronale (a), sagittale (b) et axiale (c)¹.

1. Images issues de notre base de données.

Anatomie de l'encéphale

L'encéphale est la partie du système nerveux central qui est située dans la boîte crânienne. Son anatomie est décrite dans [Netter \(2019\)](#), il se compose de plusieurs structures, quelques unes sont visibles sur la [Figure.I.3](#).

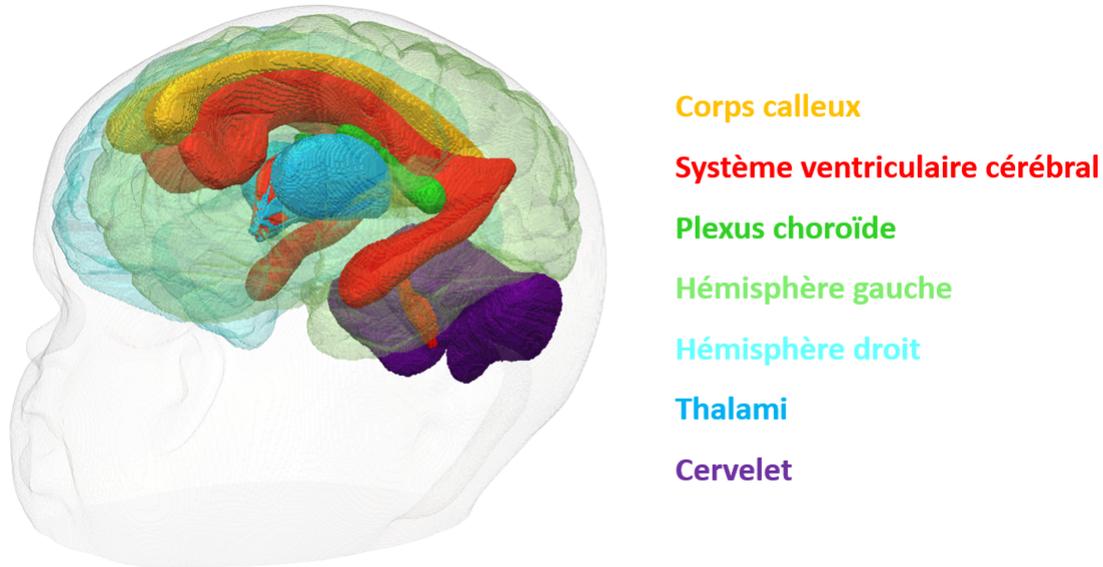


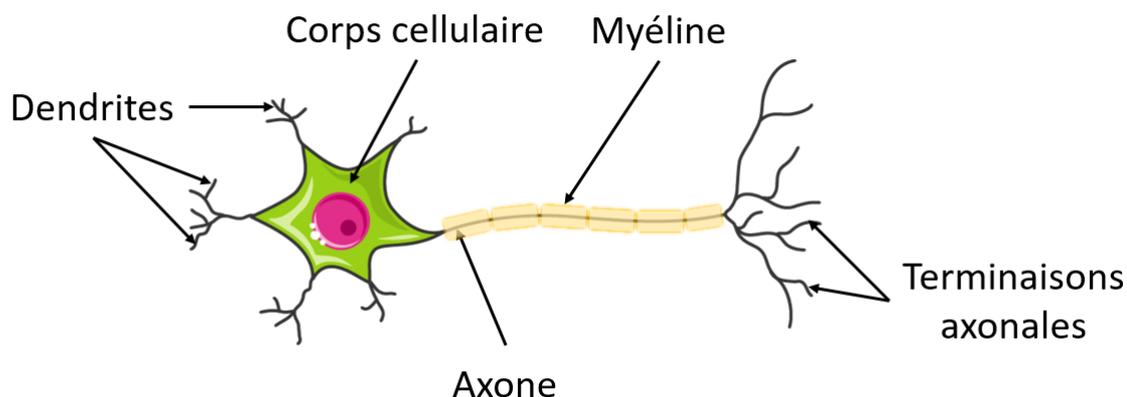
FIGURE I.3 – Anatomie simplifiée de l'encéphale².

Les hémisphères cérébraux droit et gauche sont composés de la substance blanche (SB) et de la substance grise (SG) périphérique. Le corps calleux (CC), qui est situé entre les deux hémisphères, permet à ces deux structures de communiquer entre elles. Au centre du cerveau, on trouve le système ventriculaire cérébral (SVC), qui est un ensemble de cavité contenant du Liquide CérébroSpinal (LCS). Le LCS est produit par les plexus choroïdes qui se situent dans le SVC, c'est le liquide dans lequel baigne le système nerveux central, il a pour rôle de le protéger. Cette figure montre également les thalami et le cervelet qui sont respectivement situés au centre et en dessous de l'arrière du cerveau. Ces structures jouent respectivement un rôle dans la coordination des informations sensorielles et le contrôle des mouvements.

Anatomie du neurone

L'encéphale est majoritairement composé de cellules appelées neurones ([Figure.I.4](#)). Ces derniers permettent de traiter les informations sensorielles (ouïe, vue, toucher, ...), de commander les mouvements du corps et d'acquérir des mécanismes complexes tels que le langage, les comportements sociaux, le raisonnement, ...

2. Image obtenue à partir d'un modèle 3D acheté à la société True Phantom basée à Windsor (Canada)

FIGURE I.4 – Anatomie d'un neurone³.

Les neurones reçoivent (via leurs dendrites), traitent (dans leur corps cellulaire) et transmettent (via leur axone) des informations sous forme d'influx nerveux. Les corps cellulaires des neurones composent la SG. Au niveau de l'encéphale, celle-ci se trouve en périphérie (cortex ou SG périphérique) et dans les noyaux gris centraux (thalami, hypothalamus, ...). La SB se situe entre ces deux régions et assure la connexion entre la SG périphérique et la SG centrale. Elle est constituée des axones, la propagation de l'influx nerveux par ces derniers est accélérée par la présence de myéline autour de ces fibres nerveuses. Lorsque des anomalies cérébrales empêchent ou perturbent l'échange d'informations entre les neurones, cela peut entraîner des troubles du neuro-développement.

I.1.3 Imagerie cérébrale chez le prématuré

L'imagerie cérébrale permet de diagnostiquer les anomalies cérébrales qui peuvent toucher les enfants prématurés. En pratique, deux modalités sont utilisées pour l'imagerie du cerveau : l'échographie et l'Imagerie par Résonance Magnétique (IRM).

Échographie transfontanellaire

L'échographie ou imagerie ultrasonore (US) est réalisée avec un échographe et une sonde échographique (Figure.I.5.a). Cette méthode d'imagerie repose sur l'envoi d'ondes ultrasonores, par la sonde échographique, dans le corps du patient. Ces ondes se propagent dans le corps et sont réfléchies lorsqu'elles traversent des milieux qui ont des impédances acoustiques différentes (frontière entre deux organes ou tissus, ...). Ces ondes réfléchies, ou échos, sont captées par la sonde et traitées par l'échographe pour reconstruire une image de la région dans laquelle les ultrasons se sont propagés. Chez le prématuré, l'échographie cérébrale est réalisée par la fontanelle antérieure (Figure.I.5.b) car cette zone laisse passer les ultrasons, contrairement aux os du crâne. Cet examen d'imagerie est appelé échographie transfontanellaire (ETF).

La mobilité de l'échographe est l'un des gros avantages de cette modalité d'imagerie, elle permet de réaliser l'examen au chevet des patients qui sont extrêmement fragiles et donc difficilement transportables. De plus, en comparaison à l'IRM, elle est peu coûteuse, accessible et permet de reproduire les examens autant de fois que nécessaire.

3. Image originale : <https://smart.servier.com/>

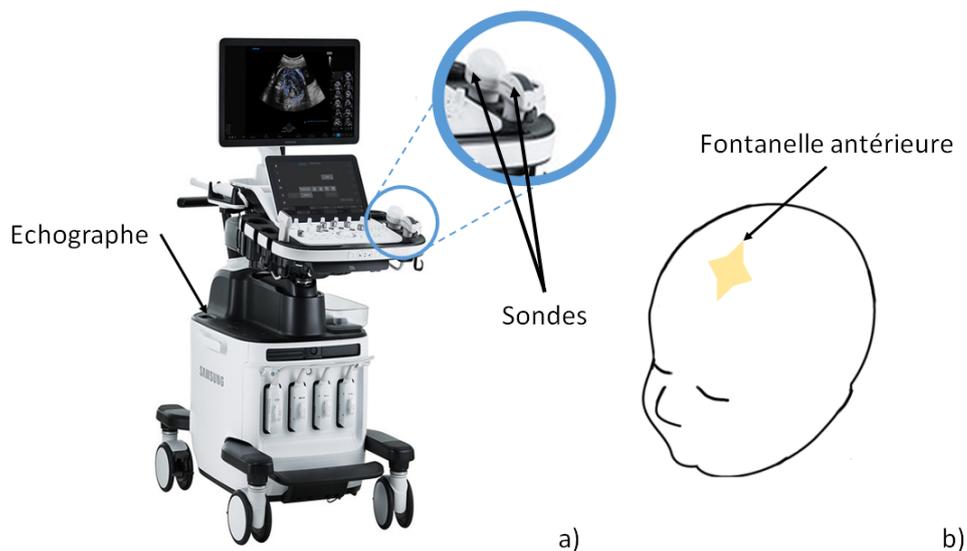


FIGURE I.5 – Échographe équipé de plusieurs sondes (a)⁴ et position de la fontanelle antérieure (b).

Elle a néanmoins plusieurs inconvénients :

1. La zone et le champ d'acquisition sont restreints par l'anatomie de la fontanelle, ce qui ne permet pas d'obtenir des images de l'intégralité de l'encéphale.
2. Les images échographiques sont acquises dans un repère lié à la sonde dont les plans ne correspondent pas aux plans anatomiques de référence, il est donc difficile de localiser les structures visualisées.
3. Le contraste entre les structures cérébrales peut être faible, ce qui peut rendre leurs distinctions difficiles.
4. La résolution et la luminosité varient dans les images en fonction de la profondeur.

Un exemple d'image 2D en coupe proche coronale est visible sur la Figure.I.6.



FIGURE I.6 – Image échographique 2D en coupe coronale provenant d'une ETE.

4. Image originale : <https://www.flickr.com/photos/samsungtomorrow/31608993508>

Cet examen est réalisé systématiquement en routine clinique. En fonction du centre de néonatalogie dans lequel séjourne un enfant prématuré, deux protocoles standards peuvent être utilisés : réalisation de deux ETFs ou réalisation d'une ETF toutes les semaines jusqu'à la sortie du centre. Lorsqu'une pathologie est détectée lors de la première ETF, des examens plus réguliers sont réalisés.

Imagerie par résonance magnétique

L'IRM (Figure.I.7) repose sur la magnétisation du corps humain. Nos tissus sont majoritairement composés de molécules d'eau qui forment chacune un dipôle électrique. Sous l'effet d'un champ magnétique, il est possible d'aligner ces dipôles dans une même direction, qui est une position d'équilibre stable. L'ajout d'un second champ magnétique permet de les faire entrer en résonance et donc de les éloigner de cette position. Lorsque ce dernier champ est coupé, les dipôles vont progressivement retrouver leur position d'équilibre. La mesure des temps de retour à l'équilibre (T1 et T2), qui dépendent de la concentration en eau des tissus, permet d'imager les organes du corps et notamment le cerveau. L'examen est réalisé en plaçant l'enfant dans l'anneau de l'IRM.



FIGURE I.7 – Appareil d'IRM⁵.

Cette modalité d'imagerie a l'avantage de permettre d'acquérir des images de la totalité de l'encéphale dans les plans anatomiques de référence. Les images obtenues sont de plus bien contrastées, ce qui facilite leur interprétation. L'IRM a néanmoins des inconvénients majeurs pour cette application :

1. L'enfant doit être transporté jusqu'à l'IRM et être immobile durant l'acquisition.
2. L'appareil est très coûteux et par conséquent peu accessible.

Des exemples d'images du cerveau acquises par IRM sont données sur la Figure.I.8 en coupe coronale, sagittale et axiale.

5. Image originale : <https://www.maxpixel.net/Magnetic-Resonance-Imaging-Diagnostics-Mri-Hospital-2813899>



FIGURE I.8 – Images IRM, en pondération T2, du cerveau du prématuré en coupe coronale (a), sagittale (b) et axiale (c).

La réalisation d'une IRM cérébrale chez un enfant prématuré dépend des centres de néonatalogie. La pratique courante est de réaliser une IRM lorsque l'enfant est né à moins de 28 SA ou quand une anomalie a été détectée lors d'une ETF. Une IRM peut aussi être pratiquée en cas d'événement neurologique aigu (convulsion, ...). En France, on estime que le pourcentage d'enfants prématurés qui reçoivent une IRM est de seulement 6% (15% des prématurés de moins de 33 SA) alors qu'ils bénéficient tous d'au moins deux ETFs pendant leur hospitalisation.

I.1.4 Anomalies cérébrales chez le prématuré

Dans la première section de ce chapitre, nous avons vu que les enfants prématurés pouvaient être touchés par des problèmes du neuro-développement. Les causes principales de ces troubles sont les hémorragies intra-ventriculaires (IVH), les dilatations ventriculaires, les anomalies de la SB et les retards de croissance des structures cérébrales. Cette section décrit ces anomalies et la façon dont elles sont dépistées et prises en charge en clinique. La prise en charge des prématurés par l'IRM a récemment fait l'objet d'une thèse de médecine (Fontanarosa (2015)).

Hémorragie intraventriculaire et ventriculomégalie

IVH. Les IVH s'expliquent par l'immaturité du système vasculaire cérébral chez le prématuré. Cette immaturité entraîne une fragilité des vaisseaux sanguins qui peuvent se rompre, notamment au niveau de la matrice germinale, qui est adjacente au plancher des ventricules cérébraux. La prise en charge clinique des IVHs est avant tout préventive, elle repose sur la prise d'un traitement neuroprotecteur par la mère de l'enfant quand la grossesse est considérée à risque. Elles sont classées, par gravité, en 4 stades (Papile *et al.* (1978)). Dans les stades II et III, l'hémorragie s'étend au SVC et provoque sa dilatation en empêchant le LCS de s'écouler normalement en dehors de celui-ci (Burstein *et al.* (1979)).

Dilatation ventriculaire. Cette complication est appelée ventriculomégalie (VM), elle apparaît dans environ 30 % des IVH. Cependant, la VM n'est pas exclusivement dû aux IVH et peut aussi être causée par des anomalies de la SB (Leviton *et Gilles* (1996)). Dans ce cas, la dilatation vient combler le volume de SB déficitaire. Cette pathologie peut également apparaître de manière spontanée sans

être associée à d'autres anomalies. Un exemple de SVC normal, dilaté et très dilaté, consécutivement à une IVH, est visible sur la Figure.I.9.

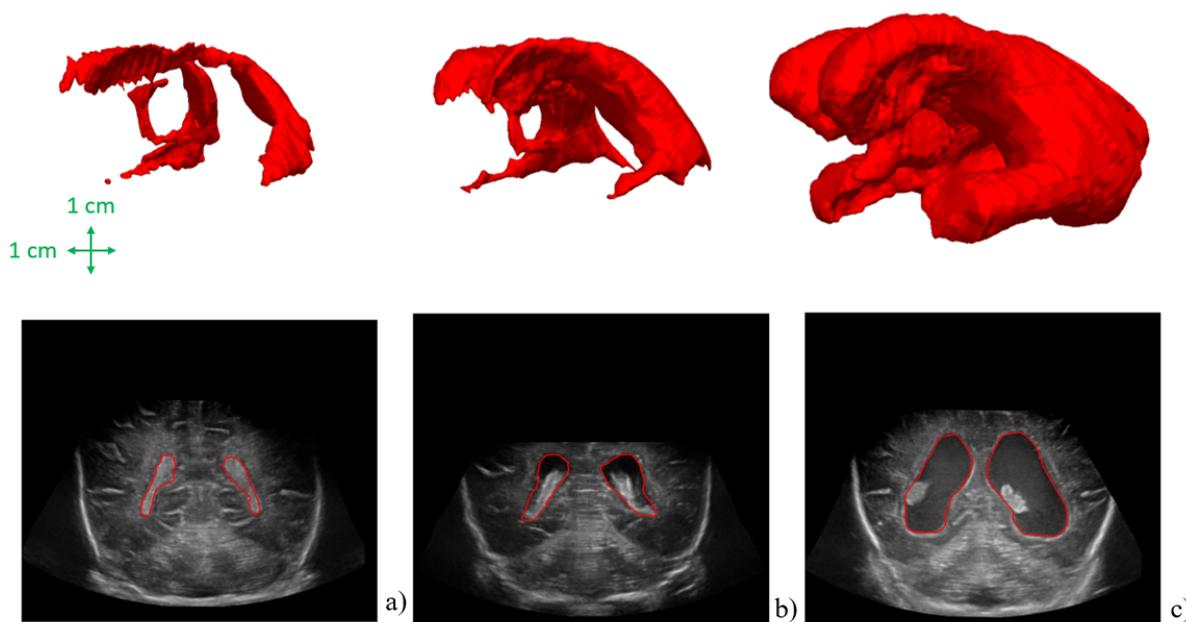


FIGURE I.9 – Segmentation 3D manuelle et images échographiques associées en coupe coronale pour des SVC normaux (a), dilatés (b) et très dilatés (c). Les contours du SVC sont en rouges sur les image US.

Quelle que soit l'étiologie de la VM ([Ment *et al.* \(1999\)](#)), il est primordial de la diagnostiquer et de suivre son évolution. En routine clinique, les cliniciens utilisent l'ETF 2D pour y parvenir. La prise en charge des VMs est passée en revue dans [Mazzola *et al.* \(2014\)](#).

Prise en charge des VMs causées par une IVH. Lorsqu'une dilatation ventriculaire est associée à une IVH, l'évolution du volume ventriculaire (VV) est suivi régulièrement. Dans le cas où la dilatation ne se résorbe pas d'elle même, elle peut devenir très importante et augmenter la pression intra-crânienne en comprimant le cerveau contre la boîte crânienne qui s'élargit alors, on parle dans ce cas d'hydrocéphalie. Cet état peut, à court terme, provoquer des convulsions et entraîner la mort de l'enfant dans les cas les plus critiques.

En pratique, les médecins estiment le VV à partir d'images échographiques 2D en réalisant des mesures manuelles telles que : l'index ventriculaire, la largeur de la corne frontale ou la distance thalamo-occipital ([Brouwer *et al.* \(2012\)](#)). Ces mesures sont utilisées pour suivre l'évolution du VV, dans le cas où la dilatation devient trop importante, elles permettent de prendre la décision d'intervenir chirurgicalement. L'opération consiste à mettre en place une dérivation ventriculaire externe pour drainer le LCS en dehors du SVC.

Prise en charge des VMs non causées par une IVH. Lorsqu'une dilatation ventriculaire n'est pas associée à une IVH, le VV est aussi régulièrement suivi et les cliniciens cherchent à déterminer si la dilatation est associée à une perte de SB. Concernant ce dernier point, un indicateur tel que l'index cérébro-ventriculaire est généralement utilisé pour approximer le rapport entre le VV et le

volume cérébral. Si la VM est uniquement associée à une perte de SB alors il n'y a pas de risque d'hydrocéphalie. Si nécessaire, les cliniciens interviennent chirurgicalement en faisant une ponction lombaire ou une dérivation ventriculaire interne.

Conséquences à long terme de la VM. L'objectif principal de la prise en charge décrite dans les deux précédents paragraphes est de réduire les risques, pour l'enfant, en période néonatale. Par ailleurs, plusieurs études menées en échographie (Ancel *et al.* (2006), Fox *et al.* (2014), Larroque *et al.* (2003), Pappas *et al.* (2018)) et en IRM (Melhem *et al.* (2000), Maunu *et al.* (2011)) ont montré que la présence d'une VM était corrélée à de futurs troubles du neuro-développement. Les mesures réalisées peuvent donc également servir comme paramètres pronostiques du neuro-développement de l'enfant.

Anomalies de la substance blanche et retard de croissance cérébrale

Les anomalies de la substance blanche regroupent les leucomalacies périventriculaires (PVL), échogènes ou cavitaires (forme la plus grave), les lésions punctiformes de la substance blanche (PWML) et les atteintes globales destructrices des substances blanche et grise. De la même manière que pour les IVHs, la prise en charge clinique de ces pathologies est avant tout préventive et repose sur l'administration de médicaments. Ces anomalies touchent principalement les grands et très grands prématurés. Elles sont fortement corrélées à l'apparition de troubles du neuro-développement (Beaino *et al.* (2010), Spittle *et al.* (2011), Jeon *et al.* (2012), Tusor *et al.* (2017)). Leurs conséquences sont d'autant plus graves que la taille des lésions est grande et que les lésions sont situées dans des zones denses en fibres nerveuses. En clinique, le diagnostic de ces anomalies est actuellement réalisé, par IRM et de manière qualitative, chez les prématurés de moins de 28 SA. Il permet d'établir un pronostic sur la probabilité d'apparition de troubles du neuro-développement.

Le retard de croissance des structures cérébrales, qui ont alors un volume plus faible que la normale, a également été corrélé à des retards du neuro-développement chez l'enfant à 2 ans : Lind *et al.* (2011), Cheong *et al.* (2016), Keunen *et al.* (2016). Par exemple, les volumes de la SB et des thalami sont de bons indicateurs pronostics du neuro-développement. Un retard de croissance des thalami est par exemple corrélé à la présence de lésions de la substance blanche (Zubiaurre-Elorza *et al.* (2012)). En clinique, l'évaluation de la croissance des structures cérébrales est réalisée par IRM de manière qualitative ou semi-quantitative. Des mesures quantitatives sont aujourd'hui réalisées uniquement dans des centres de recherche, mais ont démontré leur intérêt dans la prise en charge de ces enfants.

I.1.5 Limites et pistes d'amélioration du diagnostic des anomalies cérébrales par l'imagerie chez l'enfant prématuré

Limite du diagnostic des anomalies cérébrales par l'imagerie

La prise en charge des anomalies cérébrales chez les enfants prématurés a deux objectifs : réduire les risques que représente la ventriculomégalie en période néonatale et extraire des paramètres quantitatifs permettant de pronostiquer les troubles du neuro-développement. Ce pronostic permet de cibler les enfants qui auront besoin d'un suivi multidisciplinaire précoce adapté et ainsi de réduire les risques de handicaps.

Ventriculomégalie. L'échographie 2D permet d'estimer le VV mais pas de le calculer directement. Ces mesures manuelles demandent du temps aux cliniciens et il existe une variabilité intra et inter opérateurs qui peut rendre le suivi de l'évolution de la VM imprécis. De plus, [Kishimoto *et al.* \(2016\)](#) ont récemment montré que les mesures réalisées en échographie 2D pour estimer le VV étaient faiblement corrélées à celle obtenues par échographie 3D ([Kishimoto *et al.* \(2016\)](#)), alors que ces dernières sont bien corrélées aux mesures réalisées par IRM ([Boucher *et al.* \(2018b\)](#)), qui font office de référence. Il serait donc intéressant d'utiliser une méthode d'imagerie 3D pour cette application et de développer des outils pour mesurer automatiquement le VV.

Anomalie de la substance blanche et retard de croissance des structures cérébrales. Nous avons vu que l'IRM était utilisé pour diagnostiquer les anomalies de la SB et évaluer le retard de croissance des structures cérébrales. Les analyses réalisées par IRM sont fiables, néanmoins, par manque d'accessibilité, elles ne sont réalisées en routine clinique que chez des prématurés de moins de 28 SA. Or, il a été récemment montré qu'un nombre important de prématurés moyens étaient susceptibles de présenter les mêmes troubles du neuro-développement ([Allen \(2008\)](#)), [Arpino *et al.* \(2010\)](#)). Dans ce cas aussi, une méthode d'imagerie 3D accessible permettrait d'améliorer significativement la prise en charge des prématurés.

Pistes d'améliorations du diagnostic des anomalies cérébrales par l'imagerie

Par manque d'accessibilité, l'imagerie IRM ne peut pas être réalisée chez tous les prématurés pour donner un pronostic des troubles du neuro-développement. L'imagerie échographique 3D, de part son accessibilité en dehors des Centres Hospitaliers Universitaires et son coût réduit, serait une alternative réaliste dans les conditions cliniques actuelles qui améliorerait significativement la prise en charge en période néonatale et infantile des enfants prématurés. Elle pourrait constituer une alternative à l'IRM utilisable en routine clinique chez tous les prématurés. La forme la plus connue d'imagerie 3D est réalisée avec une sonde matricielle. Malheureusement, à notre connaissance, la qualité des images 3D obtenues avec le matériel actuellement disponible sur le marché n'est pas assez satisfaisante ([Kurian *et al.* \(2017\)](#)) pour extraire tous les paramètres pronostics dont nous avons parlé. De même, la qualité des images échographiques 2D ne permet généralement pas de distinguer toutes les structures cérébrales et donc d'extraire tous les paramètres pronostics qui nous intéressent.

En routine clinique, le centre de néonatalogie d'Avignon, qui est partenaire de cette thèse, utilise une sonde linéaire Siemens pseudo-matricielle 4-9 MHz (sonde 9L4). Cette sonde est adaptée à l'échographie fœtale, elle n'est pas adressée en première intention aux pédiatres pour l'ETF; ce qui explique qu'elle soit méconnue pour cet usage. Elle permet en pratique d'acquérir des images de très haute résolution. A titre de comparaison, des images obtenues avec une sonde néonatale 2D (Siemens 10V4) et la sonde 9L4 sont données sur la Figure.I.10.

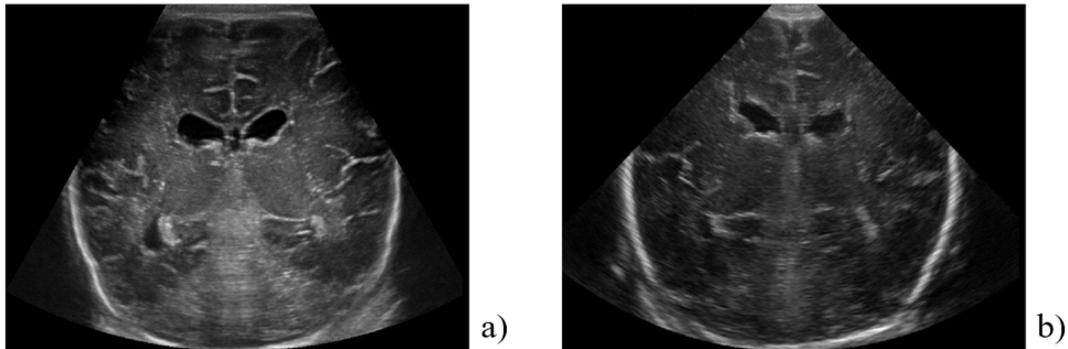


FIGURE I.10 – Comparaison d'images échographiques acquises avec une sonde 9L4 (a) et une sonde 10V4 (b).

On peut voir que les images acquises avec la sonde 9L4 sont bien plus contrastées et détaillées que celles obtenues avec la sonde 10V4. La plupart des structures cérébrales peuvent être distinguées et il semble que les PWML (lorsqu'il y en a) soient également visibles, comme cela peut être observé sur la Figure.I.11.

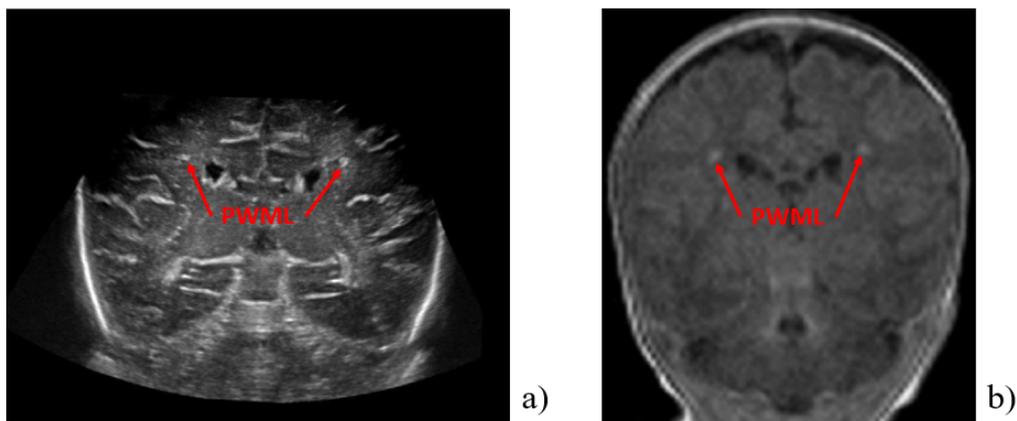


FIGURE I.11 – PWMLs visibles dans des volumes 3D US (a) et IRM en pondération T2 (b).

Les images US 2D obtenues avec la sonde 9L4 sont de meilleure qualité que celles acquises avec une sonde matricielle. Même si il n'est pas possible d'obtenir des images 3D du cerveau avec la sonde 9L4, on peut acquérir des images de la majeure partie de l'encéphale en effectuant un balayage manuel lors de l'acquisition. Cela donne du crédit au développement d'une échographie 3D quantitative dédiée à l'imagerie cérébrale chez le prématuré.

I.1.6 Conclusion

Dans cette première partie, nous avons vu que les méthodes d'imageries utilisées actuellement pour diagnostiquer les anomalies cérébrales chez le prématuré étaient limitées. D'un côté, l'IRM permet de réaliser toutes les mesures quantitatives utiles pour prendre en charge les prématurés, mais elle ne peut être pratiquée que sur 15 % des enfants prématurés nés à moins de 33 SA. De l'autre côté, l'échographie transfontanellaire 2D est réalisée chez tous les prématurés mais ne permet pas d'obtenir toutes les mesures quantitatives qui nous intéressent et de manière précise.

Pour pallier ces limitations, l'échographie 3D présente un fort potentiel. Sous réserve que les images qu'elle permet d'obtenir soient d'une qualité suffisante, elle permettrait de diagnostiquer les anomalies cérébrales et d'obtenir des mesures quantitatives à visée pronostic chez tous les prématurés.

L'objectif de cette thèse est de montrer qu'il est possible de quantifier automatiquement le volume des structures cérébrales dans des volumes US 3D, dans un temps court, compatible avec l'usage clinique quotidien, et ainsi d'ouvrir la voie à une échographie 3D quantitative dédiée aux prématurés. Elle se focalise dans un premier temps sur l'obtention de volume 3D de haute qualité à partir des acquisitions réalisées avec la sonde 9L4 puis, à titre de faisabilité, sur la segmentation du SVC et des thalami.

I.2 Problématique scientifique associée

Résoudre la problématique médicale définie dans la section précédente peut se résumer à : "Développer une échographie 3D quantitative permettant de dépister et suivre les anomalies cérébrales chez l'enfant né prématuré". Par rapport à cette problématique, l'objectif de cette thèse est la reconstitution de volumes US 3D fiables et le développement d'algorithmes qui permettent de réaliser automatiquement des mesures quantitatives dans un temps clinique. Cette thèse s'intéresse en particulier à la mesure des volumes ventriculaire et thalamique dont l'intérêt clinique a été montré dans la partie précédente. D'un point de vue technique, montrer cette faisabilité nécessite de résoudre deux problématiques :

1. Obtenir des volumes 3D de haute qualité à partir des acquisitions réalisées avec la sonde 9L4 utilisée par le centre de néonatalogie du Centre Hospitalier (CH) d'Avignon.
2. Développer des algorithmes permettant de quantifier les volumes ventriculaire et thalamique dans un temps clinique, ce qui nécessite d'attribuer une étiquette à plusieurs millions de voxels en quelques secondes.

Les deux prochaines parties de cette section s'attachent à modéliser les problèmes à résoudre pour répondre à ces deux derniers points.

I.2.1 Reconstruction 3D d'une séquence d'images 2D

Le centre de néonatalogie du CH d'Avignon réalise les ETFs avec une sonde qui permet d'obtenir des images 2D de très haute qualité. Dans la partie précédente, nous avons vu que les mesures quantitatives réalisées sur des images 2D pouvaient être faiblement corrélées aux mesures quantitatives faites sur des images 3D. Cette dernière serait donc à privilégier, néanmoins, la qualité des images qu'elle fournit est aujourd'hui insuffisante pour distinguer les différentes structures cérébrales. Une solution pour obtenir des volumes de haute qualité serait de reconstruire en 3D les séquences d'images 2D réalisées avec la sonde 9L4.

Acquisition d'une séquence d'images 2D

Pour acquérir une séquence d'images 2D de la plus grande partie possible du cerveau, le pédiatre partenaire de ce projet réalise un balayage angulaire manuel par la fontanelle antérieure (Figure.I.12.a). Ce type de balayage permet d'obtenir une séquence de T images, appelée Séquence Dynamique (SD). Elle contient des images I_1, I_2, \dots, I_T , proches d'une coupe coronale (Figure.I.12.b), acquises aux temps $t = 1, t = 2, \dots, t = T$.

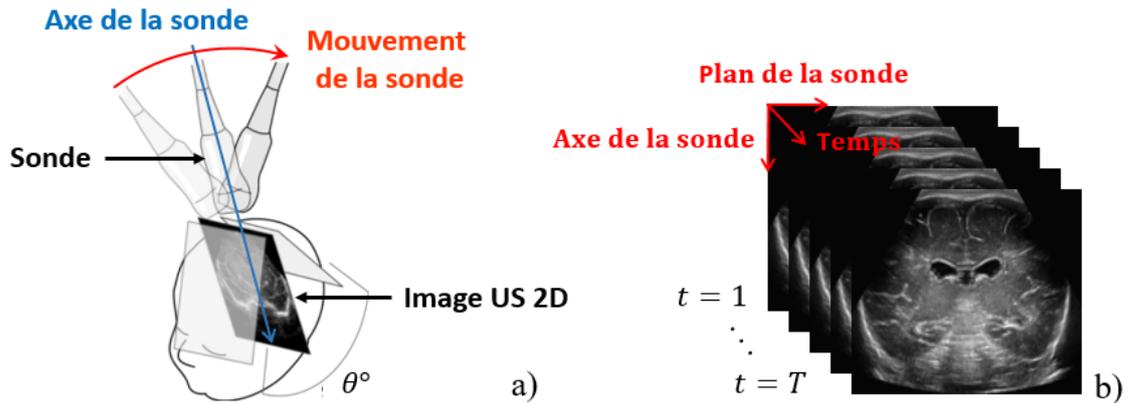


FIGURE I.12 – Balayage angulaire manuel d'angle θ réalisé par la fontanelle antérieure (a) et séquence dynamique obtenue après la réalisation d'un balayage manuel (b).

Comme cela a été mentionné dans la partie précédente, un des inconvénients de l'échographie est que les images sont acquises dans un repère lié à la sonde. Lors de l'acquisition, toutes les images sont donc acquises dans des repères différents. Le repère temps, axe de la sonde ne forme donc pas un espace de coordonnées cartésiennes, en particulier dans la direction de l'empilement (coupe sagittale visible sur Figure.I.13).

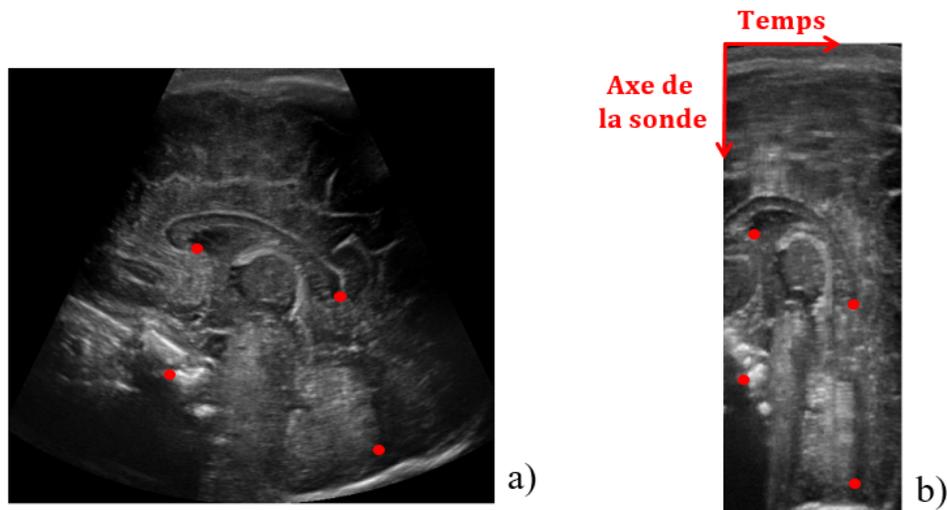


FIGURE I.13 – Image acquise en coupe sagittale avec la sonde 9L4 (a) et coupe sagittale de l'empilement des images d'une séquence dynamique (b). Les mêmes repères anatomiques (points rouges) ont été identifiés sur les deux images.

On observe une différence importante entre la position relative des points placés sur la coupe sagittale de référence (Figure.I.13.a) et sur la coupe sagittale de l'empilement d'images de la SD (Figure.I.13.b). Cela confirme que le repère temps, axe de la sonde ne correspond pas à un repère cartésien.

Repositionnement des images les unes par rapport aux autres

Une image se modélise par une matrice dans le cas 2D ou un tenseur dans le cas 3D, leurs éléments sont respectivement appelés pixel ou voxel. Dans notre cas, chaque pixel/voxel est associé à un niveau de gris qui correspond à une valeur entière comprise entre 0 et 255. Pour reconstruire un volume 3D qui soit spatialement cohérent dans toutes les directions à partir d'une SD, il faut déterminer la position relative des images de la SD les unes par rapport aux autres et les replacer dans une grille de reconstruction cartésienne. Ces positions peuvent être obtenues en déterminant la position relative de chaque image par rapport à la suivante, soit de I_t par rapport à I_{t+1} . Pour cela on munit chaque image d'un repère (O_t, i_t, j_t, k_t) où O_t est le centre de l'image et (i_t, j_t, k_t) un repère orthonormé (Figure.I.14).

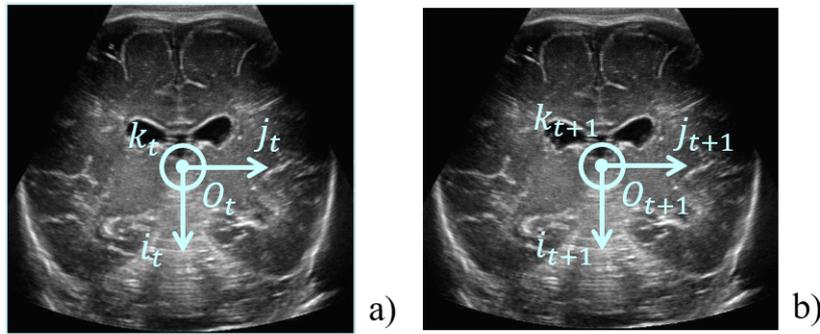


FIGURE I.14 – Association d'un repère à deux images successives : I_t (a) et I_{t+1} (b).

La position relative de I_t par rapport à I_{t+1} est parfaitement définie si on estime 3 paramètres de position, correspondant à la distance $O_t O_{t+1}$ entre les centres des images, et 3 paramètres angulaires. On notera respectivement ces paramètres d_x, d_y, d_z et $\alpha_t, \beta_t, \gamma_t$. Ils sont représentés sur la Figure.I.15.

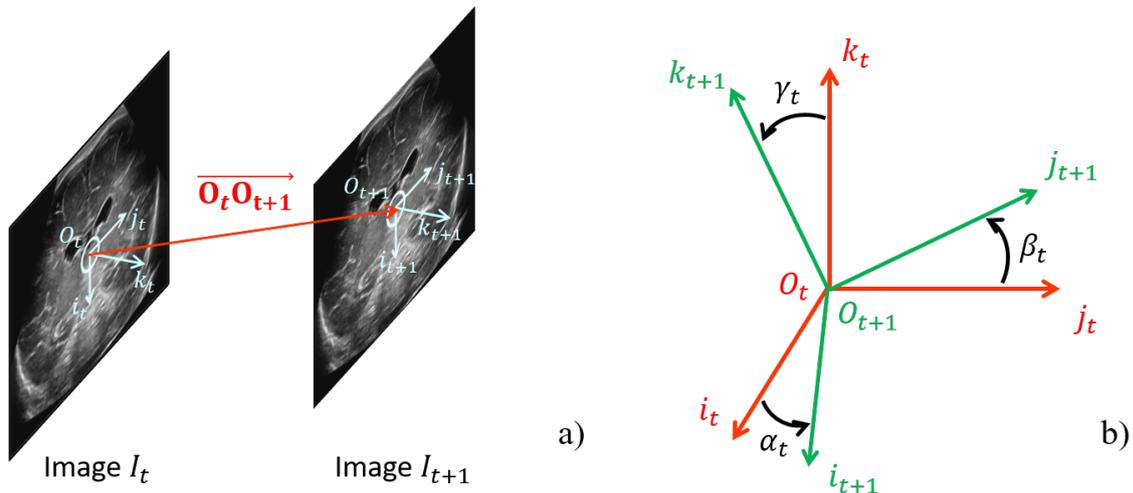


FIGURE I.15 – Distance entre les centres de deux images successives (a) et angles entre deux images successives (b).

Il faut ainsi estimer, $\forall t \in [|1, T - 1|]$:

$$\overrightarrow{O_t O_{t+1}} = (d_{x_t}, d_{y_t}, d_{z_t}) \quad (\text{I.1})$$

et :

$$\begin{aligned} \alpha_t &= \widehat{(i_t, i_{t+1})} \\ \beta_t &= \widehat{(j_t, j_{t+1})} \\ \gamma_t &= \widehat{(k_t, k_{t+1})} \end{aligned} \quad (\text{I.2})$$

Interpolation des volumes reconstruits

Une fois les paramètres définis par les équations (I.1) et (I.2) estimés, les images doivent être placées dans la grille cartésienne de reconstruction. Cette dernière peut être représentée comme une grille 3D donc chaque nœud correspond à un voxel, de coordonnées entières, auquel il faut associer un niveau de gris. Cette étape est délicate car lorsqu'on replace les images dans la grille, leurs pixels ne tombent pas exactement dans les voxels. En effet, les coordonnées des pixels après positionnement dans la grille ne sont en général pas entières. Il est donc nécessaire de déterminer une méthode pour interpoler un niveau de gris dans chaque voxel à partir des valeurs des pixels replacés dans la grille. Ce problème est illustré, en coupe sagittale par soucis de simplification, sur la Figure.I.16 :

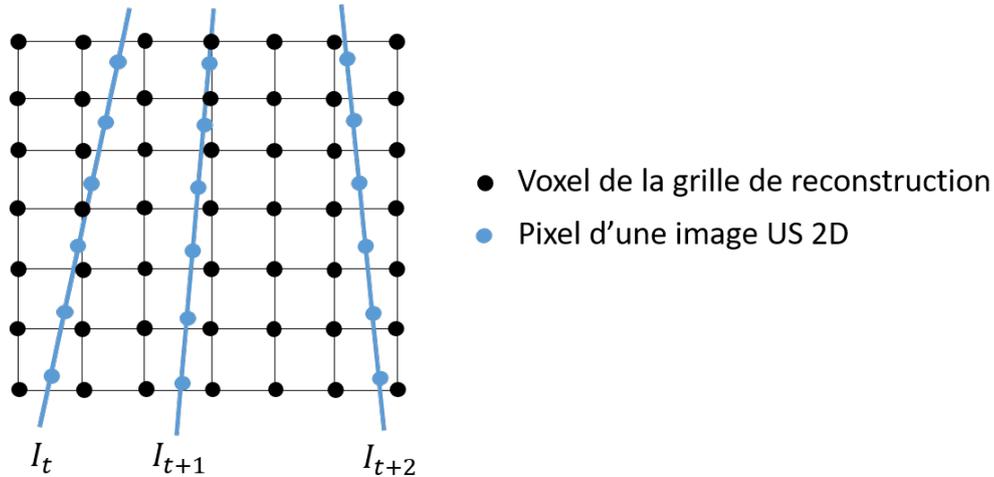


FIGURE I.16 – Grille cartésienne de reconstruction en coupe sagittale et images replacées dans cette grille.

Orientation du volume reconstruit dans les trois plans anatomiques

Une fois le volume interpolé (Figure.I.17.b), il faut orienter celui-ci dans les plans anatomiques de référence (Figure.I.17.a). Cette orientation permet de faciliter l'interprétation des volumes reconstruits et de réaliser des comparaisons avec des volumes obtenus par IRM. Il faut déterminer trois angles pour réaliser cette étape, on appellera ceux-ci φ, ϕ, ψ . Ils sont représentés sur la Figure.I.17.c :

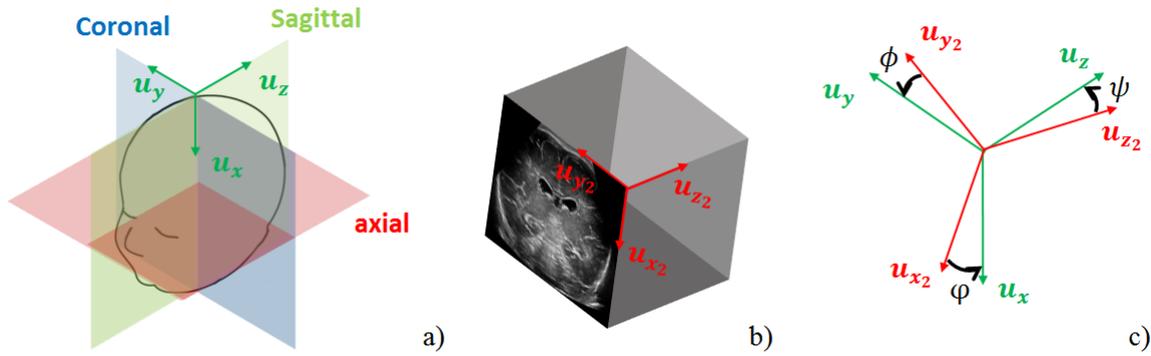


FIGURE I.17 – Plans anatomiques de référence (a), volume 3D interpolé (b) et angles à estimer pour orienter le volume 3D interpolé par rapport aux plans anatomiques de référence (c).

On a :

$$\begin{aligned}
 \varphi &= (\widehat{u_{x2}, u_x}) \\
 \phi &= (\widehat{u_{y2}, u_y}) \\
 \psi &= (\widehat{u_{z2}, u_z})
 \end{aligned}
 \tag{I.3}$$

I.2.2 Segmentation du SVC et des thalami dans des volumes reconstruits en 3D

Définition

En imagerie médicale, l'extraction de paramètres quantitatifs tels que le volume des structures cérébrales se fait à partir de leur segmentation. Segmenter une image consiste à regrouper des pixels entre eux, selon des critères pré-définis qui dépendent du problème que l'on souhaite résoudre. Dans notre cas, l'objectif est de séparer automatiquement les voxels appartenant aux SVC, aux thalami et les autres qui appartiennent à une classe que l'on appellera l'arrière-plan, le tout dans un temps clinique. Un exemple de segmentation manuelle du SVC et des thalami est visible en coupe coronale sur la Figure.I.18.

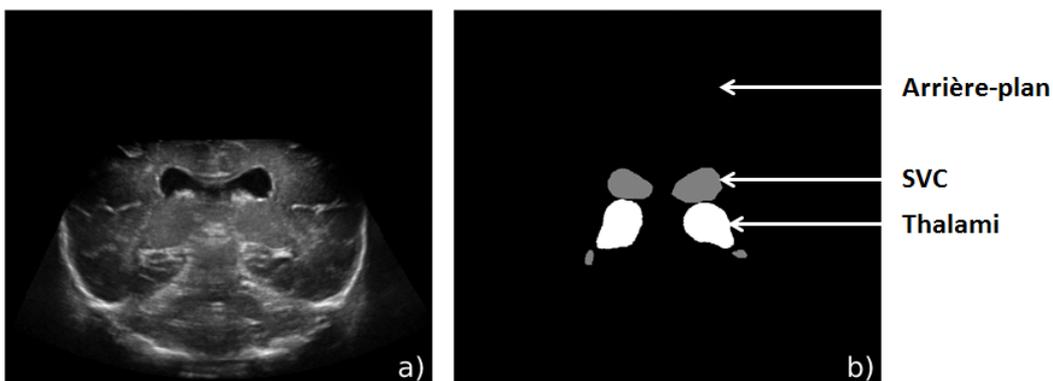


FIGURE I.18 – Image US 2D en coupe coronale (a) et segmentations manuelles de l'arrière-plan, des thalami et du SVC (b).

Il existe de nombreuses méthodes de segmentation dont le choix dépend du problème que l'on souhaite résoudre. Parmi les approches les plus classiques, on peut citer :

- la segmentation par recherche des frontières des structures à segmenter.
- la segmentation par recherche des régions correspondant aux structures à segmenter.
- la segmentation par classification un à un des pixels d'une image en fonction de leur voisinage.

Ces approches peuvent être utilisées de manière isolée ou coopérative, l'objectif final étant que chaque voxel des images 3D reconstruites reçoive une étiquette correspondant aux classes SVC, thalami ou arrière-plan.

Anatomie du SVC et des thalami et aspect dans des données US.

Les volumes anatomiques que nous souhaitons segmenter sont le SVC et le thalami. L'anatomie de ces structures et leur aspect en échographie sont décrits dans cette partie de manière à donner une idée des difficultés liées à ces problèmes de segmentation.

Anatomie et aspect des ventricules latéraux. Le SVC se situe au centre du cerveau (Figure.I.3), son anatomie est décrite dans les plans anatomiques de référence sur la Figure.I.19.

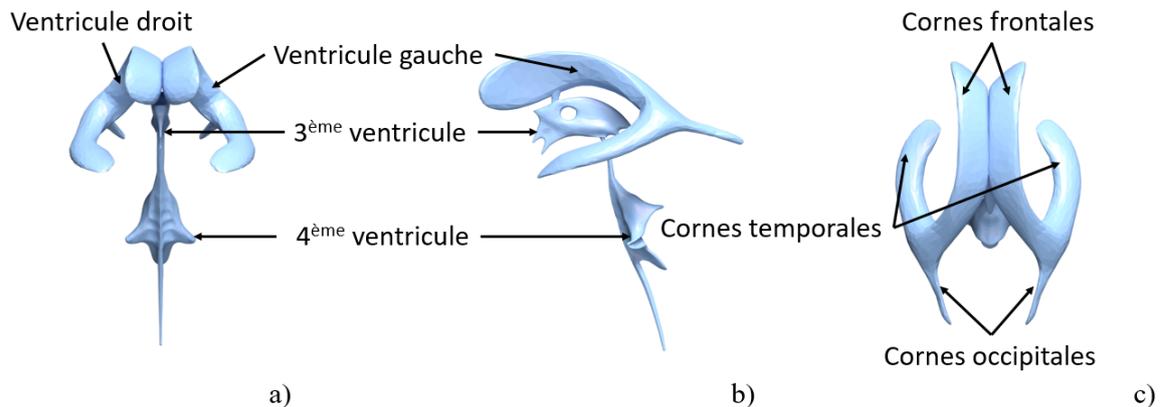


FIGURE I.19 – Anatomie du système ventriculaire cérébral en vue coronale (a), sagittale (b) et axiale (c)⁶.

Il se compose de quatre ventricules : deux ventricules latéraux (VL) qui sont appelés ventricule droit et ventricule gauche, du troisième ventricule et du quatrième ventricule. Les VL ont une forme en fer à cheval, ils sont chacun composés de trois cornes (Figure.I.19.c) : une corne temporale, une corne occipitale et une corne frontale. Cette thèse s'intéresse uniquement à la segmentation des ventricules latéraux et du troisième ventricule car le quatrième ventricule est difficilement observable, même dans des données US de haute qualité, et sa segmentation n'a pas beaucoup d'intérêt clinique à notre connaissance.

6. Images originales : <https://commons.wikimedia.org/wiki/>

L'aspect du SVC dans des volumes 3D US reconstruits peut être vu dans des coupes coronales sur la Figure.I.20.b. La profondeur en z de ces coupes est représentée sur une segmentation manuelle d'un SVC dilaté (Figure.I.19.a).

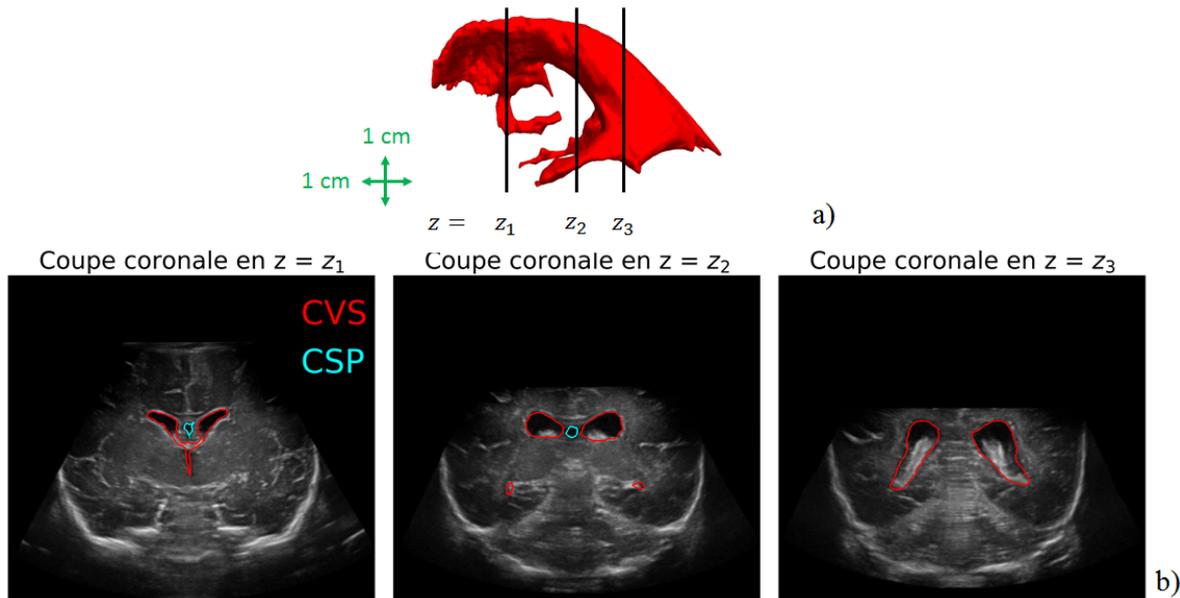


FIGURE I.20 – Vue sagittale d'une segmentation 3D manuelle du SVC (a) et , aspect du SVC et du CSP en coupes coronales, à différentes profondeurs en z , dans des données US 3D reconstruites (b).

On voit sur la Figure.I.20.b que le SVC présente à la fois un aspect hypoéchogène (faible intensité), qui correspond au LCS, et hyperéchogène (forte intensité), qui correspond au plexus choroïde. On observe également qu'une autre structure, le Cavum Septi Pellucidi (CSP), située entre les ventricules latéraux, a également un aspect hypoéchogène. De plus, la Figure.I.20.a permet de se rendre compte que certaines parties du SVC, comme les cornes temporales, sont extrêmement fines.

A partir de cette description, on peut dire que la segmentation du SVC est un problème difficile. Premièrement car la forme de la structure à segmenter est complexe, deuxièmement parce que certaines parties du SVC sont extrêmement fines, troisièmement car l'intensité du SVC est très hétérogène et enfin parce qu'une autre structure (le CSP), d'aspect semblable au SVC, se trouve proche de celui-ci. Néanmoins, le SVC a un fort contraste avec la plupart des structures qui lui sont voisines, ce qui est un avantage pour réaliser sa segmentation.

Anatomie et aspect des thalami. La forme des thalami est moins complexe que celle du SVC. Cette structure se décompose en un thalamus droit et un thalamus gauche, ils ont tous les deux une forme d'œuf et sont connectés au niveau de l'adhérence inter-thalamique. Son anatomie est visible sur la Figure.I.21.

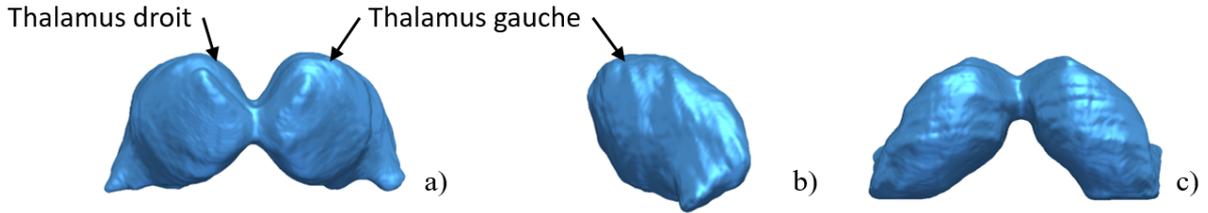


FIGURE I.21 – Anatomie des thalami en vue coronale (a), sagittale (b) et axiale (c)¹.

Comme pour le SVC, l'aspect des thalami dans des données échographiques peut être observé sur la Figure.I.22.b à différentes profondeurs représentées sur la Figure.I.22.a.

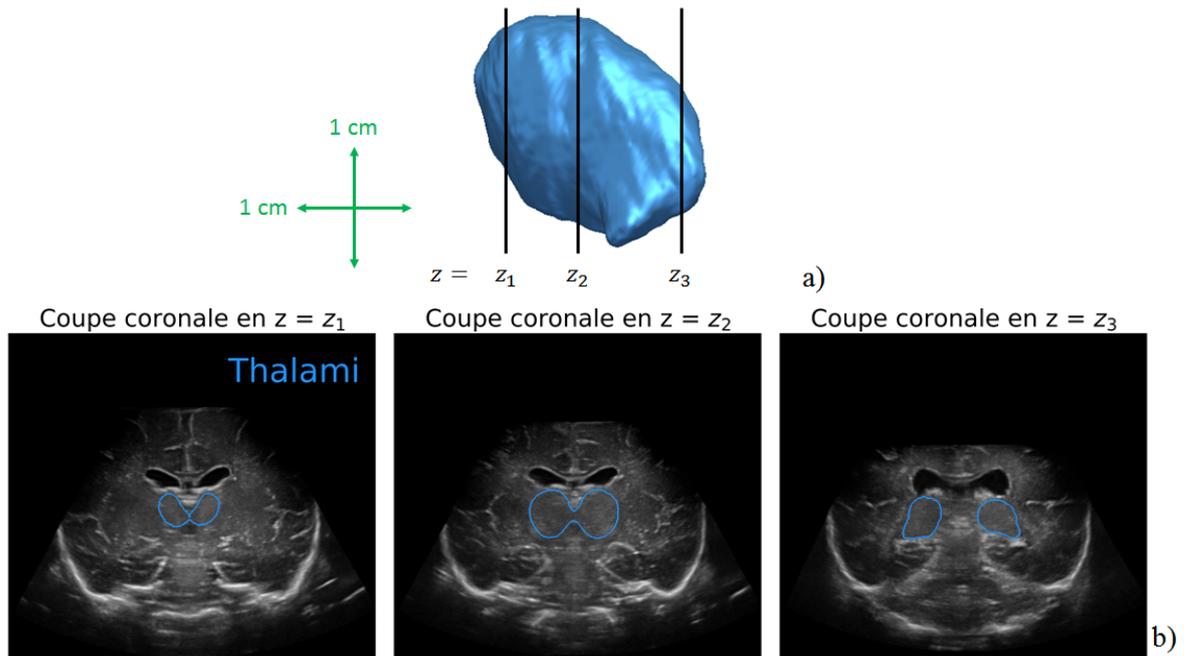


FIGURE I.22 – Vue sagittale d'une segmentation 3D manuelle des thalami (a) et aspect des thalami en coupes coronales, à différentes profondeurs en z , dans des données US 3D reconstruites (b)¹.

On peut voir que les thalami ont un aspect homogène et qu'ils apparaissent avec une intensité moyenne. Néanmoins, cette structure est extrêmement difficile à segmenter car elle a un faible contraste avec ses structures voisines. Parvenir à la segmenter serait très encourageant pour la suite du projet car les thalami sont l'un des problèmes les plus difficiles à résoudre.

Histogramme d'intensité du SVC et des thalami et gradients d'intensité. Ce paragraphe propose un histogramme d'intensité pour le SVC, les thalami et l'arrière-plan ainsi que le calcul de la magnitude du gradient sur une de nos images. Ces données permettent de mieux comprendre pourquoi ce problème de segmentation est complexe et donc pourquoi l'utilisation des méthodes de segmentation classiques, comme celles qui se basent sur un seuil d'intensité ou la magnitude du gradient, ne sont pas envisageables pour le résoudre.

Histogramme d'intensité. Les images avec lesquelles nous travaillons sont en niveau de gris. Chaque pixel est donc décrit par un nombre entier compris entre 0 et 255 qui correspond à une cou-

leur allant du noir (0) au blanc (255), ce nombre définit l'intensité du pixel. En traçant l'histogramme du SVC, des thalami et de l'arrière-plan pour un volume de notre base de données, nous avons obtenu le résultat suivant :

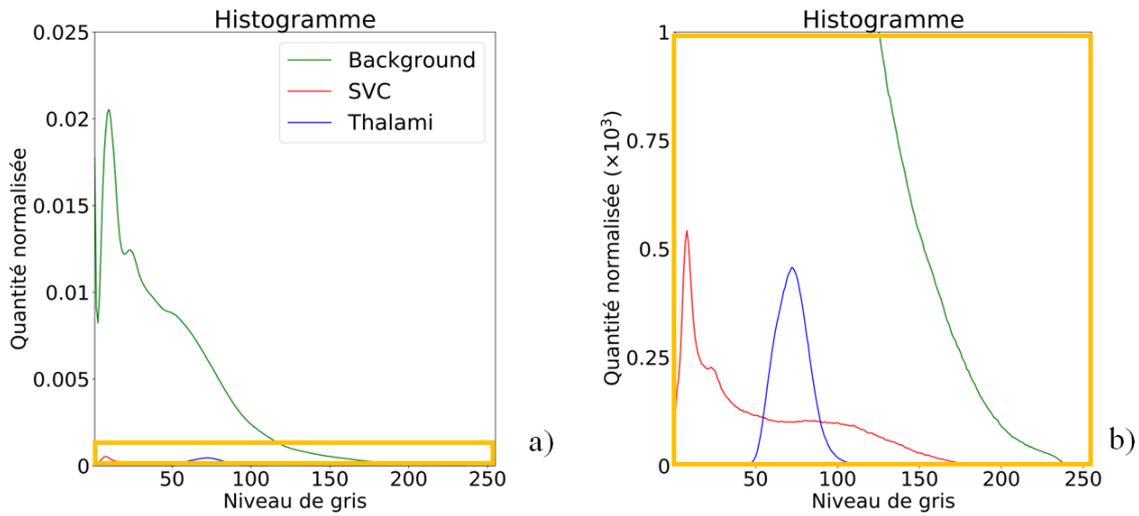


FIGURE I.23 – Histogramme d'intensité pour l'arrière-plan, les thalami et le SVC (a) et zoom sur une partie d'intérêt (b).

On observe que l'histogramme des thalami se superpose avec celui du SVC, ces deux derniers sont de plus noyés dans celui de l'arrière-plan. Partant de ce constat, on peut conclure que des méthodes basées sur des seuils d'intensité sont insuffisantes pour segmenter le SVC et les thalami.

Magnitude du gradient. Le gradient d'intensité correspond aux variations d'intensité entre les pixels voisins d'une image. Cette valeur peut être utilisée pour déterminer les bords d'une structure que l'on souhaite segmenter, elle est bien adaptée aux problèmes où les structures à segmenter contrastent avec leurs structures voisines. Un exemple de gradient d'intensité calculé sur une image US 2D est visible sur la Figure.I.24.

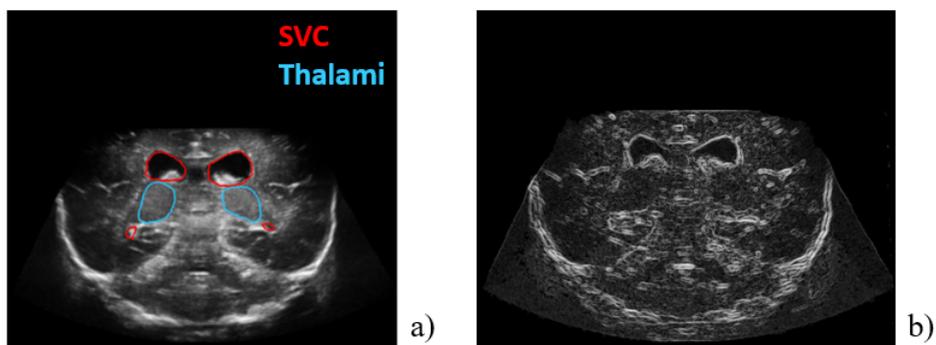


FIGURE I.24 – Image US 2D et magnitude du gradient calculé sur cette image.

On voit sur cette image que les gradients d'intensité entre les thalami et les structures qui les entourent sont extrêmement faibles. Il serait donc difficile de segmenter les thalami en utilisant ces valeurs. Concernant le SVC, il existe un fort gradient entre le LCS et les structures adjacentes, cela

rend la partie hypoéchogène du SVC identifiable. Néanmoins, il existe beaucoup d'autres endroits sur l'image où il y a de forts gradients, il faudrait donc mettre en œuvre une méthode plus complexe pour conserver uniquement les contours qui nous intéressent.

Création d'une base de données annotées

La création d'une base de données annotées est une problématique majeure dans les problèmes de segmentation en imagerie médicale. Cette étape est primordiale pour permettre d'évaluer la précision des algorithmes de segmentation développés et utiliser des méthodes d'apprentissage supervisé qui ont récemment rendu possible la segmentation de structures complexes en quelques secondes. La réalisation de ces annotations est malheureusement extrêmement complexe et coûteuse en temps. A notre connaissance, aucune base de données 3D annotées dédiées à notre problématique n'est aujourd'hui en libre accès, il a donc été nécessaire de construire notre propre base.

Réaliser ces segmentations manuelles nécessitent des connaissances poussées de l'anatomie cérébrale chez le prématuré et de savoir identifier les structures dans des coupes qui ne correspondent pas aux plans anatomiques de référence. Seul un expert médical peut donc réaliser ces segmentations extrêmement difficiles. Pour ce problème, une segmentation manuelle complète de chaque structure nécessite d'annoter au minimum une centaine de coupes par acquisition. Si l'on considère qu'il faut en moyenne 3 minutes pour annoter rigoureusement une structure sur une coupe, cela signifie qu'il faut 600 minutes (pour environ 250 coupes), soit 10h, pour annoter complètement une acquisition. En sachant qu'une base de données d'une taille correcte contient environ 20 volumes, sa construction nécessite au total au moins 200h de travail, sans compter le temps qu'il faut consacrer pour apprendre à identifier précisément les structures sur les images, en particulier les thalami qui n'ont à notre connaissance, jamais été segmentés en échographie.

Il n'est pas envisageable en pratique qu'un médecin puisse se libérer autant de temps en dehors de sa pratique médicale pour réaliser ces annotations. Il est donc impossible d'annoter toutes les coupes où les structures sont visibles dans les acquisitions. Un autre aspect de la problématique de segmentation sera donc de développer une méthode permettant de réduire le nombre de contourage manuel à réaliser pour obtenir notre base de données annotées.

I.3 Conclusion

Cette deuxième partie a défini les problématiques scientifiques et techniques auxquelles il est nécessaire de répondre pour montrer qu'il est possible de résoudre la problématique médicale associée à cette thèse. Il faudra dans un premier temps développer une méthode permettant la reconstruction 3D des balayages manuels 2D réalisés par le médecin pour imager le cerveau du prématuré, dans un deuxième temps il sera nécessaire d'annoter les SVC et les thalami sur ces données, enfin, il faudra développer un algorithme de segmentation du SVC et des thalami. Les deux chapitres suivants seront respectivement consacrés à la bibliographie en rapport avec la problématique de reconstruction 3D et la problématique de segmentation.

CHAPITRE II

ÉTAT DE L'ART : RECONSTRUCTION 3D

Contents

II.1 Estimation de la position relative de deux images	30
II.1.1 Acquisition US main-libre réalisée sans capteur	31
Mesure de la décorrélation du speckle	31
Modélisation du mouvement d'acquisition	33
Reconstruction par apprentissage supervisé	33
II.1.2 Enregistrement du mouvement d'acquisition	34
Suivi de la sonde par capteur électromagnétique	34
Suivi de la sonde par capteur optique	34
Suivi de la sonde par accéléromètre et gyromètre	35
Mécanisation de la sonde utilisée pour l'acquisition	35
II.1.3 Conclusion	36
Objectif à court terme	36
Objectif à long terme	36
II.2 Interpolation dans la grille cartésienne de reconstruction	38
II.2.1 Méthodes basées sur les pixels	38
II.2.2 Méthodes basées sur les voxels	39
II.2.3 Méthodes basées sur les fonctions	41
II.2.4 Comparaison des méthodes et conclusion	42
II.3 Repositionnement dans le repère anatomique	42

CHAPITRE II

DANS le chapitre précédent, nous avons établi qu'il fallait choisir/développer une méthode qui permette de reconstruire un volume 3D à partir d'une séquence dynamique. Ce chapitre présente l'état de l'art des méthodes qui pourraient être utilisées à cette fin.

Dans le cadre de ce projet, deux objectifs sont associés à la problématique de reconstruction :

- un objectif à court terme qui est de reconstruire en 3D les acquisitions 2D réalisées en routine clinique au centre de néonatalogie d'Avignon.
- un objectif à plus long terme qui est de permettre l'acquisition directe de volumes 3D en routine clinique.

Ce premier objectif a été fixé car nous souhaitons construire une base de données 3D annotées dans un temps raisonnable (1 à 2 ans). Cela n'aurait pas été possible si nous nous étions uniquement focalisés sur le deuxième objectif. Pour l'atteindre, il aurait fallu : développer une méthode adaptée, valider celle-ci sur fantôme, ouvrir une étude clinique pour valider la méthode in-vivo, réaliser un nombre conséquent d'acquisitions sur de nouveaux patients et enfin annoter ces données. Cela aurait probablement dépasser le cadre de cette thèse et aurait mis de côté la problématique de segmentation associée à celle-ci.

Pour répondre au premier objectif, la difficulté majeure à surmonter est que la reconstruction ne peut être effectuée qu'à partir des informations contenues dans les images ou à partir d'hypothèses sur le mouvement d'acquisition. Pour répondre au deuxième objectif, on peut en revanche envisager d'utiliser un dispositif supplémentaire permettant de suivre la sonde durant l'acquisition. Il faut dans ce dernier cas considérer plusieurs problématiques pratiques : le dispositif choisi doit permettre l'utilisation de la sonde linéaire 9L4 pour pouvoir conserver une excellente qualité d'image, il faut qu'il soit compatible avec le mouvement d'acquisition et il doit pouvoir être utilisé dans un environnement clinique (les enfants sont dans des couveuses et peuvent bouger lors des examens).

Les méthodes qui permettent d'estimer la position relative de deux images échographiques sont présentées dans la section suivante.

II.1 Estimation de la position relative de deux images

L'acquisition d'une séquence d'images 2D est généralement réalisée en effectuant l'un des trois mouvements d'acquisition suivants : une translation linéaire de la sonde, une rotation de la sonde autour d'une axe fixe ou une rotation de la sonde autour d'un axe mobile. Ces mouvements sont représentés sur la Figure.II.1

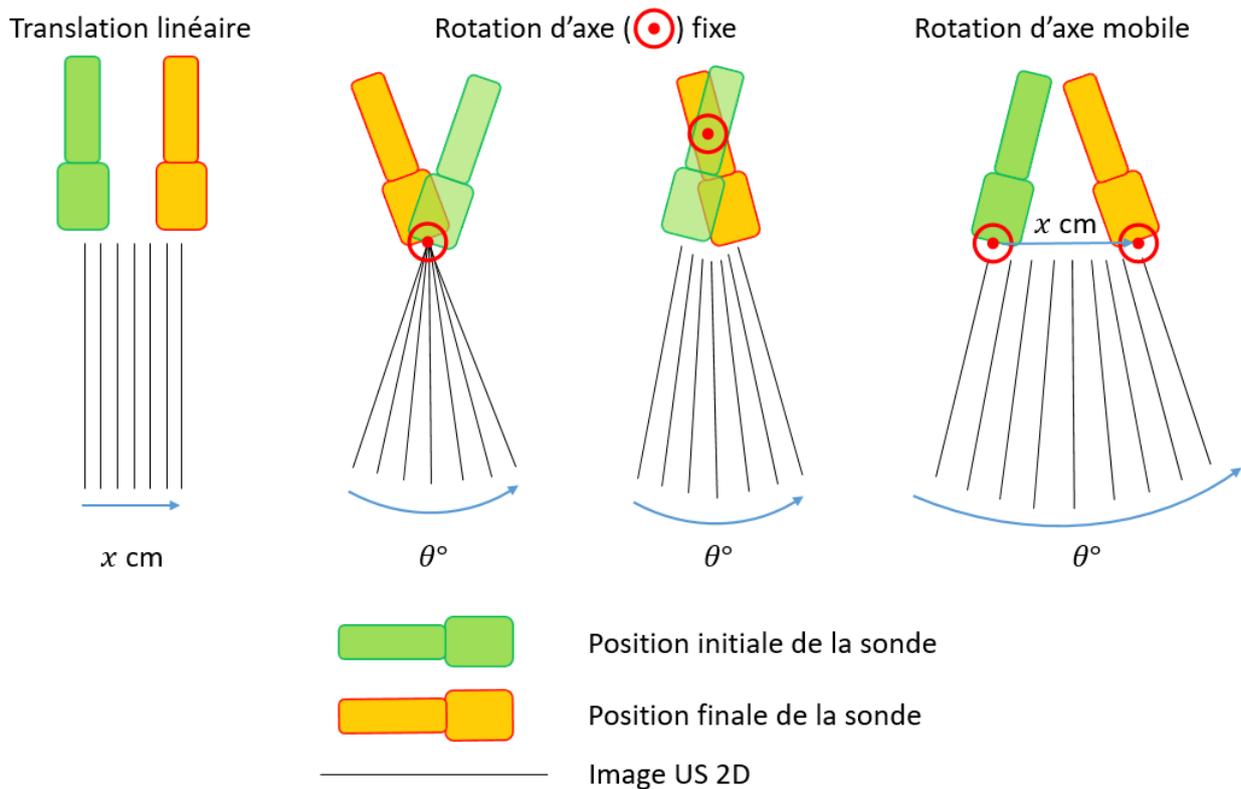


FIGURE II.1 – Les mouvements d'acquisition en imagerie US.

Le mouvement d'acquisition réalisé lors d'une ETF 2D chez le prématuré est une rotation d'axe mobile. Ce mouvement est le plus complexe, néanmoins, c'est celui qui permet de s'adapter au mieux à l'anatomie de la fontanelle et d'imager le plus de parties de l'encéphale.

Pour estimer le déplacement qui a lieu durant l'acquisition, il existe trois méthodes qui se basent uniquement sur les images échographiques ou le mouvement d'acquisition théorique :

- la mesure de la décorrélation du speckle
- la modélisation du mouvement d'acquisition
- l'apprentissage supervisé

Deux autres groupes de méthodes permettent d'enregistrer le mouvement d'acquisition :

- l'utilisation d'un capteur de position
- l'usage d'un dispositif électromécanique pour réaliser l'acquisition.

II.1.1 Acquisition US main-libre réalisée sans capteur

Cette section présente trois méthodes utilisées pour reconstruire les acquisitions US main-libre réalisées sans capteur. Celles-ci pourraient être utilisées pour répondre à nos objectifs à court terme et à long terme.

Mesure de la décorrélation du speckle

Une première approche consiste à utiliser la statistique du speckle présent dans les images échographiques. Celui-ci est causé par les diffuseurs du milieu imagé qui ont une taille inférieure à la longueur d'onde de l'onde US émise. Comme la valeur d'un pixel d'une image US est obtenue à partir des contributions des diffuseurs présents dans une certaine zone, si on réalise un petit déplacement perpendiculaire au plan d'acquisition, les régions contribuant à la valeur d'un même pixel se chevauchent (Figure.II.2). Il est alors possible d'établir un lien entre la décorrélation du speckle présent sur les deux images et la distance de déplacement (Chen *et al.* (1997)).

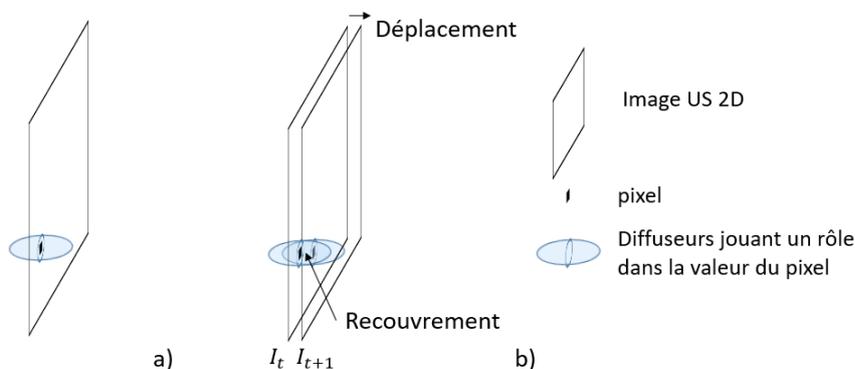


FIGURE II.2 – a) : Image US 2D et zone jouant un rôle dans la valeur d'un pixel, b) : deux images échographiques acquises successivement et recouvrement des zones jouant un rôle dans la valeur d'un pixel.

Pour déterminer le déplacement en pratique, l'image est divisée en sous-images par un damier, on estime alors le déplacement entre les sous-images se trouvant en face les unes des autres, ce qui permet de déterminer le déplacement global entre les images. Ce déplacement est appelé le mouvement hors-plan (out-of-plan motion), cette méthode est illustrée sur la Figure.II.3.

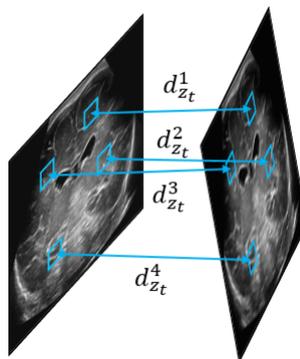


FIGURE II.3 – Estimation du mouvement hors du plan de l'image¹.

1. Images issues de notre base de données.

Pour mesurer au mieux la décorrélation du speckle, on estime au préalable le mouvement dans le plan (in-plane-motion). Ce mouvement est composé d'une rotation d'axe perpendiculaire à l'image et d'une translation dans le plan qu'elle définit. Ces paramètres sont représentés sur la Figure.II.4.

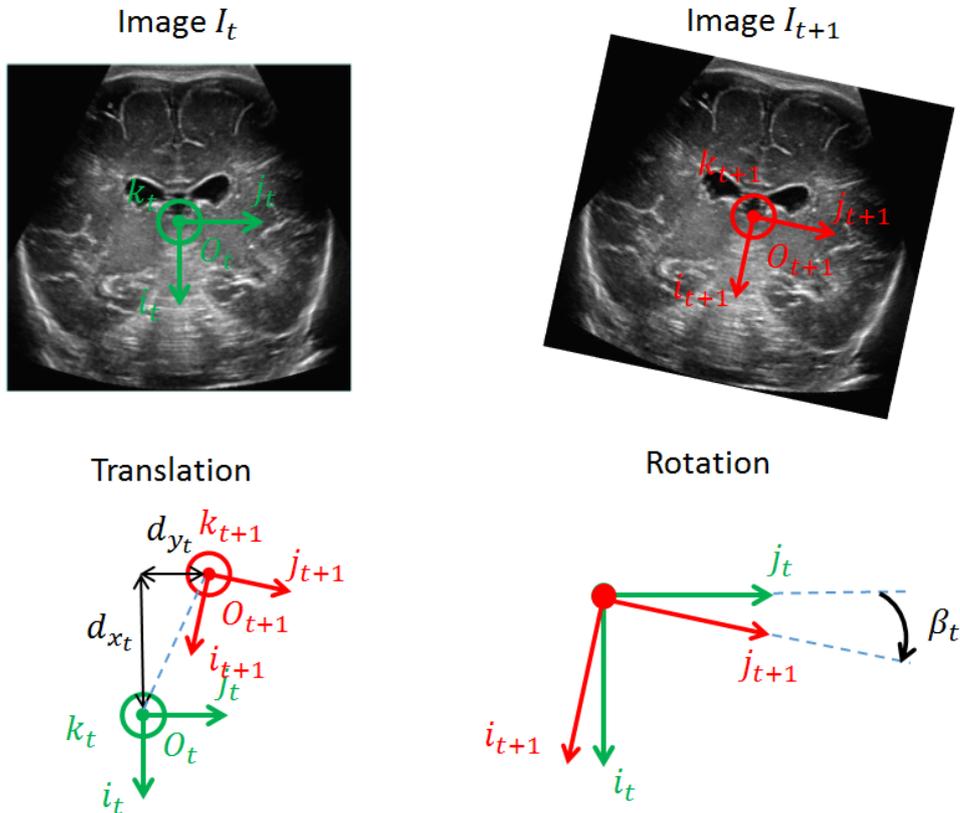


FIGURE II.4 – Définition du mouvement dans le plan de l'image.

En pratique, le mouvement dans le plan est estimé en utilisant des méthodes de recalage rigide, telle que du block-matching (Prager *et al.* (2003), Housden *et al.* (2006)), qui minimisent les moindres carrés entre les deux images. Cette étape ne pose pas de problème en pratique, le gros challenge réside dans l'estimation du mouvement hors plan.

Cette approche a été utilisée pour estimer des mouvements de translations linéaires par Prager *et al.* (2003) (sur fantôme et in-vivo), Hassenpflug *et al.* (2005) (en simulation) et Housden *et al.* (2007) (sur fantôme). Elle a également été mise en œuvre par Housden *et al.* (2008) (sur fantôme) pour estimer un mouvement de rotation d'axe fixe.

Plusieurs inconvénients limitent l'usage de cette méthode in-vivo et pour notre application. Premièrement, les modèles de décorrélation du speckle ne sont valables que pour des mouvements contraints comme la translation linéaire ou la rotation d'axe fixe. Nous n'avons pas connaissance d'un modèle existant qui serait adapté à une combinaison des deux mouvements. Deuxièmement, il n'est théoriquement possible de mesurer la décorrélation du speckle que dans des régions où celui-ci est "complètement développé". Cette condition revient à dire que le milieu doit être composé uniquement de diffuseurs répartis aléatoirement. Cette condition est rarement respectée en pra-

tique car le milieu imagé possède des structures continues qui réduisent fortement la décorrélation du speckle. En outre, l'utilisation de cette méthode nécessite que les images soient suffisamment rapprochées, ce qui n'est pas évident dans notre cas, en particulier dans les zones basses des images à cause du mouvement de rotation effectué lors de l'acquisition. Dans ces conditions, il n'est pas possible d'utiliser les modèles de décorrélation du speckle pour estimer le déplacement entre deux images de manière fiable. [Laporte et Arbel \(2011\)](#) ont récemment proposé d'adapter localement le modèle de décorrélation mais l'approche n'a été validée que pour des mouvements de translation linéaire.

En conclusion, même si cette approche est prometteuse, son application in-vivo est loin d'être évidente aujourd'hui. En particulier pour des acquisitions d'une structure complexe comme le cerveau qui sont réalisées avec un mouvement de rotation d'axe mobile.

Modélisation du mouvement d'acquisition

Une seconde approche consiste à modéliser le mouvement d'acquisition. Dans la littérature, ces modèles se limitent au cas des translations linéaires et des rotations d'axe fixe ([Fenster et al. \(2001\)](#)). Les reconstructions 3D sont réalisées en mesurant manuellement l'angle de rotation ou la distance de translation et en faisant l'hypothèse que ce paramètre est constant.

Nous avons trouvé peu de littérature sur ces méthodes, elles sont peu utilisées en pratique. Une raison qui pourrait l'expliquer est que la précision des reconstructions effectuées avec cette méthode n'est pas très bonne. On fait l'hypothèse que la vitesse de rotation ou de translation de la sonde est constante durant l'acquisition, ce qui est difficilement réalisable lors d'une acquisition main-libre. L'angle de rotation ou la distance de translation sont de plus mesurés manuellement ou fixés intuitivement, ce qui peut être une grosse source d'erreurs. Les mouvements correspondant à notre cas, une rotation d'axe mobile, semblent ne pas avoir été étudiés. Cela peut s'expliquer par le fait qu'il serait compliqué de donner à la fois une estimation de l'angle de rotation et du mouvement de l'axe de rotation.

Cette approche ne permet pas de réaliser des reconstructions 3D anatomiquement fiables d'acquisition 2D. Les reconstructions 3D réalisées avec cette méthode sont néanmoins de bonne qualité ([Fenster et al. \(2001\)](#)).

Reconstruction par apprentissage supervisé

Une troisième approche a récemment été proposée par [Prevost et al. \(2017\)](#). Celle-ci utilise des méthodes d'apprentissage supervisé pour déterminer le déplacement existant entre deux images. Leur idée est, à partir de deux images successives, d'apprendre à un réseau de neurones convolutifs à estimer le modèle de décorrélation du speckle pour être en capacité de déterminer les 6 paramètres de position permettant de replacer les images les unes par rapport aux autres. Leur approche a été validée pour un mouvement de translation linéaire. Ils ont utilisé un système de suivi optique de la sonde pour construire leur base de données annotées.

Pour leur application, cette méthode a permis d'obtenir des reconstructions 3D plus précises que celles réalisées en mesurant la décorrélation du speckle. Elle apparaît donc comme une bonne option pour réaliser la reconstruction d'acquisition main-libre sans capteur. Néanmoins, il n'est pas envisageable de l'utiliser pour répondre à notre objectif à court terme car elle nécessiterait de créer une base de données annotées et donc mettre en place un système de suivi de la sonde pour obtenir des annotations.

II.1.2 Enregistrement du mouvement d'acquisition

Cette section présente les méthodes utilisées pour enregistrer le mouvement d'acquisition. Celles-ci peuvent être envisagées pour répondre à notre objectif à long terme.

Suivi de la sonde par capteur électromagnétique

Une première approche consiste à solidariser un capteur électromagnétique à la sonde. En générant un champ magnétique à l'aide d'un dispositif extérieur, proche du capteur, il est alors possible de suivre le déplacement de la sonde.

Cette méthode a été utilisée in-vivo pour diverses applications telles que la reconstruction 3D d'acquisitions de la colonne thoracique (Dewi *et al.* (2009)) ou du seins (Huang *et al.* (2017a)). Elle a aussi été mise en œuvre, in-vivo, pour notre application par Nagdyman *et al.* (1999). Elle avait permis de mesurer précisément le VV sur fantôme.

Cette approche permet de réaliser des reconstructions précises mais elle a plusieurs limites qui réduisent son utilisation en pratique clinique. Tout d'abord, il faut générer un champ magnétique dans la zone d'acquisition à l'aide d'un dispositif extérieur qui doit être placé proche du capteur. De plus, le champ magnétique peut être perturbé par des dispositifs électroniques ou métalliques qui sont proches de la zone d'acquisition, il faut donc être vigilant car cela peut engendrer des erreurs de mesures.

Suivi de la sonde par capteur optique

Une seconde approche repose sur l'utilisation d'un capteur optique. Une première stratégie consiste à placer un repère sur la sonde et à suivre l'évolution de sa position au cours de l'acquisition à l'aide de deux caméras, une seconde stratégie est de placer un motif sur le patient et de solidariser une caméra à la sonde pour déterminer le déplacement à partir des images du motif.

La première stratégie a été utilisée sur fantôme par Treece *et al.* (2003) et in-vivo par Prevost *et al.* (2017) pour reconstruire des balayages linéaires réalisés sur des avant-bras. La deuxième stratégie a été employée sur fantôme par Poulsen *et al.* (2005) et in-vivo par Sun *et al.* (2013) pour créer des volumes 3D à partir d'acquisitions réalisées sur la cuisse. Plus récemment, Sun *et al.* (2014) ont astucieusement proposé de filmer directement la peau plutôt que d'y coller un motif car elle possède des motifs naturels.

Cette approche permet de réaliser des reconstructions précises mais a des limites pratiques pour notre application. Par rapport à la première stratégie, il faut que le champ de vu entre le repère suivi et la caméra soit dégagé durant toute l'acquisition. Cela pourrait être difficile dans notre cas car les enfants se trouvent dans des couveuses et leur tête n'est pas toujours orientée vers l'extérieur de celle-ci. La seconde stratégie a le même défaut que la première, pour réaliser le mouvement d'acquisition, qui est complexe, la main du médecin occupe la majeure partie du manche de la sonde et couvre une zone importante de la tête de l'enfant.

Suivi de la sonde par accéléromètre et gyromètre

Une dernière approche capteur a été récemment proposée par [Morgan *et al.* \(2018\)](#). Celle-ci consiste à solidariser un accéléromètre et un gyromètre à la sonde pour suivre son déplacement.

Leur approche a été validée sur fantôme et in-vivo, elle permet de réaliser des reconstructions précises. Elle a de plus l'avantage de ne pas souffrir des inconvénients cités plus haut : champ de vu dégagé concernant les capteurs optiques, étendu du champ magnétique limité et sensibilité aux objets métalliques/dispositifs électroniques dans le cas des capteurs électromagnétiques. Elle est également très peu couteuse comparée aux méthodes utilisant des capteurs optiques ou électromagnétiques. Son utilisation est néanmoins limitée à des mouvements rotations autour d'un axe fixe pour le moment.

Mécanisation de la sonde utilisée pour l'acquisition

Plutôt que d'utiliser un capteur pour suivre le mouvement d'acquisition, il est possible d'avoir recours à un dispositif électromécanique pour mettre la sonde en mouvement. L'évolution de la position de la sonde est alors parfaitement connue au cours du temps. Le dispositif électromécanique peut être externe, auquel cas l'ensemble de la sonde est mis en mouvement, ou interne, auquel cas c'est seulement le transducteur qui est mis en mouvement.

La première stratégie a été utilisée par [Kishimoto *et al.* \(2013\)](#) pour notre application. Ils ont développé et validé in-vivo un système qui se fixe sur le manche de la sonde et qui permet de la mettre en rotation autour d'un axe fixe. Cette solution s'avère particulièrement intéressante pour notre problématique, néanmoins le dispositif a été conçu pour des sondes néonatales convexes. Il n'est pas certain qu'un contact entre une sonde linéaire et la tête de l'enfant puisse être maintenu pendant toute l'acquisition avec ce dispositif. Pour réaliser des acquisitions 3D du foie, [Neshat *et al.* \(2013\)](#) ont développé un dispositif qui permet d'effectuer trois mouvements différents : des translations linéaires, des rotations d'axe fixe et des combinaisons de ces deux derniers. Ce dispositif est intéressant mais nécessite de pouvoir être posé sur un support plat, ce qui n'est pas possible dans notre cas.

La seconde stratégie a été utilisée par [Yan *et al.* \(2012\)](#) et [Bax *et al.* \(2008\)](#) pour faire de l'imagerie 3D de la prostate. Elle a également été mise en œuvre pour notre application par [Haiden *et al.* \(2005\)](#), [Gilmore *et al.* \(2001\)](#) et [Salerno *et al.* \(2000\)](#). Par rapport à la première stratégie, cette méthode a l'avantage de ne pas demander l'usage d'un dispositif extérieur à la sonde. Néanmoins, la mettre en œuvre dans notre cas requerrait le développement complet d'une sonde pour conserver la qualité

d'image obtenue avec la sonde 9L4. Cela ne pourrait se faire qu'avec l'implication d'un constructeur dans le projet, ce qui n'a pas été possible pour le moment.

II.1.3 Conclusion

Objectif à court terme

Pour atteindre notre objectif à court terme, trois méthodes permettant d'effectuer des reconstructions d'acquisition main-libre sans capteur ont été présentées. Les méthodes utilisant la statistique du speckle et l'apprentissage supervisé semblent difficiles à mettre en œuvre pour les raisons citées dans les parties qui leurs sont dédiées. Nous nous sommes donc orientés vers la modélisation du mouvement d'acquisition, même si cette méthode est réputée pour produire des reconstructions peu précises. Les données médicales ont aujourd'hui une grande valeur et il est intéressant de développer des outils pour créer des données US 3D à moindre coût. Nous nous sommes attachés à réduire les limitations de cette méthode en développant un modèle de rotation autour d'une axe mobile et en estimant les paramètres de la transformation plutôt qu'en les mesurant manuellement. Ces contributions seront présentées dans le chapitre 4.

Objectif à long terme

Pour atteindre notre objectif à long terme, on peut envisager d'utiliser n'importe quelles méthodes présentées dans cette partie. Celles qui sont basées sur le suivi du mouvement de la sonde ont néanmoins un inconvénient majeur : elles ne peuvent pas prendre en compte un déplacement entre deux images qui serait causé par un mouvement de la tête de l'enfant et non de la sonde. Cela arrive régulièrement pendant les acquisitions et pourrait compromettre l'obtention d'un volume 3D cohérent, à moins d'effectuer plusieurs fois les acquisitions. Partant de ce constat, nous pensons qu'il est préférable d'utiliser une méthode qui se base uniquement sur les images pour reconstruire les acquisitions ou d'utiliser une méthode combinant le suivi de la sonde avec un capteur et l'estimation du déplacement sur les images.

L'utilisation de la statistique du speckle n'étant pas possible pour des acquisitions in-vivo, faire de l'apprentissage supervisé avec des réseaux de neurones convolutifs nous paraît être la méthode la plus intéressante si on se base uniquement sur les images pour effectuer la reconstruction. Néanmoins, il est nécessaire de construire une base de données annotées et donc de réaliser des acquisitions avec un capteur pour pouvoir entraîner un réseau de neurones convolutifs. Il faut donc dans tous les cas choisir une méthode qui utilise un capteur pour suivre le mouvement d'acquisition de manière à obtenir une vérité terrain. Étant donné que la méthode utilisée par [Prevost *et al.* \(2017\)](#) n'a été validée que pour des mouvements de translation linéaire, qui ne correspondent pas à notre cas, la méthode choisie pour enregistrer le mouvement d'acquisition doit être la plus adaptée possible à notre problématique dans le cas où la reconstruction par réseau de neurones convolutifs ne fonctionnerait pas.

Parmi les méthodes enregistrant le mouvement d'acquisition, nous avons rejeté celles qui utilisent un dispositif électromécanique interne ou externe à la sonde car, avec ces méthodes, il se-

rait difficile de s'adapter à l'anatomie de la fontanelle pour acquérir la plus grande partie possible de l'encéphale. Les méthodes qui se servent de capteurs optiques ou électromagnétiques nous paraissent difficiles à mettre en œuvre en pratique pour les raisons citées plus haut. Nous nous sommes donc orientés vers la solution utilisant un accéléromètre et un gyromètre pour suivre le mouvement de la sonde : elle est économique, le capteur peut facilement être fixé sur la sonde via un adaptateur fabriqué avec une imprimante 3D et elle ne souffre pas de problèmes liés au champ de vue ou à la perturbation d'un champ magnétique.

Pour développer cette dernière méthode pour notre application et si possible constituer une base de données pour réaliser de l'apprentissage supervisé, nous avons établi un partenariat avec l'université américaine de Duke, où le dispositif a été développé.

II.2 Interpolation dans la grille cartésienne de reconstruction

Le choix d'une méthode d'interpolation a été fait par rapport à notre objectif à court terme. Dans ce cas, il faut que la méthode choisie respecte les contraintes suivantes :

- Préserver au maximum la qualité des images.
- Avoir un temps d'interpolation raisonnable (quelques minutes).
- Être adaptée à un mouvement de rotation d'axe mobile.

Il existe trois groupes de méthodes permettant d'interpoler les valeurs des voxels d'une grille de reconstruction. Elles ont été passées en revue dans [Solberg *et al.* \(2007\)](#), [Miller *et al.* \(2012\)](#) et [Mozafari *et Lee* \(2017\)](#). On distingue les méthodes basées sur les pixels (PBM), sur les voxels (VBM) et sur les fonctions (FBM). Celles-ci vont être présentées dans les trois sections suivantes. Une quatrième section sera dédiée aux forces et aux faiblesses de ces méthodes vis à vis de notre application.

II.2.1 Méthodes basées sur les pixels

Une première méthode d'interpolation se base sur les pixels. Elle consiste à parcourir tous les pixels des images replacés dans la grille de reconstruction et à utiliser leur valeur pour en donner une aux voxels.

L'approche la plus basique est de faire contribuer la valeur de chaque pixel au voxel le plus proche (PNN). Il est possible que deux pixels contribuent au même voxel, dans ce cas, leurs valeurs sont en général moyennées. Cette approche est illustrée sur la [Figure II.5](#) :

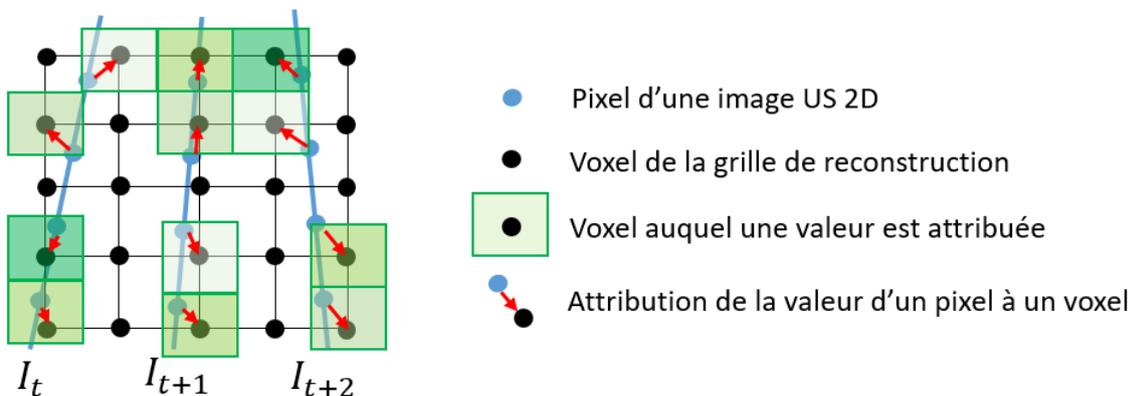


FIGURE II.5 – Distribution des valeurs de chaque pixel au voxel le plus proche.

Après cette première étape de distribution (DS), en fonction de l'espacement entre les images acquises et de la taille de la grille de reconstruction, il est possible qu'aucune valeur n'ait été attribuée à certains voxels : la grille de reconstruction contient des trous. Cette étape est donc suivie par une étape de remplissage (HFS) qui permet d'attribuer une valeur à tous les voxels. Une première manière de faire est de parcourir tous les voxels vides et de leur attribuer une valeur à partir de celles des voxels non vides qui se trouvent dans un certain voisinage. Des méthodes associant la valeur maximum des voxels, la valeur pondérée moyenne dans un certain voisinage ([Nelson *et Pretorius* \(1997\)](#)) ou encore la moyenne de la valeur des deux voxels les plus proches ont été proposées. Elles

sont illustrées sur la Figure.II.6. Une seconde manière de faire est de parcourir les voxels qui ont déjà une valeur et d'utiliser celle-ci pour en attribuer une aux voxels vides. Des méthodes utilisant des noyaux gaussien sphériques (San José-Estépar *et al.* (2003)) ou des convolutions normalisées ont aussi été utilisées à cette fin.

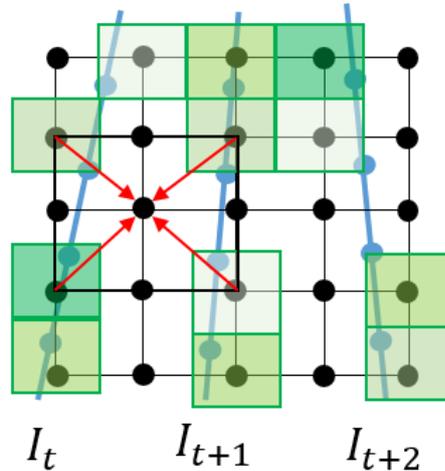


FIGURE II.6 – Interpolation de la grille à partir des voxels non vides dans un certain voisinage.

Une seconde approche consiste à faire contribuer la valeur d'un pixel à plusieurs voxels dans son voisinage. La méthode la plus commune est l'utilisation d'un noyau sphérique ou Gaussien avec un poids inversement proportionnel à la distance entre le voxel et le pixel, celle-ci a été proposée par Barry *et al.* (1997) et est illustrée sur la Figure.II.7.

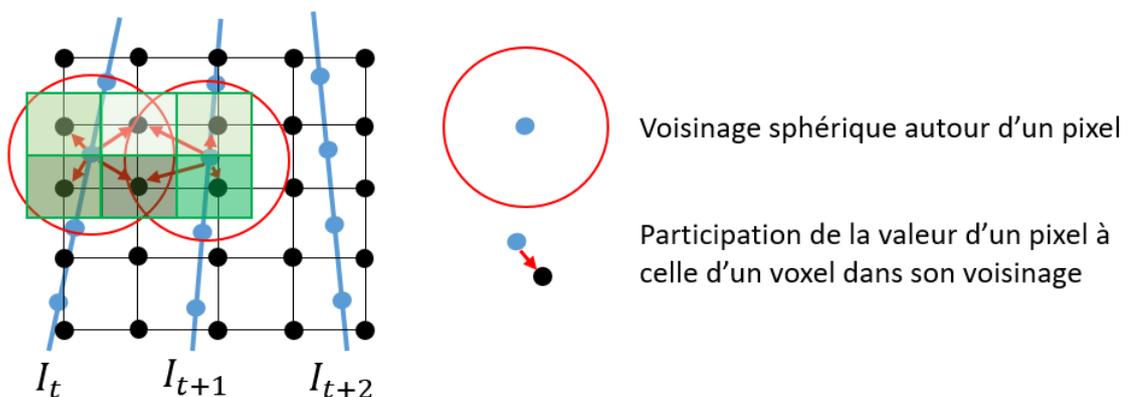


FIGURE II.7 – Distribution de la valeur d'un pixel à plusieurs voxels de son voisinage.

Cette dernière approche a l'avantage de permettre de réaliser la DS et la HFS en même temps. De plus, elle utilise les vraies coordonnées des pixels replacés dans la grille, plutôt que les coordonnées des pixels après la DS, au moment de pondérer leurs valeurs pour en attribuer aux voxels, ce qui est plus précis.

II.2.2 Méthodes basées sur les voxels

Une seconde méthode d'interpolation consiste à parcourir un à un les voxels de la grille de reconstruction et à leur attribuer une valeur en fonction de celle des pixels qui se trouvent dans un

certain voisinage.

L'approche la plus basique consiste à attribuer à un voxel la valeur du pixel le plus proche (VNN), il peut donc exister des zones où les voxels voisins ont tous la même valeur. Cette approche est illustrée sur la Figure.II.8.

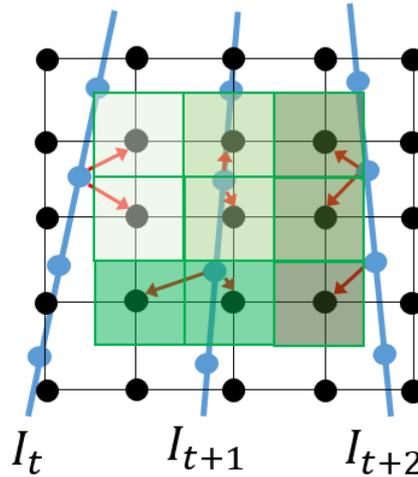


FIGURE II.8 – Attribution à chaque voxel de la valeur du pixel le plus proche.

Une autre approche consiste, pour chaque voxel, à utiliser les valeurs de plusieurs pixels contenu dans un certain voisinage, par exemple sphérique. Ces valeurs sont en général pondérées par l'inverse de leur distance au voxel. Cette méthode est illustrée sur la Figure.II.9

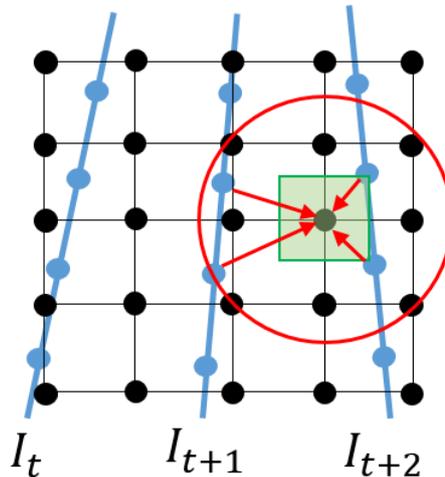


FIGURE II.9 – Attribution de la valeur d'un voxel à partir des pixels contenus dans un certain voisinage.

Enfin, une dernière approche utilise la valeur d'un certain nombre de pixels sur les deux images adjacentes à chaque voxel. Elle est représentée sur la Figure.II.10

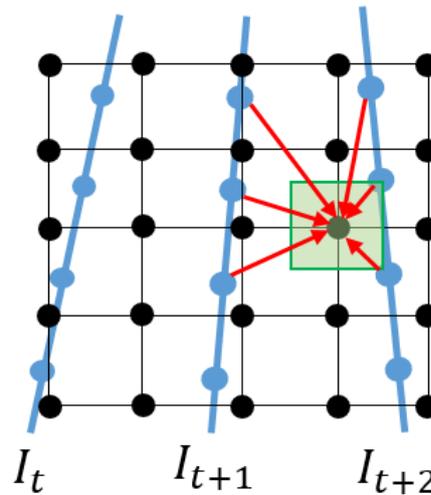


FIGURE II.10 – Attribution de la valeur d'un voxel à partir de deux images adjacentes.

Cette dernière méthode a l'avantage de toujours utiliser le même nombre de pixel pour attribuer une valeur à un voxel, ce qui n'est pas le cas lorsque que l'on définit un voisinage autour d'un voxel car l'espacement entre les images peut varier.

II.2.3 Méthodes basées sur les fonctions

La dernière méthode d'interpolation se base sur les fonctions. Tout d'abord, une fonction qui prend la valeur des pixels en chacune de leur coordonnées est déterminée. Puis, la grille de reconstruction est complétée en évaluant la valeur de cette fonction en la coordonnée de chaque voxel. Cette méthode est illustrée sur la Figure.II.11.

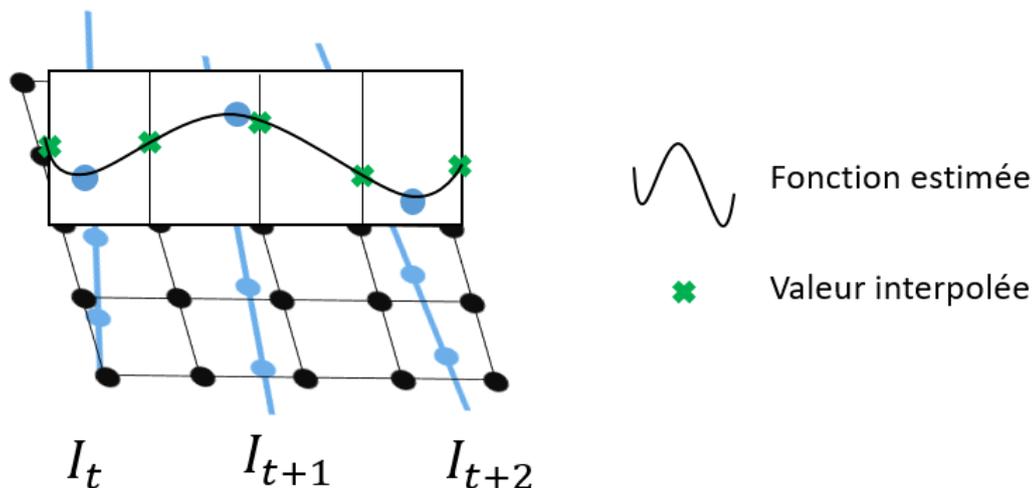


FIGURE II.11 – Interpolation de la grille de reconstruction à partir de fonctions.

Plusieurs type de fonctions ont été proposés pour cette méthode. Il est par exemple possible de recourir à des fonctions polynomiales, des fonctions de base radiale (Zhang *et al.* (2004)) ou encore des courbes de Béziérs (Huang *et al.* (2014)). Ces méthodes permettent des reconstructions de bonne

qualité mais sont très lentes : plusieurs heures pour interpoler un de nos volumes. La méthode proposée par [Huang *et al.* \(2014\)](#) a permis de réduire drastiquement le temps nécessaire pour réaliser l'interpolation (entre 90.6 % et 97.2 %) mais ne s'applique que pour des balayages linéaires.

II.2.4 Comparaison des méthodes et conclusion

Notre objectif est de réaliser l'interpolation des volumes de manière à obtenir une bonne qualité d'image dans un temps raisonnable (quelques minutes).

Parmi les trois familles de méthodes existantes pour arriver à ces fins, on peut directement exclure les FBM car le temps d'interpolation serait trop long pour nos volumes dont la taille est généralement de plusieurs millions de voxels. De plus, compte tenu de notre mouvement d'acquisition, sa mise en œuvre nécessiterait l'utilisation d'un ensemble de fonctions bien particulier.

Parmi les PBM et les VBM, les méthodes PNN et VNN sont les plus rapides, elles sont généralement utilisées pour des interpolations temps réel. Elles donnent en revanche les reconstructions de moins bonne qualité. Les méthodes plus complexes réalisant l'interpolation à partir de plusieurs valeurs sont plus lentes mais permettent de réaliser les interpolations dans un temps raisonnable. Elles permettent de plus de bien conserver la qualité des images 2D initiales, nous nous sommes donc orientés vers ces méthodes qui seront comparées dans le chapitre 4.

II.3 Repositionnement dans le repère anatomique

Pour repositionner un volume US dans le repère anatomique, des méthodes de recalage sur des données provenant d'une autre modalité sont en générales utilisées. Pour cette application, seules des images IRM, acquises dans le repère anatomique, peuvent être utilisées. L'objectif est alors de repositionner le volume US sur le volume IRM en cherchant à minimiser une métrique adaptée au problème. On considère deux cas de figure, soit on dispose de données IRM des patients de notre base US soit on cherche des acquisitions IRM réalisées sur des patients différents. Dans le premier cas de figure, [Fuerst *et al.* \(2014\)](#) ont par exemple utilisé la métrique de combinaison linéaire de corrélations locales (LC^2) pour recaler des volumes US du cerveau du fœtus sur des IRMs cérébrales acquises au préalable sur ces patients. Cette idée a été reprise dans le deuxième cas de figure par [Boucher *et al.* \(2018b\)](#). Ils disposaient de volumes 3D US du cerveau d'enfants âgés de 2 à 8 mois qu'ils ont recalés sur des atlas IRM d'autres patients. Les atlas qu'ils ont utilisés étaient :

- Un atlas d'IRM construit à partir d'acquisition faites sur des enfants âgés de 1 an du centre d'imagerie du Cincinnati ([Altaye *et al.* \(2008\)](#))
- Un atlas du centre d'imagerie McConnel ([Fonov *et al.* \(2009\)](#)) réalisé à partir d'acquisitions faites sur des enfants de 2 à 5 mois.
- Neuf volumes IRM extraits de l'atlas ALBERTs ([Gousias *et al.* \(2012\)](#)) contenant des IRM d'enfants prématurés.

Comme nous ne disposions pas d'IRMs associées à nos données US, nous avons essayé de recalculer nos volumes en utilisant des volumes IRMs de l'atlas Albert. Nous ne sommes néanmoins pas parvenus à recalculer nos données de manière satisfaisante avec cet Atlas. Ce recalage est en effet extrêmement difficile du fait que les volumes US reconstruits ne permettent pas de visualiser l'intégralité de l'encéphale, que l'aspect des images échographiques et IRM est très différent et qu'il existe une grande variabilité anatomique entre les enfants à cet âge. Nous avons finalement choisi d'orienter les volumes US dans le repère anatomique en définissant manuellement les angles de rotations, en discussion avec le pédiatre partenaire de ce projet. Plus de détails seront donnés dans le Chapitre 4.

CHAPITRE III

ÉTAT DE L'ART : SEGMENTATION

Contents

III.1 Segmentation des structures cérébrales chez le nouveau né en IRM	47
III.1.1 Segmentation multi-atlas	47
III.1.2 Méthode paramétrique utilisant des atlas probabilistes	48
III.1.3 Segmentation du SVC : ventriculomégalie et hydrocéphalie	50
III.1.4 Classification	50
Classification à partir d'une définition manuelle des caractéristiques	51
Representation learning	51
III.2 Segmentation en échographie	53
III.2.1 Segmentation du SVC par contours actifs	53
III.2.2 Segmentation du SVC par deep learning	57
III.3 Conclusion	58

CHAPITRE III

LA segmentation du SVC et des thalami est un problème difficile pour les raisons évoquées dans le Chapitre 1. Pour y répondre, la méthode de segmentation développée/choisie doit en outre être adaptée à des besoins cliniques contraignants : segmentation automatique d'un gros volume de données en quelques secondes. D'après l'analyse des données réalisée en introduction, les méthodes basiques qui utilisent des histogrammes d'intensité ou la magnitude du gradient pour segmenter une image ne seront pas assez performantes pour résoudre ce problème. Ce chapitre fait l'état de l'art des méthodes de segmentation qui ont été utilisées pour segmenter les structures cérébrales chez le prématuré. Il commence par présenter celles qui ont été utilisées pour la segmentation des tissus cérébraux en IRM et termine par décrire celles qui ont été employées pour la segmentation du SVC en imagerie US.

III.1 Segmentation des structures cérébrales chez le nouveau né en IRM

Notre problématique de segmentation est très peu traitée en échographie 3D. Néanmoins, elle a été étudiée ces dernières années en IRM, ce qui offre une bibliographie à partir de laquelle il est possible de s'inspirer pour développer une méthode adaptée à notre problème. Ce dernier est également difficile à résoudre en IRM. Par rapport au problème de segmentation d'IRM cérébrale chez l'adulte il y a : un manque de base de données annotées, une grande variabilité anatomique entre les patients et un faible rapport contraste sur bruit dans les images. La plupart des méthodes reportées dans la littérature se concentre sur la segmentation de la SB, de la SG et du LCS et parfois également sur les noyaux gris centraux et le SVC. Ces méthodes ont été passées en revue récemment par [Devi et al. \(2015\)](#) et [Makropoulos et al. \(2018\)](#). Les plus utilisées sont les méthodes multi-atlas, les méthodes paramétriques et les méthodes de classification, elles feront respectivement l'objet des trois prochaines parties.

III.1.1 Segmentation multi-atlas

Pour limiter les interactions avec les utilisateurs en milieu clinique, les méthodes de segmentations développées doivent être automatiques. Pour répondre à ce besoin pour ce problème difficile, des méthodes de segmentations multi-atlas ont été proposées.

Par atlas, nous désignerons dans tout ce chapitre l'association d'un volume IRM et d'un volume comportant des annotations. Le type d'atlas le plus classique est celui où les annotations correspondent à des segmentations manuelles. Un exemple typique, pour cette application, est le multi-atlas ALBERT ([Gousias et al. \(2012\)](#)). Celui-ci est composé de 15 atlas 3D réalisés chez des enfants

prématurés et 5 chez des enfants nés à terme. Dans chaque volume, 50 structures cérébrales sont segmentées manuellement. En libre accès, c'est l'un des plus gros atlas IRM dédié aux enfants nés prématurés avec celui proposé pour le challenge MICCAI NeoBrainS12.

La méthodologie la plus typique pour réaliser une segmentation multi-atlas est celle proposée par Gousias *et al.* (2013). Elle comprend les étapes suivantes : recalcr chacun des atlas 3D (ALBERTs) sur le volume que l'on souhaite segmenter en utilisant des transformations rigides puis affines puis élastiques, appliquer ces transformations aux annotations de l'atlas, fusionner les cartes de segmentations transformées pour obtenir la segmentation finale. Cette méthode est représentée sur la Figure.III.1. Il existe différentes stratégies pour fusionner les segmentations après leur transformation (Artaechevarria *et al.* (2009)), celle utilisée par Gousias *et al.* (2013) est décrite dans Kittler (1998).

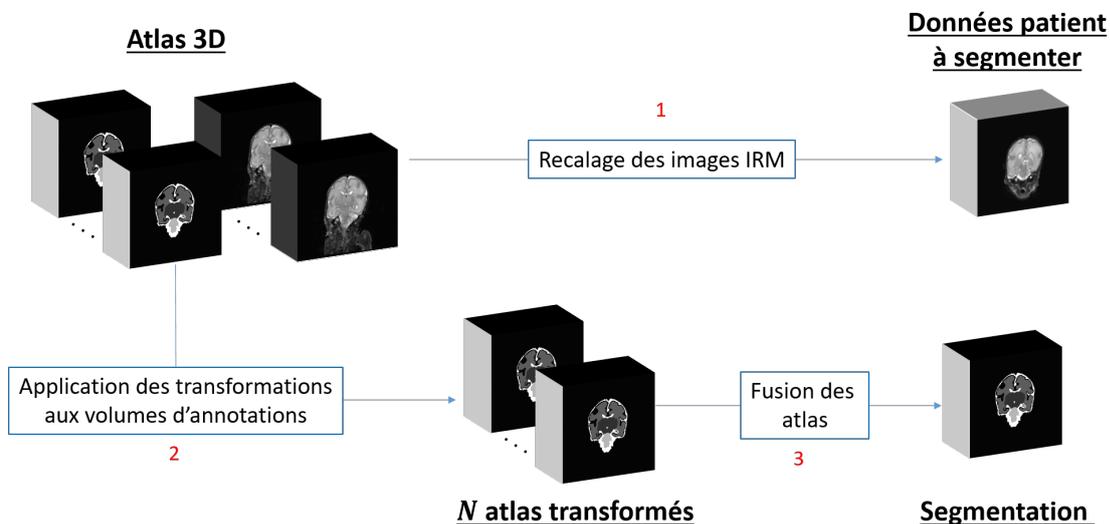


FIGURE III.1 – Segmentation par recalage multi-atlas¹.

Cette approche a également été utilisée par Srhoj-Egekher *et al.* (2013), Weisenfeld *et Warfield* (2009) et Kim *et al.* (2015). Une variante de cette méthode est d'effectuer la segmentation des volumes par patch. Dans ce cas là, on cherche au préalable les patches de l'atlas qui sont les plus similaires à la zone que l'on souhaite segmenter (Wright *et al.* (2014)).

Cette méthode a deux inconvénients majeurs : le premier est que le recalage des atlas peut être peu précis car l'anatomie du cerveau varie beaucoup chez les enfants prématurés, le second est que sa mise en œuvre nécessite le recalage de plusieurs volumes d'une taille importante, ce qui est coûteux en temps. Pour palier ces limitations, des méthodes paramétriques utilisant des atlas probabilistes et des modèles d'intensités des tissus peuvent être utilisées. Ces méthodes font l'objet de la prochaine section.

1. Images issues de notre base de données.

III.1.2 Méthode paramétrique utilisant des atlas probabilistes

Pour résoudre ce problème de segmentation, les méthodes paramétriques se basent sur un modèle d'intensité pour chaque structure, souvent gaussien, ainsi que des a priori spatiaux, généralement obtenus grâce à un atlas probabiliste. Un atlas probabiliste correspond à un volume IRM, en général obtenu en fusionnant plusieurs examens, associé à un volume dont chacun des voxels contient les probabilités d'appartenir aux classes à segmenter.

Cette approche a été utilisée par [Makropoulos *et al.* \(2014\)](#) pour segmenter le cerveau d'enfants prématurés en 50 régions. La segmentation est réalisée par les étapes suivantes : recaler les images de l'atlas ALBERTs sur le volume à segmenter, appliquer les transformations déterminées aux annotations de l'atlas pour créer un atlas probabiliste, initialiser la segmentation des tissus à partir de l'atlas probabiliste et la raffiner par la méthode des k-means, utiliser l'algorithme espérance-maximisation ([Van Leemput *et al.* \(1999\)](#)), abrégé comme EM, pour estimer le maximum de vraisemblance par rapport à l'hypothèse que les intensités de chaque structure ont une distribution gaussienne et enfin effectuer un post-traitement, qui utilise une régularisation par champ de Markov ([Fischl *et al.* \(2002\)](#)), pour supprimer les faux-positifs isolés. Cette méthode est représentée sur la Figure.III.2.

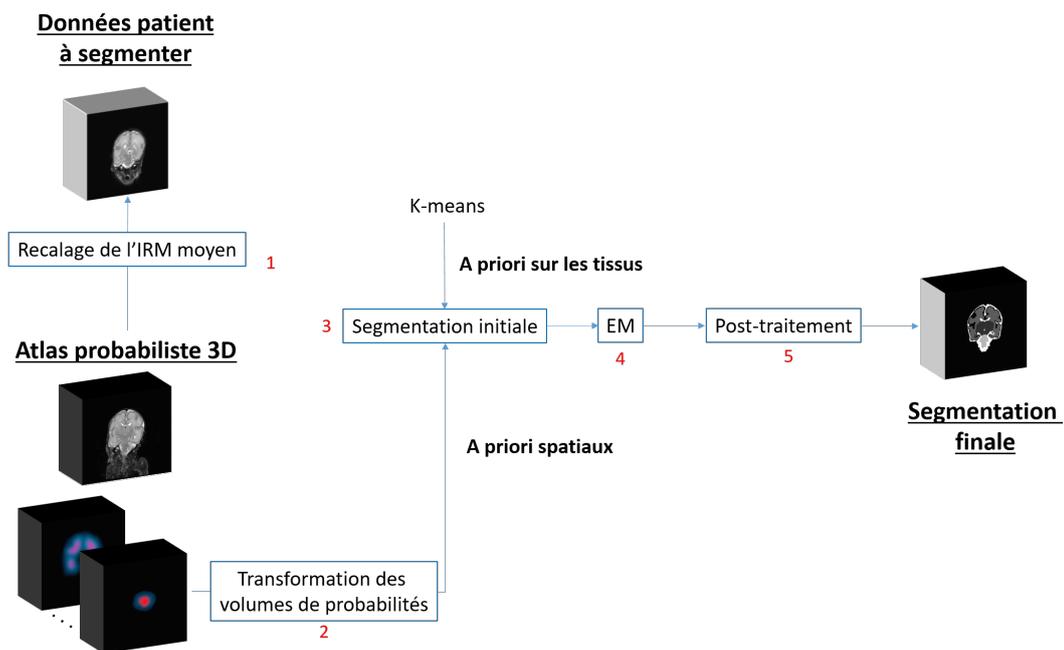


FIGURE III.2 – Segmentation à partir d'un modèle paramétrique et d'a priori spatiaux obtenus grâce à un atlas probabiliste.

Des approches similaires ont été proposées par [Prastawa *et al.* \(2005\)](#), [Melbourne *et al.* \(2012\)](#), [Song *et al.* \(2006\)](#), [Liu *et al.* \(2016b\)](#) et [Cardoso *et al.* \(2013\)](#). D'autres méthodes que l'algorithme EM, comme les graph-cuts ([Song *et al.* \(2006\)](#)), peuvent être utilisées pour affiner la segmentation initiale. Certains auteurs utilisent aussi directement des atlas probabiliste en libre accès, plutôt que d'en créer à partir de plusieurs atlas d'annotations manuelles.

III.1.3 Segmentation du SVC : ventriculomégalie et hydrocéphalie

Les algorithmes présentés dans les deux sections précédentes étaient dédiés à la segmentation de tissus ne présentant pas d'anomalies. Or nous avons vu que des dilatations ventriculaires pouvaient survenir chez le prématuré. Il est donc important de développer des algorithmes capables de segmenter des SVC dilatés.

Ce cas a été étudié par [Qiu *et al.* \(2015b\)](#). Ceux-ci ont créé leur propre base de données, elle est composée d'images d'enfants prématurés parmi lesquels 19 sont atteints de VM et 7 d'hydrocéphalie. La méthode proposée pour segmenter le SVC comprend deux étapes principales : une première où un recalage multi-atlas est effectué pour obtenir une segmentation initiale et une deuxième durant laquelle le contour de cette segmentation évolue en utilisant une méthode basée sur les level-set ([Osher *et Sethian* \(1988\)](#)). Du leave-one-out a été utilisée pour valider leur méthode, c'est à dire que la segmentation de chaque patient a été effectuée en considérant le reste de la base de données comme le multi-atlas permettant d'obtenir la segmentation initiale.

Cette méthode a permis de réaliser la segmentation automatique du SVC chez des patients présentant une dilatation ventriculaire. Elle est néanmoins trop lente pour être utilisée en clinique : les temps de segmentation reportés vont de 1,3 à 0,5 heures respectivement chez les patients atteints de VM et d'hydrocéphalie.

III.1.4 Classification

Des méthodes de classification ont aussi été utilisées pour segmenter les tissus cérébraux chez le prématuré. Ces méthodes attribuent une étiquette à chaque pixel d'une image en se basant sur un ensemble de caractéristiques qui sont à définir. L'attribution d'une étiquette à chaque pixel est fait par un classifieur, qui est une fonction mathématique définie par :

$$f: \mathcal{X} \rightarrow \mathcal{Y} \quad (\text{III.1})$$

\mathcal{Y} est l'espace de sortie du classifieur, ses éléments correspondent à une ou plusieurs valeurs à partir desquelles on peut attribuer une étiquette. Par exemple, un classifieur peut donner les probabilités d'appartenir à chaque classe, dans un problème à C classes on aura alors $\mathcal{Y} = \{0, 1\}^C$. \mathcal{X} correspond à l'ensemble des caractéristiques qui sont utilisées pour réaliser la classification, il peut par exemple être égal à \mathbb{R}^N .

Pour obtenir un classifieur qui puisse être utilisé en clinique, il faut au préalable l'entraîner sur un jeu de données qui constitue une base d'entraînement. L'objectif étant que le classifieur soit capable de généraliser son apprentissage à de nouveaux cas. Ces méthodes ont récemment gagnées en popularité. On peut distinguer deux approches : une première où les caractéristiques utilisées par le classifieur sont définies manuellement et une deuxième où les caractéristiques sont apprises automatiquement à partir des données à segmenter, on parle dans ce dernier cas de representation learning. Ces deux approches sont présentées dans les parties suivantes.

Classification à partir d'une définition manuelle des caractéristiques

Pour entraîner un classifieur, l'approche la plus classique est de définir manuellement les caractéristiques définissant son espace d'entrée. Comme il n'est pas possible de connaître, à l'avance, les caractéristiques qui seront utiles à un classifieur pour segmenter une image, on définit un grand nombre de caractéristiques en première intention et on effectue une sélection par la suite. Pour ce problème, les caractéristiques typiquement définies sont : des images T1 et T2 du cerveau, des caractéristiques de position (coordonnées par rapport à la boîte englobante définie par le cerveau et distance cartésienne à celle-ci) et les dérivées des images T1 et T2 dans toutes les directions. Différents classifieurs ont été utilisés pour réaliser la segmentation des tissus cérébraux chez le prématuré, parmi ceux-ci, on retrouve l'algorithme des K plus proches voisins (Anbeek *et al.* (2013), Anbeek *et al.* (2008), Moeskops *et al.* (2015)), les machines à vecteurs de support (Sanroma *et al.* (2016)) et les forêts aléatoires (Wang *et al.* (2015)). Cette approche est représentée sur la Figure.III.3.

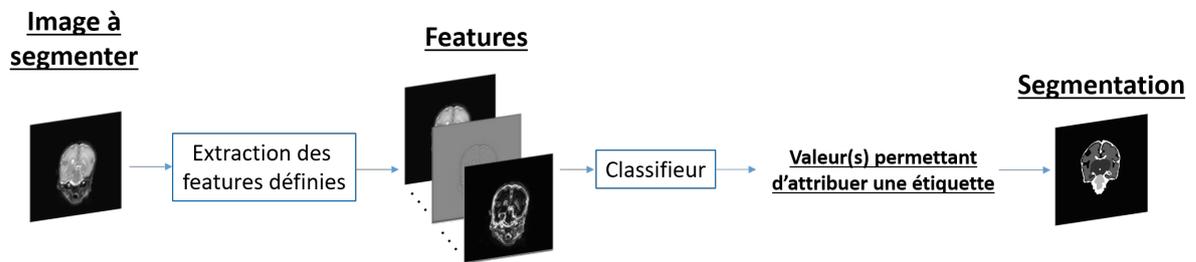


FIGURE III.3 – Méthode classique pour segmenter une image avec un classifieur.

Une fois le classifieur entraîné, cette méthode est complètement automatique. L'inconvénient de cette approche est que les caractéristiques sont définies manuellement, on est donc jamais certain d'avoir défini celles qui sont les plus pertinentes.

Representation learning

Pour ne pas avoir à définir manuellement les caractéristiques qui représentent les données, il est possible d'utiliser des méthodes de representation learning. Ces méthodes permettent d'apprendre automatiquement des représentations, plus ou moins abstraites, des données que l'on souhaite segmenter. Celles-ci sont ensuite utilisées par le classifieur pour attribuer une étiquette à chaque voxel. Cette approche est représentée sur la Figure.III.4 :

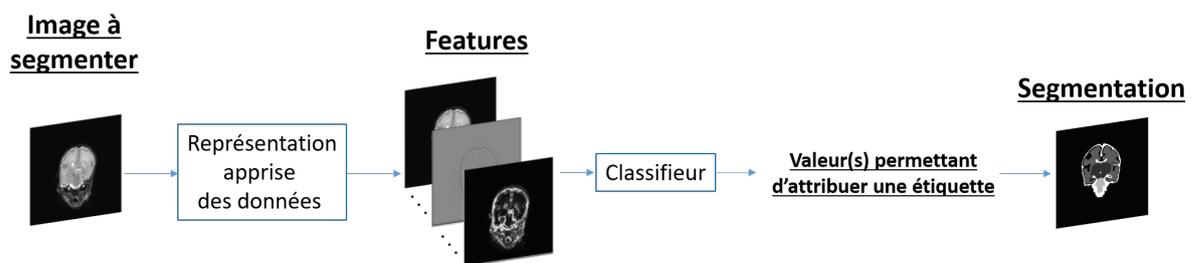


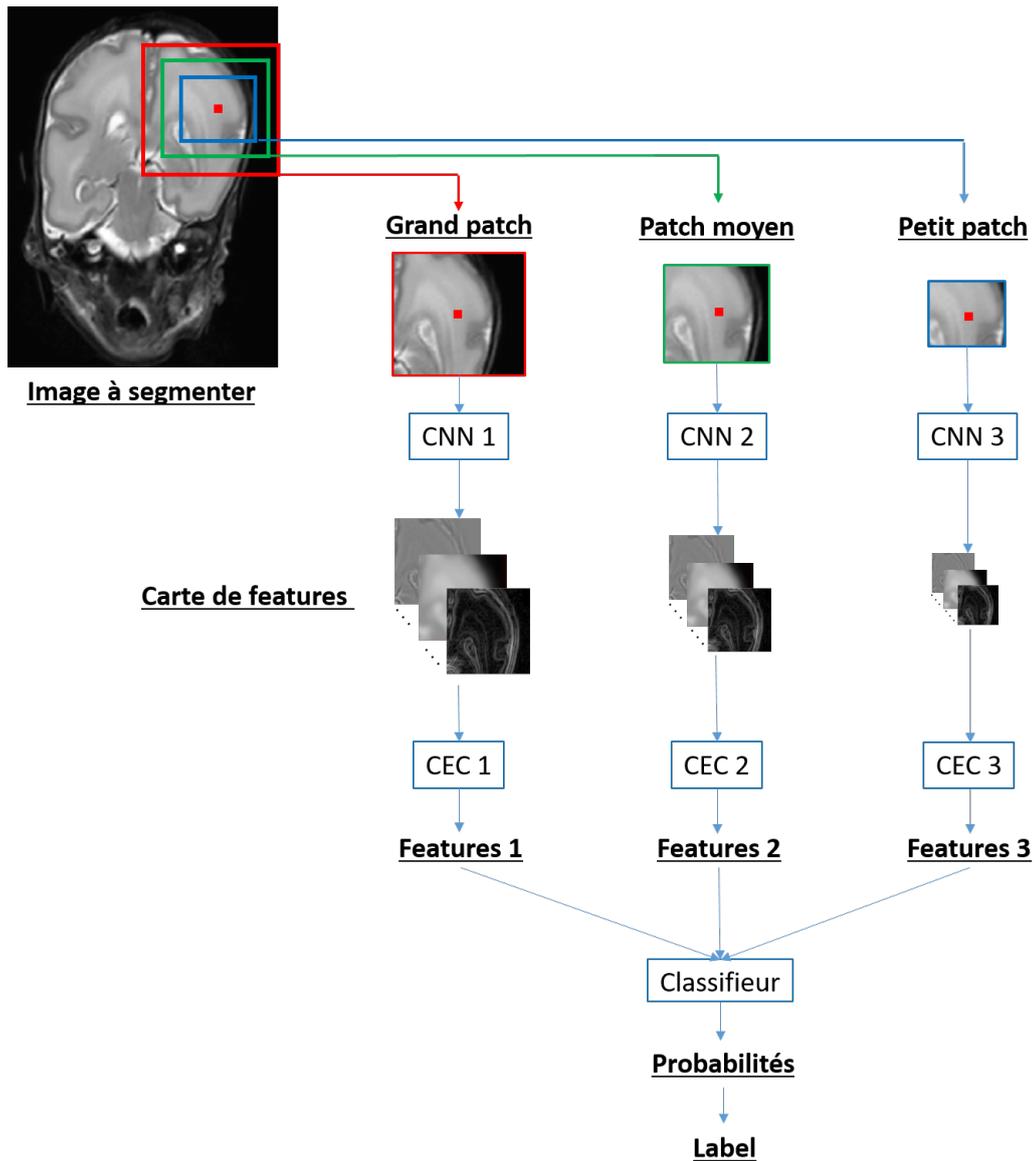
FIGURE III.4 – Approche de representation learning pour segmenter une image avec un classifieur.

En particulier, les méthodes de Deep learning, qui permettent d'apprendre des représentations

très abstraites, ont été grandement utilisées ces dernières années. Parmi ces méthodes, les réseaux de neurones convolutifs (CNN) ont permis de segmenter automatiquement des images complexes avec une précision extrêmement élevée par rapport à l'état de l'art ([Badrinarayanan *et al.* \(2015\)](#), [Long *et al.* \(2015\)](#)), et ce, en particulier dans le domaine de l'imagerie médicale ([Ronneberger *et al.* \(2015\)](#), [Milletari *et al.* \(2016\)](#)).

Pour résoudre ce problème de segmentation, cette approche a été utilisée par [Moeskops *et al.* \(2016\)](#). Dans ce papier, il est proposé un algorithme de segmentation multi-échelle par patch utilisant des CNNs. Le principe de cette approche est d'attribuer une étiquette au pixel qui se situe au centre des patches, mis en entrée du réseau, à partir des caractéristiques obtenues avec plus ou moins de contexte. Les données d'entrées de leur réseau sont trois patches de taille différente (tous centrés sur le même voxel). A partir de ces entrées, des cartes de caractéristiques sont produites en parallèle par trois CNNs. Ces cartes de caractéristiques sont vectorisées par une couche de neurones entièrement connectée (CEC) permettant d'obtenir de nouvelles caractéristiques. Elles sont ensuite utilisées par un classifieur qui calcule la probabilité qu'à le pixel central d'appartenir à chaque structure, ce qui permet de lui attribuer une étiquette (celle correspondant à la plus grande probabilité). La segmentation complète d'une image est réalisée en utilisant cette méthode sur chaque pixel de l'image. Leur réseau est représenté sur la Figure.III.5.

Les réseaux de neurones convolutifs sont actuellement parmi les méthodes les plus performantes pour résoudre les problèmes de segmentation en imagerie médicale. Ils ont également été utilisés pour segmenter le cerveau en IRM chez l'adulte ([de Brebisson *et Montana* \(2015\)](#), [Chen *et al.* \(2018\)](#), [Milletari *et al.* \(2017\)](#))). Elles sont rapides, précises, ne nécessitent pas la définition manuelle de caractéristiques et sont capables d'apprendre des représentations très abstraites des données d'entrée. Il faut néanmoins rester prudent lors de leur utilisation. Les caractéristiques sont apprises automatiquement par le réseau, elles pourraient correspondre à des biais présent dans les images et non à des caractéristiques des structures que l'on veut segmenter. Il faut donc, dans la mesure du possible, toujours visualiser les cartes de caractéristiques pour être sûr que le réseau n'apprenne pas à produire des segmentations à partir de biais présents dans les images (artefacts, ...).

FIGURE III.5 – CNN multi-échelle proposé par Moeskops *et al.* (2016).

III.2 Segmentation en échographie

En échographie, des algorithmes de segmentation automatique ont été proposés uniquement pour le SVC. Il n'y a donc pas de bibliographie sur la segmentation des thalami et des autres structures du cerveau. Cette partie présente les méthodes qui ont été utilisées pour segmenter le SVC.

III.2.1 Segmentation du SVC par contours actifs

Les contours actifs constituent un groupe de méthodes de segmentation, leur objectif est de faire converger un contour fermé vers les bords de la structure que l'on souhaite segmenter. En pratique, un contour est initialisé au niveau de la zone que l'on souhaite segmenter, celui-ci est ensuite modifié de manière à minimiser une fonction de coût. Cette dernière contient en général des termes qui caractérisent le contour par rapport à : l'intensité à l'intérieur et à l'extérieur de celui-ci, sa ré-

gularité et sa localisation. Ces méthodes sont utilisées pour les problèmes mono-classes comme ce problème. Pour la segmentation du SVC en échographie 3D, cette méthode a été utilisée par [Qiu et al. \(2017\)](#) (version automatisée de [Qiu et al. \(2015a\)](#)) et [Boucher et al. \(2018b\)](#) (version étendue de [Boucher et al. \(2018a\)](#)), ces approches vont être présentées dans les deux paragraphes suivants.

Dans [Qiu et al. \(2017\)](#) la base de données US 3D contient à la fois des patients non pathologiques, touchés par une VM ou présentant une hydrocéphalie, leur âge est compris entre 37 et 42 SA. Pour ne pas avoir à initialiser manuellement le contour du SVC comme dans [Qiu et al. \(2015a\)](#), les auteurs se sont inspirés des méthodes de segmentation multi-atlas utilisées en IRM, ce qui leur a permis d'obtenir automatiquement un contour initial. Pour jouer le rôle du multi-atlas, ils ont utilisé leur propre base de données car il n'existe pas d'atlas dédiés à cette problématique en échographie et les atlas IRMs ne contiennent pas de patients pathologiques. Il était composé de segmentations manuelles du SVC et de volumes de congruence de phase ([Kovesi et al. \(1997\)](#)) obtenus à partir des données US 3D. De la même manière que pour la segmentation du SVC en IRM ([Qiu et al. \(2015b\)](#)), ils ont effectués du leave-one-out pour valider leur méthode.

La méthode employée pour segmenter le SVC est la suivante : déterminer les transformations permettant de recalculer les atlas de congruence de phase sur le volume de congruence de phase du patient à segmenter, appliquer ces transformations aux annotations manuelles en sélectionnant 50% des meilleurs atlas, fusionner les segmentations transformées pour obtenir une segmentation initiale, extraire son contour, optimiser ce contour par une méthode de contour actif (level-set) et enfin remplir ce contour pour obtenir la segmentation finale. Cette méthode est représentée sur la [Figure.III.6](#).

Cet article est le premier à proposer une méthode pour segmenter automatiquement le SVC chez le prématuré dans des échographies 3D. Ces segmentations ont été obtenues avec une bonne précision. Néanmoins, la méthode proposée reste trop lente pour qu'il puisse être envisagé de l'utiliser en clinique : un temps de segmentation de 54 minutes par volume a été reporté, la grande majorité de celui-ci correspond au temps nécessaire au recalage des atlas.

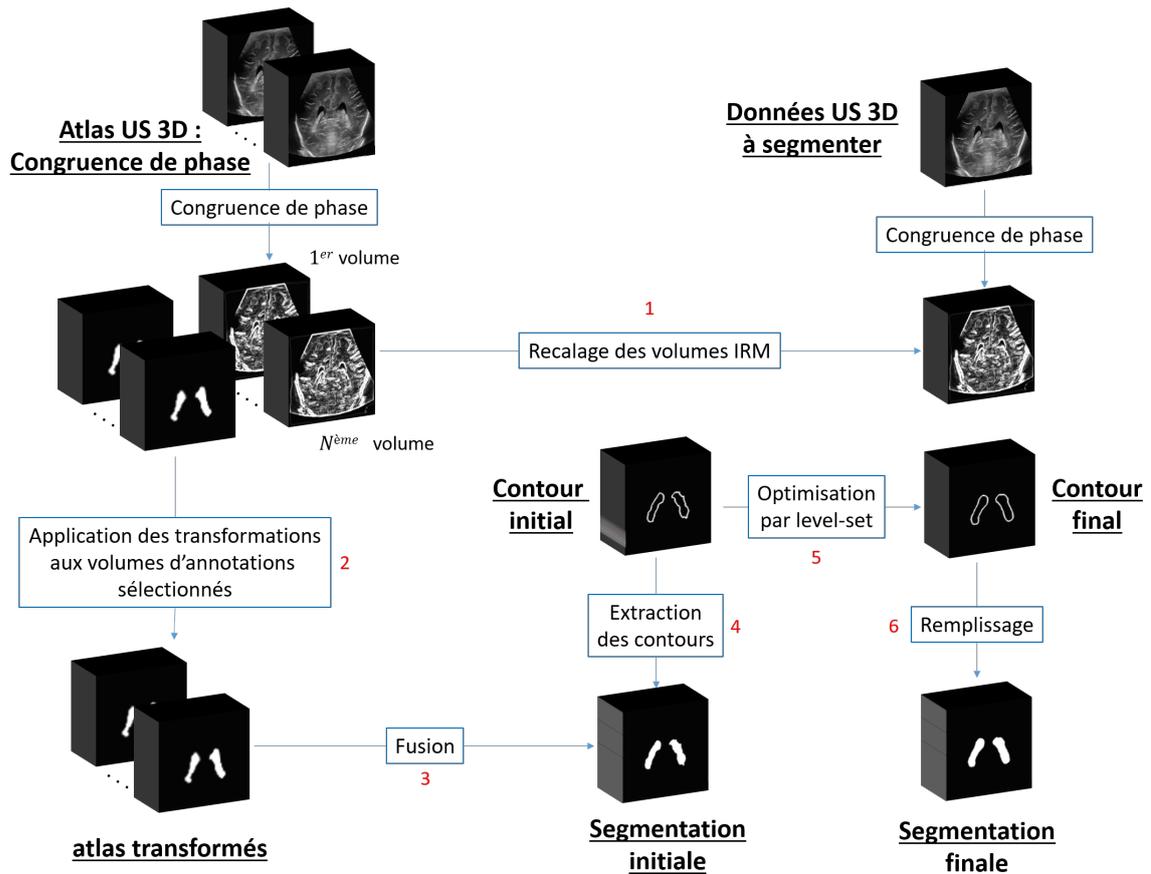


FIGURE III.6 – Segmentation du SVC dans des volumes US 3D, méthode proposée par Qiu *et al.* (2017).

Boucher *et al.* (2018b) ont utilisé une base de données US 3D qui contient 33 patients âgés de 3 à 11 mois. De la même manière que Qiu *et al.* (2017), ils ont utilisé une approche de recalage multi-atlas pour initialiser une segmentation du SVC qui sera affinée par la suite. Le multi-atlas qu'ils ont utilisés était un regroupement de plusieurs atlas IRM : le multi-atlas ALBERTs (Gousias *et al.* (2012)), un atlas dédié aux patients de un an (Altaye *et al.* (2008)) et un atlas réalisé pour les patients de 2 à 5 mois (Fonov *et al.* (2009)). Leur base de données semble contenir moins de variabilité qui serait due à des dilations ventriculaires que la base de Qiu *et al.* (2017), il est probable qu'ils se soient servis d'un multi-atlas IRM plutôt que d'en construire un à partir leur base de donnée US pour cette raison.

Les étapes suivantes ont été mises en œuvre pour réaliser les segmentations : recalcr (transformation rigide puis élastique) les atlas IRM sur le patient à segmenter et sélectionner les quatre meilleurs recalages (par rapport à la métrique utilisée), appliquer les transformations aux volumes d'annotations manuelles, fusionner ces volumes transformés pour créer une carte de probabilité, appliquer un seuil de 0.8 à la carte de probabilité pour obtenir une segmentation initiale et utiliser des opérations morphologiques pour la lisser, créer un maillage à partir de cette segmentation, optimiser ce dernier par rapport à un critère d'énergie, remplir le maillage pour obtenir une segmentation finale. Cet algorithme est présenté sur la Figure.III.7.

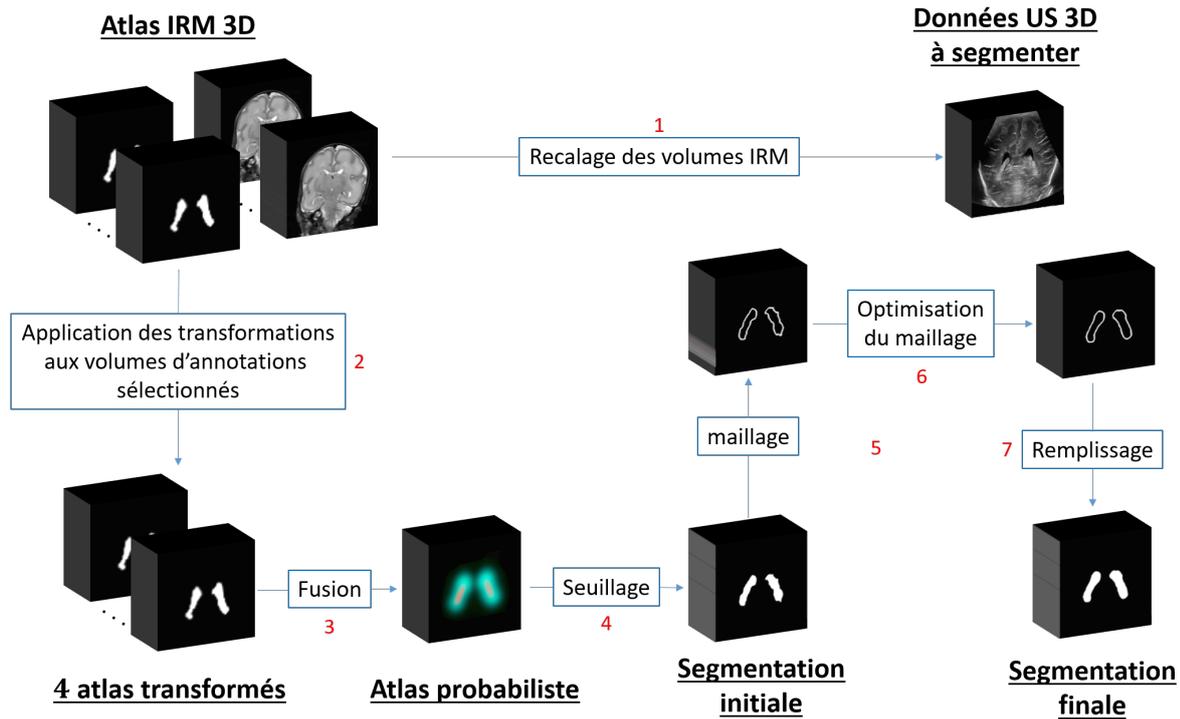


FIGURE III.7 – Segmentation du SVC dans des volumes US 3D, méthode proposée par Boucher *et al.* (2018b).

Cette méthode a permis d'obtenir des segmentations précises du SVC dont le volume est en accord avec les mesures réalisées en IRM. Néanmoins, d'après les résultats fournis dans leur article, ceux-ci n'avaient pas de cas de patients présentant une ventriculomégalie, au moment de l'examen, dans leur base de données. Il est donc possible que leur méthode, notamment l'étape de recalage multi-atlas, ne puissent pas être utilisée pour segmenter des cas pathologiques. De plus, aucun temps de segmentation n'a été reporté dans leur article.

Les articles de Boucher *et al.* (2018b) et Qiu *et al.* (2017) sont, au regard de notre problématique, les travaux les plus aboutis que l'on trouve dans la littérature : les méthodes proposées permettent de segmenter automatiquement le SVC dans des volumes US 3D. D'autres articles sur la segmentation du SVC existent dans la littérature. En échographie 3D, nous avons proposé, au début de cette thèse, une méthode de segmentation semi-automatique qui est basée sur les level-set (Sciolla *et al.* (2016)). Elle nous a permis d'obtenir des segmentations précises mais n'est pas automatique : elle doit être initialisée manuellement par l'utilisateur et nécessite l'ajustement de plusieurs hyperparamètres. En échographie 2D, la segmentation automatique des VL a été réalisée par Tabrizi *et al.* (2018) en utilisant une méthode basée sur les contours actifs. L'approche proposée se limite néanmoins à la segmentation des cornes temporales dans des images acquises en coupe coronale où le foramen de Monro est visible. Cette méthode n'est donc pas adaptée pour une segmentation du SVC dans des données 3D. Ces deux dernières méthodes sont moins abouties que celles proposées par Boucher *et al.* (2018b) et Qiu *et al.* (2017), c'est pourquoi nous ne sommes pas rentrés dans les détails de leur mise en œuvre.

III.2.2 Segmentation du SVC par deep learning

En échographie 2D, une approche basée sur de l'apprentissage profond a récemment été proposée par [Wang *et al.* \(2018\)](#) pour segmenter le SVC.

La base de données qu'ils ont utilisées était composée de 687 images échographiques 2D. Pour effectuer la segmentation du SVC, ils ont mis en œuvre une méthode de transfer learning ([Pan *et Yang* \(2009\)](#)). Leur CNN est composé de deux parties : la première est un réseau dense-net ([Huang *et al.* \(2017b\)](#)) pré-entraîné sur une base contenant des millions d'images, il est utilisé pour extraire des caractéristiques de hauts niveaux (abstraites) des image US en entrée du réseau, la deuxième partie est un CNN dont les paramètres ont été optimisés à partir de la base de données US, son rôle est de décoder les caractéristiques fournies par le dense-net pour produire une carte de probabilité qui permettra d'obtenir la segmentation du CVS. Cette méthode est schématisée sur la Figure.III.8.

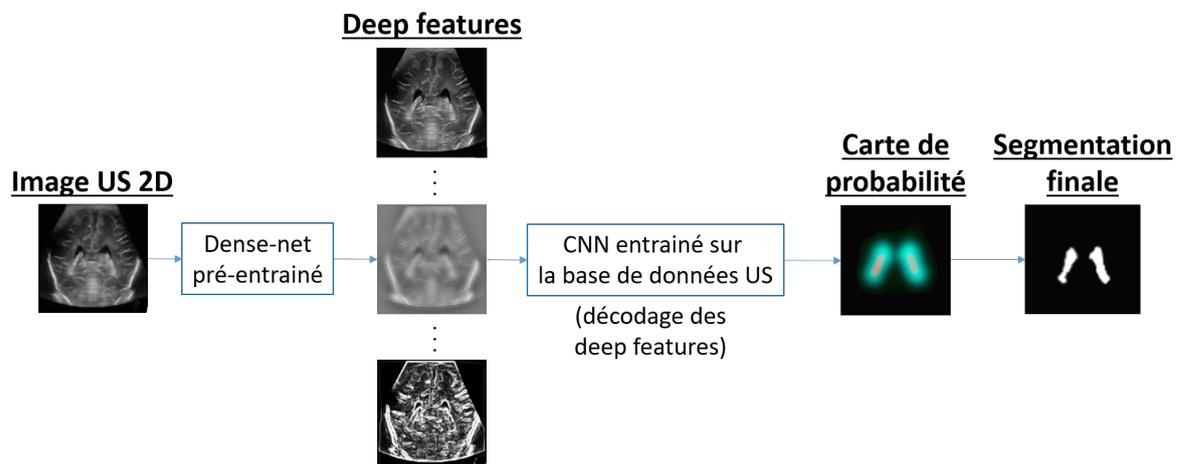


FIGURE III.8 – Segmentation du SVC dans des images échographiques 2D, méthode proposée par [Wang *et al.* \(2018\)](#).

Leur méthode est précise et extrêmement rapide : un temps de 22 ms a été reporté pour segmenter une image de taille 600×600 . Il est néanmoins difficile d'évaluer clairement la précision de leur méthode car peu d'informations sont disponibles sur leur base de données, (patients pathologiques, nombre de patients, âge des patients, répartition des patients entre les jeux de données d'entraînement et de test, ...). L'approche reste néanmoins très intéressante pour segmenter des images 2D. En revanche, il ne serait pas possible d'extraire des caractéristiques à partir de volume 3D avec cette méthode car le dense-net est pré-entraîné et fonctionne sur des bases d'images 2D.

III.3 Conclusion

Dans ce chapitre, nous avons présenté les algorithmes de segmentation qui ont été utilisés pour segmenter les tissus cérébraux chez les enfants prématurés. La première section se focalisait sur la segmentation d'images obtenues par IRM et la deuxième sur celle des échographies. Dans l'introduction, nous avons rappelé que notre objectif était de développer une méthode qui permette de segmenter automatiquement le SVC et les Thalami dans des images échographiques 3D, dans un temps clinique. En outre, la méthode choisie doit si possible pouvoir être étendue à la segmentation des autres tissus cérébraux dans le futur. L'utilisation potentielle des algorithmes présentés est discutée dans les prochains paragraphes.

Les approches de segmentation multi-atlas nous paraissent difficiles à mettre en œuvre pour répondre à nos objectifs. Comme évoqué précédemment, il n'existe pas d'atlas en libre accès en échographie 3D. De plus, nous avons dans notre base de données des patients ayant une dilatation ventriculaire, il serait donc difficile d'utiliser des atlas IRM, comme [Boucher *et al.* \(2018b\)](#)), pour mettre en œuvre cette méthode. Il faudrait donc que l'on crée notre propre atlas, de la même manière que [Qiu *et al.* \(2017\)](#). Néanmoins, nous avons vu que cette méthode était très lente et donc pas adaptée aux contraintes cliniques. Pour ces raisons nous avons décidé de ne pas nous orienter vers ces méthodes.

L'usage de méthodes paramétriques nous paraît également compliqué pour notre application. Premièrement car les a priori spatiaux qu'elles utilisent sont obtenus grâce à un atlas probabiliste, dont l'utilisation n'est pas souhaitable comme nous l'avons dit dans le paragraphe précédent. Deuxièmement car il n'existe pas de modèle d'intensité pour les structures cérébrales en échographie 3D. Il serait possible de rechercher une distribution statistique parmi celles classiquement utilisées en US ([Destrempe *et Cloutier* \(2010\)](#)), cela aurait peut être permis d'identifier une distribution connue, mais nous ne nous sommes pas orienté dans cette voie.

En revanche, les méthodes de classification nous paraissent adaptées à notre problème. Leur utilisation nécessite la définition de caractéristiques à partir desquelles la segmentation pourra être obtenue. Compte tenu de la difficulté de notre problème et de la capacité du deep learning à apprendre des représentations abstraites, nous nous sommes orientés vers les méthodes utilisant des réseaux de neurones convolutionnels. Sous réserve de déterminer une architecture qui soit adaptée à notre problème, ces méthodes sont extrêmement puissantes et permettent aujourd'hui de réaliser la segmentation de gros volume de données en quelques secondes.

Pour un problème de segmentation du cerveau, il nous a tout de même paru important d'introduire des a priori spatiaux. L'anatomie du cerveau peut varier énormément d'un individu à l'autre mais les structures sont toujours situées dans les mêmes régions de l'encéphale. Nous souhaitons éviter de recourir à un atlas et avons privilégié la piste d'un apprentissage de ces a priori spatiaux par les CNNs. Ces travaux seront présentés dans la suite de ce mémoire (Chapitre VI).

CHAPITRE IV

RECONSTRUCTION 3D D'ÉCHOGRAPHIE TRANSFONTANELLAIRE 2D

Contents

IV.1 Introduction	63
IV.1.1 Données cliniques utilisées pour créer les volumes échographiques	63
IV.1.2 Hypothèses sur le mouvement d'acquisition	64
Rappel des paramètres à estimer	64
Formulation d'hypothèses sur le mouvement d'acquisition	65
IV.1.3 Travail réalisé sur les fichiers DICOM	66
IV.2 Recalage deux à deux des images contenues dans les séquences dynamiques	68
IV.2.1 Méthode	68
Recalage d'un couple d'image (I_t, I_{t+1})	68
Recalage deux à deux des images d'une séquence dynamique	69
Optimisation du temps de calcul	70
Résumé des étapes permettant le recalage deux à deux des images d'une SD.	70
IV.2.2 Expériences et résultats	71
IV.2.3 Conclusion	72
IV.3 Positionnement des images dans la grille de reconstruction	73
IV.3.1 Grille de reconstruction	73
Définition de la grille de reconstruction	73
Placement des images dans la grille de reconstruction	74
IV.3.2 Modèle d'acquisition paramétrique	75
IV.3.3 Optimisation des paramètres des modèles paramétriques	78
Définition de I_{sag}	78
Minimisation de la distance aux points de référence	81
Distance entre les ensembles de points de repère après optimisation	83
Valeurs des paramètres après optimisation de f_1	84
Résolution du problème tel qu'il était formulé initialement	87
IV.3.4 Validation des modèles de reconstruction	87
Validation à partir des images I_{ref}	87
Validation sur fantôme	92

IV.3.5 conclusion	94
IV.4 Interpolation des volumes reconstruits	95
IV.4.1 Méthodes	95
PNN	96
VNN	96
Interpolation à partir d'un voisinage sphérique	97
Interpolation à partir des plus proches voisins	97
IV.4.2 Temps d'interpolation et résultats qualitatifs	98
Temps d'interpolation	99
Résultats qualitatifs	100
IV.4.3 Conclusion	103
IV.5 Positionnement dans le repère anatomique	104
IV.5.1 Méthode	104
IV.5.2 Résultats qualitatifs	105
IV.6 Débruitage	108
IV.7 Créations de volume échographiques à partir d'ETFs réalisées avec d'autres sondes	110
IV.8 Conclusion et perspectives	112

CHAPITRE IV

Ce chapitre présente l'algorithme de reconstruction 3D d'échographie transfontanellaire 2D que nous avons développé pendant cette thèse. Cet algorithme nous a permis d'obtenir des données échographiques 3D de haute qualité, dans lesquelles il est possible de distinguer la plupart des structures cérébrales, comme c'est le cas sur les images obtenues avec la sonde 9L4. Ce sont ces volumes qui ont été utilisés pour travailler à la résolution des problèmes de segmentation du SVC et des thalami.

Les données mises en entrée de cet algorithme sont les données brutes fournies par le CH d'Avignon. A partir de celles-ci, la création d'un volume échographique orienté selon les plans anatomiques de référence est réalisée en six étapes (Figure.IV.1) : une étape d'extraction de la zone où se situe les échographie dans les images brutes (1), une étape de recalage deux à deux des images de la SD (2), une étape de positionnement des images dans la grille de reconstruction à partir d'un modèle d'acquisition (3), une étape d'interpolation de la valeur des voxels de la grille de reconstruction (4), une étape de positionnement dans un repère anatomique, orienté selon les plans anatomiques de référence (5) et pour terminer une étape de débruitage (6).

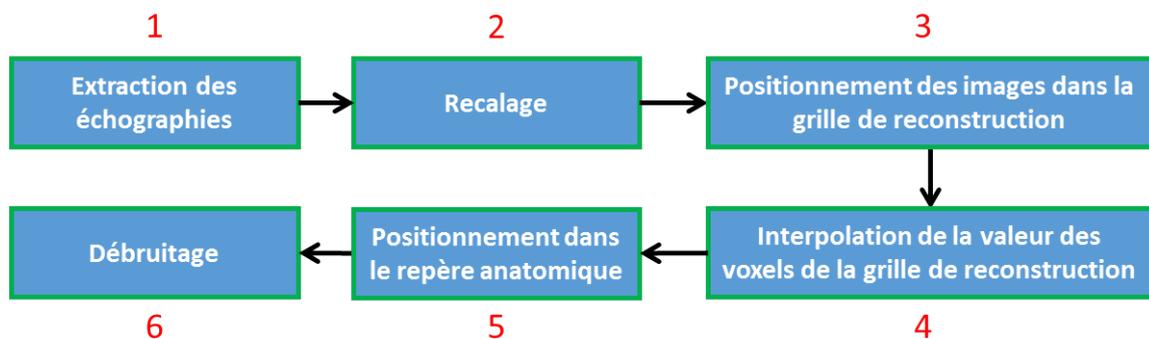


FIGURE IV.1 – Étapes de l'algorithme permettant de créer un volume échographique à partir d'une ETF 2D.

A notre connaissance, c'est la première fois que ce type de méthode est utilisé pour créer des volumes échographiques du cerveau du prématuré. Il présente les avantages suivants :

- La reconstruction des volumes échographiques ne requiert pas l'utilisation de matériel supplémentaire lors de la réalisation des ETFs.
- Les paramètres de l'algorithme ne dépendent pas de la sonde utilisée pour réaliser l'acquisition, il peut donc, a priori, être utilisé pour créer des volumes échographiques à partir d'ETFs qui ne sont pas réalisées avec la sonde 9L4.

Les contributions principales présentées dans ce chapitre sont :

- Le développement d'un algorithme permettant de créer un volume échographique du cerveau du prématuré à partir d'une ETF 2D, qui ne nécessite pas l'utilisation de matériel supplémentaire lors des acquisitions et qui peut être utilisé sur des acquisitions réalisées avec différentes sondes.
- La développement d'un modèle d'acquisition correspondant à une rotation d'axe mobile et d'une méthode d'optimisation de ses paramètres.

La première partie de ce chapitre présente les données cliniques qui ont été utilisées pour créer les volumes échographiques et les hypothèses que nous avons fait sur le mouvement d'acquisition de manière à réduire le nombre de paramètres à estimer pour positionner les images les unes par rapport aux autres. Les parties suivantes sont dédiées à la description et à la validation des méthodes utilisées lors des six étapes permettant la création des données échographiques 3D.

IV.1 Introduction

IV.1.1 Données cliniques utilisées pour créer les volumes échographiques

Notre base d'images échographiques 3D a été créée à partir de 25 SDs obtenues par ETF 2D. Ces acquisitions ont toutes été réalisées avec une sonde 9L4 sur des enfants nés prématurés au centre de néonatalogie d'Avignon. Comme nous l'avons vu au chapitre 1, les images ont été acquises à travers la fontanelle antérieure en effectuant un balayage angulaire. Cet examen et les données échographiques qu'il permet d'obtenir sont respectivement représentés sur la Figure.IV.2.a et la Figure.IV.2.b.

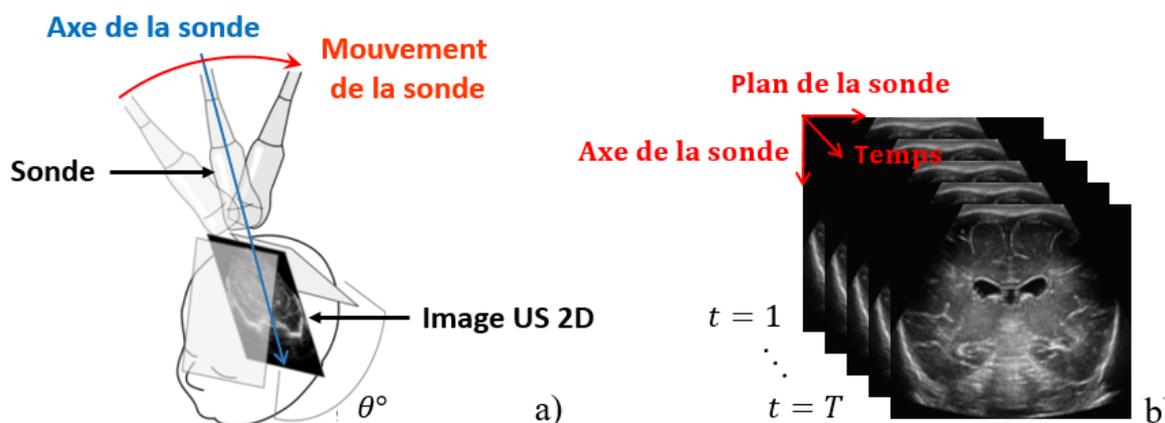


FIGURE IV.2 – Balayage angulaire manuel d'angle θ réalisé au niveau de la fontanelle antérieure (a) et séquence dynamique obtenue après la réalisation de ce balayage (b)¹.

Plus précisément, le mouvement d'acquisition réalisé lors d'une ETF correspond à une rotation d'axe mobile. Il permet de s'adapter au mieux à l'anatomie de la fontanelle et d'obtenir des images, dans des coupes proches du plan coronal, de la majeure partie de l'encéphale.

Les images contenues dans les SDs ont toutes la même taille : 567×763 pixels. Ces deux valeurs correspondent respectivement au nombre de pixels contenus dans la hauteur et dans la largeur des images. En revanche, le nombre d'images dans une SD varie d'une acquisition à l'autre. Il est en moyenne de $T = 249 \pm 69$, au minimum de 139 et au maximum de 482.

Un repère orthonormé $R_1 = (O_1, u_{x_1}, u_{y_1}, u_t)$ a été associé à chaque SD de manière à se repérer au sein de celles-ci. Il est représenté sur la Figure.IV.3. Son origine $O_1 = (0, 0, 0)$ est définie canoniquement par la première image de la SD et les coordonnées des pixels dans cette image, l'axe x_1 est orienté dans le sens crânio-caudal, l'axe y_1 dans le sens gauche-droite (par rapport à la tête de l'enfant) et l'axe t dans le sens antéro-postérieur.

1. Images issues de notre base de données.

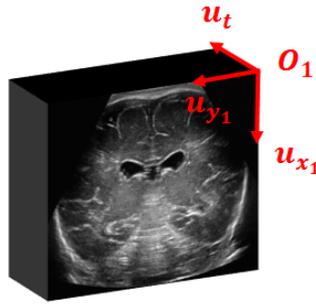


FIGURE IV.3 – Définition d'un repère orthonormé R_1 dans chaque SD.

IV.1.2 Hypothèses sur le mouvement d'acquisition

Rappel des paramètres à estimer

Dans le Chapitre I, nous avons vu qu'estimer les 3 paramètres angulaires ($\alpha_t, \beta_t, \gamma_t$) et les 3 paramètres de position ($(d_x, d_y, d_z) = \overrightarrow{O_t O_{t+1}}$), représentés sur la Figure IV.4, était suffisant pour déterminer la position relative de deux images successives appartenant à une SD. Il a également été établi que déterminer les 3 paramètres angulaires ϕ, φ, ψ , visibles sur la Figure IV.5.c, était suffisant pour orienter les volumes échographiques obtenus, selon les plans anatomiques de référence (Figure IV.5.b), dans la grille de reconstruction.

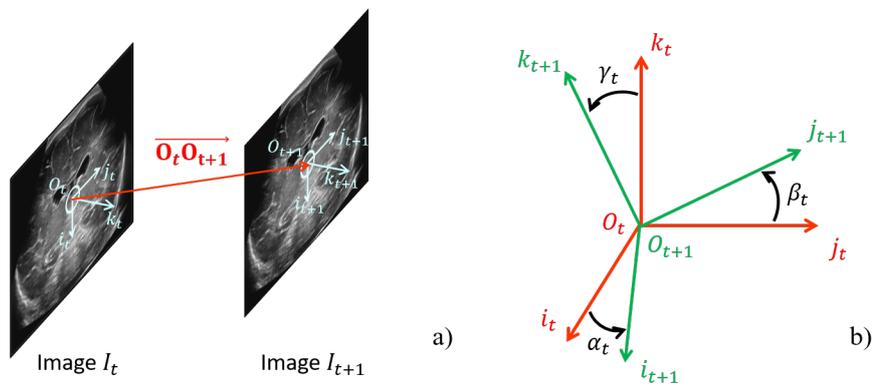


FIGURE IV.4 – Angles entre deux images successives d'une SD (b) et distance entre leur centre respectif (a).

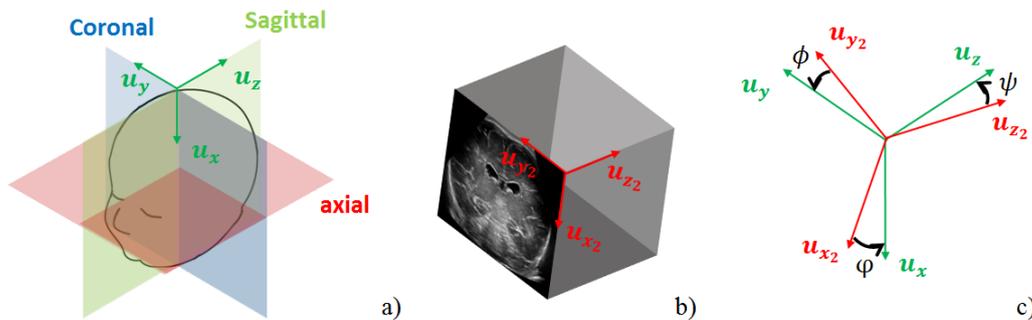


FIGURE IV.5 – Plans anatomiques de référence (a), volume échographique interpolé (b) et angles à estimer pour orienter le volume par rapport aux plans anatomiques de référence (c).

Les bases orthonormées (u_x, u_y, u_z) (Figure.IV.5.a) et $(u_{x_2}, u_{y_2}, u_{z_2})$ (Figure.IV.5.b) permettent respectivement de décrire les volumes orientés selon les plans anatomiques de références et les volumes obtenus dans la grille de reconstruction.

Formulation d'hypothèses sur le mouvement d'acquisition

De manière à réduire le nombre de paramètres à estimer pour placer les images dans la grille de reconstruction, nous avons fait deux hypothèses sur le mouvement d'acquisition :

1. Il n'y a pas de rotation autour de l'axe de la sonde pendant l'acquisition.
2. Les images acquises par la sonde sont orthogonales au plan sagittal.

Ces hypothèses et leurs conséquences sont décrites dans les deux paragraphes suivants.

Première hypothèse sur le mouvement d'acquisition. L'axe de la sonde (Figure.IV.2) est colinéaire avec l'axe défini par i_t (Figure.IV.5) pour chaque image acquise lors d'une ETF. L'hypothèse 1 signifie donc qu'il n'y a pas de rotation autour de cet axe entre l'acquisition de deux images. Des mouvements d'acquisition qui respectent et ne respectent pas cette hypothèse sont respectivement représentés sur les Figure.IV.6.a et Figure.IV.6.b.

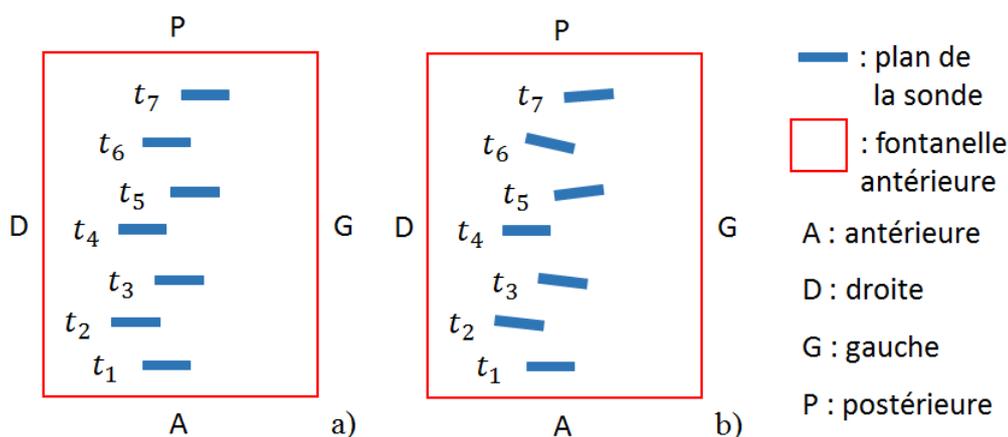


FIGURE IV.6 – Mouvements d'acquisition respectant l'hypothèse 1 (a) et ne respectant pas l'hypothèse 1 (b).

En faisant cette hypothèse, on considère que des rotations sont uniquement possibles autour des axes définis par j_t et k_t entre l'acquisition de deux images. On note γ_t l'angle de rotation autour de l'axe défini par j_t et β_t l'angle de rotation autour de l'axe défini par k_t . En effectuant ces deux rotations successivement : $\beta_t = \widehat{(j_t, j_{t+1})}$, $\gamma_t = \widehat{(k_t, k_{t+1})}$ et α_t s'exprime uniquement en fonction de β_t et γ_t . Cette première hypothèse permet donc de faire passer de 3 à 2 le nombre de paramètres angulaires à estimer pour déterminer la position relative de deux images.

De plus, si on fait l'hypothèse que les angles γ_t et β_t sont petits, on peut les estimer séparément, respectivement dans un plan défini par les couples de vecteurs (i_t, k_t) et (i_t, j_t) .

Deuxième hypothèse sur le mouvement d'acquisition. Des mouvements d'acquisition qui respectent et ne respectent pas la deuxième hypothèse sont respectivement représentés sur les Figure.IV.7.a et Figure.IV.7.b.

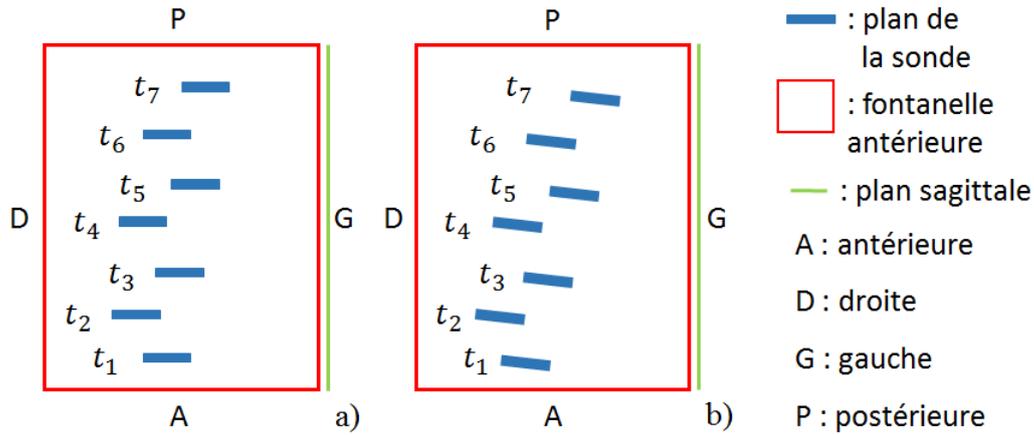


FIGURE IV.7 – Mouvements d’acquisition respectant l’hypothèse 2 (a) et ne respectant pas l’hypothèse 2 (b).

Cette hypothèse signifie que le déplacement global de la sonde s’effectue dans une direction définie par un vecteur qui est orthogonal à u_y , il n’est donc pas nécessaire d’effectuer une rotation autour de l’axe défini par u_{x_2} pour orienter le volume interpolé selon les plans anatomiques de référence. De façon similaire à la première hypothèse, si on appelle respectivement ϕ et ψ les angles de rotations autour des axes définis par u_{z_2} et u_{y_2} , on obtient $\phi = (\widehat{u_y, u_{y_2}})$, $\psi = (\widehat{u_z, u_{z_2}})$ et φ qui s’exprime uniquement en fonction de ϕ et ψ . Cette deuxième hypothèse fait donc également passer de 3 à 2 le nombre de paramètres angulaires à estimer pour orienter le volume échographique créé selon les plans anatomiques de référence.

En faisant l’hypothèse que l’angle de rotation autour de l’axe défini par u_{z_2} (ϕ) est proche de zéro, on peut estimer ψ dans un plan défini par les vecteurs u_{x_2} et u_{z_2} . On estime ensuite ϕ dans un plan défini par les vecteurs u_{y_2} et u_{x_2} après avoir effectué un rotation d’angle ψ (il n’est en général pas proche de zéro).

IV.1.3 Travail réalisé sur les fichiers DICOM

Les données qui ont été fournies par le centre de néonatalogie du CH d’Avignon sont au format DICOM. C’est le format standard international dans lequel sont enregistrés les examens médicaux réalisés en clinique. Dans ce cas, les données contiennent une image dans laquelle se trouve l’acquisition réalisée et des métadonnées qui fournissent des informations sur l’examen (dimensions des pixels, date de l’examen, ...).

La première étape de l’algorithme (Figure.IV.1) permet d’extraire la zone qui contient l’échographie des images brutes contenues dans les fichiers DICOM. En outre, les champs qui se trouvent dans la zone extraite et qui ne correspondent pas à des données échographiques ont été éliminés. Cette première étape a permis d’extraire une séquence d’images échographiques (SD) qui sera transformée en volume 3D par les étapes suivantes de l’algorithme.

En résumé, cette première partie a présenté les données qui ont été utilisées pour créer les volumes échographiques de notre base de donnée ainsi que les hypothèses qui nous ont permis de faire passer de 9 à 7 le nombre de paramètres à estimer pour créer ces volumes. Elle a aussi présenté le travail préparatoire effectué sur les fichiers DICOM pour obtenir des SDs. Les prochaines parties seront dédiées aux autres étapes de l'algorithme qui nous ont permis de créer ces données 3D.

IV.2 Recalage deux à deux des images contenues dans les séquences dynamiques

Lors de la réalisation d'une ETF, des mouvements latéraux peuvent survenir, ces derniers sont dûs aux déplacements de la main du médecin ou aux mouvements de la tête de l'enfant. Lors de la visualisation d'une SD qui a été acquise avec ce type de mouvements, on observe des rotations et des translations dans le plan des images. Cette étape a pour objectif de corriger ces mouvements latéraux en estimant, par recalage, le mouvement dans le plan entre chaque couple d'images (I_t, I_{t+1}) . La prochaine partie présente la méthode de recalage que nous avons utilisée pour estimer ce mouvement, la suivante explique comment nous avons optimisé la précision et le temps de recalage.

IV.2.1 Méthode

Recalage d'un couple d'image (I_t, I_{t+1})

Dans le Chapitre II, nous avons vu que le mouvement dans le plan était défini par les 3 paramètres représentés sur la Figure IV.8.

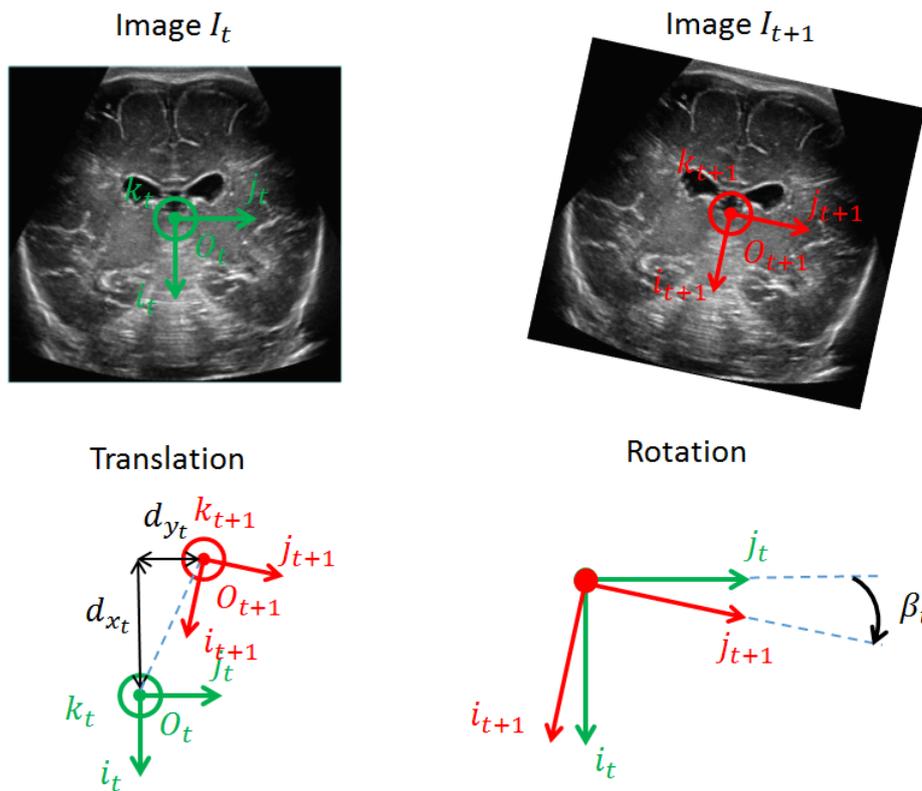


FIGURE IV.8 – Paramètres caractérisant le mouvement dans le plan de l'image.

Pour estimer ces paramètres, un recalage rigide est classiquement utilisé. Comme les images à recaler proviennent de la même modalité, le recalage est effectué de sorte à minimiser les moindres carrés entre les images. Lors d'un recalage rigide, on considère une image fixe (I_t) et une image mobile (I_{t+1}) que l'on va recaler sur la première en effectuant une rotation et une translation. La com-

binaison de ces deux opérations correspond à une transformation affine du plan qui peut s'écrire à l'aide d'une matrice A_{t+1} (d'application associée a_{t+1}) comme définie par l'Équation.IV.1.

$$A_t = \begin{pmatrix} \cos(\beta_t) & -\sin(\beta_t) & d_{x_t} \\ \sin(\beta_t) & \cos(\beta_t) & d_{y_t} \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{IV.1})$$

Cette matrice permet de transformer les coordonnées de chaque pixel de l'image I_{t+1} de manière à la recaler sur I_t , on obtient alors une nouvelle image $a_t(I_{t+1})$. Cette matrice est déterminée en résolvant le problème défini par l'Équation.IV.2.

$$\operatorname{argmin}_{\beta_t, d_{x_t}, d_{y_t}} \sum_{(i,j) \in E_{t+1}} (I_t(i,j) - a_t(I_{t+1})(i,j))^2 \quad (\text{IV.2})$$

L'ensemble E_{t+1} (Équation.IV.2), sur lequel les moindres carrés sont minimisés, est composé des coordonnées des pixels de I_t et $a_t(I_{t+1})$ qui se superposent après recalage. Cet ensemble est représenté sur la Figure.IV.9.d.

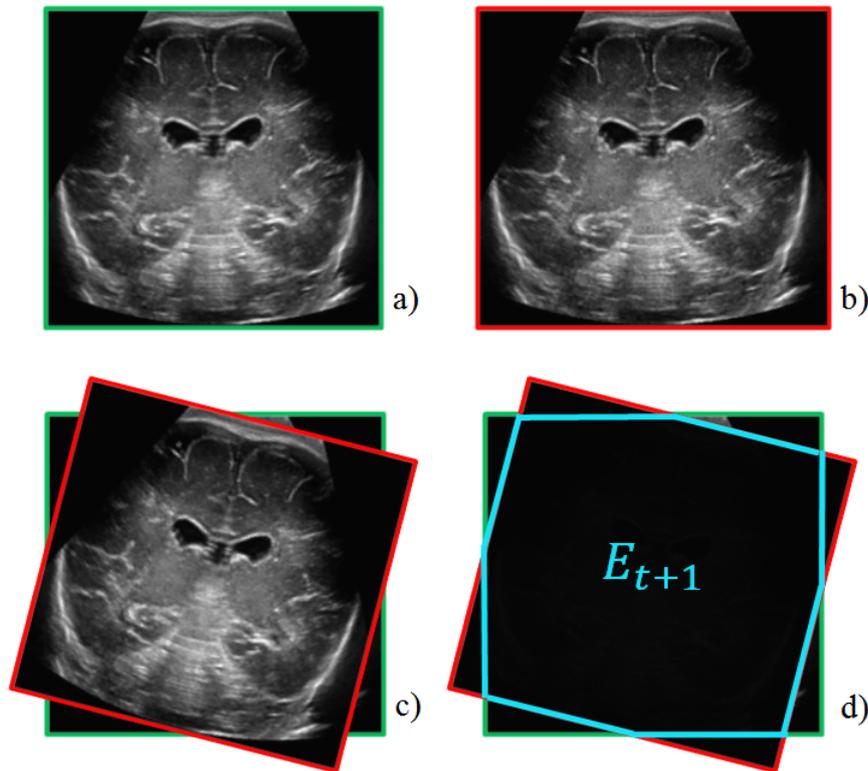


FIGURE IV.9 – a) : image I_t , b) : image I_{t+1} , c) : superposition des images I_t et $a_t(I_{t+1})$, d) : ensemble E_{t+1} correspondant à la zone où I_t et $a_r(I_{t+1})$ se superposent.

Recalage deux à deux des images d'une séquence dynamique

Pour recalage deux à deux les images contenues dans une SD, il faut tout d'abord définir une image de référence I_t puis recalage chaque couple d'images à partir de celle-ci, en prenant en compte les

recalages précédents. Par exemple, si I_1 est choisie comme image de référence, on commence par recaler I_2 sur I_1 , puis I_3 sur $a_1(I_2)$, puis I_4 sur $a_2(I_3)$, et ainsi de suite.

En tant qu'image de référence, nous avons choisi l'image centrale de chaque SD, d'indice $c = E(\frac{T}{2})$ où E désigne la fonction partie entière. Le recalage deux à deux des images de la séquence a ensuite été réalisé de part et d'autre de I_c , en recalant I_{t+1} sur $a_{t-1}(I_t)$ si $t \geq c$ et I_{t-1} sur $a_{t-1}(I_t)$ si $t \leq c$.

I_c a été choisie comme image de référence plutôt que I_1 ou I_T car ces dernières contiennent, en général, moins de structures qui permettent d'effectuer un recalage précis et qu'il y a parfois d'important mouvement latéraux en bout de séquence. De plus, comme un recalage peu précis entre un couple d'images entraîne une erreur qui se répercute lors des recalages suivants, on limite ce phénomène en choisissant I_c comme référence.

Optimisation du temps de calcul

Les images à recaler ont une taille importante (567×763 pixels) et les SDs contiennent un grand nombre d'images (250 en moyenne). Le temps nécessaire pour recaler deux à deux toutes les images d'une SD est donc important (> 30 minutes). De manière à réduire ce temps de calcul, nous avons sous-échantillonné les images à recaler. Leurs dimensions ont ainsi été multipliées par un facteur s (compris entre 0 et 1) et les images ont ensuite été reconstruites en utilisant une interpolation bicubique. Des matrices de transformation $A_{t,s}$ ont ensuite été déterminées à partir de ces images.

Néanmoins, on souhaite que les volumes échographiques créés soient de la plus grande qualité possible. Pour cela, le facteur s a été intégré au niveau des termes de translations des matrices $A_{t,s}$ de manière à obtenir des matrices de transformation A_t qui puissent être utilisées pour transformer les images de dimensions initiales. L'expression des matrices A_t incluant le facteur s est donnée par l'Équation.IV.3.

$$A_{t,s} = \begin{pmatrix} \cos(\beta_t) & \sin(\beta_t) & \frac{d_{x_t}}{s} \\ -\sin(\beta_t) & \cos(\beta_t) & \frac{d_{y_t}}{s} \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{IV.3})$$

Résumé des étapes permettant le recalage deux à deux des images d'une SD.

Les différentes étapes qui permettent le recalage deux à deux des images d'une SD sont données sur la Figure.IV.10.

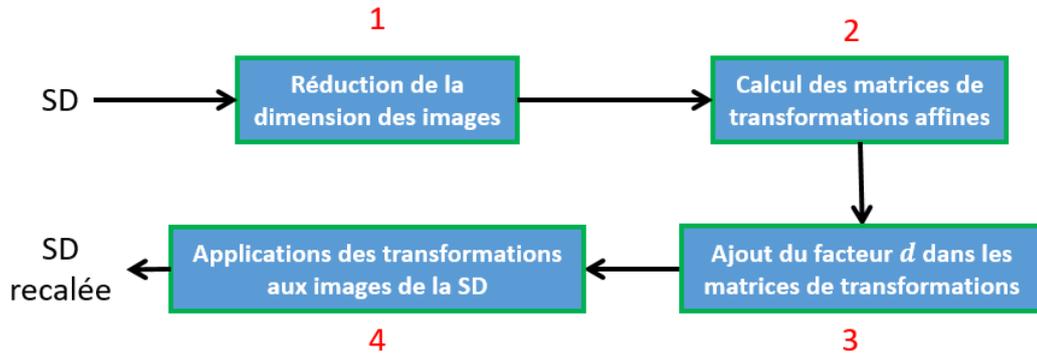


FIGURE IV.10 – Recalage des images d'une séquence dynamique.

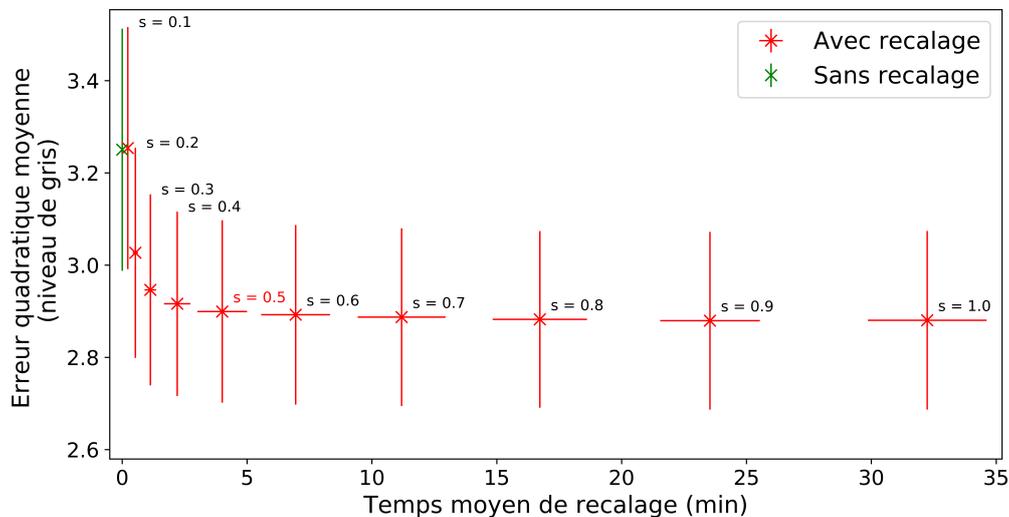
IV.2.2 Expériences et résultats

Cette partie a pour objectifs de montrer que la méthode présentée dans la partie précédente permet bien de réduire la distance quadratique entre les images et d'expliquer comment la valeur du facteur s a été choisie.

Le choix du facteur s a été réalisé de manière à permettre le recalage de toutes les images d'une SD dans un temps raisonnable (quelques minutes) tout en conservant la précision du recalage (point de comparaison à $s=1.0$). La précision du recalage a été évaluée avec l'erreur quadratique moyenne (EQM) définie par l'Équation.IV.4 où $|E_{t+1}|$ désigne le cardinal de E_{t+1} .

$$EQM = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T-1} \frac{1}{|E_{t+1}|} \sum_{(i,j) \in E_{t+1}} (a_{t-1}(I_t(i,j)) - a_t(I_{t+1})(i,j))^2} \quad (IV.4)$$

Pour déterminer la valeur optimale du facteur s , l'étape de recalage a été réalisée sur tous les examens de la base de données avec s allant de 0.1 à 1 par pas de 0.1. L'EQM moyenne et le temps moyen de recalage obtenus sur l'ensemble de la base de données sont affichés sur la Figure.IV.11 en fonction de s . L'EQM est à interpréter sachant que le niveau de gris est compris entre 0 et 255.


 FIGURE IV.11 – Distance quadratique moyenne et temps de recalage moyen en fonction de s .

On peut constater que cette étape de recalage diminue bien l'EQM moyenne entre les images de chaque SD, cette distance est en effet de 3.25 sans recalage et de 2.88 avec recalage sans réduire la dimension des images.

En comparant l'EQM moyenne et le temps de recalage moyen pour différentes valeurs de s , on constate que le temps de recalage moyen diminue fortement entre $s = 1.0$ et $s = 0.5$ (de plus de 30 minutes à moins de 5 minutes) et que l'EQM moyenne reste stable (de 2.88 à 2.90). En comparaison, le temps de recalage moyen diminue peu entre $s = 0.5$ et $s = 0.1$ (de moins de 5 minutes à quelques secondes) et l'EQM moyenne augmente fortement de 2.90 à 3.25 (valeur obtenue sans recalage).

Nous avons donc fixé la valeur du facteur s à 0.5, qui nous semble être une valeur permettant un bon compromis entre le temps et la précision du recalage. Les valeurs du temps de recalage moyen et de l'EQM moyenne sont données dans la Table.IV.1 pour $s = 1.0$, $s = 0.5$ et sans recalage.

TABLE IV.1 – Erreur quadratique moyenne calculée sur les SDs et les SDs recalées avec $s = 0.5$ et $s = 1.0$.

	Erreur quadratique moyenne (niveau de gris)	Temps de recalage (minutes)
SDs sans recalage	3.25 ± 0.18	0
SDs recalées avec $s = 0.5$	2.9 ± 0.2	4.0 ± 1.0
SDs recalées avec $s = 1.0$	2.88 ± 0.19	32.2 ± 2.4

Le temps de recalage moyen pour $s = 0.5$ est de 4 ± 1 minutes, ce qui est tout à fait raisonnable. Des résultats qualitatifs, sous forme de vidéos, sont fournis en complément de ce manuscrit. On peut y voir une SD dynamique sans recalage ([sans recalage.avi](#)) et après recalage ([donnees recalees.avi](#)) avec $s = 0.5$.

IV.2.3 Conclusion

Cette deuxième étape de la chaîne nous a permis de compenser les mouvements latéraux qui pouvaient survenir lors des acquisitions. A ce stade, les paramètres β_t et d_{y_t} ont été définitivement estimés, une première estimation de d_{x_t} a été obtenue mais sa valeur sera modifiée lors de la prochaine étape de la chaîne. Les paramètres qui restent à estimer pour placer toutes les images dans la grille de reconstruction sont γ_t , d_{x_t} et d_{z_t} , leur valeur vont être estimées lors de la prochaine étape de la chaîne.

IV.3 Positionnement des images dans la grille de reconstruction

Cette étape a pour objectif de placer les images de la séquence obtenue après l'étape de recalage dans une grille cartésienne de reconstruction. Cette partie présente dans l'ordre : la grille de reconstruction, le modèle d'acquisition paramétrique développé ainsi que la méthode d'optimisation de ses paramètres et enfin comment la précision du placement des images dans la grille de reconstruction a été validée in-vivo et sur fantôme.

IV.3.1 Grille de reconstruction

La création de données échographiques 3D à partir d'une séquence d'images 2D nécessite de définir une grille de reconstruction dans laquelle seront placées les images. Classiquement, une fois cette grille définie, il faut fixer la position d'au moins une image dans celle-ci et estimer les paramètres γ_t , d_{x_t} et d_{z_t} entre chaque couple d'images successives pour être en mesure de placer toute la séquence dans la grille de reconstruction.

Définition de la grille de reconstruction

Les paramètres d_{x_t} et d_{z_t} correspondent à un déplacement dans le plan sagittal (Figure.IV.8) et les rotations d'angle γ_t s'effectuent autour d'un axe perpendiculaire au plan sagittal (Figure.IV.8), les transformations permettant de placer chaque coupe sagittale de la séquence d'images dans la grille de reconstruction sont donc les mêmes. Il suffit donc d'estimer ces paramètres à partir d'une coupe sagittale pour être en mesure de placer toute la séquence dans la grille de reconstruction. Ainsi, lors de chaque examen, une image de référence en coupe sagittale notée I_{ref} a été acquise pour estimer les paramètres d_{x_t} , d_{z_t} et γ_t . La Figure.IV.12.b montre une séquence dynamique en coupe sagittale qui sera transformée par comparaison à l'image I_{ref} (Figure.IV.12.a).

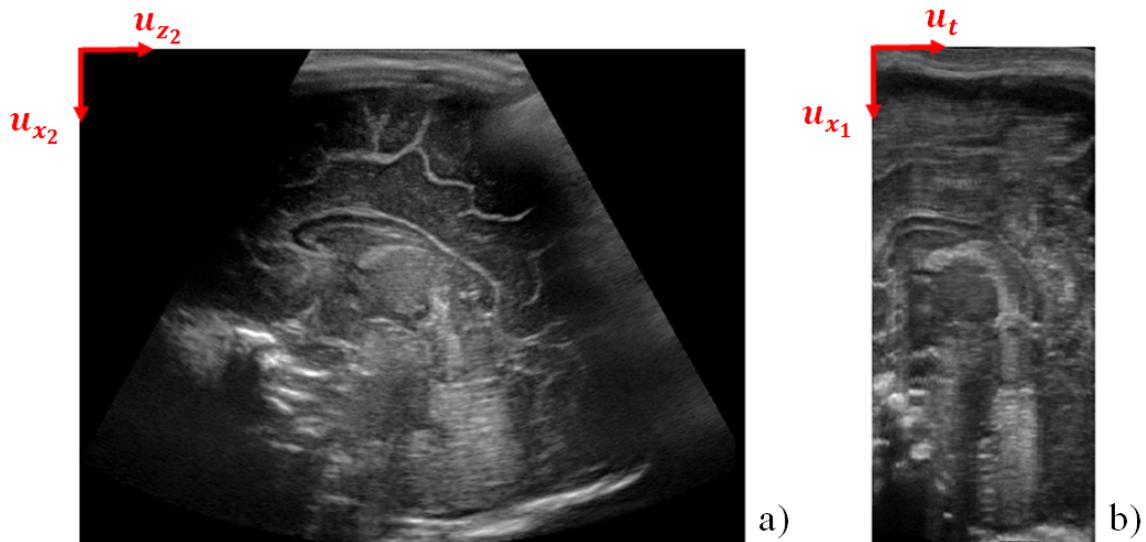


FIGURE IV.12 – Image acquise en coupe sagittale (a) et coupe sagittale d'une séquence d'images après l'étape de recalage (b).

Cette image définit la grille de reconstruction pour une coupe sagittale de la séquence d'image

obtenue après l'étape de recalage. Étant donné que le mouvement que l'on cherche à estimer est le même dans toutes les coupes sagittales, la grille de reconstruction a été définie comme un espace ayant les dimensions de I_{ref} en coupe sagittale et le nombre de coupe sagittale de la séquence d'image pour troisième dimension. Les reconstructions sont ainsi effectuées dans une grille de reconstruction à laquelle est associée un repère $R_2 = (O_2, \vec{u}_{x_2}, \vec{u}_{y_2}, \vec{u}_{z_2})$ défini sur la Figure.IV.13.a où le nombre de pixel selon \vec{u}_{y_2} est le même que selon \vec{u}_{y_1} (Figure.IV.13.b).

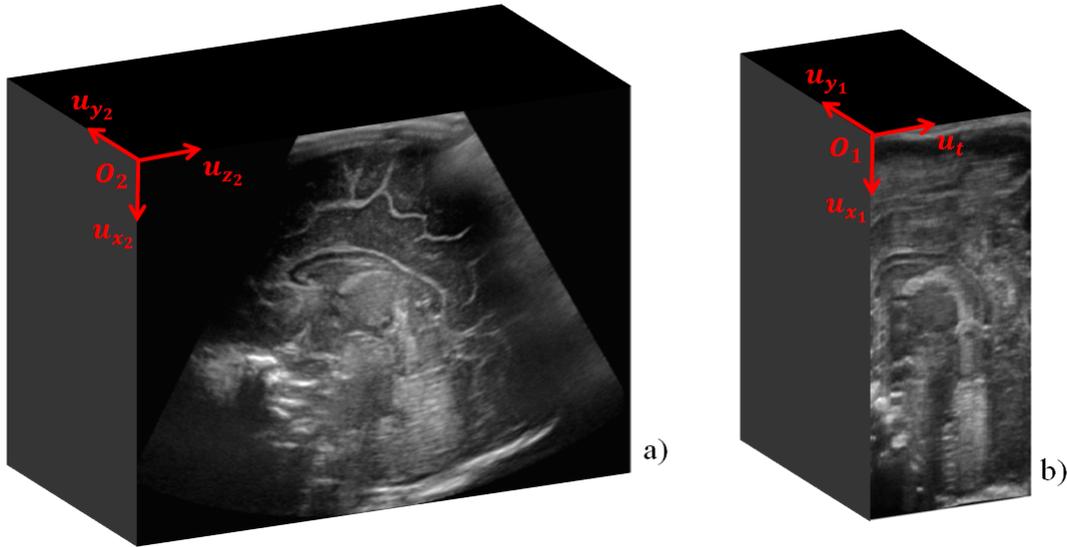


FIGURE IV.13 – Grille de reconstruction dans le repère R_2 a) et séquence dynamique dans le repère R_1 .

Placement des images dans la grille de reconstruction

Lorsqu'un volume échographique 3D est créé à partir d'un modèle d'acquisition, en général, la position d'une image est fixée dans la grille d'acquisition puis les paramètres du modèle sont estimés ou fixés manuellement pour déterminer γ_t , d_{x_t} et d_{z_t} et replacer les images dans la grille. Dans notre cas, la grille de reconstruction est définie à partir d'une image de référence acquise en coupe sagittale. Plutôt que d'estimer les paramètres du modèle d'acquisition manuellement, nous les avons estimés par rapport à I_{ref} , ce qui permet d'éviter les erreurs dues à des mesures manuelles.

Pour réaliser cette estimation, on choisit la coupe sagittale de la SD, que l'on appellera I_{sag} , qui est la plus semblable à I_{ref} . Un exemple est visible sur la Figure.IV.12.b. Ensuite, grâce à des points de repères anatomiques placés dans I_{ref} et dans I_{sag} , les paramètres du modèle sont estimés de manière à déplier l'empilement d'images pour que I_{sag} soit la plus semblable possible à I_{ref} . L'utilisation de points de repères anatomiques permet à la fois d'estimer, via le modèle paramétrique, γ_t , d_{x_t} et d_{z_t} , et de placer la séquence d'image dans la grille de reconstruction.

IV.3.2 Modèle d'acquisition paramétrique

D'après nos observations et nos discussions avec le pédiatre partenaire du projet, le mouvement d'acquisition réalisé lors d'une ETF 2D est un mouvement de rotation d'axe mobile. Dans ce type de mouvement, la sonde et en particulier, le plan d'acquisition de la sonde, est dans un référentiel en mouvement dans le temps par rapport au référentiel du patient (défini par I_{ref} et lié au référentiel terrestre). Pour définir un modèle du mouvement d'acquisition, on suppose que le mouvement de la sonde est un mouvement paramétrique : plus précisément, on fait l'hypothèse que le repère de la sonde a un mouvement de rotation uniforme dont l'axe suit un mouvement de translation rectiligne et uniforme. Dans cette hypothèse, un point de coordonnées (x_1, y_1) dans le référentiel de la sonde a des coordonnées (x_2, y_2, z_2) dans le référentiel du patient, qui s'expriment comme trois fonctions du temps, des paramètres et de (x_1, y_1) . Le modèle paramétrique va être défini dans la suite de cette partie.

On appelle C_t le centre de rotation de la sonde. Au cours d'une acquisition, le référentiel de la sonde $R = (C_t, \vec{i}_t, \vec{j}_t, \vec{k}_t)$ est en mouvement dans le référentiel du patient. Sa position par rapport à celui-ci à l'instant t est donné par $\overrightarrow{O_2C_t}$ et $\theta_t = \widehat{(u_{x_2}, \vec{i}_t)}$ qui sont définis par l'Équation.IV.5 :

$$\begin{aligned} \overrightarrow{O_2C_t} &= (C_{x_2} + v_{x_2} t) \vec{u}_{x_2} + (C_{z_2} + v_{z_2} t) \vec{u}_{z_2} \\ \theta_t &= \theta_0 + \omega t \end{aligned} \quad (IV.5)$$

On considère un point K dans le plan acquis par la sonde, ses coordonnées dans le référentiel R s'écrivent $\overrightarrow{C_tK} = k_x \vec{i}_t + k_y \vec{j}_t$. La position de ce point peut être donné au cours du temps dans le référentiel R_2 avec l'Équation.IV.6 :

$$\overrightarrow{O_2K} = \overrightarrow{O_2C_t} + \overrightarrow{C_tK}. \quad (IV.6)$$

En utilisant le fait que $\vec{i}_t = \cos(\theta_t) \vec{u}_{x_2} + \sin(\theta_t) \vec{u}_{z_2}$ et que $\vec{j}_t = u_{y_2}$, $\overrightarrow{C_tK}$ s'exprime dans le repère R_2 avec l'Équation.IV.7.

$$\overrightarrow{C_tK} = k_x \cos(\theta_t) \vec{u}_{x_2} + k_y \vec{u}_{y_2} + k_x \sin(\theta_t) \vec{u}_{z_2} \quad (IV.7)$$

On a donc :

$$\begin{aligned} \overrightarrow{O_2K} &= \overrightarrow{O_2C_t} + \overrightarrow{C_tK} \\ &= (C_{x_2} + v_{x_2} t + k_x \cos(\theta_t)) \vec{u}_{x_2} + k_y \vec{u}_{y_2} + (C_{z_2} + v_{z_2} t + k_x \sin(\theta_t)) \vec{u}_{z_2} \\ &= (C_{x_2} + v_{x_2} t + k_x \cos(\theta_0 + \omega t)) \vec{u}_{x_2} + k_y \vec{u}_{y_2} + (C_{z_2} + v_{z_2} t + k_x \sin(\theta_0 + \omega t)) \vec{u}_{z_2} \end{aligned} \quad (IV.8)$$

De plus $k_x = \Delta + x_1$ et $k_y = y_1$ où (x_1, y_1) sont les coordonnées de K dans le repère R_1 et Δ la distance signée entre C_t et le plan de coordonnées $x_1 = 0$. Le mouvement d'acquisition peut finalement se modéliser par la fonction f_1 donnée par l'équation.IV.9.

$$\begin{aligned} f_1 : \quad I_{sag} &\rightarrow I_{ref} \\ (x_1, y_1, t)_{R_1} &\mapsto \begin{pmatrix} C_{x_2} + v_{x_2} t + (x_1 + \Delta) \cos(\omega t + \theta_0) \\ y_1 \\ C_{z_2} + v_{z_2} t + (x_1 + \Delta) \sin(\omega t + \theta_0) \end{pmatrix}_{R_2} \end{aligned} \quad (IV.9)$$

La coordonnées en y_1 étant la même dans les deux référentiels, le modèle retenu ne tient pas compte de celle-ci, le modèle final est donné par l'Equation.IV.10.

$$f_1 : \begin{matrix} I_{sag} & \rightarrow & I_{ref} \\ (x_1, t)_{R_1} & \mapsto & \left(\begin{matrix} C_{x_2} + v_{x_2} t + (x_1 + \Delta) \cos(\omega t + \theta_0) \\ C_{z_2} + v_{z_2} t + (x_1 + \Delta) \sin(\omega t + \theta_0) \end{matrix} \right)_{R_2} \end{matrix} \quad (IV.10)$$

L'espace d'arrivée I_{ref} correspond au référentiel du patient qui est défini par la coupe sagittale de référence, l'espace de départ I_{sag} est défini par la coupe sagittale choisie dans la séquence d'images acquise dans le référentiel de la sonde. Cette fonction permet d'attribuer de nouvelles coordonnées à chaque pixel de I_{sag} dans l'espace défini par I_{ref} , notre objectif étant que $f_1(I_{sag})$ soit la plus semblable possible à I_{ref} . Le sens physique de chaque paramètre est donné dans le tableau.IV.2.

TABLE IV.2 – Sens physique des paramètres du modèle défini par l'Équation.IV.10

Paramètre	Sens physique
C_{x_2}	Position initiale x_2 du centre de rotation de la sonde
C_{z_2}	Position initiale z_2 du centre de rotation de la sonde
v_{x_2}	Vitesse suivant l'axe x_2 du centre de rotation de la sonde
v_{z_2}	Vitesse suivant l'axe z_2 du centre de rotation de la sonde
Δ	Distance entre le centre de rotation de la sonde et les images
ω	Vitesse de rotation de la sonde
θ_0	Position angulaire initiale de la sonde par rapport à l'axe x_2

Le point (C_{x_2}, C_{z_2}) correspond à la position initiale du centre de rotation dans l'espace défini par I_{ref} , entre deux acquisitions, celui-ci translate d'une distance $v_{x_2}(\times 1)$ selon l'axe défini par x_2 et $v_{z_2}(\times 1)$ suivant l'axe défini par z_2 . Ce centre de rotation se situe à une distance Δ du début des images, cette distance peut être positive ou négative en fonction de si l'image se situe en dessous ou au dessus du centre de rotation. Le paramètre θ_0 est l'angle entre l'axe de la sonde et l'axe x_2 au début de l'acquisition, $\omega(\times 1)$ correspond à l'angle de rotation de la sonde, dans le plan sagittal, entre l'acquisition de deux images. Ces paramètres sont représentés graphiquement sur la Figure.IV.14.

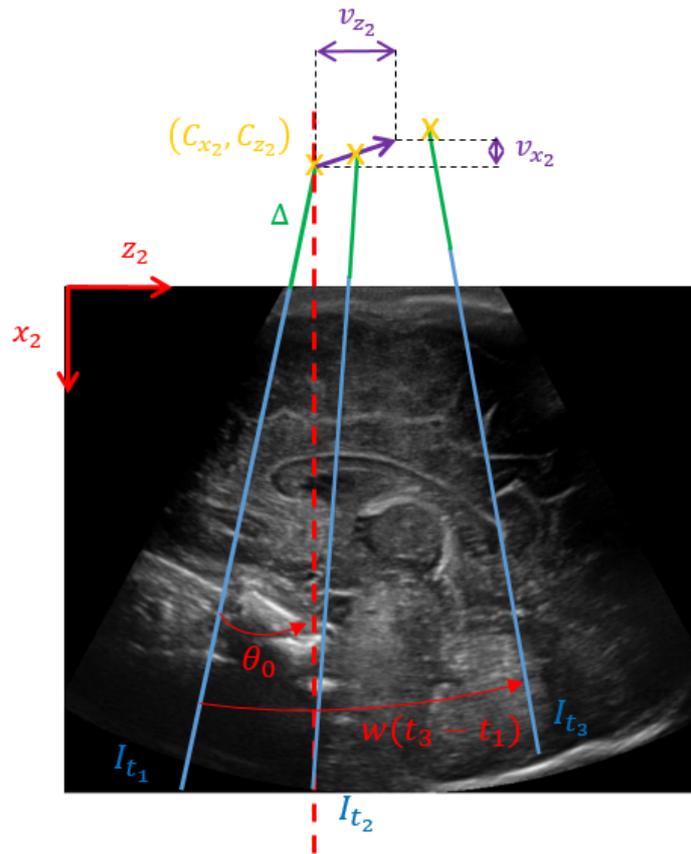


FIGURE IV.14 – Représentation graphique des paramètres du modèle d'acquisition de rotation d'axe mobile.

Ce modèle fait deux hypothèses importantes sur le mouvement d'acquisition :

- La vitesse de rotation de la sonde est constante durant l'examen.
- La translation du centre de rotation est linéaire et se fait à vitesse constante.

Les données dont nous disposons ont été acquises par le pédiatre de manière à ce que ces hypothèses soient respectées au mieux.

Pour valider la pertinence de ce modèle d'acquisition, nous avons également défini des modèles correspondant à un mouvement d'acquisition de rotation d'axe fixe (f_2 , donné par l'équation.IV.11) et de translation linéaire (f_3 , donné par l'équation.IV.12).

$$\begin{aligned}
 f_2: \quad I_{sag} &\rightarrow I_{ref} \\
 (x_1, t) &\mapsto \begin{pmatrix} C_{x_2} + (x_1 + \Delta) \cos(\omega t + \theta_0) \\ C_{z_2} + (x_1 + \Delta) \sin(\omega t + \theta_0) \end{pmatrix}
 \end{aligned} \tag{IV.11}$$

$$\begin{aligned}
 f_3: \quad I_{sag} &\rightarrow I_{ref} \\
 (x_1, t) &\mapsto \begin{pmatrix} C_{x_2} + v_{x_2} t + (x_1 + \Delta) \cos(\theta_0) \\ C_{z_2} + v_{z_2} t + (x_1 + \Delta) \sin(\theta_0) \end{pmatrix}
 \end{aligned} \tag{IV.12}$$

Les paramètres de ces deux modèles ont le même sens physique que les paramètres de f_1 . Le modèle de rotation d'axe fixe a été construit en supprimant les termes de translation du centre de rotation (v_{x_2} et v_{z_2}) et le modèle de translation linéaire a été défini en enlevant le terme de rotation w .

La prochaine partie détaille la méthode qui a été utilisée pour estimer les paramètres de ces modèles.

IV.3.3 Optimisation des paramètres des modèles paramétriques

Une fois l'espace de reconstruction et les modèles permettant de placer les images dans celui-ci définis, il faut optimiser les paramètres de ces modèles. Pour cela, nous avons placé des repères anatomiques se correspondant sur les images I_{ref} et I_{sag} . L'optimisation des paramètres des modèles d'acquisitions a ensuite été réalisée de manière à minimiser la distance euclidienne entre les points de repère anatomique placés sur I_{ref} et la transformation de ceux qui avaient été placés sur I_{sag} .

Définition de I_{sag}

I_{sag} a été définie en choisissant une coupe de la séquence dynamique dans laquelle les points de repères anatomiques, que l'on souhaite placés sur I_{ref} , étaient visibles. Cette étape est complexe car il est en général difficile de trouver une coupe sagittale de la séquence d'images qui réponde à ce critère. En pratique, il est impossible de garantir que I_{ref} et les coupes en y_1 de la séquence d'images correspondent parfaitement à des coupes sagittales. Il peut donc être difficile de déterminer une correspondance satisfaisante entre ces coupes. Pour palier cette limitation, I_{sag} a été définie de sorte à ce qu'elle puisse correspondre à une coupe quelconque de la séquence d'images.

Pour définir I_{sag} , des repères anatomiques sont dans un premier temps identifiés sur I_{ref} . Dans un second temps, des coupes de la séquence d'images selon y_1 , dans lesquelles ces repères apparaissent sont choisies. Cette méthode est représentée sur la Figure.IV.15, I_{ref} et les repères anatomiques définis sont visibles sur la Figure.IV.15.a, les deux coupes sagittales de la séquence d'images dans lesquelles apparaissent ces repères anatomiques sont visibles sur la Figure.IV.15.b. Le nombre de coupes en y_1 nécessaire pour être en mesure d'identifier tous les repères anatomiques dépend des examens.

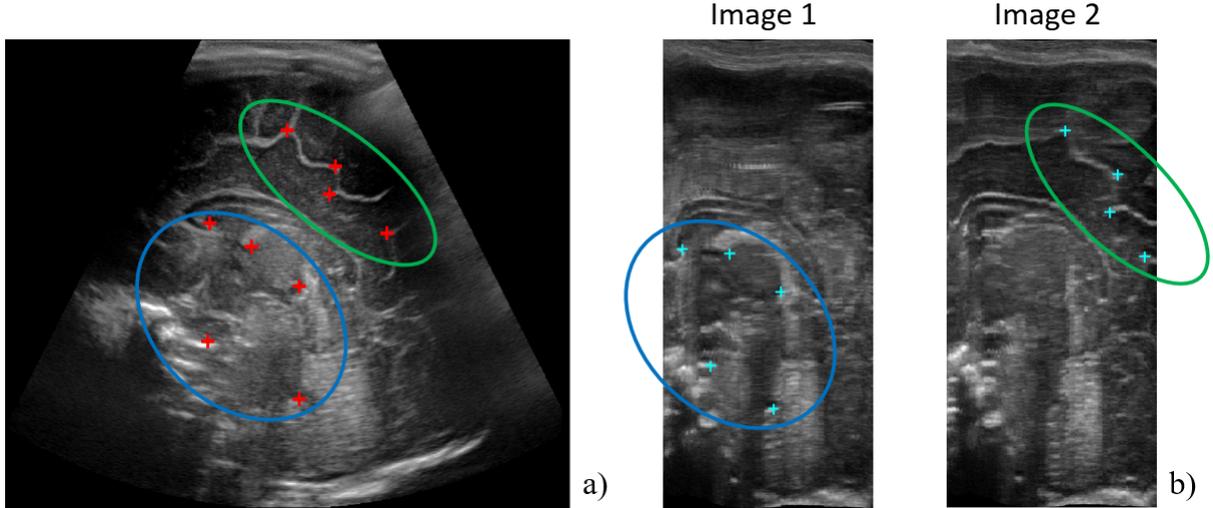


FIGURE IV.15 – Image de référence en coupe sagittale (a) avec quelques repères anatomiques (points rouges) que l'on souhaite retrouver (point bleus) dans deux images en coupe sagittales (b).

Une fois les repères anatomiques placés dans les coupes sagittales de la séquence d'image, on cherche à déterminer une unique coupe dans laquelle il soit possible de tous les visualiser. Pour cela, on cherche le plan P dont la distance à l'ensemble défini par les points de repère anatomiques est minimale. Dans $(0, x_1, y_1, t)$, ce plan est défini par l'Équation.IV.13 :

$$(P) : ax_1 + by_1 + ct + d_p = 0 \quad (\text{IV.13})$$

Où $a, b, c, d_p \in \mathbb{R}$. En pratique, on souhaite pouvoir visualiser le plan P pour y placer des points de références anatomiques. Pour cela, on cherche à déterminer une base orthonormée dont l'un des vecteurs est orthogonal à P , en l'occurrence $n = (a, b, c)$. Dans ce cas, le paramètre d n'a pas d'intérêt, on peut donc normaliser l'Équation.IV.13 en divisant par $-d$, ce qui permet de la réécrire sous la forme de l'Équation.IV.14 :

$$(P_2) : a_2x_1 + b_2y_1 + c_2t = 1 \quad (\text{IV.14})$$

où $a_2 = -\frac{a}{d_p}$, $b_2 = -\frac{b}{d_p}$ et $c_2 = -\frac{c}{d_p}$. Le vecteur $n_2 = (a_2, b_2, c_2)$ orthogonal à P_2 est aussi orthogonal à P car n_2 et n sont colinéaires, on cherchera donc à calculer n_2 plutôt que n . Déterminer l'équation de l'hyperplan P_2 dont la distance à l'ensemble des N_{sag} points de repères anatomiques est minimale est équivalent à résoudre le problème défini par l'équation.IV.15.

$$\operatorname{argmin}_{a_2, b_2, c_2} \sum_{i=1}^{N_{sag}} (a_2x_i + b_2y_i + c_2t_i - 1)^2 \quad (\text{IV.15})$$

Le problème défini par l'Équation.IV.15 correspond à un problème de régression linéaire, il peut se réécrire sous forme matricielle en utilisant les Équations.IV.16, IV.18 et IV.17.

$$A = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \dots & & \\ x_{N_{sag}} & y_{N_{sag}} & z_{N_{sag}} \end{pmatrix} \quad (\text{IV.16})$$

$$Y = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \quad (\text{IV.17})$$

$$X = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} \quad (\text{IV.18})$$

Résoudre le problème défini par l'Équation.IV.15 est équivalent à résoudre l'équation $AX = Y$ en X . Cette dernière n'a pas de solution car les points de repères anatomiques ne sont pas coplanaires, on la résout en se servant de l'équation normale (Équation.IV.19), utilisée classiquement en statistiques pour résoudre les problèmes de régression linéaire.

$$X = (A^T A)^{-1} A^T Y \quad (\text{IV.19})$$

Où T désigne l'opérateur transposé et $^{-1}$ désigne l'inverse d'une matrice. Un vecteur n_2 orthogonal au plan P est ainsi obtenu en normalisant X . Étant donné que l'orientation de P est proche d'une coupe sagittale, n_2 est presque colinéaire à y_1 . On définit donc les vecteurs de la nouvelle base dans laquelle on souhaite visualiser la séquence d'image en utilisant l'Équation.IV.20 et l'Équation.IV.21.

$$n_3 = \frac{(0, -c, b)}{\sqrt{b^2 + c^2}} \quad (\text{IV.20})$$

$$n_1 = n_2 \wedge n_3 \quad (\text{IV.21})$$

où \wedge correspond au produit vectoriel. On obtient ainsi une nouvelle base orthonormée (n_1, n_2, n_3) dans laquelle visualiser la séquence d'image.

Pour visualiser P dans la base (n_1, n_2, n_3) , il faut déterminer sa coordonnée en n_2 . Le point $A = (0, 1/b, 0)$, dont les coordonnées sont données dans (x_1, y_1, t) , appartient à P_2 . En utilisant la matrice de passage de la base (x_1, y_2, t) à (n_1, n_2, n_3) , on obtient les coordonnées de A dans (n_1, n_2, n_3) et donc sa coordonnée en n_2 . Le plan P obtenu à partir de cette coordonnée est visible sur la Figure.IV.16.b.

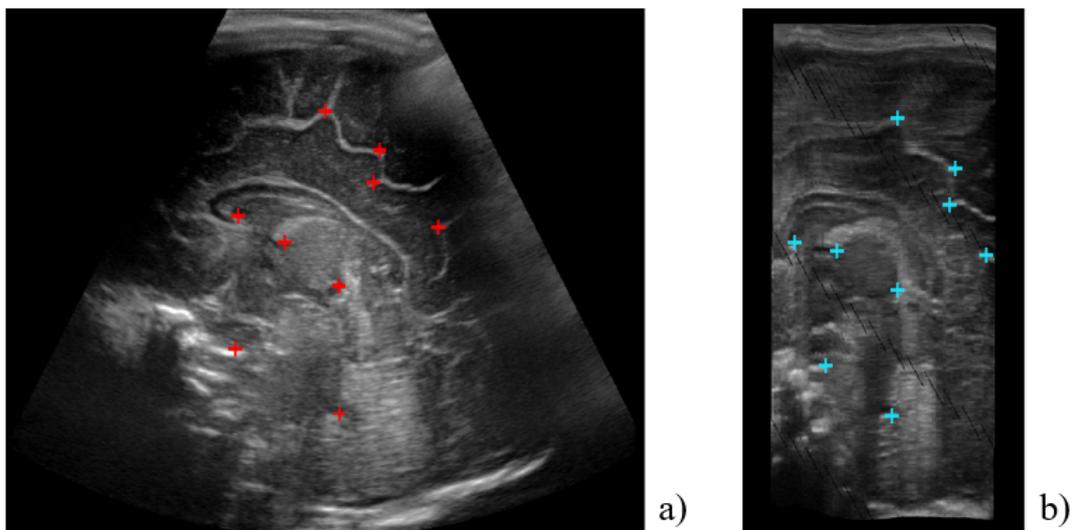


FIGURE IV.16 – I_{ref} avec quelques repères anatomiques (a) et plan P (I_{sag}) dans lequel on peut voir tous les repères anatomiques (points bleus) placés sur I_{ref} (b).

On peut voir que tous les points de repères anatomiques définis sur la Figure.IV.16.a peuvent être placés dans le plan P (Figure.IV.16.b).

Minimisation de la distance aux points de référence

Pour optimiser les paramètres des modèles f_1 , f_2 et f_3 , donnés respectivement par les Équations.IV.10, IV.11 et IV.12, un maximum de points de repères anatomiques se correspondant sont placés sur I_{ref} et sur I_{sag} . Un exemple est donné sur la Figure.IV.17.

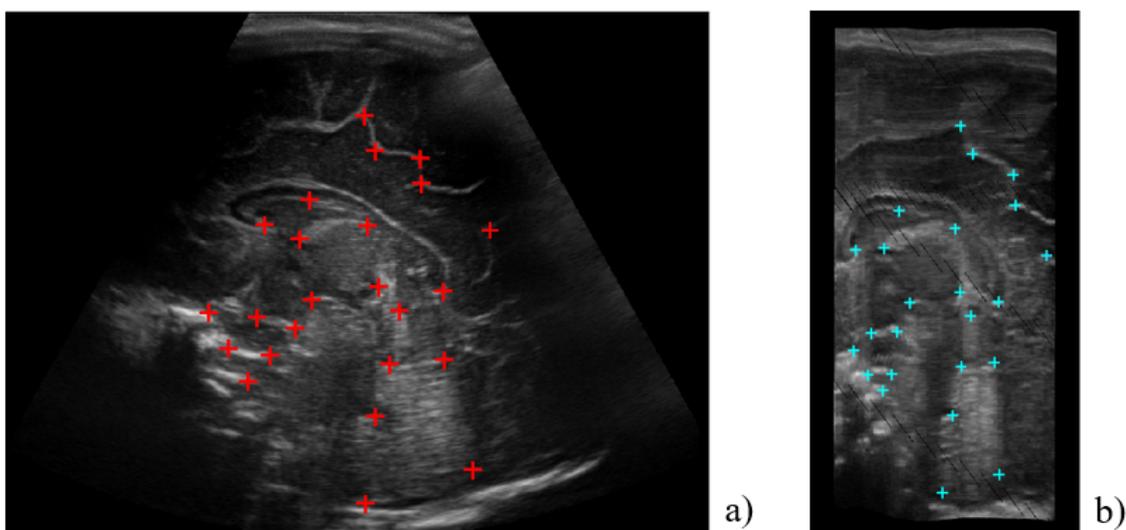


FIGURE IV.17 – a) : I_{ref} avec repères anatomiques (points rouges). b) : I_{sag} avec les repères anatomiques correspondants à ceux placés sur I_{ref} .

Dans le cas du modèle f_1 , les N points de repères anatomiques ont été utilisés pour résoudre le problème défini par l'Équation.IV.22.

$$\underset{C_{x_2}, C_{z_2}, v_{x_2}, v_{z_2}, \Delta, \omega, \theta_0}{\operatorname{argmin}} \sum_{i=1}^N d(f(x_{1_i}, t_i), (x_{2_i}, z_{2_i}) \times \frac{\delta_{ref}}{\delta_{SD}}) \quad (\text{IV.22})$$

On cherche les paramètres $C_{x_2}, C_{z_2}, v_{x_2}, v_{z_2}, \Delta, \omega, \theta_0$ qui minimisent la distance euclidienne d entre les points placés sur I_{ref} et l'image par f_1 des points placés sur I_{sag} . Les paramètres δ_{ref} et δ_{SD} correspondent respectivement aux dimensions des pixels de I_{ref} et de la séquence d'images. Ces dimensions sont à prendre en compte car elles sont en général différentes. On peut de plus remarquer que le problème défini par l'équation.IV.22 ne prend pas en compte les coordonnées en y_2 et y_1 des points de repères anatomiques. Ce choix a été fait par soucis de simplification. Premièrement car nous n'avons pas la coordonnée en y_2 des points placés sur I_{ref} et deuxièmement car la coordonnée en y_1 varie peu d'un point à un autre par rapport aux coordonnées en x_1 et z_1 .

Ce problème (Équation.IV.22) a été résolu en utilisant l'algorithme à régions de confiance (Coleman *et Li* (1996), Coleman *et Li* (1994)) implémenté dans la fonction "lsqnonlin" de Matlab. Les valeurs initiales et les bornes des paramètres du modèle qui ont été utilisées lors de l'optimisation sont données dans la Table.IV.3.

TABLE IV.3 – Valeurs initiales et bornes des paramètres à optimiser dans le modèle f_1 .

Paramètre	Valeur initiale	Borne inférieure	Borne supérieure
C_{x_2} (pixel)	0	-10000	10000
C_{z_2} (pixel)	350	-10000	10000
v_{x_2} (pixel.acquisition ⁻¹)	0	-10000	10000
v_{z_2} (pixel.acquisition ⁻¹)	0	-10000	10000
Δ (pixel)	0	-10000	10000
ω (radian)	0	0	$\frac{\pi}{T}$
θ_0 (radian)	0	$-\frac{\pi}{2}$	$\frac{\pi}{2}$

Au niveau des unités, la taille des pixels était variable d'une acquisition à l'autre mais était comprise entre 0.15 et 0.2 mm et la cadence d'acquisition était d'environ 13 images par seconde. Les paramètres $C_{x_2}, v_{x_2}, v_{z_2}, \Delta$ et θ_0 ont été initialisés à 0 car ils peuvent a priori être négatif ou positif. C_{z_2} a été initialisé à 350 car c'est une valeur proche de la taille de I_{ref} en z_2 et ω a été initialisé à 0 car, même si ce paramètre est toujours positif, sa valeur est généralement faible.

Les bornes inférieures et supérieures des paramètres $C_{x_2}, C_{z_2}, v_{x_2}, v_{z_2}$ et Δ ont été définies à -10000 et 10000, ce qui est très large, pour que la position du centre de rotation du modèle et sa vitesse de translation puisse être déterminés librement. Les bornes de ω ont été fixées à 0 et $\frac{\pi}{T}$ car la vitesse de rotation est toujours positive et que l'angle de balayage total ne peut pas excéder 180°. Les bornes de θ_0 ont été définies à $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ car la sonde est obligatoirement orientée vers la tête de l'enfant au début de l'acquisition.

Dans le cas des modèles d'acquisition f_2 et f_3 , un problème d'optimisation analogue à l'Équa-

tion.IV.22 a été défini pour chaque modèle. Les valeurs données dans la Table.IV.3 ont été utilisées en tant que valeurs initiales et bornes pour les paramètres de ces modèles.

Distance entre les ensembles de points de repère après optimisation

Résultats qualitatifs. Des points de repère anatomique placés sur I_{sag} et I_{ref} sont visibles sur les Figure.IV.18.a et Figure.IV.18.c. L'image de ces points après optimisation du modèle f_1 est visible sur la Figure.IV.18.b (points verts) avec ceux qui étaient ciblés (points rouges).

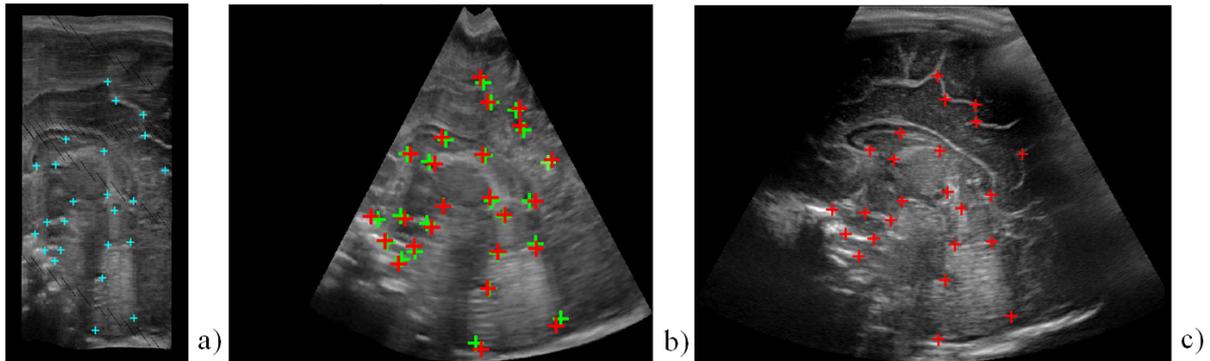


FIGURE IV.18 – a) : I_{sag} et points de repères anatomiques, b) coupe sagittale après placement des images dans la grille de reconstruction avec les points de référence placés sur I_{ref} (point rouges) et l'image par f_1 des points de référence placés sur I_{sag} (points verts), c) : I_{ref} et points de repères anatomiques.

On voit que les deux nuages de points visibles sur la Figure.IV.18.b sont très proches, ce qui montre que la méthode utilisée donne des résultats cohérents. Des images similaires sont données sur la Figure.IV.19 pour les trois modèles d'acquisition.

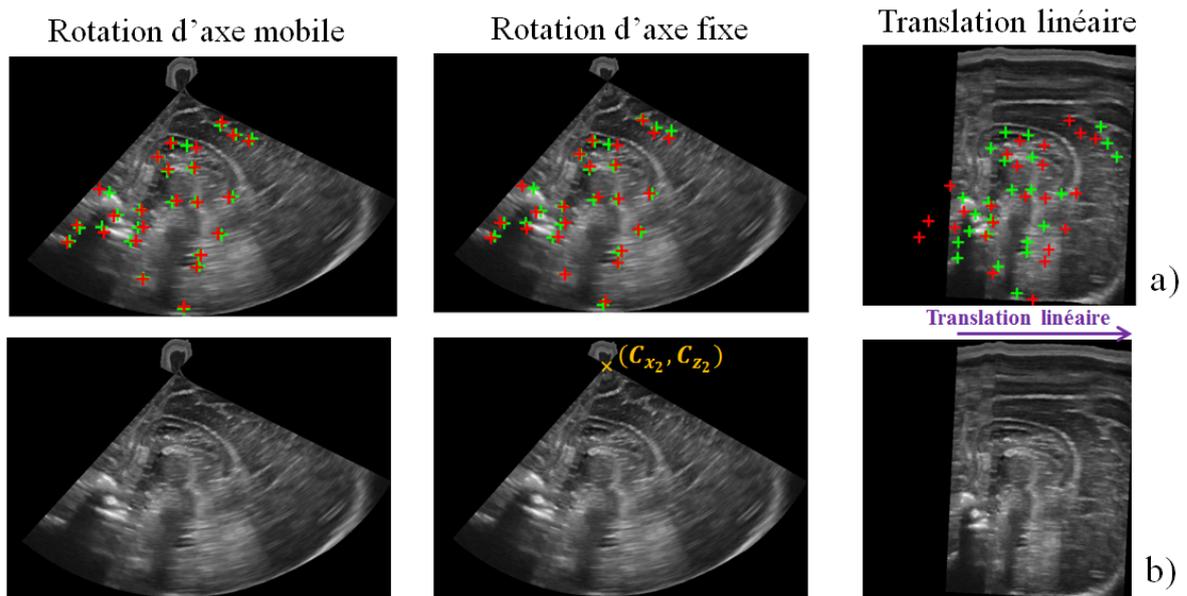


FIGURE IV.19 – Coupe sagittale des volumes échographiques avec les repères anatomiques placés sur I_{ref} (points rouges) et l'image des points placés sur I_{sag} après transformation par les différents modèles (points verts) (a) et sans ces points de repères anatomiques (b).

Les images présentées sur la Figure.IV.19.b montrent que les volumes échographiques obtenus sont en accord avec les modèles d'acquisition utilisés pour les créer. On observe un centre de rotation fixe (croix orange) et une zone échographique en éventail dans le cas du modèle de rotation d'axe fixe. Dans le cas du modèle de translation linéaire on voit que l'échographie a une forme rectangulaire ou en parallélogramme. De même, dans le cas du modèle de rotation d'axe mobile, on observe que l'échographie a une forme en éventail mais qu'on ne peut pas identifier un centre de rotation fixe comme c'était le cas pour le modèle de rotation d'axe fixe.

Sur la Figure.IV.19.a, on voit les nuages de points de références placés sur I_{ref} (points rouges) et l'image des points placés sur I_{sag} (points verts) par chacun des modèles d'acquisition. Ces nuages sont proches dans le cas des modèles de rotation d'axe fixe et d'axe mobile, en revanche, ils sont éloignés dans le cas du modèle de translation linéaire. Les deux premiers modèles semblent donc plus adaptés pour résoudre notre problème que le troisième.

Résultats quantitatifs. Après avoir optimisé les paramètres de f_1 , f_2 et f_3 pour chacun des examens de notre base de données, la distance euclidienne moyenne entre les points placés sur I_{ref} et l'image, par les modèles, des points placés sur I_{sag} a été calculée. La distance euclidienne moyenne obtenue en optimisant les trois modèles sur tous les examens est donnée dans la Table.IV.4.

TABLE IV.4 – Distance moyenne entre les points de repère anatomique après optimisation des modèles d'acquisition.

Modèle	Translation linéaire	Rotation d'axe fixe	Rotation d'axe mobile
Distance moyenne (mm)	3.63 ± 1.26	1.58 ± 0.35	1.5 ± 0.33

La meilleure correspondance (i.e la plus petite distance) a été obtenue par le modèle de rotation d'axe mobile (1.5 ± 0.33 mm), ce qui n'est pas surprenant car il a plus de degré de liberté que les deux autres modèles. Cette erreur est néanmoins proche de celle obtenue avec modèle de rotation d'axe fixe (1.58 ± 0.35 mm). Ces deux derniers modèles sont meilleurs que le modèle de translation linéaire.

Valeurs des paramètres après optimisation de f_1 .

Un modèle d'acquisition f_1 a été optimisé pour chaque examen de la base de donnée. Les valeurs moyennes, minimales et maximales obtenues pour chacun des paramètres sont données dans la Table.IV.5.

TABLE IV.5 – Valeur des paramètres du modèle f_1 après optimisation.

Paramètre	Valeur moyenne	Valeur minimale	Valeur maximale
C_{x_2} (pixel)	-148 ± 256	-981	142
C_{z_2} (pixel)	391 ± 41	292	489
$v_{x_2} \times \delta_{SD} \times T$ (mm)	3.0 ± 2.8	-1.0	9.3
$v_{z_2} \times \delta_{SD} \times T$ (mm)	-1.8 ± 19.0	-46.9	36.0
Δ (pixel)	155 ± 258	-158	984
$\omega \times T$ (°)	61.8 ± 18.6	35.4	111.7
θ_0 (°)	-25.0 ± 10.5	-47.1	-8.7

Ces paramètres sont représentés graphiquement dans le cas où le centre de rotation de la sonde est situé au-dessus du plan de la sonde (Figure.IV.20) et en-dessous du plan de la sonde (Figure.IV.21) lors de l'acquisition.

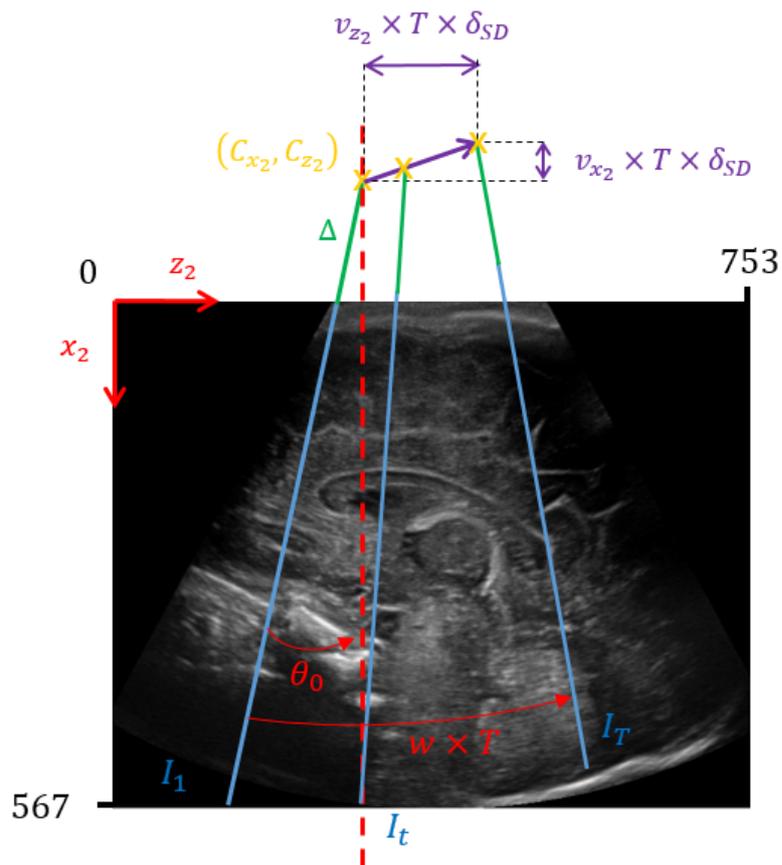


FIGURE IV.20 – Représentation graphique des paramètres du modèle f_1 dans le cas où le centre de rotation se situe au-dessus du plan de la sonde.

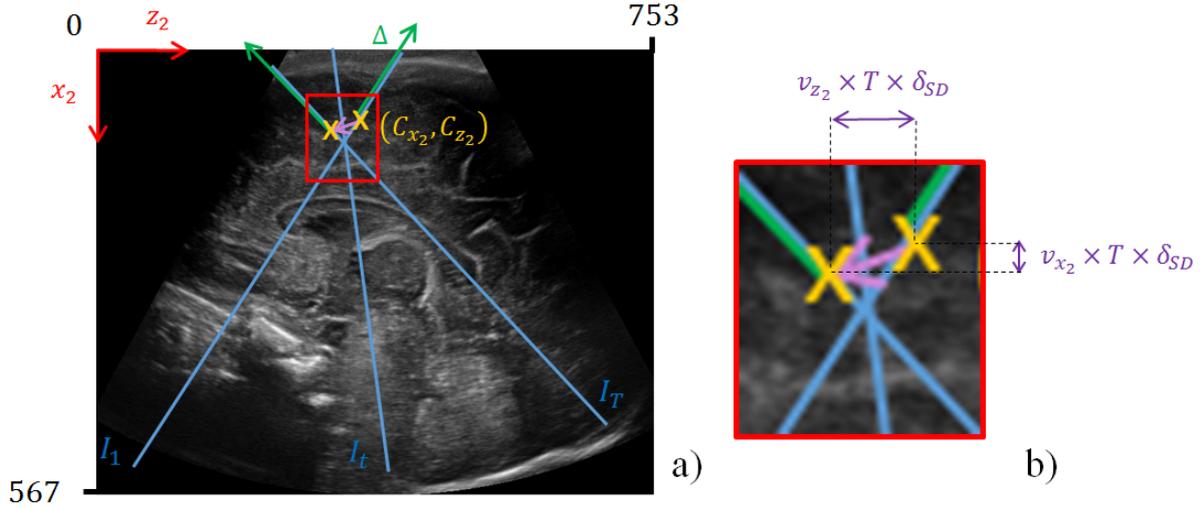


FIGURE IV.21 – Représentation des paramètres du modèle f_1 dans le cas où le centre de rotation se situe en dessous du plan de la sonde.

La Table.IV.5 donne les distances de translations du centre de rotation et l'angle du balayage plutôt que les vitesses qui sont associées à ces paramètres car ces valeurs ont plus de sens physique. Toutes les valeurs moyennes obtenues par les paramètres paraissent cohérentes.

Au début de l'acquisition, le centre de rotation est en moyenne situé sur le point $(-148, 391)$, c'est à dire au dessus du plan de la sonde ($C_{x_2} = -148$) et au niveau de la fontanelle ($C_{z_2} = 391$) qui est proche de la moitié de la taille de l'image en z_2 . Ce cas moyen correspond à celui qui est représenté sur la Figure.IV.20. Lorsque C_{x_2} est positif (valeur maximale de 142), le centre de rotation est situé sous le plan de la sonde comme cela est représenté sur la Figure.IV.21.

Les translations du centre de rotation suivant x_2 ($v_{x_2} \times \delta_{SD} \times T$) et suivant z_2 ($v_{z_2} \times \delta_{SD} \times T$) sont en moyenne faibles : 3.0 ± 2.8 mm et -1.8 ± 19.0 mm. La distance moyenne de translation du centre de rotation suivant x_2 semble cohérente car le déplacement de la sonde a lieu principalement selon l'axe antéro-postérieur (proche de l'axe défini par z_2), cette translation suivant l'axe x_2 est plutôt une conséquence de la courbure de la tête de l'enfant. $v_{z_2} \times \delta_{SD} \times T$ est en moyenne faible mais son écart-type est important, ce paramètre prend la valeur minimale de -46.9 mm et la valeur maximale de 36 mm. Cela peut s'expliquer par le fait que le médecin doit parfois s'adapter à l'anatomie de la fontanelle durant l'acquisition et que cette translation peut être réalisée dans le sens antéro-postérieur ou postéro-antérieur.

La valeur moyenne obtenue pour Δ est de 155 pixels, sa valeur minimum est de -158 et sa valeur maximum de 984. Ces résultats sont cohérents avec les valeurs trouvées pour C_{x_2} , le début des images doit avoir une coordonnée en x_2 qui est proche de 0. Il peut ainsi prendre des valeurs négatives lorsque C_{x_2} est positif, et donc situé sous le plan de la sonde.

La valeur moyenne de l'angle de balayage est de $61.8 \pm 18.6^\circ$ avec un maximum de 111.7° et un minimum de 35.4° , ce qui est cohérent pour ce type d'acquisition.

La valeur moyenne de θ_0 est de $-25 \pm 10.5^\circ$ avec un minimum de -47.1° et un maximum de -8.7° ce qui semble cohérent car la sonde est orientée vers l'avant de la tête quand l'acquisition débute.

Résolution du problème tel qu'il était formulé initialement

L'objectif de cette étape était initialement d'estimer les paramètres γ_t , d_{x_t} et d_{z_t} entre chaque couple d'images successives de la séquence d'images à placer dans la grille de reconstruction. Nous allons montrer que ces trois paramètres peuvent être déterminés à partir des paramètres des modèles paramétriques.

Si l'on considère le modèle f_1 , entre deux images I_t et I_{t+1} , on a $\gamma_t = \omega \times 1$ et $(d_{x_t}, d_{z_t}) = (f_1(x_m, t + 1) - f_1(x_m, t))$ où x_m désigne la coordonnée, suivant l'axe défini par x_1 , du centre de chaque image de la séquence. Cette étape a donc bien permis d'estimer les trois derniers paramètres qu'il fallait estimer pour être en mesure de positionner théoriquement toutes les images d'une SD les unes par rapport aux autres.

IV.3.4 Validation des modèles de reconstruction

Pour déterminer le modèle d'acquisition le plus adapté à ce problème, la précision du placement des images dans la grille de reconstruction a été évaluée sur I_{ref} et sur un fantôme de tête de nouveau-né.

Validation à partir des images I_{ref}

Choix de la structure cérébrale utilisée pour valider le modèle. Pour évaluer la précision du placement des images dans la grille de reconstruction, nous avons décidé d'utiliser une structure cérébrale qui était à la fois visible sur I_{ref} et dans une coupe sagittale de la séquence d'image après l'étape de recalage. Dans l'idéal, la précision du placement des images aurait dû être évaluée sur les thalami ou le CVS car ce sont les structures que l'on souhaite segmenter. Néanmoins, ces structures n'étant pas toujours visibles sur I_{ref} , une autre structure a été choisie : le corps calleux (CC). Ce choix a été fait car le CC est toujours visible sur I_{ref} et que sa forme est à peu près constante en coupe sagittale, ce qui permet de choisir plus facilement une coupe qui correspond à ce qu'on observe sur I_{ref} dans la séquence d'image.

Segmentations du corps calleux. Pour comparer le CC visible sur I_{ref} au CC de la séquence d'image après placement dans la grille de reconstruction, nous avons choisi la coupe sagittale de la séquence d'image où le corps calleux et son environnement étaient les plus proches de celui que l'on pouvait observer sur I_{ref} . Pour chaque examen, des segmentations manuelles du CC ont été réalisées sur I_{ref} (Figure.IV.22.a) et sur la coupe sagittale choisie dans la séquence d'images (Figure.IV.22.b).

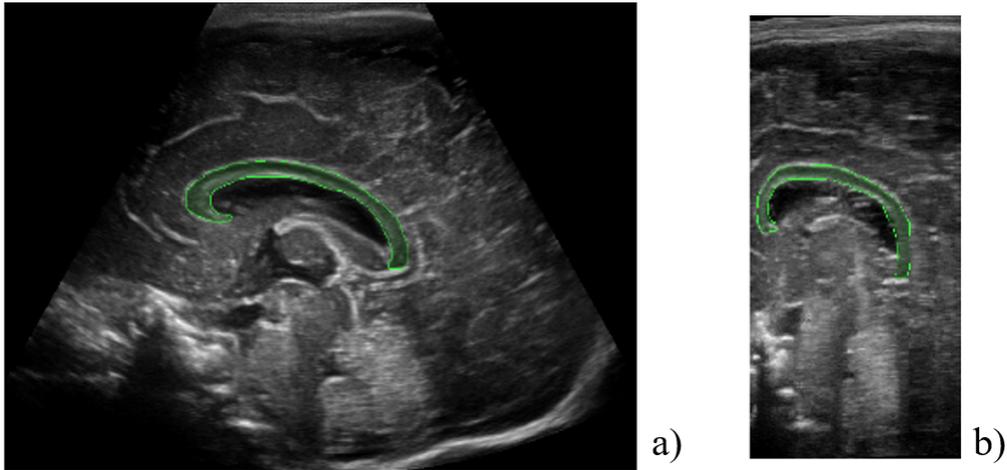


FIGURE IV.22 – a) : segmentation du corps calleux sur I_{ref} et b) : segmentation du corps calleux sur une coupe sagittale de la séquence d'images.

Les segmentations montrées sur la Figure.IV.22 ont été réalisées sur tous les examens où le CC était entièrement visible en coupe sagittale et pouvait être distingué du reste du cerveau. Lorsque ce n'était pas le cas, les examens ont été exclus de la base utilisée pour cette étape de validation. Des exemples d'examens où le CC est incomplet ou difficile à distinguer sont respectivement visibles sur la Figure.IV.23.a et la Figure.IV.23.b.

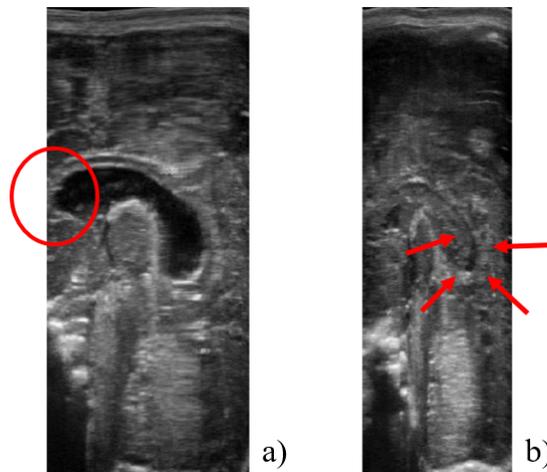


FIGURE IV.23 – Coupe sagittale où le corps calleux est tronqué a) et coupe sagittale sur laquelle il est difficile de déterminer où se situe les bords du corps calleux b).

De plus, lorsque plusieurs examens correspondaient à un même patient, nous n'en avons conservé qu'un pour ne pas biaiser les résultats. Finalement, en conservant uniquement les examens qui répondaient aux critères qui ont été énoncés, nous avons pu utiliser 13 examens pour évaluer la précision du placement dans la grille de reconstruction.

1. Images issues de notre base de données.

Pour obtenir des segmentations du CC dans la grille de reconstruction, les modèles d'acquisition ont été utilisés pour transformer la segmentation manuelle du CC effectuée sur une coupe sagittale de la séquence d'image en segmentation dans l'espace de reconstruction. Des images des segmentations obtenues en utilisant les trois modèles de reconstructions sont visibles sur les Figures IV.24.c.d.e. A titre de comparaison, la segmentation manuelle réalisée sur I_{ref} est visible sur la Figure IV.24.a et la segmentation manuelle réalisée sur une coupe sagittale de la séquence d'image est visible sur la Figure IV.24.b.

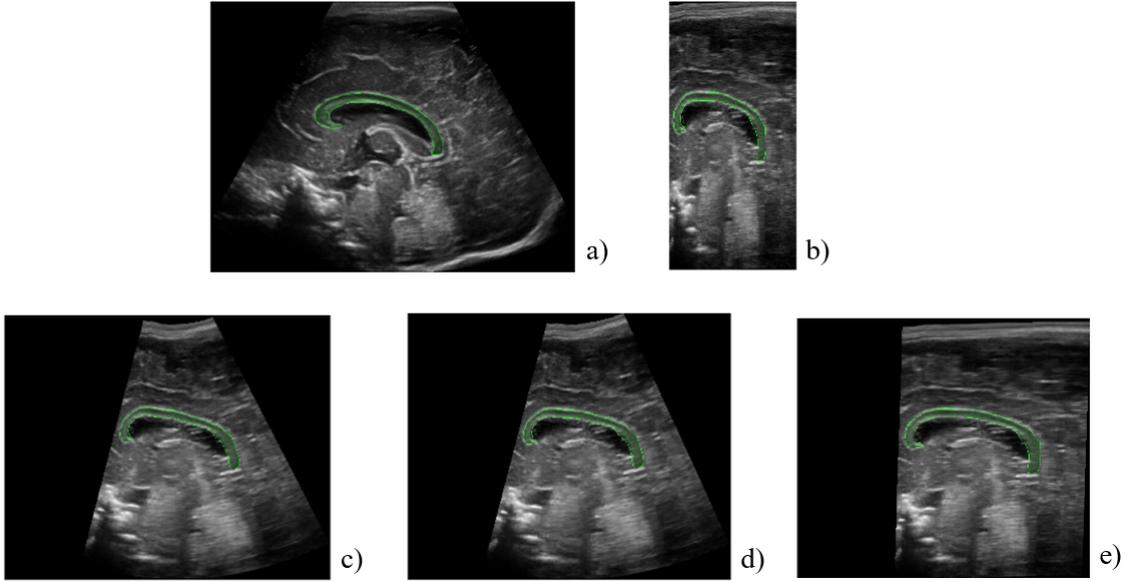


FIGURE IV.24 – Segmentations manuelles du corps calleux réalisées sur I_{ref} a) et sur une coupe sagittale de la séquence d'images b). Segmentations reconstruites du corps calleux à partir des modèles de rotation d'axe mobile c), de rotation d'axe fixe d) et de translation linéaire e).

Métriques. La précision du placement des images dans la grille de reconstruction a été évaluée par rapport aux segmentations du CC en utilisant trois métriques : la distance moyenne (MAD : Équation.VI.5), l'indice de Dice (Équation.VI.4) et la différence absolue de surface (ASD : Équation : IV.25) entre les segmentations.

$$Dice = \frac{2 \times |Y_{CC_{ref}} \cap Y_{CC_{3D}}|}{|Y_{CC_{ref}}| + |Y_{CC_{3D}}|} \quad (IV.23)$$

$Y_{CC_{ref}}$ et $Y_{CC_{3D}}$ désignent les segmentations du CC sur I_{ref} et après reconstruction par l'un des modèles d'acquisition. \cap désigne l'opérateur intersection d'ensembles et $|\cdot|$ la fonction cardinal.

$$MAD = \frac{1}{2} \left(\frac{\sum_{x \in \partial Y_{CC_{ref}}} \min d(x, \partial Y_{CC_{3D}})}{|Y_{CC_{ref}}|} + \frac{\sum_{x \in \partial Y_{CC_{3D}}} \min d(x, \partial Y_{CC_{ref}})}{|\partial Y_{CC_{3D}}|} \right) \quad (IV.24)$$

d désigne la distance euclidienne dans \mathbb{R}^2 et l'opérateur ∂E le bord d'un ensemble E .

$$ASD = \left\| \frac{|Y_{CC_{ref}}| \times \delta_{ref}^2 - |Y_{CC_{3D}}| \times \delta_{SD}^2}{|Y_{CC_{ref}}| \times \delta_{ref}^2} \right\|_1 \quad (IV.25)$$

$\|\cdot\|_1$ désigne la fonction valeur absolue dans \mathbb{R} .

Résultats Évaluée sur les 13 examens sélectionnés, la précision du placement dans la grille de reconstruction, par rapport aux métriques définies dans le paragraphe précédent, est donnée dans la Table.IV.6. La surface moyenne du CC sur les images I_{ref} était en moyenne de $1,17 \pm 0.23 \text{ cm}^2$.

TABLE IV.6 – Précision du placement des images dans la grille de reconstruction, évaluée par rapport au CC.

	MAD (mm)	Dice	ASD (%)
Référence	0	1	0
Rotation d'axe mobile	0.69 ± 0.14	0.63 ± 0.08	15.2 ± 10.1
Rotation d'axe fixe	0.69 ± 0.16	0.64 ± 0.09	15.4 ± 10.1
Translation linéaire	1.34 ± 0.41	0.4 ± 0.13	30.0 ± 21.2

Par rapport à toutes les métriques, les modèles de rotation d'axe mobile et de rotation d'axe fixe sont meilleurs que le modèle de translation linéaire. De plus, la précision obtenue par ces deux derniers modèles selon ces trois métriques semble correcte compte tenu de la petitesse du CC. Ils ont tous les deux obtenus une *MAD* moyenne de 0.69 mm, le meilleur Dice moyen est obtenu par le modèle de rotation d'axe fixe (0.64), cette valeur peut paraître faible mais compte tenu de la taille de la structure segmentée, elle est très correcte. Enfin, la plus petite *ASD* moyenne de 15.2% a été obtenue par le modèle de rotation d'axe mobile. Les modèles de rotation d'axe fixe et d'axe mobile paraissent donc adaptés pour placer les images dans la grille de reconstruction, au moins au niveau du CC. Il est en revanche difficile de dire si l'un de ces deux modèles est plus précis que l'autre, les valeurs obtenues selon les trois métriques étant très proches. De plus, ces valeurs sont tantôt en faveur du modèle de rotation d'axe mobile, tantôt en faveur du modèle de rotation d'axe fixe.

En utilisant le modèle de rotation d'axe mobile, l'examen ayant obtenu le pire Dice et la pire *MAD* est montré sur la Figure.IV.25 et les examens ayant respectivement obtenus le meilleur Dice et la meilleure *MAD* sont visibles sur la Figure.IV.26.a et la Figure.IV.26.b.

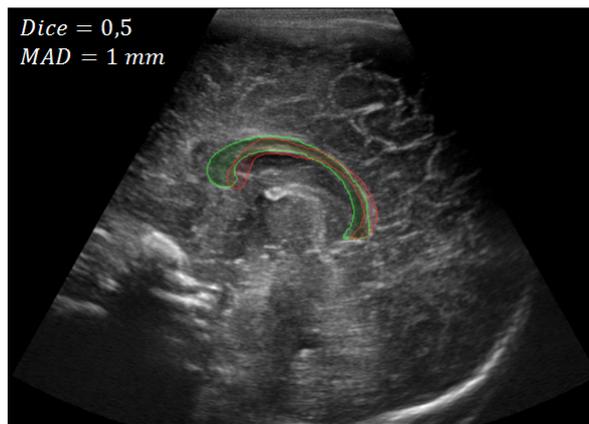


FIGURE IV.25 – Examen pour lequel la MAD et le Dice, calculés sur les segmentations du CC (référence en vert et reconstruction en rouge), étaient les moins bons.

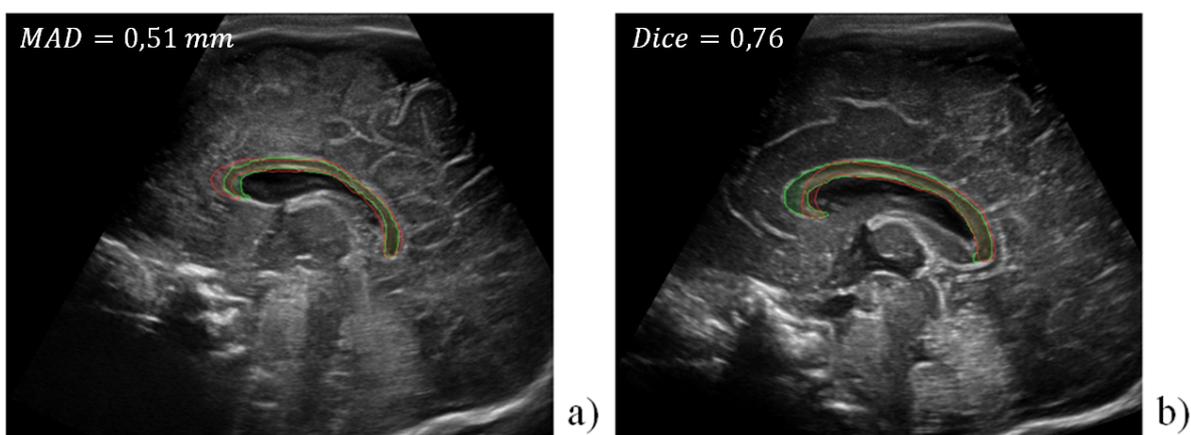


FIGURE IV.26 – Examens pour lesquels la MAD (a) et le Dice (b), calculés sur les segmentations du CC (référence en vert et reconstruction en rouge), étaient les meilleurs.

On voit que les segmentations du CC sur les examens ayant obtenu le meilleur Dice et la meilleure MAD sont très proches. Dans le cas ayant donné la pire MAD et le pire Dice, on peut voir que la forme globale du CC reconstruit est correcte, néanmoins cette segmentation est légèrement décalée, ce qui fait que la valeur obtenue par chacune de ces métriques est mauvaise.

Discussion Ces résultats montrent que la précision n'est pas aussi bonne que dans la plupart des modalités d'imagerie à référentiel fixe (IRM, CT, ...). Dans notre cas, il y a des déformations variables, de presque 1 mm (Table.IV.6). En comparaison à l'IRM, la précision anatomique est bien moindre. En revanche, pour des images acquises en main libre, les coordonnées sont reconstruites avec une excellente précision, notre objectif consistant à obtenir des données 3D de haute-résolution est donc complètement atteint.

Dans l'avenir, on pourra envisager de réaliser des reconstructions plus précises et reproductibles en équipant la sonde d'un accéléromètre et d'un gyromètre. Il sera également nécessaire de savoir si les images sont placées avec précision dans les toutes les régions du cerveau, même si les images

obtenues dans le grille de reconstruction semblent globalement cohérentes avec ce qu'on observe sur I_{ref} . Pour évaluer la précision sur l'ensemble du cerveau, nous avons essayé de valider le modèle de reconstruction à partir d'un fantôme de tête de nouveau-né.

Validation sur fantôme

Pour évaluer de façon plus précise, et dans des données 3D, les performances des différentes étapes de la chaîne de reconstruction et plus particulièrement celle de l'étape de placement des images dans la grille de reconstruction, nous avons fait l'acquisition d'un fantôme de cerveau de nouveau-né. Ce fantôme (**Newborn Head (Complex)**) a été acheté à la société True Phantom, basée à Windsor au Canada. Il nous est parvenu, dans une première version, en Juin 2019. Celui-ci contient plusieurs structures cérébrales qui sont visibles sur la Figure.IV.27.

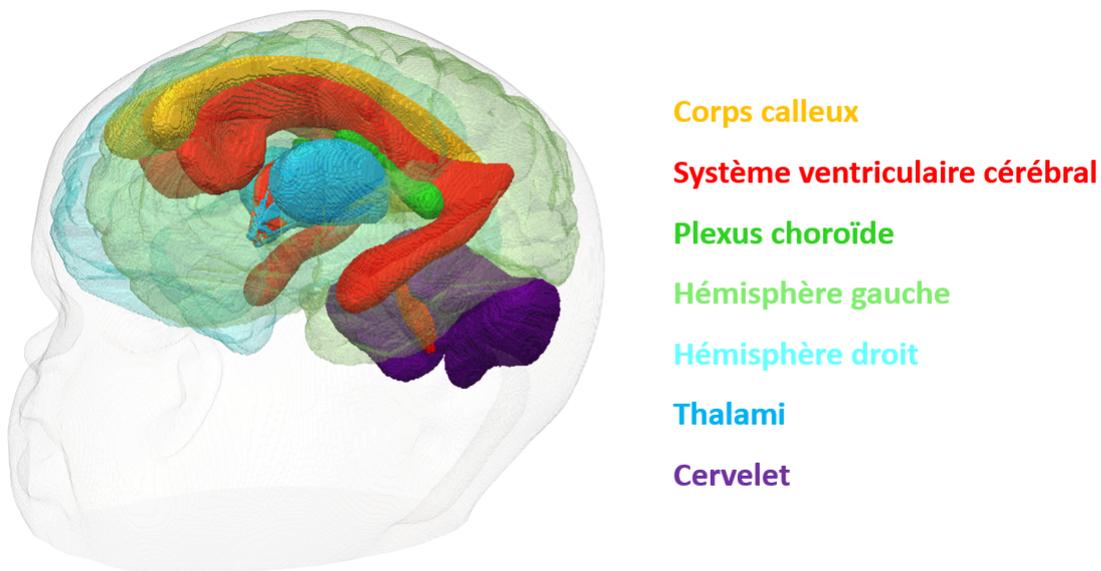


FIGURE IV.27 – Structures cérébrales du fantôme utilisé pour valider les modèles d'acquisition ².

En plus de ce fantôme, nous avons fait l'acquisition du modèle CAO utilisé pour créer celui-ci, dans le but de valider le placement des images dans la grille de reconstruction par rapport à ce modèle CAO.

Sept SDs ont été réalisées sur ce fantôme au CH d'Avignon avec la sonde 9L4. Nous avons essayé de créer des volumes échographiques à partir de celles-ci mais les résultats obtenus ne nous ont pas paru cohérents. Plusieurs raisons peuvent expliquer cela :

1. L'anatomie de certaines structures cérébrales, comme le SVC et le CC, est disproportionnée par rapport à la réalité dans le fantôme (Figure.IV.28). Comme on souhaite voir ces structures cérébrales entièrement sur I_{ref} et dans la SD, le médecin est obligé d'adapter son mouvement d'acquisition qui ne correspond plus exactement au mouvement qu'il effectue classiquement.
2. La matière dans laquelle est fait le fantôme est plus dure que la fontanelle antérieure, ce qui gêne la réalisation du mouvement d'acquisition.

2. Image créée à partir du modèle CAO acheté à la société True Phantom basée à Windsor au Canada.

3. Il y a moins de repères anatomiques identifiables sur l'image de référence acquise en coupe sagittale sur le fantôme que dans la réalité.
4. Le contraste entre les structures est moins bon que sur les images acquises in-vivo (Figure.IV.29)
5. Le fantôme ne contient pas de boîte crânienne. Celle-ci est une région hyperéchogène très utile lors de l'étape de recalage, son absence pourrait rendre cette étape moins performante et la définition d'une coupe I_{sag} plus difficile.

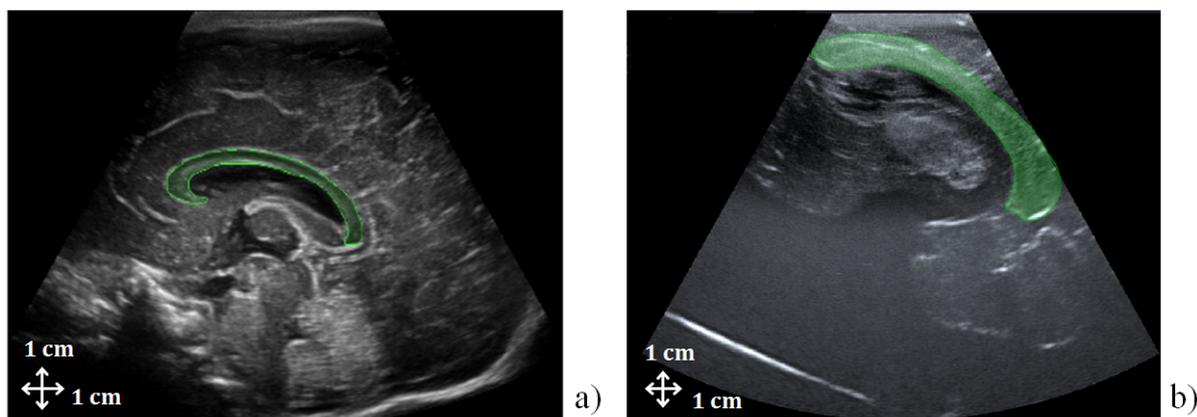


FIGURE IV.28 – Segmentations du CC (zone verte) sur une image I_{ref} (a) et sur une coupe sagittale acquise sur le fantôme (b).

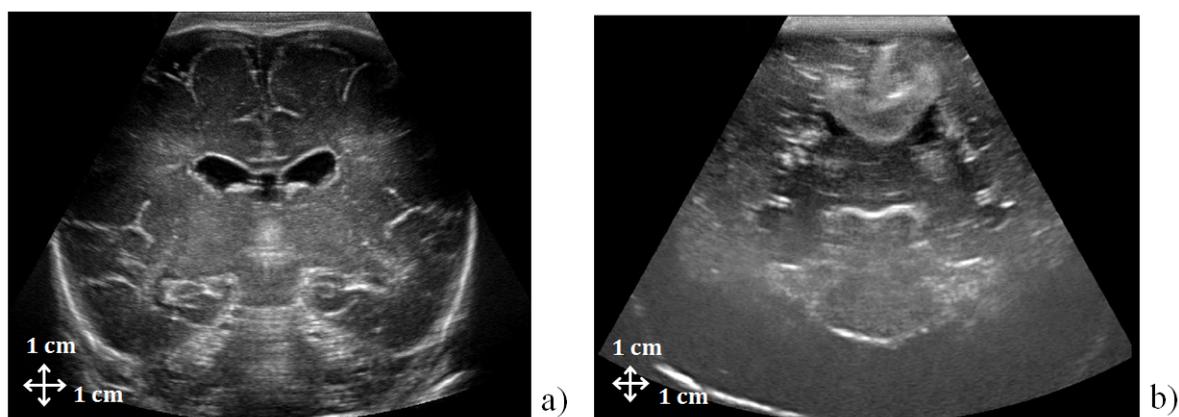


FIGURE IV.29 – Images acquises en coupe coronale sur un enfant prématuré (a) et sur le fantôme de cerveau (b).

Cette première étude nous a permis, en attendant d'être en possession du fantôme définitif, de nous rendre compte des difficultés que représentent l'acquisition de SDs sur celui-ci. Pour pouvoir utiliser ce fantôme pour valider les modèles d'acquisitions présentés dans cette partie, il faudra au préalable déterminer une manière de réaliser une acquisition de l'intégralité du SVC et des autres structures centrales de l'encéphale tout en effectuant le mouvement d'acquisition classiquement utilisé en clinique. On pourrait par exemple réaliser des acquisitions où le fantôme serait immergé dans un bac d'eau ou de gel échographique.

IV.3.5 conclusion

Cette partie a présenté comment les séquences d'images obtenues après l'étape de recalage ont été placées dans une grille de reconstruction de manière à obtenir des données échographiques 3D. Cette étape a été réalisée en optimisant des modèles paramétriques représentant le mouvement d'acquisition réalisé par le pédiatre. La validation de ces modèles a été effectuée sur des segmentations du CC en coupe sagittale, elle a conclu une bonne précision des modèles de rotation d'axe mobile et d'axe fixe. Une validation plus complète des modèles, dans des données 3D et sur l'ensemble du cerveau, devra néanmoins être menée de manière à avoir plus de recul sur leur précision et si l'un des deux modèles est plus adapté que l'autre. En attendant de pouvoir réaliser cette validation, les volumes de la base de données ont été créés en utilisant le modèle de rotation d'axe mobile.

IV.4 Interpolation des volumes reconstruits

Après avoir placé les images dans la grille de reconstruction, il a été nécessaire d'utiliser une méthode d'interpolation pour attribuer une valeur à chacun des voxels de la grille. Cette étape est délicate pour plusieurs raisons :

- Les coordonnées des pixels des images ne correspondent pas exactement aux coordonnées des voxels de la grille de reconstruction.
- La répartition des pixels dans la grille de reconstruction n'est pas homogène à cause du mouvement de rotation effectué durant l'acquisition (la densité de pixels est plus grande lorsqu'on se rapproche de la sonde).
- Certaines zones de la grille de reconstruction sont peu denses en pixels, notamment lorsqu'on s'éloigne du plan de la sonde.

Le choix d'une méthode d'interpolation permettant d'attribuer une valeur à chaque voxel de la grille de reconstruction a été fait pour répondre aux objectifs suivants : préserver la qualité des images (structures cérébrales ayant des bords réguliers et speckle préservé) obtenues avec la sonde 9L4 et être en mesure de construire un volume 3D dans un temps raisonnable (quelques minutes).

IV.4.1 Méthodes

Nous avons mis en œuvre plusieurs méthodes pour réaliser cette étape de reconstruction 3D. Parmi les méthodes les plus basiques, nous avons utilisé les méthodes PNN et VNN. Parmi les méthodes plus élaborées, nous nous sommes focalisés sur celles qui se basent sur les voxels (VBM) pour interpoler les valeurs de la grille de reconstruction, pour les raisons évoquées dans le Chapitre II. Parmi ces méthodes, nous avons retenu celles-ci :

- Interpolation de la valeur des voxels à partir de la valeur, pondérée, de chaque pixel se trouvant dans un voisinage sphérique autour de chaque voxel (interpolation à partir d'un voisinage sphérique).
- Interpolation de la valeur des voxels à partir de la valeur des pixels qui sont les plus proches voisins dans les deux images adjacentes à chaque voxel (interpolation à partir des plus proches voisins).

Les sections suivantes détaillent ces méthodes et les paramètres qui ont été utilisés lors de leur mise en œuvre.

PNN

La méthode PNN consiste à parcourir chaque pixel des images et à attribuer sa valeur au voxel de la grille de reconstruction qui lui est le plus proche. Cette méthode est représentée sur la Figure.IV.30.

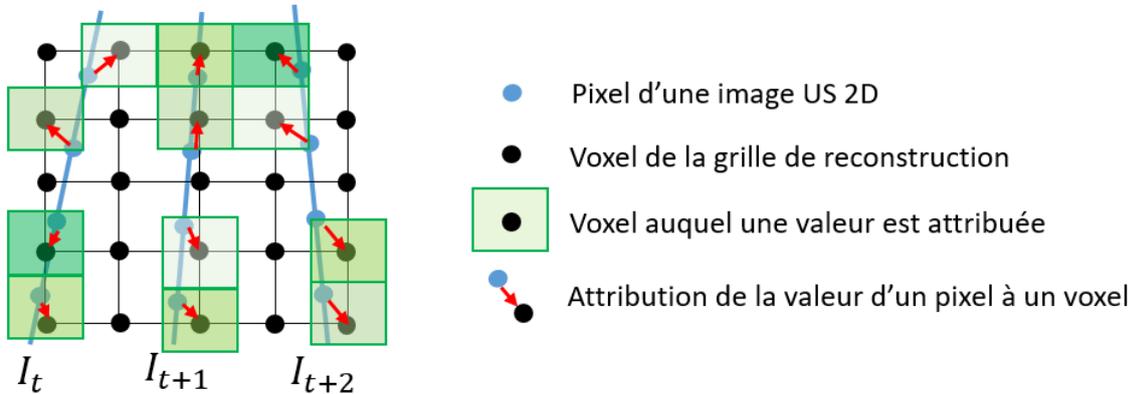


FIGURE IV.30 – Interpolation de la valeur de chaque voxel de la grille de reconstruction par la méthode PNN.

Elle a l'avantage de permettre l'obtention d'un volume échographique en quelques secondes, son inconvénient majeur est qu'elle laisse en général des voxels vides dans la grille de reconstruction, comme on peut le voir sur la Figure.IV.30.

VNN

La méthode VNN consiste à parcourir chaque voxel de la grille de reconstruction et à lui attribuer la valeur du pixel qui lui est le plus proche, cette méthode est représentée sur la Figure.IV.31.

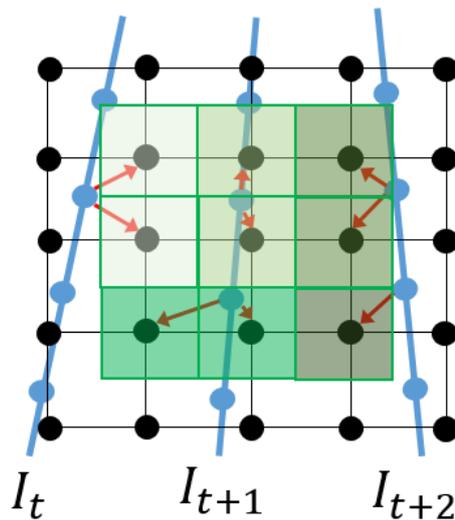


FIGURE IV.31 – Interpolation de la valeur de chaque voxel de la grille de reconstruction par la méthode VNN.

Elle a l'avantage de ne pas laisser de voxels vides dans la grille de reconstruction et de permettre l'obtention d'un volume échographique en quelques secondes. Néanmoins, la qualité de la reconstruction n'est en général pas très bonne car cette méthode a tendance à former des groupes de voxels adjacents de même valeur, comme le voit sur la Figure.IV.31.

Interpolation à partir d'un voisinage sphérique

Cette méthode consiste à parcourir chaque voxel de la grille de reconstruction, à rechercher les pixels qui sont dans un voisinage sphérique de rayon r autour de celui-ci et à lui attribuer l'intensité moyenne de ces pixels en les pondérant par l'inverse de leur distance au voxel. Ainsi l'intensité $G(v)$ d'un voxel v de la grille de reconstruction G est donnée par l'Équation.IV.26.

$$G(v) = \frac{\sum_{p \in S_r(v)} \frac{1}{d(p,v)} SI(p)}{\sum_{p \in S_r(v)} \frac{1}{d(p,v)}} \quad (\text{IV.26})$$

Où $S_r(v)$ désigne la sphère de rayon r et de centre v , $d(p, v)$ la distance euclidienne entre p et v et $SI(p)$ l'intensité du pixel p de la séquence d'image (SI) placée dans la grille de reconstruction. Cette méthode est représentée graphiquement, en 2D par soucis de simplification, sur la Figure.IV.32.

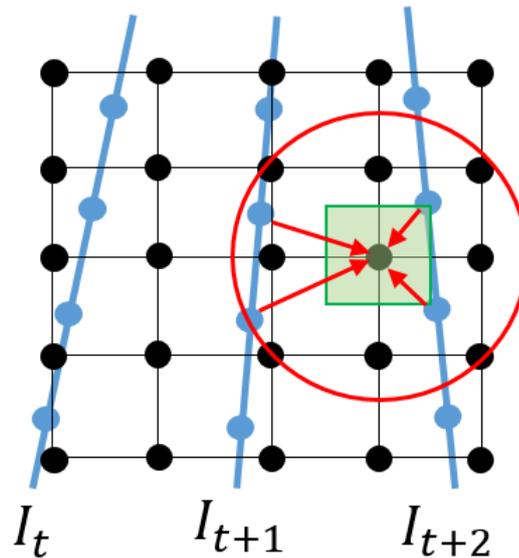


FIGURE IV.32 – Interpolation de la valeur de chaque voxel de la grille de reconstruction à partir de celles des pixels se situant dans un voisinage sphérique.

Cette méthode a l'avantage de ne pas laisser de trous dans la grille de reconstruction si le rayon r est bien choisi. De plus, elle utilise les vraies coordonnées de plusieurs pixels pour attribuer une valeur à chaque voxel. Cette méthode est plus précise que les méthodes PNN et VNN.

Ces principaux inconvénients sont que le rayon r optimal peut varier d'un examen à l'autre en fonction de la vitesse de la sonde et que, comme la répartition des pixels n'est pas homogène dans la grille de reconstruction, le nombre de pixels utilisé pour interpoler la valeur d'un voxel varie.

Nous avons implémenté cette méthode en choisissant $r = 3$ et $r = 5$.

Interpolation à partir des plus proches voisins

Cette méthode consiste à parcourir chaque voxel de la grille de reconstruction, à chercher ses k plus proches pixels voisins sur les images I_t ($kNNI_t$) et I_{t+1} ($kNNI_{t+1}$) qui lui sont adjacentes, puis à lui attribuer la moyenne, pondérée par l'inverse de la distance, des intensités des pixels. Dans

notre cas, nous avons aussi cherché les k plus proches voisins (lorsque $k > 1$) qui se trouvent dans les deux coupes sagittales adjacentes à la coupe dans laquelle se trouvait le voxel v auquel on cherchait à attribuer une valeur; ce qui permet de réaliser une interpolation en 3D. L'intensité $G(v)$ d'un voxel de la grille de reconstruction est donnée par l'Équation.IV.27 dans le cas où $k = 3$.

$$G(v) = \frac{\sum_{p=(x,y,z) \in_k NNI_t(v) \cup_k NNI_{t+1}(v)} \frac{1}{d(p,v)} SI(p) + \frac{1}{\sqrt{d(p,v)^2+1}} (SI(x,y-1,z) + SI(x,y+1,z))}{\sum_{p \in S_r(v)} \frac{1}{d(p,v)} + \frac{2}{\sqrt{d(p,v)^2+1}}} \quad (\text{IV.27})$$

et par l'Équation.V.3 dans le cas où $k = 1$.

$$G(v) = \frac{\sum_{p \in_k NNI_t(v) \cup_k NNI_{t+1}(v)} \frac{1}{d(p,v)} SI(p)}{\sum_{p \in S_r(v)} \frac{1}{d(p,v)}} \quad (\text{IV.28})$$

Elle est représentée, en 2D par soucis de simplification, dans le cas où $k = 3$ sur la Figure.IV.33.

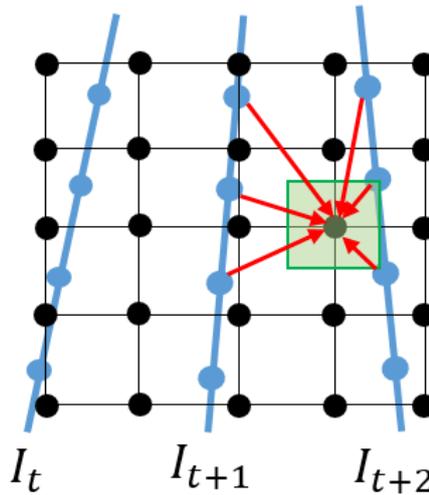


FIGURE IV.33 – Interpolation de la valeur de chaque voxel de la grille de reconstruction à partir des valeurs des pixels se situant sur les images adjacentes.

Cette méthode a l'avantage de ne pas laisser de trous dans la grille de reconstruction et de toujours utiliser le même nombre de pixels pour interpoler la valeur d'un voxel. De manière à réduire les temps de calcul, la recherche des plus proches voisins a été réalisée en construisant des kd-tree Bentley (1975) avec la fonction KDTreeSearcher de Matlab.

Dans la prochaine partie, ces méthodes sont comparées en temps de calcul et par des résultats qualitatifs.

IV.4.2 Temps d'interpolation et résultats qualitatifs

Les performances des quatre méthodes décrites précédemment ont été évaluées en reconstruisant tous les examens de la base de données à partir du modèle de rotation d'axe mobile.

Temps d'interpolation

De manière à limiter le temps de reconstruction, les paramètres permettant d'effectuer l'interpolation ont été déterminés dans une coupe sagittale de la grille de reconstruction puis ont été appliqués à toutes les autres coupe sagittale car les pixels y ont tous les mêmes coordonnées. De plus, dans le cas des méthodes basées sur les voxels, un masque des voxels auxquels on cherchait à attribuer une valeur a été créée, à partir du modèle d'acquisition, de manière à ce que les algorithmes d'interpolation ne considèrent que ces pixels. Les temps moyens d'interpolation sont donnés pour chacune des quatre méthodes dans la Table.IV.7.

TABLE IV.7 – Temps moyen d'interpolation de la valeur des voxels de la grille de reconstruction.

Méthode	Temps de reconstruction (s)
PNN	7.0 ± 1.0
VNN	5.0 ± 1.0
Sphérique ($r = 3$)	321.0 ± 90.0
Sphérique ($r = 5$)	656.0 ± 181.0
1 NN	60.0 ± 21.0
3 NN	140.0 ± 39.0

Comme attendu, les méthodes PNN et VNN sont les plus rapides et permettent de reconstruire ces volumes de taille importante en quelques secondes. Les méthodes utilisant les plus proches voisins sont relativement rapides et mettent entre 1 et 3 minutes pour reconstruire un volume. Les méthodes d'interpolation dans un voisinage sphérique sont plus lentes que les méthodes basées sur les plus proches voisins, le temps de reconstruction moyen est légèrement supérieur à 5 minutes avec $r = 3$ (ce qui est raisonnable) et d'environ 11 minutes avec $r = 5$ (ce qui est un peu trop long). Si on se base uniquement sur les temps de reconstruction, toutes les méthodes, sauf l'interpolation dans un voisinage sphérique de rayon $r = 5$, sont adaptées pour répondre à nos objectifs.

Résultats qualitatifs

Cette partie présente les résultats qualitatifs que nous avons obtenus avec chaque méthode.

Les résultats obtenus avec la méthode PNN sont visibles sur la Figure.IV.34.a en coupe sagittale et sur la Figure.IV.34.b en coupe coronale.

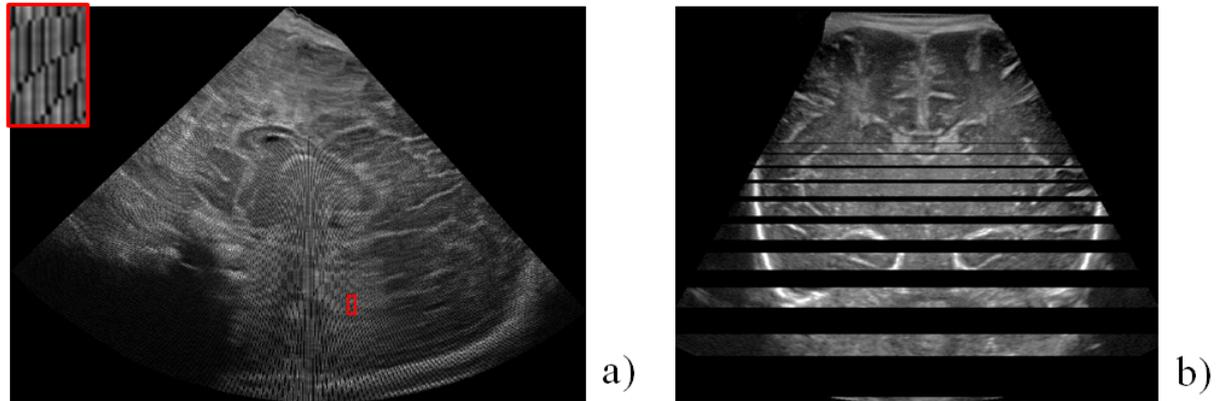


FIGURE IV.34 – Exemple de volume interpolé avec la méthode PNN en coupe sagittale (a) et coronale (b). Un zoom sur une zone où la grille de reconstruction présente des valeurs nulles est affiché sur (a).

Comme attendu, cette méthode laisse des voxels vides dans la grille de reconstruction. Elle ne répond donc pas aux objectifs que nous avons fixés. Les résultats obtenus avec la méthode VNN sont présentés sur la Figure.IV.35.

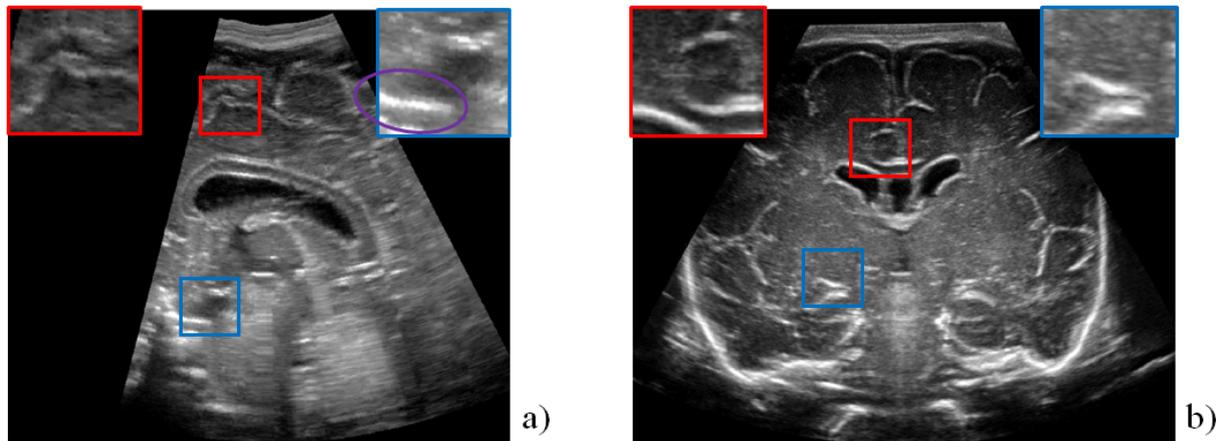


FIGURE IV.35 – Coupe sagittale (a) et coronale (b) d'un volume échographique après interpolation des valeurs de la grille de reconstruction avec la méthode VNN.

La qualité des images obtenues semble qualitativement correcte en coupe sagittale et en coupe coronale. Le speckle paraît particulièrement bien conservé en coupe coronale. On observe néanmoins de légères irrégularités en coupe sagittale au niveau de la zone entouré en violet sur la Figure.IV.35.a qui pourraient être causées par des plages de voxels adjacents de même valeur.

Des résultats qualitatifs obtenus avec la méthode d'interpolation dans un voisinage sphérique sont présentés sur la Figure.IV.36 avec $r = 3$ et sur la Figure.IV.37 avec $r = 5$.

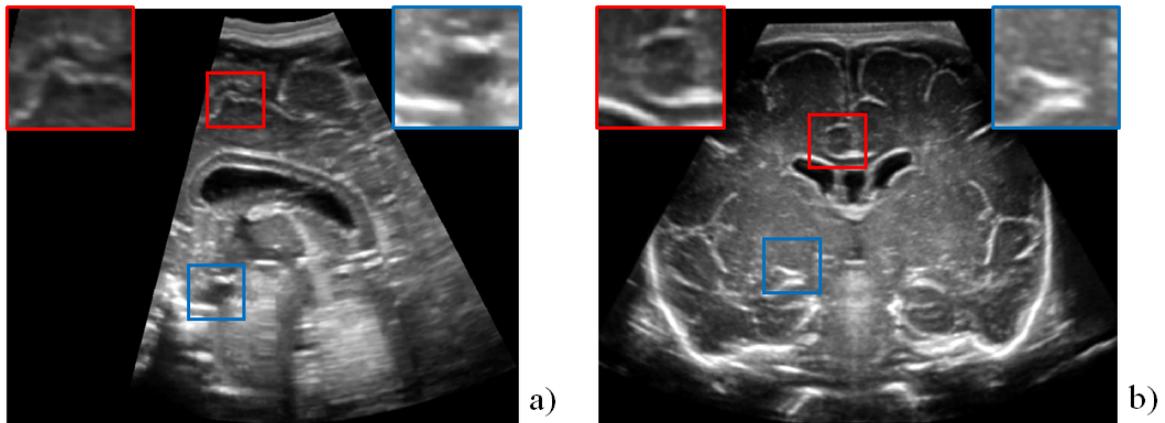


FIGURE IV.36 – Coupe sagittale (a) et coronale (b) d'un volume échographique après interpolation des valeurs de la grille de reconstruction en utilisant un voisinage sphérique de rayon 3 autour de chaque voxel.

1

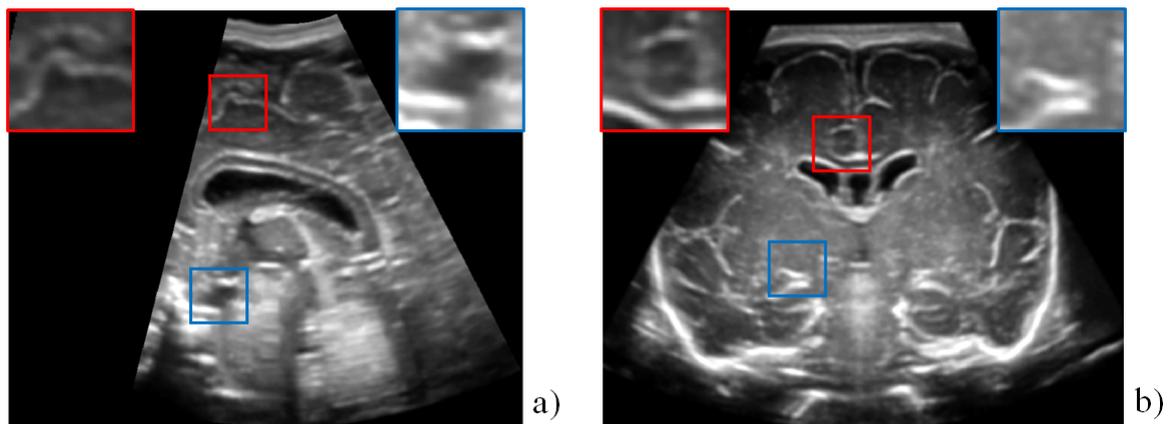


FIGURE IV.37 – Coupe sagittale (a) et coronale (b) d'un volume échographique après interpolation des valeurs de la grille de reconstruction en utilisant un voisinage sphérique de rayon 5 autour de chaque voxel.

En utilisant un rayon $r = 3$, les structures paraissent continues sur les images obtenues et on observe toujours du speckle même si celui-ci est réduit par rapport aux images obtenues avec la sonde 9L4. Ce rayon permet d'attribuer une valeur à tous les voxels de la grille dans la plupart des cas, des trous y étaient visible après interpolation pour 1 seul examen. Lorsqu'on fixe $r = 5$, le bord des structures cérébrales apparaît continue mais le speckle n'est presque plus visible sur les images, ce qui est à l'inverse de l'objectif recherché.

Des résultats qualitatifs obtenus par la méthode d'interpolation à partir des k plus proches voisins sont présentés sur la Figure.IV.38.a avec $k = 1$ et sur la Figure.IV.39.b avec $k = 3$.

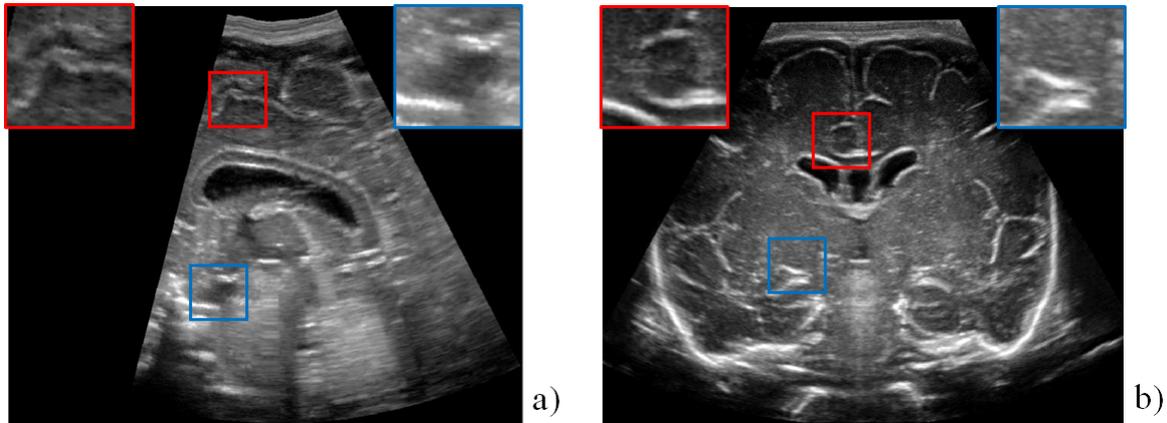


FIGURE IV.38 – Coupe sagittale (a) et coronale (b) d'un volume échographique après interpolation des valeurs de la grille de reconstruction en utilisant la valeur la plus proche sur chacune des images adjacentes à chaque voxel.

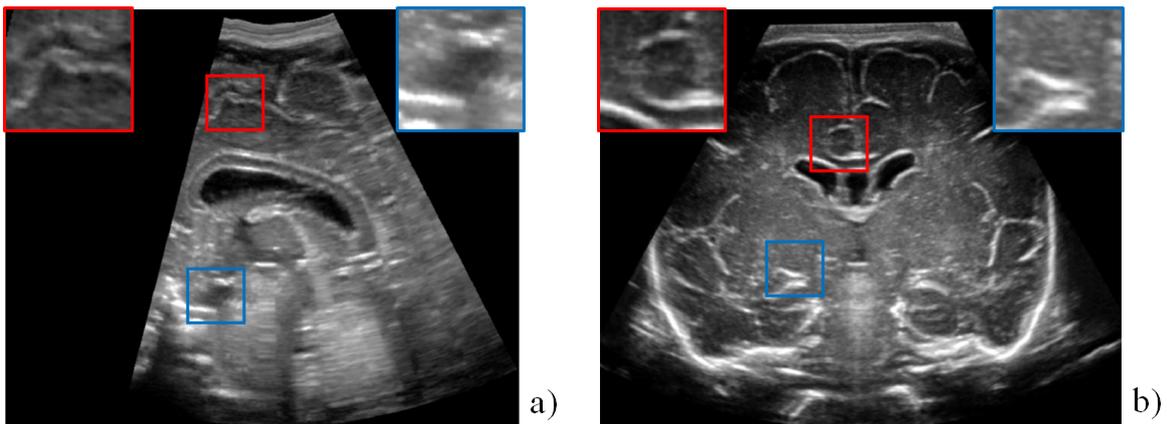


FIGURE IV.39 – Coupe sagittale (a) et coronale (b) d'un volume échographique après interpolation des valeurs de la grille de reconstruction en utilisant les trois valeurs les plus proches sur chacune des images adjacentes à chaque voxel.

Le speckle semble bien préservé sur les images obtenues avec $k = 1$, on observe néanmoins de légères irrégularités sur les bords des structures en coupe sagittale, comme avec la méthode VNN. Avec $k = 3$, le speckle est un peu dégradé par rapport aux images obtenues avec $k = 1$, mais il reste bien visible. Les irrégularités observées en coupe sagittale sont en revanche moins marquées.

La Figure.IV.40 compare des résultats, en coupe sagittale (cadre bleu) et coronale (cadre rouge), obtenus avec les différentes méthodes d'interpolation.

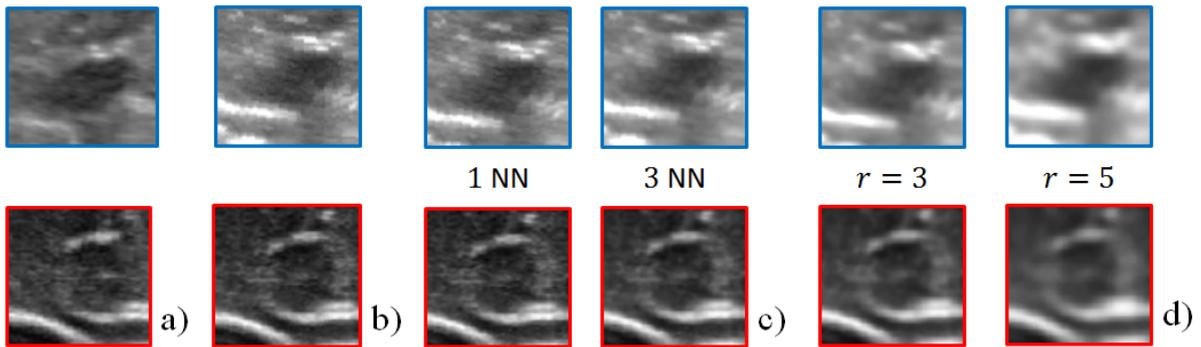


FIGURE IV.40 – Extraits d’images en coupe sagittale (cadre bleu) et coronale (cadre rouge) acquises avec les sondes 9L4 (a) et après interpolation avec la méthode VNN (b), la méthode d’interpolation dans un voisinage sphérique (c) et la méthode d’interpolation à partir des plus proches voisins (d).

En comparaison aux images de référence acquises avec la sonde 9L4 (Figure.IV.40.a), la méthode qui semble réaliser le meilleur compromis entre régularité des structures et préservation du speckle est la méthode d’interpolation à partir des plus proches voisins avec $k = 3$ (Figure.IV.40.c).

Dans les images obtenues avec la méthode d’interpolation dans un voisinage sphérique, les structures sont régulières mais le speckle n’est pratiquement plus visible. A contrario, sur les images obtenues avec la méthode VNN et la méthode d’interpolation à partir des plus proches voisins avec $k = 1$, le speckle est bien préservé mais certaines structures sont irrégulières en coupe sagittale.

IV.4.3 Conclusion

Dans cette partie, plusieurs méthodes d’interpolation ont été comparées. Leur objectif était d’interpoler une valeur à chaque voxel de la grille de reconstruction pour obtenir un volume échographique. Pour choisir la méthode la plus adaptée à notre problématique, elles ont été comparées en temps de calcul et de manière qualitative sur des images en coupe coronale et sagittale. L’objectif était d’obtenir des structures ayant des bords réguliers tout en préservant au mieux le speckle. D’après les résultats qui ont été présentés, la méthode qui répond le mieux à ces critères est la méthode d’interpolation à partir des plus proches voisins avec $k = 3$. Le temps moyen obtenu pour reconstruire un volume avec cette méthode a été mesuré à 2 minutes et 20 secondes (± 39 s), les structures cérébrales reconstruites avaient un aspect régulier en coupe sagittale et coronale et le speckle était correctement préservé dans ces deux plans.

IV.5 Positionnement dans le repère anatomique

Une fois l'étape de reconstruction réalisée, les volumes obtenus doivent être orientés selon les plans anatomiques de référence. Ce positionnement permet de faciliter l'interprétation des volumes échographiques et d'effectuer des comparaisons avec des volumes obtenus par IRM.

IV.5.1 Méthode

Ce problème peut être vu comme un problème de recalage, la problématique est alors de recalculer les volumes échographiques sur des volumes orientés selon les plans anatomiques de référence. Comme cela a déjà été évoqué, il n'existe pas de base de données échographiques 3D en libre accès, ce recalage ne peut donc être effectué que sur des volumes IRM. Nous avons essayé de faire ce recalage sur l'atlas Alberts (Gousias *et al.* (2012)) en utilisant le logiciel Elastix (<http://elastix.isi.uu.nl/>) mais ne sommes pas parvenus à obtenir des résultats convenables. Néanmoins, nous avons choisi l'information mutuelle, qui est classiquement utilisée dans les problèmes de recalage multimodal, comme métrique pour recalculer les volumes. Il est possible qu'utiliser la métrique LC^2 (Fuerst *et al.* (2014)) comme dans Boucher *et al.* (2018b) donne de meilleurs résultats.

Avant d'orienter les volumes selon les plans anatomiques de référence, une origine commune (O_{CC}) a été définie dans chaque volume, dans le but de pouvoir les superposer de manière cohérente. Le splénium du CC, qui est visible sur la Figure IV.41 a été choisi comme origine commune.

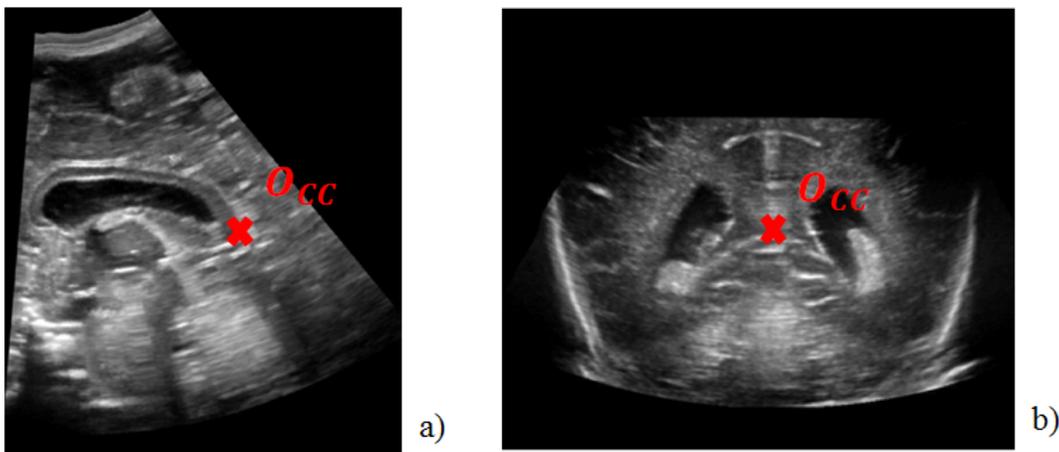


FIGURE IV.41 – Emplacement du splénium du corps calleux en coupe sagittale (a) et coronale (b).

Notre choix s'est porté sur cette structure car elle est située près du centre de l'encéphale et peut être déterminée sans équivoque dans tous les examens.

Après la définition de cette origine commune, les volumes ont été orientés selon les plans anatomiques de référence en choisissant les angles ϕ et ψ pour chaque examen. L'angle de rotation dans le plan sagittal (ϕ) a été choisi à l'aide d'une grille superposée à une image en coupe sagittale à laquelle une rotation avait été appliquée (angle compris entre 0° et 30° par pas de 1°). Pour que l'image soit correctement orientée, il faut que le CC soit à plat, c'est à dire que la grille superposée à l'image

soit tangente au bourlet, au splénium et au dos du corps calleux. Ce choix d'orientation du CC a été fait en discussion avec le pédiatre, il correspond à la convention utilisée en IRM pour orienter les volumes lors des acquisitions. Cette étape est visible sur la Figure. IV.42. Après rotation, les volumes ont été reconstruits avec une interpolation bilinéaire.

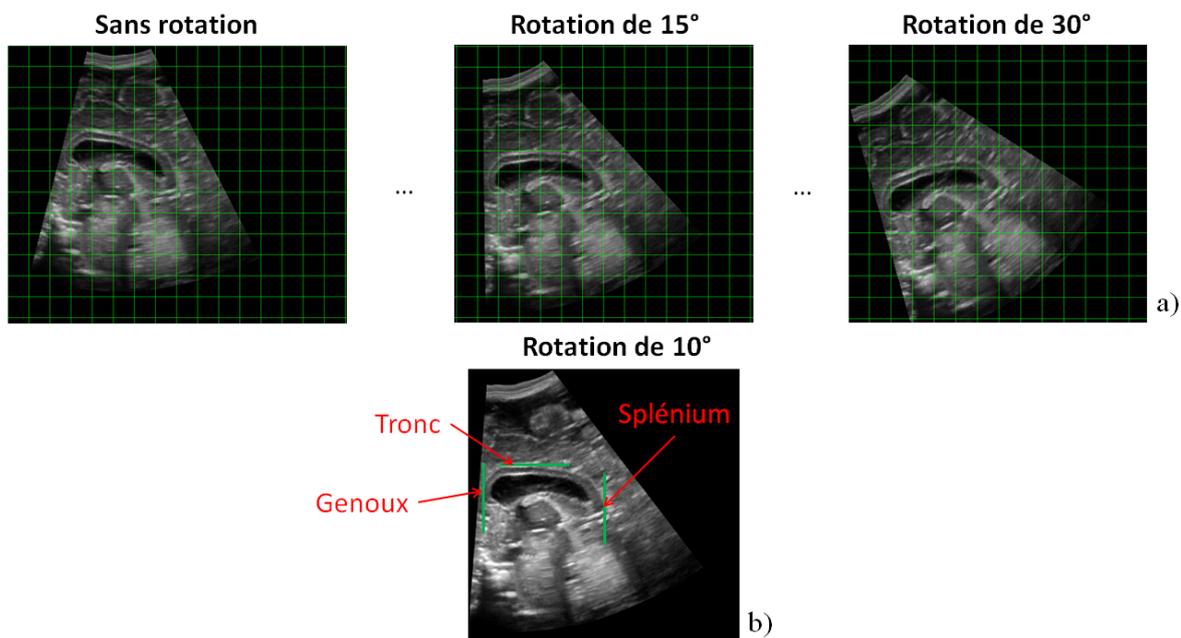


FIGURE IV.42 – Coupe sagittale à différents angles de rotation et grille permettant de déterminer l'orientation du CC (a). Angle de rotation choisi, pour lequel le CC est à plat (b).

Un angle de rotation de 10° a par exemple été choisi sur la Figure. IV.42. b pour orienter le volume reconstruit. La même méthode a été utilisée pour déterminer l'angle de rotation dans le plan coronal (ψ).

Cette étape nous a permis d'orienter les volumes selon les plans anatomiques de références et de tous les centrer en un même points O_{CC} . Les valeurs moyennes des angles de rotation utilisés ainsi que des résultats qualitatifs sont donnés dans la partie suivante.

IV.5.2 Résultats qualitatifs

Les valeurs moyennes des angles de rotations déterminées lors de cette étape sont données dans la Table. IV.8

TABLE IV.8 – Valeurs des angles de rotation en coupe sagittale et coronale

Angle de rotation	Valeur moyenne ($^\circ$)	Valeur minimale ($^\circ$)	Valeur maximale ($^\circ$)
Coupe coronale	-0.1 ± 1.2	-3.0	2.0
Coupe sagittale	15.7 ± 5.6	4	30

On constate que l'angle de rotation dans le plan sagittal est assez important et que l'angle de rotation dans le plan coronal est faible, comme nous l'avions supposé lors de la formulation de la

deuxième hypothèse sur le mouvement d'acquisition. La valeur moyenne de ϕ est cohérente car la fontanelle antérieure est située à l'avant de la tête des nouveaux-nés, celle de ψ l'est aussi car les images sont acquises dans une coupe proche d'une coupe coronale. Des exemples de volumes échographiques avant et après positionnement dans le repère commun sont visibles sur la Figure.IV.43.

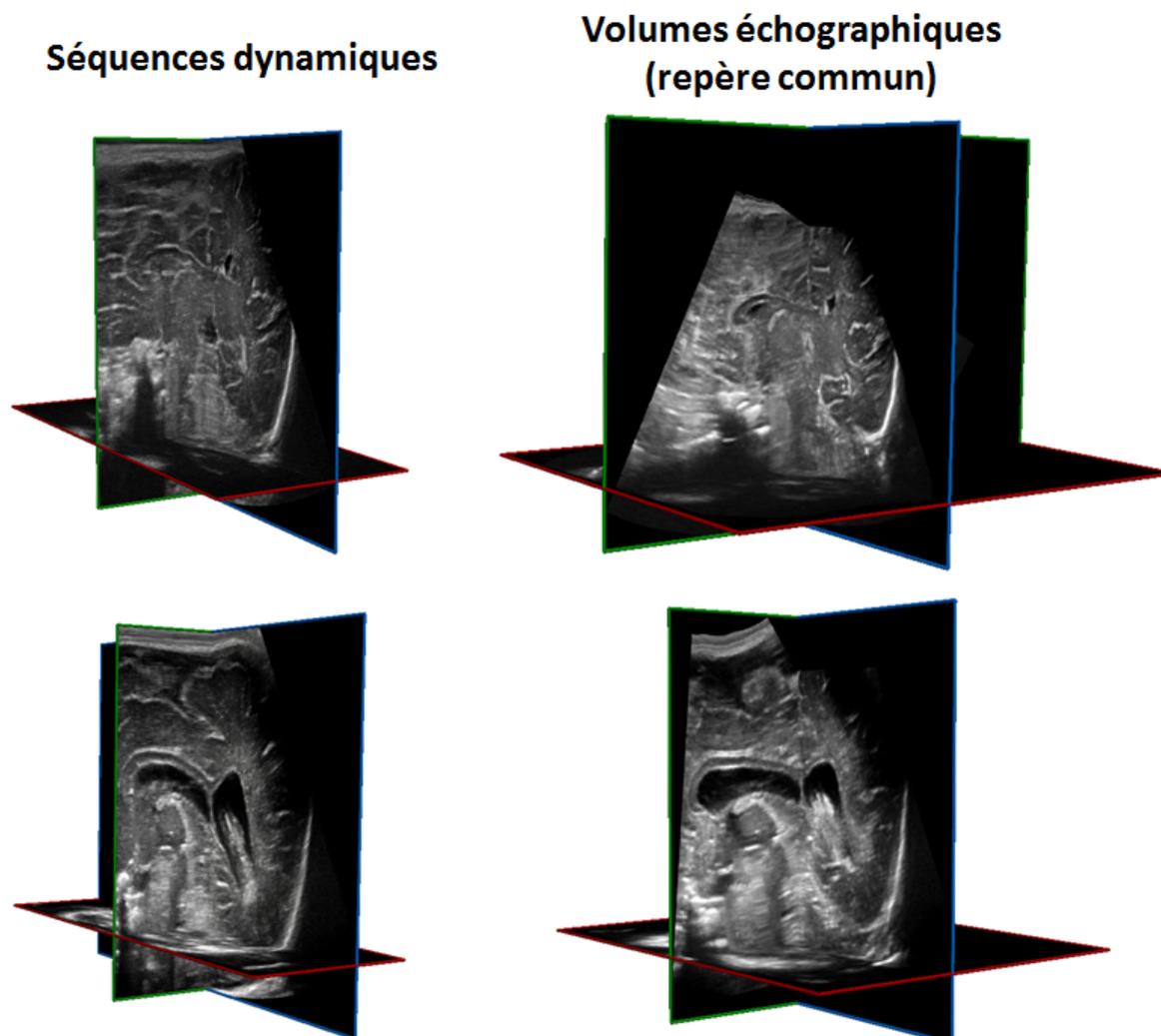


FIGURE IV.43 – Comparaison de deux séquences dynamiques et des volumes échographiques associés, placés dans le repère commun.

Une comparaison, dans les trois plans anatomiques, entre un volume échographique dans le repère commun et une acquisition IRM est montrée sur la Figure.IV.44

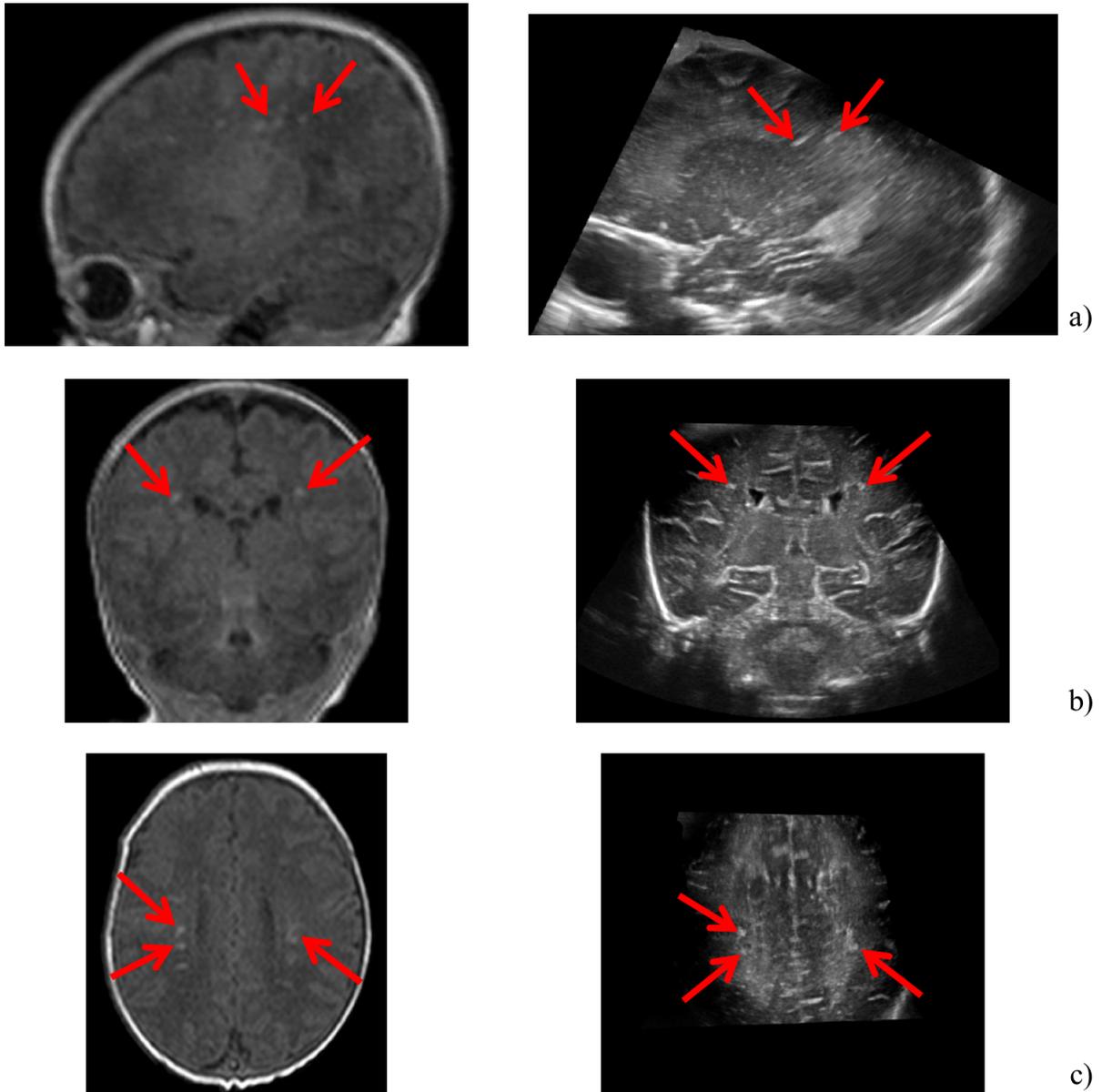


FIGURE IV.44 – Comparaison, en coupe sagittale (a), coronale (b) et axiale (c), d'images issues d'une IRM et d'un volume échographique replacé dans le repère commun. Les PWML sont indiquées par des flèches rouges.

Les examens comparés ont été réalisés sur le même patient à des jours différents. En coupe axiale (Figure.IV.44.c), on voit que les PWML ont une forme similaire sur l'IRM et l'échographie. Des observations similaires peuvent être faites en coupe coronale (Figure.IV.44.b) et sagittale (Figure.IV.44.a). Ces observations donnent du crédit à la qualité des volumes échographiques obtenus jusqu'à cette étape.

IV.6 Débruitage

La dernière étape de la chaîne de reconstruction est une étape de débruitage. Son objectif est de rendre l'interprétation des volumes échographiques plus facile, en leur donnant un aspect plus proche de celui d'IRM, et d'augmenter le contraste de certaines structures de manière à pouvoir les segmenter manuellement plus facilement.

Les volumes échographiques ont été débruités à partir d'une décomposition dans une base d'ondelette (Sendur *et Selesnick* (2002)) et de l'application d'un seuil $h = 14$. Ce débruitage a été réalisé avec le code fourni à ce lien : <http://eeweb.poly.edu/iselesni/WaveletSoftware/dt3D.html>. Des images, dans les trois plans anatomiques, avant et après débruitage sont visibles sur la Figure.IV.45.

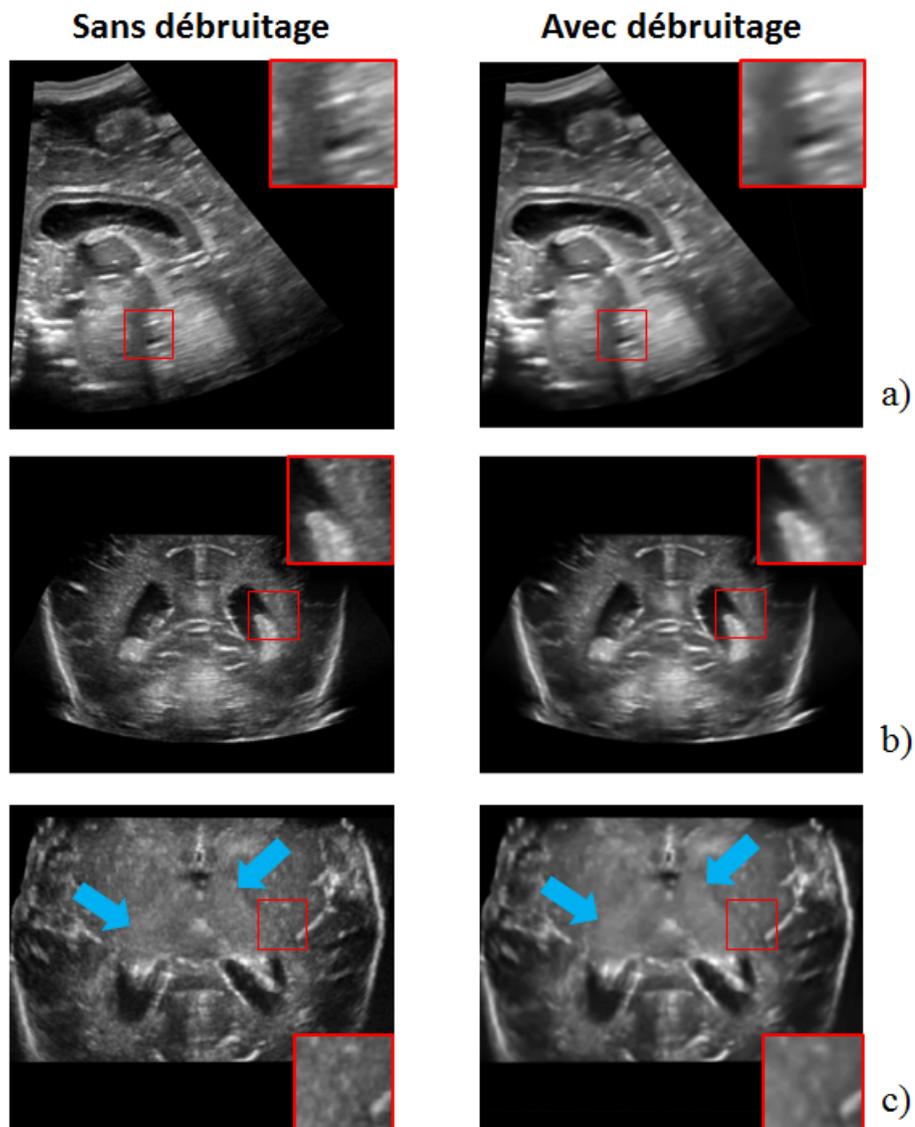


FIGURE IV.45 – Volumes échographiques dans le repère commun avec et sans débruitage, en coupe sagittale (a), coronale (b) et axiale (c), le bord des thalami est indiqué par des flèches bleues en coupe axiale.

On observe que le speckle a pratiquement été éliminé dans tous les plans, ce qui donne un aspect proche d'IRM à échographies. De plus, on voit sur la Figure.IV.45.c que le bord des thalami est nettement plus visible en coupe axiale après débruitage.

IV.7 Créations de volume échographiques à partir d'ETFs réalisées avec d'autres sondes

Cette partie a pour objectif de montrer qu'il est possible d'utiliser l'algorithme présenté dans ce chapitre pour créer des données échographiques 3D à partir d'examen réalisés avec une autre sonde que la 9L4. Pour donner une preuve de faisabilité, des ETFs 2D ont été réalisées par le pédiatre du CH d'Avignon avec les sondes suivantes pour créer des volumes échographiques :

- Sonde Philips eL18-4 PureWave.
- Sonde Siemens multi-D 10L4.
- Sonde Siemens 14L5.

Pour ces trois sonde, les résultats obtenus après positionnement (modèle de rotation d'axe mobile) et interpolation dans la grille de reconstruction sont respectivement visibles sur les Figure.IV.46.a, IV.46.b et IV.46.c.

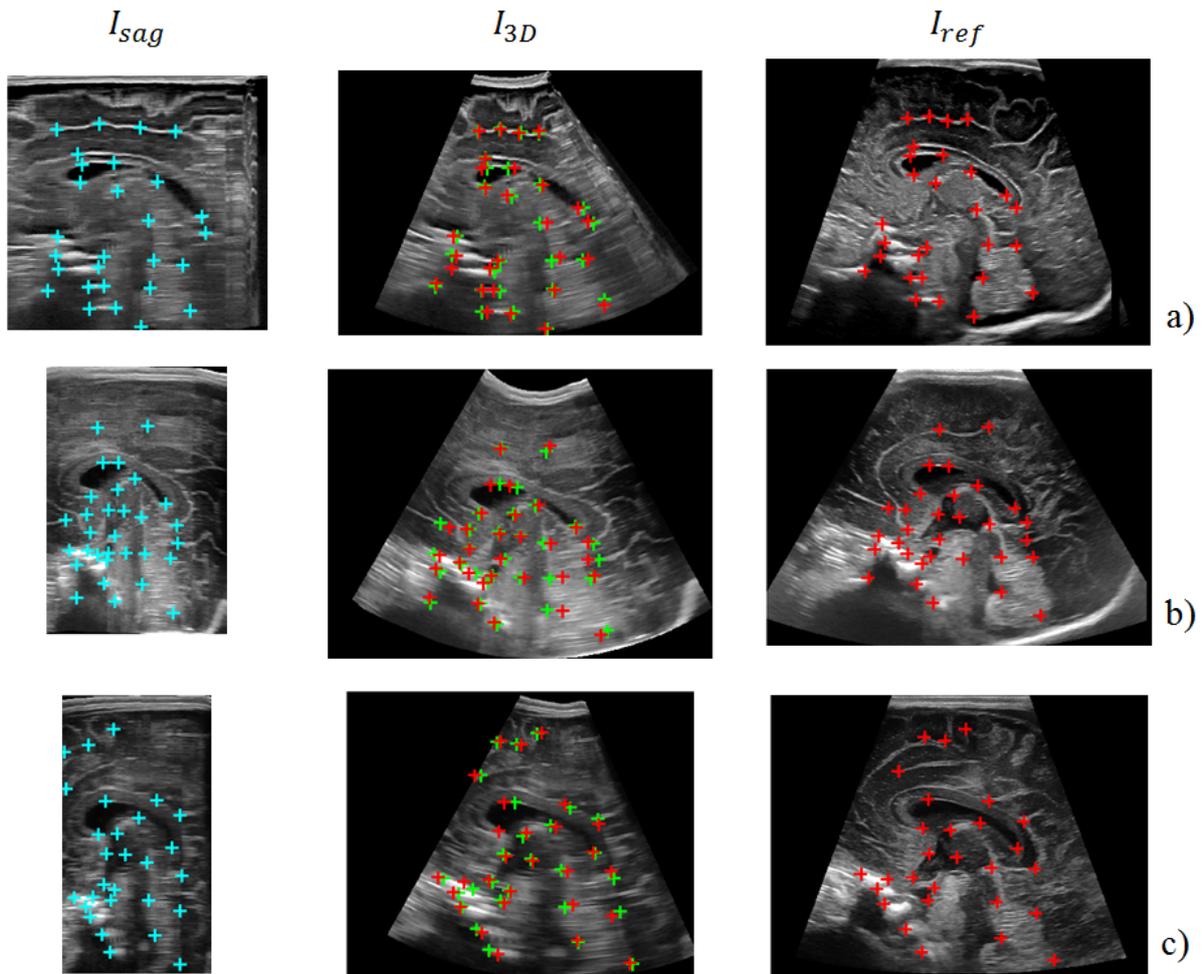


FIGURE IV.46 – Création d'un volume échographique 3D à partir d'une ETF 2D réalisée avec une sonde Philips eL18-4 PureWave (a), Siemens multi-D 10L4 (b) et Siemens 14L5 (c). Les points de repères anatomiques placés sur I_{sag} (en bleu) et sur I_{ref} (en rouge) ainsi que les points après reconstruction sur I_{3D} (en vert) sont également visibles.

Ces reconstructions paraissent cohérentes lorsqu'on compare les images I_{3D} aux images I_{ref} , le nuage de points de référence et le nuage de points après reconstructions semblent également très proches. La distance euclidienne moyenne entre ces derniers est donnée dans la Table.IV.9.

TABLE IV.9 – Distance moyenne entre les points de référence après reconstruction d'ETFs réalisées avec différentes sondes sur un examen.

Sonde	Philips eL18-4 PureWave	Siemens multi-D 10L4	Siemens 14L5
Distance moyenne (mm)	1.23 ± 0.73	1.66 ± 1.11	1.5 ± 0.79

La précision des reconstructions effectuées avec ces différentes sondes est comparable avec la précision obtenue avec la sonde Siemens multi-D 9L4 (Table.IV.4) : 1.5 ± 0.33 mm. En se basant sur ces résultats qualitatifs et quantitatifs, on peut dire que l'algorithme développé peut être utilisé pour créer des données échographiques 3D à partir d'ETFs qui n'ont pas été réalisées avec la sonde 9L4. A priori, il devrait pouvoir être utilisé pour créer des données échographiques 3D à partir d'ETF 2D acquise avec n'importe quelle sonde, sous réserve que la qualité d'image soit suffisante et que le mouvement d'acquisition corresponde au modèle de reconstruction.

IV.8 Conclusion et perspectives

Résumé du travail réalisé. Ce chapitre a présenté les étapes de l'algorithme qui nous a permis d'obtenir des données échographiques 3D à partir des ETFs 2D acquises au CH d'Avignon. Celui-ci a été conçu pour préserver au mieux la qualité des images obtenues avec la sonde 9L4 tout en permettant de créer des volumes échographiques dans un temps raisonnable. Toutes les méthodes utilisées ont été choisies pour répondre à ce compris entre qualité et temps de calcul. Finalement, si on ne prend pas en compte les temps d'interactions avec l'utilisateur (positionnement des repères anatomiques, choix des angles de rotation, ...), le temps de calcul moyen pour obtenir un volume échographique 3D à partir d'une ETD 2D est de 6 minutes 45 secondes (± 1 min 42 s). Ce qui est raisonnable pour un algorithme actuellement réservé à la recherche. En outre, la qualité des volumes obtenus est largement supérieure à celle fournie par les sondes matricielles actuellement sur le marché. Les écarts entre ces volumes échographiques et la vérité terrain restent encore à évaluer. La précision anatomique des reconstructions est sans doute inférieure à ce qu'on peut obtenir par IRM, néanmoins les volumes obtenus répondent parfaitement aux objectifs que nous nous étions fixés.

Résumé des contributions. Les contributions majeures présentées dans ce chapitre sont :

- Le développement d'un algorithme permettant de créer un volume échographique du cerveau du prématuré à partir d'une ETF 2D, qui ne nécessite pas l'utilisation de matériel supplémentaire lors des acquisitions et qui peut être utilisé sur des acquisitions réalisées avec différentes sondes.
- La développement d'un modèle d'acquisition correspondant à une rotation d'axe mobile et d'une méthode d'optimisation de ses paramètres.
- L'implémentation de chaque étape de la chaîne de reconstruction sous la forme d'un bloc qui peut être modifié de manière indépendante du reste de la chaîne.
- L'optimisation des étapes de recalage et d'interpolation en temps de calcul tout en préservant la qualité des images.

Notre objectif qui consistait à créer des données échographiques 3D cohérentes, de haute-résolution, dans un temps raisonnable et sans utiliser de dispositif supplémentaire durant les acquisitions est donc complètement atteint. Ce travail a été valorisé par une présentation orale au congrès IUS ([Martin et al. \(2018\)](#)), il est prévu qu'un article de journal soit écrit dès lors que les modèles d'acquisitions auront pu être validés sur fantôme.

Perspectives. Dans l'avenir, il sera nécessaire de valider les modèles d'acquisition, en 3D, sur l'ensemble des structures du cerveau et non pas seulement, en 2D, sur le CC. Cette validation pourra être réalisée sur le fantôme de tête de nouveau-né, après avoir déterminé une méthode d'acquisition adéquate, ou avec des IRMs acquises le même jours que des ETFs 2D. Cela permettra d'avoir plus de recul sur la précision de cet algorithme et de déterminer si l'un des modèles d'acquisitions proposé est plus adapté que les autres.

Pour aller vers un usage clinique de cet algorithme, plusieurs étapes devront être automatisées :

- Étape 1 : éliminer automatiquement les champs de l'image qui ne correspondent pas à l'échographie sans avoir à définir leur position au préalable.
- Étape 3 : définir I_{sag} et identifier des points de repères anatomiques se correspondant sur I_{sag} et sur I_{ref} .
- Étape 5 : définir automatiquement l'origine du repère commun et la transformation à effectuer pour orienter le volume échographique selon les plans anatomiques de référence.

L'étape la plus difficile dans la liste ci-dessus est celle qui correspond au point 2.

L'un des inconvénients de la méthode proposée est qu'elle ne permet pas de créer des volumes échographiques en ligne, c'est à dire en même temps que l'acquisition. Pour atteindre cet objectif et être en mesure de réaliser des échographies 3D en routine clinique, on pourrait envisager d'utiliser la méthode, basée sur le suivi de la sonde par un accéléromètre et un gyromètre, proposé par [Morgan et al. \(2018\)](#). Une étude de faisabilité sera effectuée en partenariat avec l'université de Duke, aux Etats-Unis, dans la suite de ce projet.

CHAPITRE V

CRÉATION DE SEGMENTATION 3D À PARTIR D'ANNOTATIONS MANUELLES 2D

Contents

V.1	Création des segmentations 3D du SVC	118
V.1.1	Segmentations 2D manuelles des séquences dynamiques	118
V.1.2	Recalage des images segmentées manuellement et interpolation des annotations dans les images non annotées	120
	Recalage des images segmentées manuellement	120
	Interpolation des annotations sur les images non annotées	120
	Validation de la méthode d'interpolation	122
V.1.3	Positionnement dans la grille de reconstruction et création d'une segmentation 3D	124
	Validation de l'interpolation des segmentations	125
V.1.4	Positionnement dans le repère commun	126
V.1.5	Validation globale de l'algorithme de création des segmentations 3D	126
V.1.6	Conclusion	129
V.2	Création des segmentations 3D des thalami	130
V.2.1	Segmentations 2D manuelles	130
V.2.2	Interpolation des annotations sur les images non annotées et positionnement dans le repère commun	131
	Interpolation des annotations sur les images non annotées	131
	Positionnement dans le repère commun	132
V.2.3	Validation de la méthode d'interpolation	132
V.2.4	Conclusion	134
V.3	Conclusion	135
	Résumé du travail réalisé	135
	Résumé des contributions	136
	Perspectives	136

CHAPITRE V

C E chapitre présente les algorithmes que nous avons développés pour créer des segmentations 3D à partir de segmentations manuelles 2D réalisées de manière discontinues (toutes les images où les structures d'intérêt étaient visibles n'ont pas été segmentées). Le développement de ces algorithmes était nécessaire de manière à avoir le temps d'aborder la problématique de segmentation pendant cette thèse. En effet, segmenter toutes les images des données échographiques dont nous disposons pour créer une base de données aurait pris trop de temps. Ces annotations manuelles sont indispensables pour aborder les problèmes de segmentation, elles permettent d'évaluer la précision des résultats obtenues avec les algorithmes développés et constituent une base de données pour utiliser des méthodes d'apprentissage profond. Pendant ces 3 ans, il a été possible de développer une base de données annotées pour le SVC et pour les thalami.

Les contributions présentées dans ce chapitre sont :

- Un algorithme permettant de créer des segmentations 3D à partir d'annotations 2D réalisées avec un certain espacement.
- La création de la première base de données, à notre connaissance mondiale, échographiques 3D annotée des thalami (16 volumes) chez l'enfant prématuré.
- La création d'une base de données échographiques 3D annotées du SVC contenant 25 volumes chez l'enfant prématuré.

Ce chapitre est composé de deux parties, la première détaille les étapes qui ont permis de créer les segmentations 3D du SVC et la seconde les étapes qui ont permis de créer les segmentations 3D des thalami.

V.1 Création des segmentations 3D du SVC

V.1.1 Segmentations 2D manuelles des séquences dynamiques

Les segmentations manuelles du SVC ont été réalisées sur les SDs car l'algorithme de création de volumes échographiques 3D n'avait pas encore été développé au moment où cette base a commencé à être créée. Ces segmentations ont été réalisées sur 25 examens, elles ont été dessinées dans des coupes proches du plan coronal. De manière à réduire le temps consacré à ces segmentations, celles-ci ont en moyenne été espacées de 4.1 ± 2.1 images, en fonction des variations plus ou moins rapides de la forme du SVC. Ce qui a permis, en théorie, de diviser le temps consacré à ces segmentations par 5. Si on suppose qu'il faut 5 heures pour annoter chaque coupe d'une séquence dynamique, cela représente un gain de 4 heures par volumes segmentées. Cette méthode est illustrée sur la Figure.V.1.

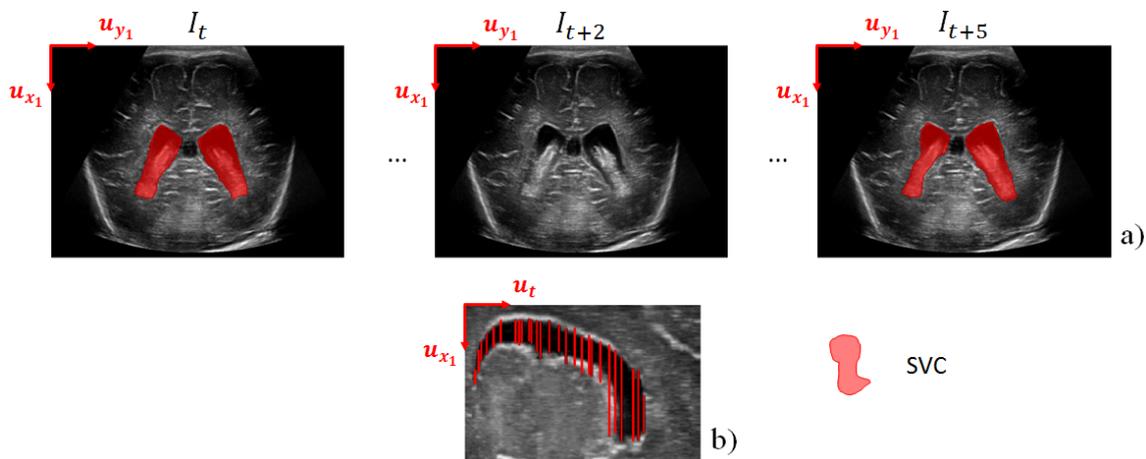


FIGURE V.1 – Segmentations manuelles, espacées de 4 images, du SVC dans une SD (a). Coupe sagittale d'une SD sur laquelle on voit les segmentations réalisées en coupe coronale (b)¹.

Comme le SVC n'a pas été segmenté sur toutes les images des SDs, des segmentations ont été interpolées entre les images qui avaient été annotées. De manière à ce que cette interpolation ait bien lieu sur toutes les images où le SVC était visible, les images contenant le début ou la fin de certaines parties du SVC ont été systématiquement segmentées manuellement (Figure.V.2).

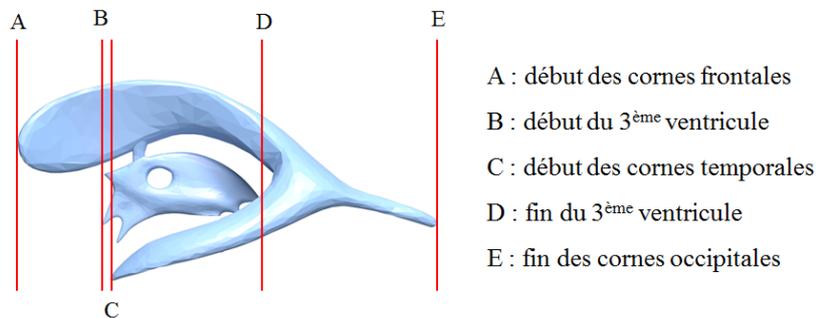


FIGURE V.2 – Parties du SVC qui ont systématiquement été segmentées manuellement dans les SDs².

1. Images issues de notre base de données.
 2. Image originale : <https://commons.wikimedia.org/wiki/>

Ces segmentations manuelles ont été réalisées par le pédiatre du centre de néonatalogie du CH d'Avignon avec le logiciel libre MITK (<http://mitk.org/wiki/MITK>). Cet outil a l'avantage de permettre la segmentation des structures, et leur visualisation, dans n'importe quel plan anatomique de référence. De manière à rendre les segmentations plus précises et leur réalisation plus rapide et moins pénible, celles-ci ont été dessinées avec un stylet (10 points de contact) sur une tablette Microsoft Surface Studio (32 Go de RAM) acquise par le CH d'Avignon pour ce projet. Un exemple de contourage des thalami est visible sur la Figure.V.3.

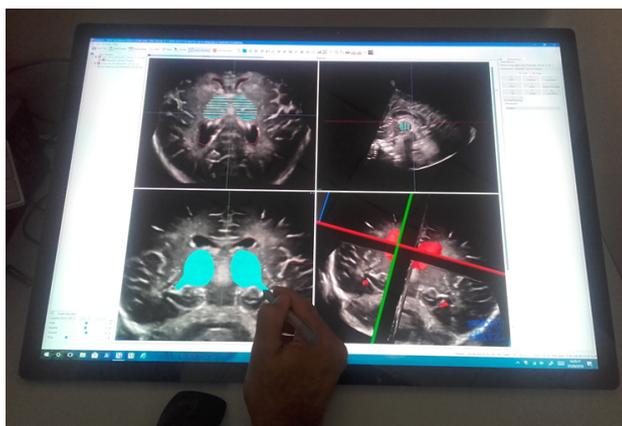


FIGURE V.3 – Segmentation manuelle des thalami sur tablette avec le logiciel MITK³.

A partir de ces annotations 2D réalisées de manière discontinues, les segmentations 3D ont été obtenues avec la chaîne représentée sur la Figure.V.4. Celle-ci est analogue à la chaîne qui a été utilisée pour créer les données échographiques 3D (Figure.IV.1), elle possède néanmoins une étape d'interpolation supplémentaire (3) et la méthode utilisée pour interpoler les segmentations dans la grille reconstruction (4) n'utilise pas les mêmes paramètres. En revanche, les étapes 2 et 5 sont identiques à celle utilisée par l'autre chaîne, elles reprennent donc les paramètres précédemment déterminés.

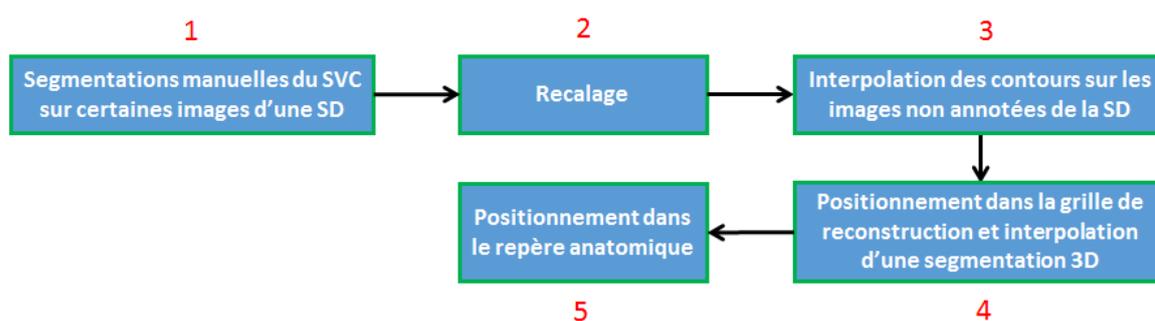


FIGURE V.4 – Étapes ayant permis de créer les segmentations 3D du SVC.

La réalisation des segmentations manuelles du SVC (1) a été décrite dans cette partie. Les prochaines parties sont consacrées, dans l'ordre : à l'étape de recalage et à l'interpolation de ces annotations sur les images non segmentées (2 et 3), à leur positionnement et leur interpolation dans la grille de reconstruction pour obtenir une segmentation 3D (4), au positionnement de la segmentation 3D obtenue dans le repère commun et à l'évaluation de l'erreur commise par cet algorithme.

3. Photo prise au centre de néonatalogie du CH d'Avignon.

V.1.2 Recalage des images segmentées manuellement et interpolation des annotations dans les images non annotées

Recalage des images segmentées manuellement

La deuxième étape de cette chaîne consiste à recalibrer les images segmentées de la même manière que durant l'étape de recalage qui a servi à créer le volume échographique associé à la SD segmentée.

Pour cela, les matrices de transformations A_t , déterminées lors de l'étape de recalage de l'algorithme de création des données échographiques 3D, sont appliquées aux segmentations Y_t (segmentations associées aux images I_t). En compensant les mouvements latéraux qui peuvent survenir durant l'acquisition, cette étape permet de réaligner au mieux les structures cérébrales et donc de faciliter l'interpolation des segmentations du SVC sur les images qui n'ont pas été segmentées.

Interpolation des annotations sur les images non annotées

La troisième étape de cette chaîne a pour objectif d'interpoler des segmentations du SVC sur les images qui n'ont pas été segmentées manuellement.

Dans un premier temps, une carte de distance (CD_t) aux bords du SVC est calculée pour chaque image I_t sur laquelle une segmentation manuelle Y_t a été réalisée. La valeur de chaque pixel de coordonnées $(x_1, y_1) \in CD_t$ est donnée par l'Équation.V.1.

$$CD_t(x_1, y_1) = \begin{cases} \min_{(x_1, y_1) \in I_t} d((x_1, y_1), \partial Y_t) & \text{si } (x_1, y_1) \in Y_t \\ - \min_{(x_1, y_1) \in I_t} d((x_1, y_1), \partial Y_t) & \text{si } (x_1, y_1) \notin Y_t \end{cases} \quad (V.1)$$

Où d désigne la distance euclidienne. La Figure.V.5 montre une segmentation Y_t du SVC et l'image CD_t qui lui est associée.

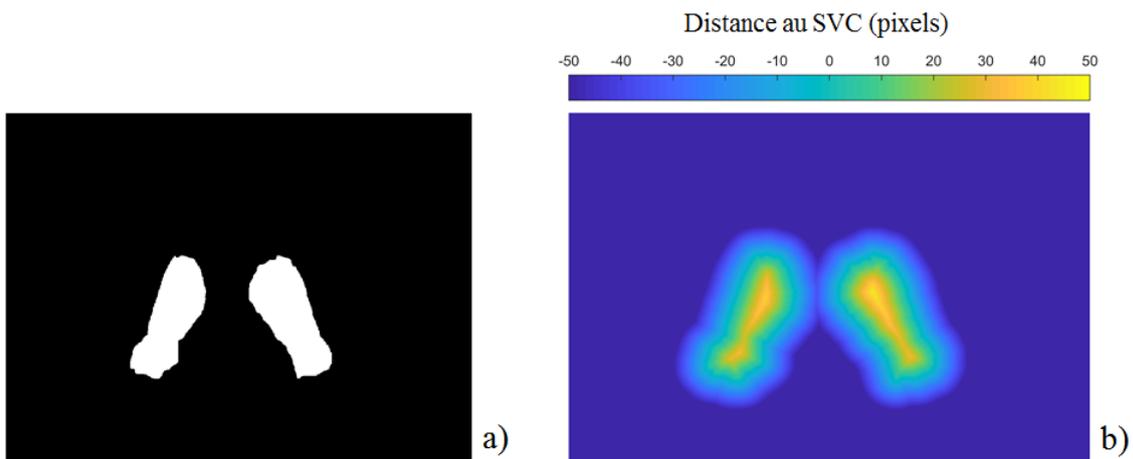


FIGURE V.5 – Image binaire correspondant à une segmentation du SVC (a) et carte de la distance aux bords du SVC (b).

Dans un second temps, une interpolation linéaire est réalisée entre les cartes de distances, de manière à obtenir une carte de distance sur chacune des images se trouvant entre deux segmen-

tations. Ainsi, si on note $\{t_1, t_2, \dots, t_{N_{SVC}}\}$ les indices des N_{SVC} images sur lesquelles des segmentations manuelles ont été réalisées, l'étape d'interpolation consiste à réaliser les opérations décrites par l'Équation.V.2.

$$\forall i \in [1, N_{SVC} - 1], \forall t \in [t_i, t_{i+1}] : CD_t(x_1, y_1) = \frac{t_{i+1} - t}{t_{i+1} - t_i} CD_{t_i}(x_1, y_1) + \frac{t - t_i}{t_{i+1} - t_i} CD_{t_{i+1}}(x_1, y_1) \quad (V.2)$$

Après avoir calculé les cartes de distance aux bords du SVC sur toutes les images où celui-ci est visible, il est possible d'obtenir sa segmentation en considérant les pixels dont la valeur est supérieure à un seuil $h = 0$. Néanmoins, les cartes de distance ont été conservées à cette étape car elles seront utilisées lors de la prochaine étape de l'algorithme. Un résumé de la méthode, avec obtention d'une segmentation finale, est proposé sur la Figure.V.6.

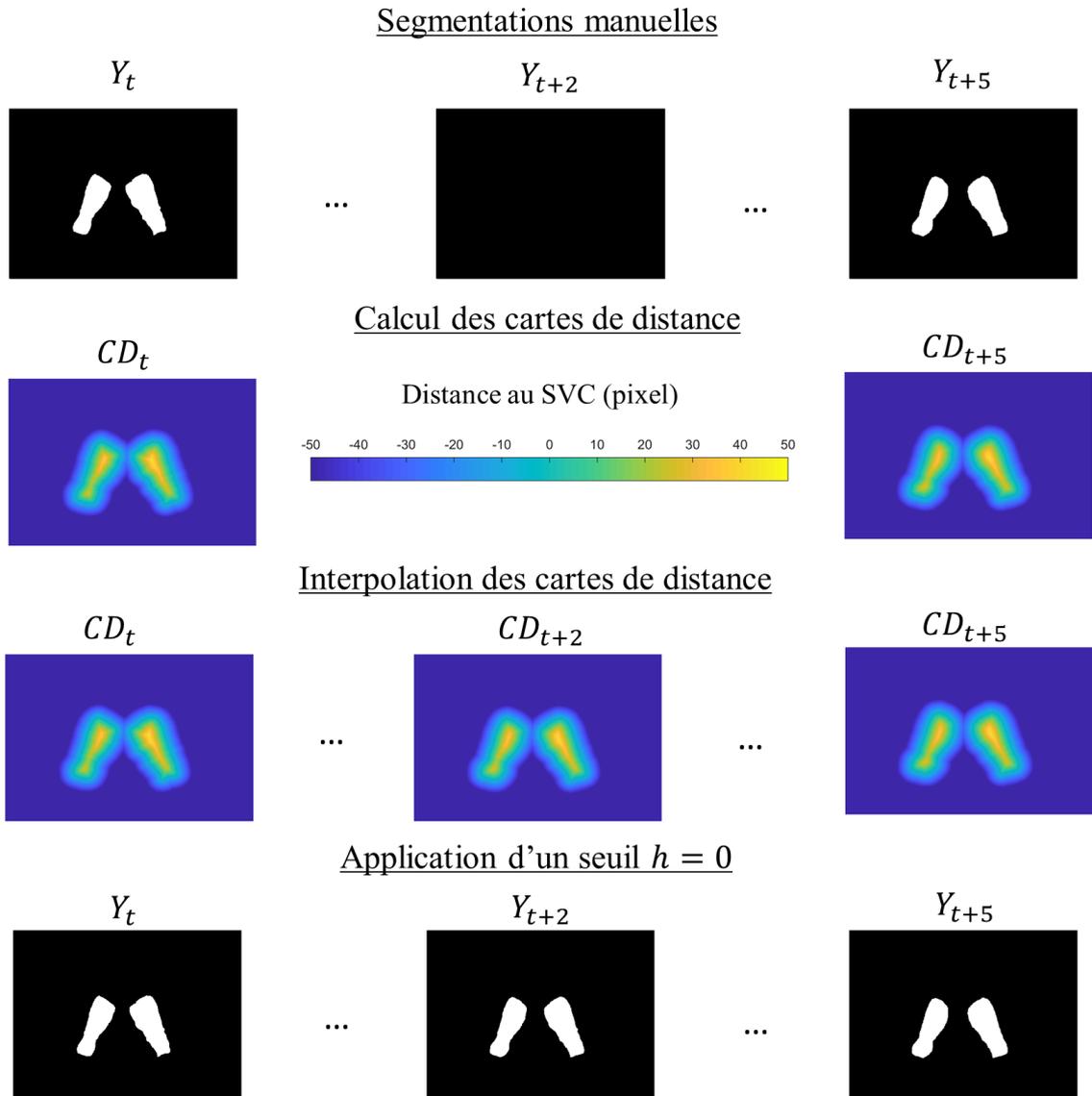


FIGURE V.6 – Segmentations manuelles (a), cartes de la distance aux bords du SVC (b), cartes de distance interpolées (c) et segmentations interpolées obtenues après application d'un seuil $h = 0$.

La Figure.V.7 montre une coupe sagittale dans laquelle on observe les segmentations manuelles du SVC réalisées en coupe coronale et la segmentation obtenue avec interpolation et seuillage des cartes de distance.

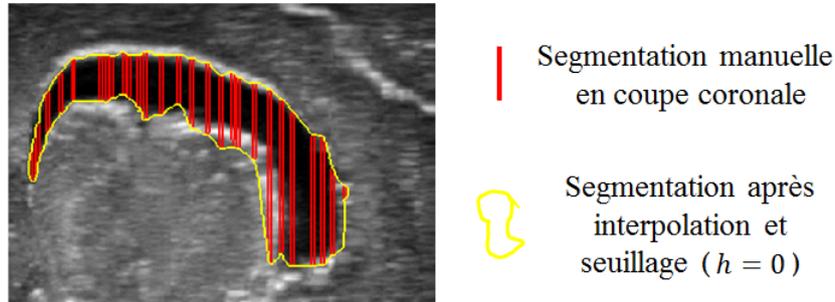


FIGURE V.7 – Segmentations manuelles du SVC et segmentation obtenue après interpolation et seuillage.

On peut voir que la segmentation obtenue semble suivre correctement le contour du SVC : région hypoéchogène (LCS) et hyperéchogène (plexus choroïde).

Validation de la méthode d'interpolation

Segmenter manuellement le SVC de manière discontinue et interpoler les segmentations restantes permet de réduire le temps consacré à ces segmentations mais diminue également leur précision par rapport à des segmentations réalisées sur toutes les images.

L'erreur commise lors de cette étape a été estimée en comparant des segmentations de référence, équivalentes à des segmentations manuelles réalisées sur toutes les coupes de SDs (comme si on se situait après l'étape 3), à des segmentations interpolées avec la méthode présentée dans cette partie. Pour obtenir ces segmentations de référence, nous sommes parti d'un modèle CAO 3D du SVC orienté dans le repère anatomique. Dans un premier temps, le modèle CAO a été mis au même dimensions que chacun des sept SVCs qui constituaient notre base de données de test qui a été utilisée pour évaluer les performances des algorithmes de segmentation. Ensuite, il a été modifié en utilisant les transformations inverses de celles correspondant aux étapes de positionnement dans le repère commun et de positionnement dans la grille de reconstruction, utilisées lors de la création des volumes échographiques. Cette méthode est illustrée sur la Figure.V.8, est nous a permis d'obtenir sept segmentations de référence.

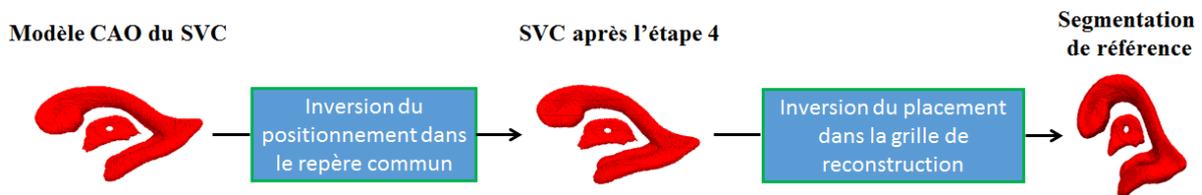


FIGURE V.8 – Obtention d'une segmentation de référence à partir d'un modèle CAO du SVC.

L'obtention d'une segmentation interpolée à partir d'une segmentation de référence a été réalisée en deux temps. Dans un premier temps, des annotations ont été supprimées sur certaines images d'une segmentation de référence de manière à obtenir une segmentation discontinue. Dans un second temps, des annotations ont été interpolées sur les images où elles avaient été supprimées de manière à obtenir une segmentation interpolée. Cette dernière a ensuite été comparée à la segmentation de référence pour estimer l'erreur commise lors de l'étape d'interpolation. Cette méthode est représentée sur la Figure.V.9.

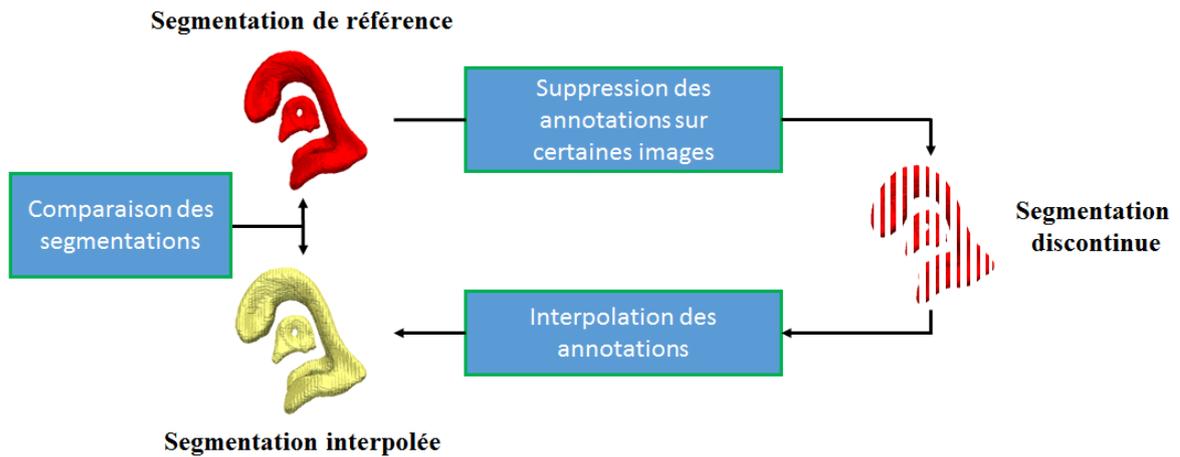


FIGURE V.9 – Méthode utilisée pour évaluer la précision de l'interpolation des contours du SVC dans les SDs.

Cette expérience a été réalisée pour chacune des sept segmentations de référence en supprimant 0,1,2,3,4,5,10,15 ou 20 images entre les images annotées. Le Dice moyen et la MAD moyenne obtenus, en fonction du nombre d'images entre deux images annotées, sont visibles sur la Figure.V.10.

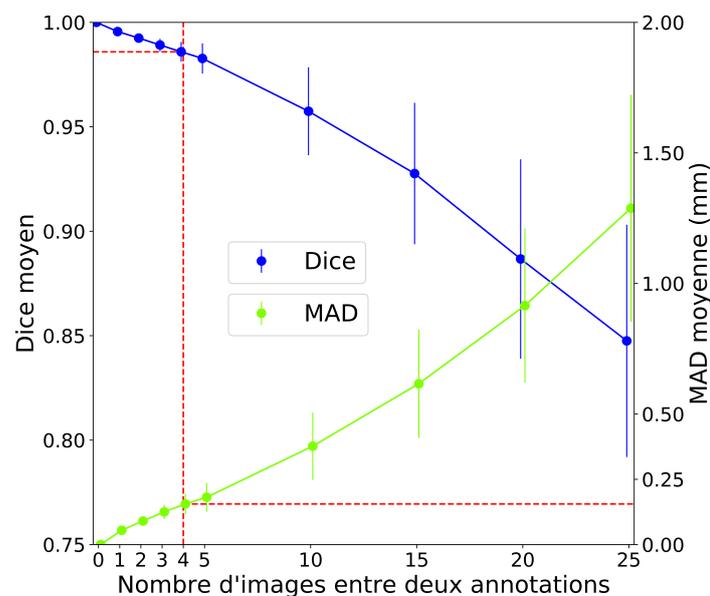


FIGURE V.10 – Précision de l'interpolation du SVC en fonction du nombre d'images laissées sans annotations entre deux segmentations.

On voit que le Dice moyen et la MAD moyenne diminue et augmente respectivement en fonction

du nombre d'images sans annotation entre deux images annotées. Ce comportement est cohérent car moins il y a d'annotations manuelles dans une SD moins la segmentation interpolée sera précise.

On observe que le Dice est raisonnable (> 0.95) jusqu'à un espacement d'environ 10 images entre les annotations et que la MAD est raisonnable (< 0.25 mm) jusqu'à un espacement d'environ 20 images entre les annotations.

Dans notre cas (représenté par la ligne pointillée rouge sur la Figure.V.10), les annotations étaient en moyenne espacées de 4 images, ce qui correspond à un Dice de 0.986 ± 0.005 et une MAD de 0.15 ± 0.04 mm. D'après ces résultats, dans les conditions où les annotations manuelles ont été réalisées, la segmentation interpolée obtenue est extrêmement réaliste par rapport à une segmentation réalisée sur toutes les images d'une SD.

Sur la Figure.V.11.b, des segmentations interpolées sont présentées pour plusieurs espacements (noté e) entre les images annotées.

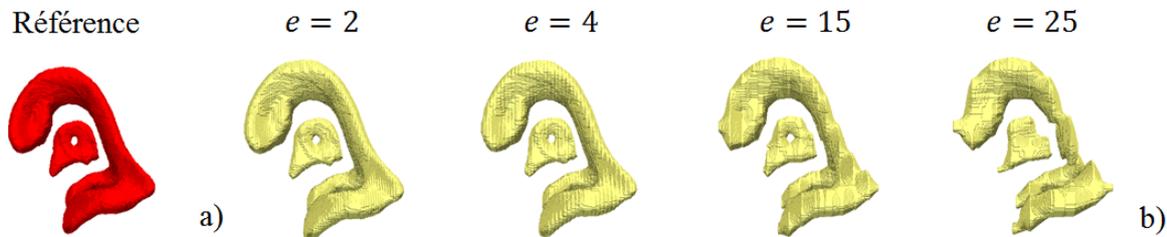


FIGURE V.11 – Segmentation de référence (a) et segmentations interpolées à partir d'annotations différemment espacées d'un nombre d'image e (b).

On observe bien que plus l'espacement est important entre les annotations, plus la forme de la segmentation interpolée s'éloigne de la segmentation de référence (Figure.V.11.a).

V.1.3 Positionnement dans la grille de reconstruction et création d'une segmentation 3D

La quatrième étape a pour objectif de positionner les cartes de distance dans la grille de reconstruction et d'interpoler une segmentation 3D du SVC à partir de celles-ci.

Pour chaque examen segmenté manuellement, les cartes de distances obtenues à l'étape précédente ont été placées dans la grille de reconstruction avec le modèle de rotation d'axe mobile utilisé pour créer le volume échographique associé à l'examen. La distance aux bords du SVC a ensuite été interpolée en chaque voxel ν de la grille en utilisant la valeur du pixel p le plus proche sur chaque image adjacente que l'on notera ${}_1NNCD_t(\nu)$ et ${}_1NNCD_{t+1}(\nu)$. Dans ce cas, seul le plus proche voisin a été utilisé pour interpoler la grille de reconstruction car les valeurs recherchées sont des distances, utiliser un plus grand nombre de voisins n'apporterait pas beaucoup plus d'information. L'interpolation d'une distance $G(\nu)$ en chaque voxel de la grille de reconstruction G a été réalisée en utilisant l'Équation.V.3.

$$G(v) = \frac{\sum_{p \in \cup_1 NNCD_t(v) \cup_1 NNCD_{t+1}(v)} \frac{1}{d(p,v)} SCD(p)}{\sum_{p \in \cup_1 NNCD_t(v) \cup_1 NNCD_{t+1}(v)} \frac{1}{d(p,v)}} \quad (V.3)$$

où $SCD(p)$ est la distance du pixel p , de la séquence de carte de distance SCD , aux bords du SVC, calculée sur les images 2D. $d(p, v)$ correspond à la distance euclidienne entre p et v . Après cette interpolation, un seuil $h > 0$ a été appliqué à la grille de reconstruction pour obtenir une segmentation 3D du SVC. La Figure.V.12 montre, en coupe sagittale, une carte de distance et la segmentation obtenue après application du seuil.

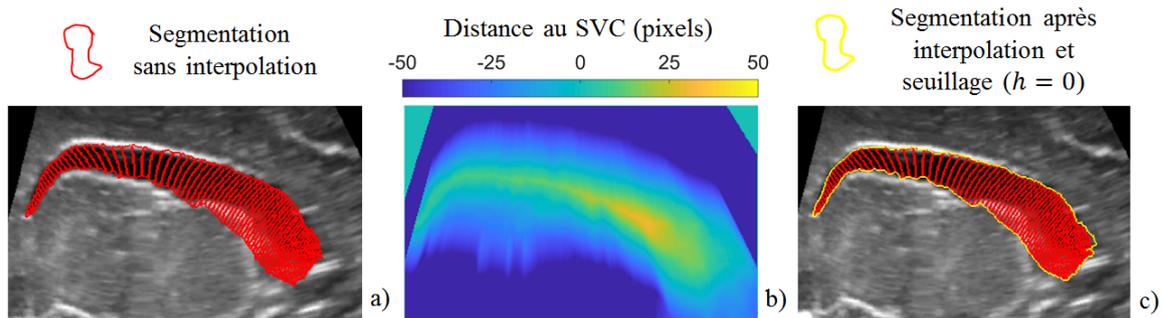


FIGURE V.12 – Segmentation du SVC après application d'un seuil $h > 0$ sans interpoler les cartes de distance dans la grille de reconstruction (a), carte de distance obtenues après interpolation (b) et segmentation obtenue après interpolation et application d'un seuil $h > 0$ (c).

On voit sur cette image que la segmentation obtenue semble suivre le bord du SVC de manière cohérente.

Validation de l'interpolation des segmentations

L'erreur commise lors de cette étape a été estimée de manière analogue à ce qui a été fait pour valider l'étape d'interpolation. Les segmentations de référence ont été obtenues à partir du modèle CAO du SVC, en lui appliquant les transformations inverses de celles utilisées pour positionner les sept volumes tests dans le repère commun. Cette étape est représentée sur la Figure.V.13.

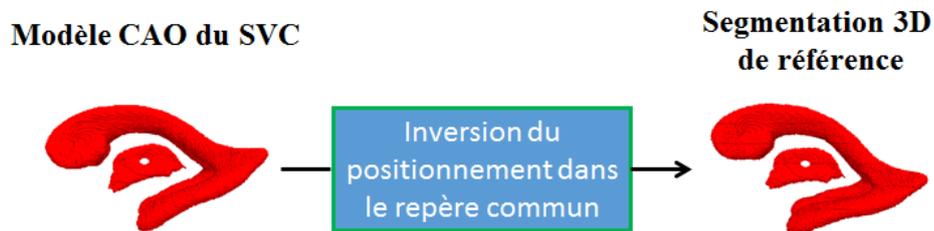


FIGURE V.13 – Obtention d'une segmentation 3D de référence.

Les sept segmentations 3D interpolées ont été obtenues en deux temps. Tout d'abord, les transformations inverses de celles utilisées pour placer les images dans la grille de reconstruction ont été appliquées aux segmentations 3D de référence, puis, les étapes décrites dans cette partie ont été appliquées à la segmentation obtenue. L'erreur a ensuite été estimée en comparant les segmentations 3D de référence aux segmentations 3D interpolées. Cette méthode est illustrée sur la Figure.V.14.

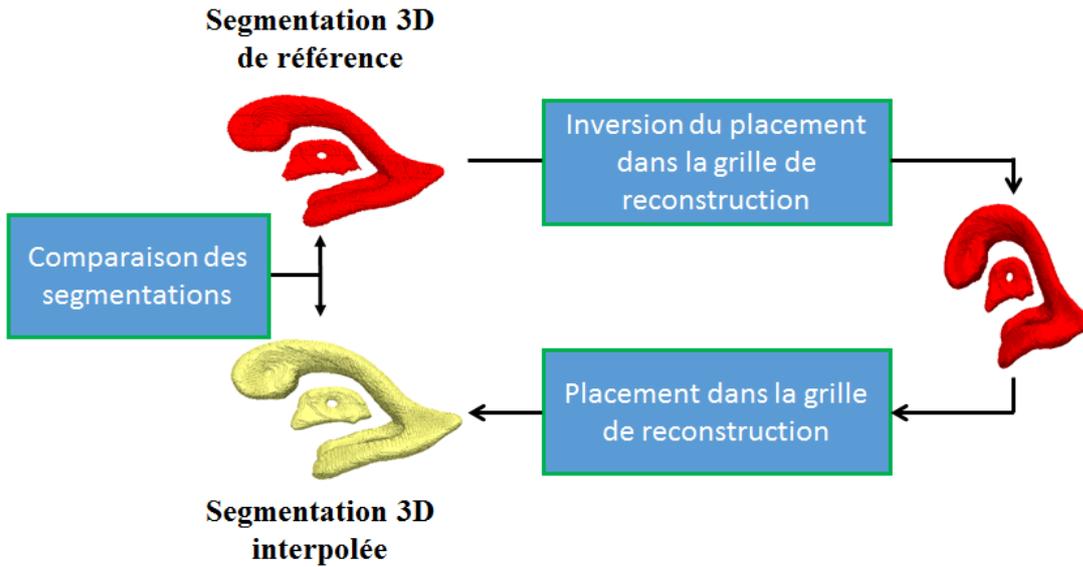


FIGURE V.14 – Méthode utilisée pour estimer l’erreur commise lors de la création d’une segmentation 3D.

Après avoir réalisée cette expérience sur les sept volumes tests, le Dice moyen obtenu était de 0.977 ± 0.002 et la MAD moyenne de 0.12 ± 0.01 mm. Cette étape permet donc de créer des segmentations 3D interpolées avec précision.

V.1.4 Positionnement dans le repère commun

La dernière étape a pour objectif de placer les segmentations 3D dans le repère commun.

Ce positionnement a été effectué en utilisant l’origine O_{CC} (Figure.IV.41) et les angles de rotations ϕ et ψ (Figure.IV.5) déterminés lors de la création des volumes échographiques.

V.1.5 Validation globale de l’algorithme de création des segmentations 3D

Cette partie est dédiée à l’estimation de l’erreur commise en enchainant les deux étapes d’interpolation décrites précédemment.

Jusqu’à présent, les erreurs commises lors de l’étape d’interpolation des annotations sur toutes les images d’une SD et lors de l’étape d’interpolation d’une segmentation 3D dans la grille de reconstruction avaient été estimées séparément. Nous allons maintenant estimer l’erreur commise en enchainant ces deux étapes. La méthode utilisée à cette fin est analogue à celles utilisées lors des validations précédentes. Les sept segmentations 3D de référence ont été obtenues avec la méthode représentée sur la Figure.V.13.

L’obtention des segmentations interpolées a été réalisée en 4 étapes. Premièrement, les transformations inverses de celles correspondant à l’étape de positionnement dans la grille de reconstruction ont été appliquées aux segmentations 3D de référence pour obtenir des segmentations complètes de SDs. Deuxièmement, des annotations ont été supprimées dans ces séquences d’images segmentées. Ensuite, les deux étapes d’interpolation décrite précédemment ont été mise en œuvre

pour obtenir des segmentations 3D interpolées. Ces dernières ont finalement été comparées aux segmentations 3D de référence de manière à estimer la précision des segmentations créées. Cette méthode est représentée sur la Figure.V.15.

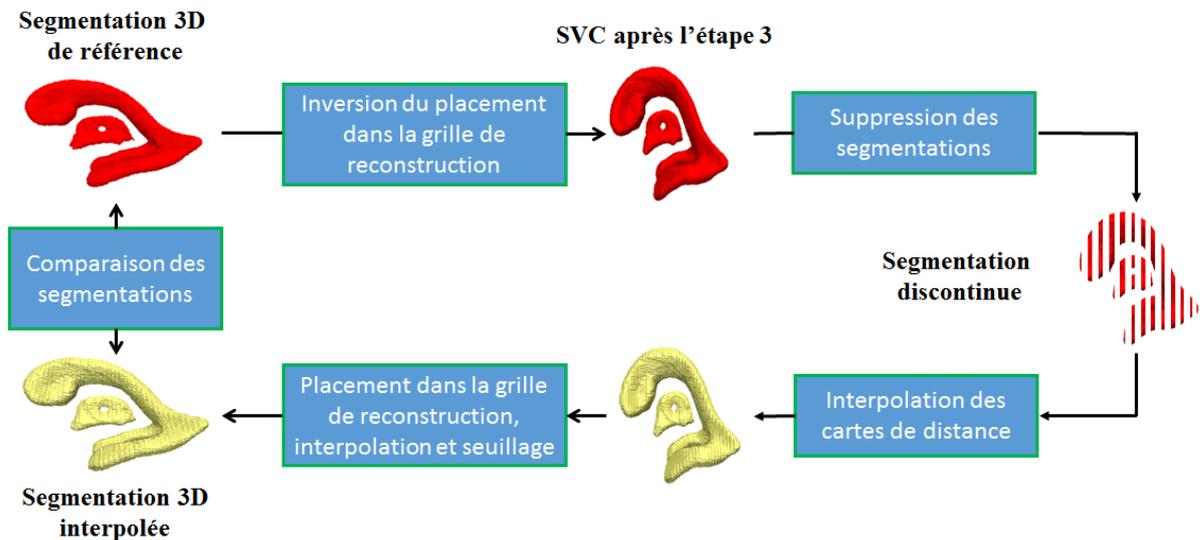


FIGURE V.15 – Méthode utilisée pour estimer l'erreur commise lors de la création d'une segmentation 3D.

Cette expérience a été réalisée pour chacune des sept segmentations de référence en supprimant 0, 1, 2, 3, 4, 5, 10, 15, 20 ou 25 images entre les images annotées. Le Dice moyen et la MAD moyenne obtenus, en fonction du nombre d'images entre deux images annotées, sont visibles sur la Figure.V.16.

On voit que le Dice moyen et la MAD moyenne diminue et augmente respectivement lorsque l'espacement entre les annotations augmente, ce qui est cohérent.

Le Dice moyen et la MAD moyenne obtenus sont raisonnables (respectivement > 0.95 et < 0.25 mm) jusqu'à un espacement, entre les images annotées, situé entre 5 et 10 images.

Dans notre cas, qui est représenté par une ligne pointillée rouge sur la Figure.V.16, un Dice et une MAD moyenne théoriques de 0.969 ± 0.006 et de 0.16 ± 0.04 mm ont été obtenus. En espaçant les annotations de 4 images, on obtient donc des segmentations 3D interpolées de très bonne qualité.

On voit qu'une bonne précision aurait été conservée en espaçant les segmentations de 5 coupes mais que ça n'aurait pas été le cas en les espaçant de 10 coupes. Un espacement de 4 images entre les segmentations semblent donc correct de manière à avoir une marge par rapport à ces valeurs limites. Des images de segmentations 3D interpolées sont visibles sur la Figure.V.17.b pour différents espacements (e) entre les images annotées sur les SDs.

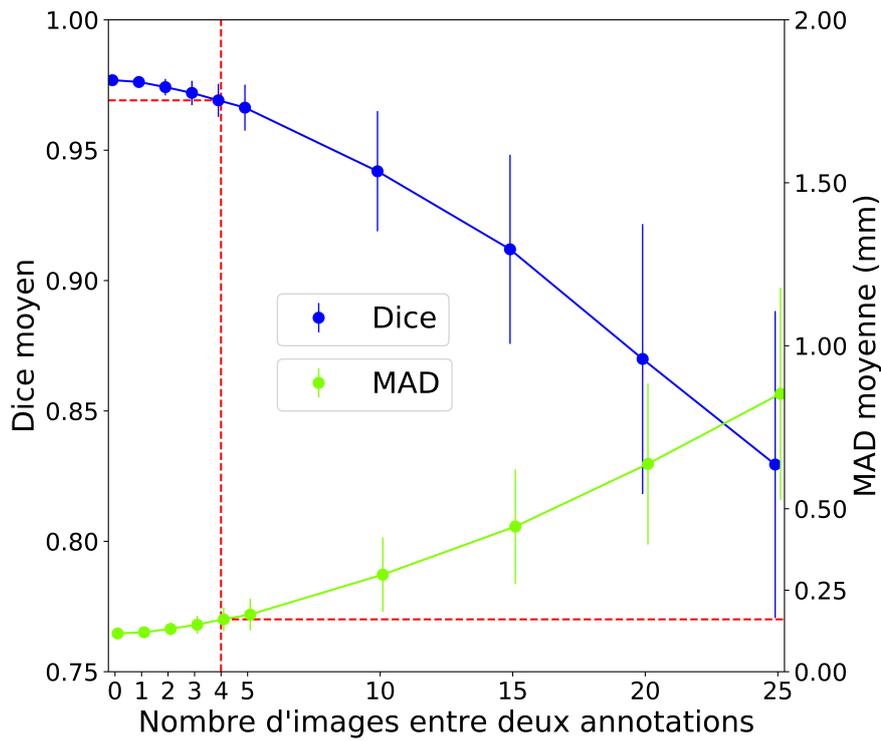


FIGURE V.16 – Précision de la création d'une segmentation 3D en fonction du nombre d'images laissées sans annotations entre deux segmentations.

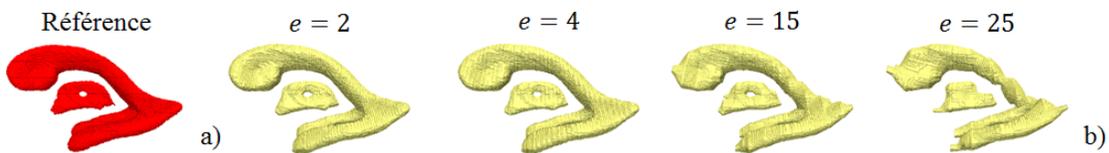


FIGURE V.17 – Précision de la création des segmentations 3D du SVC en fonction du nombre d'images non annotées entre deux images annotées.

On voit que plus l'écart en les images annotées augmente, plus la segmentation 3D obtenue après interpolation se dégrade, notamment en régularité.

Le temps moyen, calculé sur les 25 volumes constituant la base de données, nécessaire pour obtenir une segmentation 3D à partir des annotations 2D était de $1 \text{ min } 31 \pm 28 \text{ s}$. Ce temps est extrêmement rapide en comparaison au temps qui aurait été nécessaire pour annoter toutes les images. En espaçant les annotations de 4 coupes, si on considère qu'il faut 5 heures pour segmenter rigoureusement le SVC dans toutes les images, l'algorithme développé permet de réduire le temps nécessaire pour obtenir une segmentation dans toutes les coupes d'environ 4 heures par volume. Sur l'ensemble de la base de données, cet algorithme a permis de réduire d'environ 100 h le temps à consacrer aux segmentations manuelles, ce qui aurait sinon représenté des mois de travail pour un clinicien.

V.1.6 Conclusion

Cette partie a présenté l'algorithme qui a permis d'obtenir des segmentations 3D du SVC dans nos volumes échographiques à partir de segmentations 2D manuelles réalisées de manière discontinue sur des images d'ETFs 2D. Les annotations réalisées ont en moyenne été espacées de 4 images. L'algorithme développé comprend deux étapes majeures permettant d'obtenir des segmentations 3D : une étape d'interpolation des annotations sur toutes les images des SDs et une étape d'interpolation d'une segmentation 3D dans la grille de reconstruction. La précision des segmentations interpolées a été estimée, lors de chaque étape et en les enchainant, à partir d'un modèle CAO du SVC. Pour un écart moyen de 4 images entre les annotations, un Dice moyen et une MAD moyenne de 0.969 ± 0.006 et de 0.16 ± 0.04 mm ont été obtenus.

Cet algorithme nous a permis d'obtenir des segmentations 3D de référence du SVC dans 25 examens. De plus, en permettant de créer des segmentations 3D en un peu plus d'une minute, à partir d'annotations espacées de 4 coupes, cet algorithme a permis de diviser le temps nécessaire à la création de cette base de données par 5. D'après la validation effectuée à partir du modèle CAO du SVC, les segmentations 3D ont été interpolées avec une bonne précision (Dice > 0.95 et MAD < 0.25 mm). Cette étude était extrêmement importante car les valeurs théorique obtenues pour le Dice et la MAD constituent une valeur limite qu'on ne pourra pas dépasser lorsque l'on évaluera les performances d'un algorithme de segmentation. En effet, si l'algorithme réalise une segmentation parfaite, les segmentations de référence ayant été obtenues par interpolation, on obtiendra pas un Dice de 1 et une MAD de 0 mm.

De manière à améliorer la précision des segmentations 3D interpolées, deux possibilités peuvent être envisagées : réduire l'espacement entre les annotations réalisées sur les SDs ou améliorer la méthode utilisée pour interpoler les segmentations 3D. La première option n'est pas souhaitable car elle augmenterait le temps à consacrer aux segmentations manuelles. En ce qui concerne la deuxième option, on peut reprocher à la distance interpolée aux bords du SVC de correspondre à la distance calculée sur les images 2D, alors que les segmentations que l'on souhaite obtenir sont 3D. Il serait intéressant de développer une nouvelle méthode pour interpoler la distance aux bords du SVC, ce qui pourrait permettre de gagner en précision.

V.2 Création des segmentations 3D des thalami

V.2.1 Segmentations 2D manuelles

Les segmentations manuelles des thalami ont été réalisées dans les données échographiques 3D débruitées. Cette structure était plus facile à identifier dans ces données car elles étaient orientées dans les plans anatomiques de référence et que le débruitage augmentait le contraste entre les thalami et la SB. Ces segmentations ont été réalisées dans 16 volumes, qui ont constitué notre base de données annotées pour travailler sur le problème de segmentation des thalami. Les annotations des images ont été réalisées en coupe coronale ou en coupe axiale selon celles où les thalami étaient les plus visibles. En coupe coronale les annotations ont en moyenne été espacées de 5.4 ± 1.8 coupes et en coupes axiales de 3.2 ± 0.4 coupes. Cela a permis de diviser environ par 6 et 4 le temps consacré à la segmentation des thalami en coupe coronale et axiale.

Comme pour le SVC, ces segmentations ont été interpolées de manière à obtenir des annotations sur toutes les images où les thalami étaient visibles. Pour que l'interpolation ait bien lieu sur toutes les coupes contenant les thalami, le début et la fin de cette structure ont systématiquement été annotés en coupe coronale ou en coupe axiale, en fonction des cas. Cette méthode est illustrée sur la Figure.V.18.

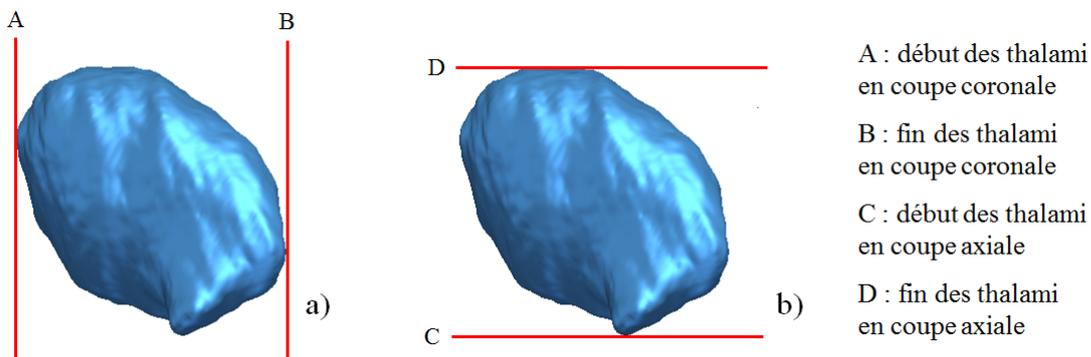


FIGURE V.18 – Parties des thalami qui ont systématiquement été segmentées manuellement en coupe coronale (a) ou axiale (b).

Ces annotations ont également été réalisées sur tablette. Les segmentations 3D ont été obtenues à partir de la chaîne représentée sur la Figure.V.19.



FIGURE V.19 – Étapes ayant permis de créer les segmentations 3D des thalami.

Cette première partie a décrit la méthode utilisée pour segmenter manuellement les thalami (1). La prochaine partie sera consacrée à l'interpolation de ces annotations sur les images non segmentées (2) et au positionnement de la segmentation 3D obtenue dans le repère commun (3). La partie qui

la suite sera dédiée à la validation de la méthode d'interpolation des segmentations des thalami.

V.2.2 Interpolation des annotations sur les images non annotées et positionnement dans le repère commun

Interpolation des annotations sur les images non annotées

Cette étape a pour objectif d'interpoler des segmentations des thalami dans les images où ils n'ont pas été segmentés manuellement.

Dans un premier temps, des cartes de distance aux bords des thalami ont été calculées sur les images où des segmentations avaient été réalisées. Ces cartes ont été obtenues avec l'Équation.V.4 ou l'Équation.V.5 en fonction de si les segmentations manuelles avaient été réalisées en coupe coronale ou axiale.

$$G_{z_1}(x_1, y_1) = \begin{cases} \min_{(x_1, y_1) \in G_{z_1}} d((x_1, y_1), \partial Y_{z_1}) & \text{si } (x_1, y_1) \in Y_{z_1} \\ - \min_{(x_1, y_1) \in G_{z_1}} d((x_1, y_1), \partial Y_{z_1}) & \text{si } (x_1, y_1) \notin Y_{z_1} \end{cases} \quad (\text{V.4})$$

$$G_{x_1}(y_1, z_1) = \begin{cases} \min_{(y_1, z_1) \in G_{x_1}} d((y_1, z_1), \partial Y_{x_1}) & \text{si } (y_1, z_1) \in Y_{x_1} \\ - \min_{(y_1, z_1) \in G_{x_1}} d((y_1, z_1), \partial Y_{x_1}) & \text{si } (y_1, z_1) \notin Y_{x_1} \end{cases} \quad (\text{V.5})$$

Des cartes de distance ont ensuite été interpolées dans les images qui n'avaient pas été annotées. Si on note $\{z_1, z_2, \dots, z_{N_{TL}}\}$ (respectivement $\{x_1, x_2, \dots, x_{N_{TL}}\}$) les indices des N_{TL} images coronales (respectivement axiales) sur lesquelles des segmentations manuelles ont été réalisées, l'étape d'interpolation consiste à effectuer les opérations décrites par l'Équation.V.6 (respectivement l'Équation.V.7). Ces équations sont analogues à l'Équation.V.2, ce qui les différencie sont les données dans lesquelles l'interpolation a lieu et la direction d'interpolation dans le cas de l'Équation.V.5.

$$\forall i \in [1, N_{TL} - 1], \forall z \in [z_i, z_{i+1}] : G_z(x_1, y_1) = \frac{z_{i+1} - z}{z_{i+1} - z_i} G_{z_i}(x_1, y_1) + \frac{z - z_i}{z_{i+1} - z_i} G_{z_{i+1}}(x_1, y_1) \quad (\text{V.6})$$

$$\forall i \in [1, N_{TL} - 1], \forall x \in [x_i, x_{i+1}] : G_x(y_1, z_1) = \frac{x_{i+1} - x}{x_{i+1} - x_i} G_{x_i}(y_1, z_1) + \frac{x - x_i}{x_{i+1} - x_i} G_{x_{i+1}}(y_1, z_1) \quad (\text{V.7})$$

Où G_z (respectivement G_x) désigne la coupe de profondeur z (respectivement x) du volume G dans lequel les segmentations manuelles ont été dessinées. Une fois des cartes de distance aux bords des thalami interpolées sur toutes les images où la structure étaient visibles, un seuil $h > 0$ a été appliqué pour obtenir une segmentation 3D interpolée dans la grille de reconstruction. La Figure.V.20.a montre une coupe sagittale dans laquelle on observe les segmentations manuelles réalisées en coupe coronale et la segmentation obtenue après interpolation et seuillage des cartes de distances. La Figure.V.20.b montre un résultat qualitatif analogue pour les segmentations manuelles réalisées en coupe axiale.

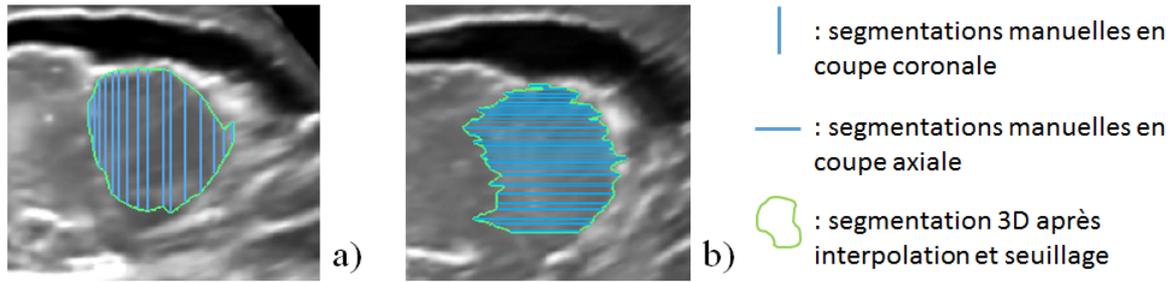


FIGURE V.20 – Segmentations manuelles des thalami en coupe coronale (a) et axiale (b) et segmentations obtenues après interpolation.

Positionnement dans le repère commun

Une fois les segmentations 3D des thalami obtenues dans la grille de reconstruction, celles-ci ont été positionnées dans le repère commun avec la même méthode que celle utilisée pour réaliser cette étape sur le SVC.

V.2.3 Validation de la méthode d'interpolation

L'erreur commise par cet algorithme de création de segmentation 3D a été estimée de manière analogue à ce qui a été effectué pour le SVC. Pour créer des segmentations de référence, un modèle CAO 3D des thalami, orienté dans le repère anatomique, a été utilisé. Celui-ci a été transformé par les transformations inverses de celles utilisées pour positionner les volumes échographiques dans le repère commun. Au total, cinq segmentations de référence ont été obtenues, les transformations utilisées pour les obtenir correspondent à celles utilisées pour créer les cinq volumes de test des algorithmes de segmentation des thalami. Parmi ceux-ci, deux volumes ont été segmentés en coupe axiale et trois volumes ont été segmentés en coupe coronale. Cette méthode est illustrée sur la Figure.V.21.



FIGURE V.21 – Étapes ayant permis de créer les segmentations 3D de référence des thalami.

La méthode utilisée pour obtenir des segmentations interpolées et estimer la qualité de celles-ci est la même que celle utilisée pour le SVC. Elle est illustrée, pour les thalami, sur la Figure.V.22

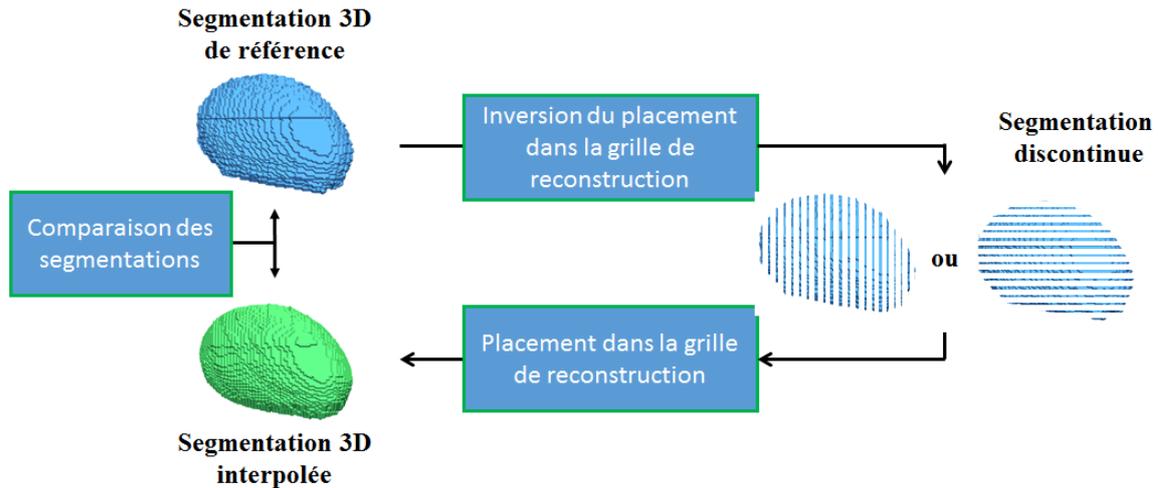


FIGURE V.22 – Méthode utilisée pour estimer l'erreur commise lors de l'interpolation d'une segmentation 3D des thalami.

Cette expérience a été réalisée pour chacune des cinq segmentations 3D de référence, en supprimant 0, 1, 2, 3, 4, 5, 10, 15, 20 ou 25 images. Le Dice moyen et la MAD moyenne obtenus, en fonction du nombre d'images non annotées entre chaque segmentation, sont donnés, pour les segmentations réalisées en coupe coronale, sur la Figure.V.23.a et sur la Figure.V.23.b pour les segmentations réalisées en coupe axiale.

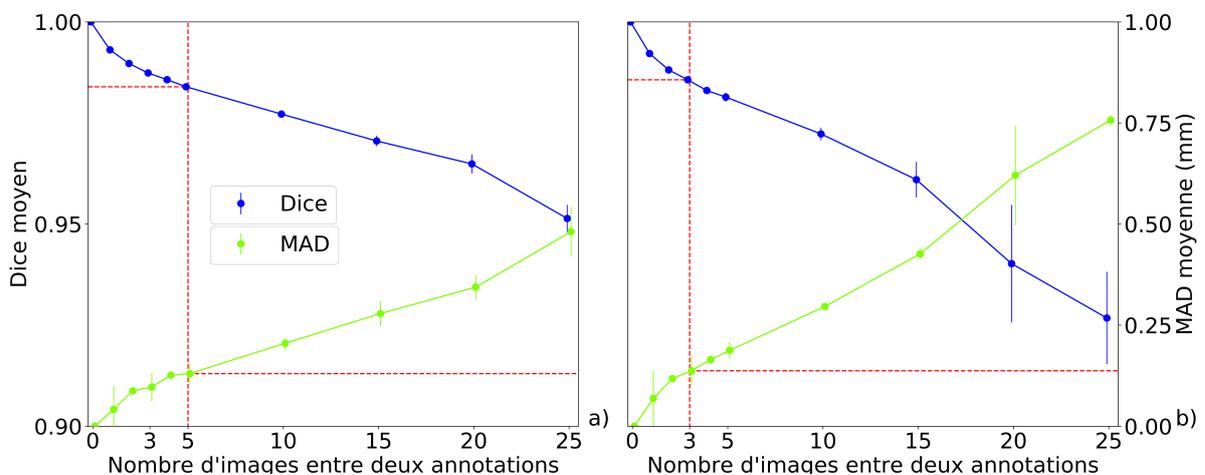


FIGURE V.23 – Précision de la création d'une segmentation 3D des thalami en fonction du nombre d'images laissées sans annotations entre deux segmentations en coupe coronale (a) et axiale (b).

On observe que le Dice moyen et la MAD moyenne diminue et augmente respectivement en fonction du nombre d'images non annotées entre deux annotations, qu'elles aient été réalisées en coupe coronale ou axiale, ce qui est cohérent. On peut aussi remarquer que la précision diminue plus fortement quand les segmentations ont été réalisées en coupe axiale.

Le Dice reste raisonnable (> 0.95) jusqu'à un espacement de plus de 25 images entre les annotations lorsque celles-ci ont été réalisées en coupe coronale. En coupe axiale, il reste raisonnable jusqu'à un espacement compris entre 15 et 20 images entre les annotations.

La MAD reste raisonnable (< 0.25 mm) jusqu'à un espacement compris entre 10 et 15 images

lorsque que les segmentations ont été réalisées en coupe coronale et jusqu'à un espacement compris entre 5 et 10 images dans le cas où les annotations ont été réalisées en coupe axiale.

Si on considère l'espacement moyen utilisé pour réaliser les segmentations manuelles des thalami (environ 5 images en coupe coronale et 3 images en coupe axiale), on peut dire que les segmentations 3D interpolées qui constituent la base de données des thalami sont de très bonne qualité : $Dice = 0.984 \pm 0.0$ et $MAD = 0.13 \pm 0.028$ en coupe coronale et $Dice = 0.982 \pm 0.001$ et $MAD = 0.159 \pm 0.022$ en coupe axiale. De plus d'après les résultats de cette expérience, il semble préférable d'effectuer les segmentations des thalami en coupe coronale et il serait possible d'augmenter l'espacement entre les annotations tout en conservant une précision raisonnable. Il faut néanmoins relativiser ce dernier point compte tenu du faible nombre de volumes, segmentés en coupe coronale ou axiale, sur lesquelles la précision de cet algorithme a été évaluée.

Le temps moyen, calculé sur les 16 volumes constituant la base de données, nécessaire pour interpoler une segmentation 3D a été estimé à 41 ± 5 s. Cela est extrêmement rapide par rapport au temps qui aurait été nécessaire pour segmenter manuellement les images où une segmentation a été interpolée. Comme pour le SVC, si on considère qu'il faut 5 heures pour annoter rigoureusement l'intégralité des thalami dans un volume, cet algorithme a permis d'éviter plusieurs heures de segmentations manuelles par volume. Ce qui a permis d'accélérer la constitution de la base de données de plusieurs mois.

V.2.4 Conclusion

Cette partie a présenté l'algorithme utilisé pour obtenir des segmentations 3D des thalami dans nos données échographiques 3D à partir de segmentations manuelles réalisées dans des volumes échographiques débruités. Les segmentations ont été réalisées en coupe coronale ou axiale avec un espacement moyen d'environ 5 et 3 images. L'étape majeure de cet algorithme consiste à effectuer une interpolation linéaire des contours sur les images qui n'ont pas été annotées. La précision de l'algorithme de segmentation a été évalué à partir d'un modèle CAO des thalami. Pour un écart moyen de 5 images en coupe coronale (respectivement axiale) un Dice moyen et une MAD moyenne de 0.984 ± 0.0 (respectivement 0.982 ± 0.001) et 0.13 ± 0.028 (respectivement 0.159 ± 0.022).

Cet algorithme a permis d'obtenir 16 segmentations 3D de référence des thalami d'enfants prématurés dans des données échographiques 3D, ce qui constitue à notre connaissance une première mondiale. Il permet de plus d'obtenir des segmentations 3D en moins d'une minute et ainsi de réduire de plusieurs heures le temps nécessaire pour obtenir ces segmentations. D'après la validation effectuée sur le modèle CAO, les segmentations ont été interpolées avec une bonne précision. Cette étude donne une idée des valeurs maximums et minimums du Dice et de la MAD que l'on pourrait obtenir lors de l'évaluation des algorithmes de segmentation.

V.3 Conclusion

Résumé du travail réalisé

Ce chapitre a décrit l'algorithme qui nous a permis d'obtenir des segmentations 3D des thalami et du SVC à partir de segmentations manuelles 2D réalisées de manière discontinue. La précision des segmentations 3D obtenues a été évaluée par rapport à des modèles CAO de ces structures. Cette expérience a conclu que les segmentations 3D obtenues étaient de très bonne qualité. L'espacement utilisé entre les annotations et l'espacement optimal que l'on pourrait utiliser sont donnés dans la Table.V.1.

TABLE V.1 – Précision des reconstruction 3D pour les espacements utilisés et espacements optimaux pour les annotations 2D du SVC et des thalami.

Structure	SVC	Thalami (coronale)	Thalami (axiale)
Echantillon (examens)	7	3	2
Espacement utilisé (images)	4.1 ± 2.1	5.4 ± 1.8	3.2 ± 0.4
Dice	0.969 ± 0.006	0.984 ± 0.0	0.982 ± 0.001
MAD (mm)	0.16 ± 0.04	0.13 ± 0.03	0.16 ± 0.02
Espacement optimal (intervalle d'images)	5 à 10	10 à 15	5 à 10

D'après ces résultats, l'espacement entre les annotations pourrait être augmenté en conservant une bonne précision des segmentations 3D obtenues. Pour déterminer précisément un espacement optimal, l'évaluation de la précision devra être réalisée sur toute la base de données et pour chaque espacement compris dans l'intervalle optimal donné dans la Table.V.1. Il faudra prendre une marge par rapport à la valeur obtenue et garder en mémoire que l'espacement doit s'adapter aux variations plus ou moins rapides de l'anatomie des structures dans les images lors des segmentations manuelles.

La précision des segmentations 3D interpolées obtenues (Table.V.1) constitue une valeur limite que l'on ne pourra pas dépasser lors de la segmentation de ces structures. Il faut également considéré que, une échographie étant une image que l'on interprète, les segmentations manuelles sont sujettes à une variabilité intra-opérateur et inter-opérateur. Cela signifie que même si toutes les images d'un volume 3D étaient annotées, il serait impossible d'obtenir en pratique un $Dice = 1$ et une $MAD = 0$ mm. La précision maximale que l'on pourra obtenir se situe donc en dessous des valeurs de Dice et de MAD données dans la Table.V.1.

Résumé des contributions

Les contributions apportées par ce chapitre sont les suivantes :

- Un algorithme permettant de créer des segmentations 3D à partir d'annotations 2D réalisées de manière discontinue.
- Une base de données annotées du SVC comprenant 25 volumes échographiques 3D.
- La première base de données annotée des thalami, à notre connaissance mondiale, comprenant 16 volumes échographiques 3D.

Une base de donnée annotées des PWML a également été créée avec un algorithme analogue. Ce qui est également une première mondiale à notre connaissance. La création de cette base n'a pas été présentée dans ce chapitre car la problématique de détection de ces lésions n'a pas été abordée durant cette thèse.

Perspectives

Ces base de données, en particulier celle des thalami, pourront être enrichies avec de nouvelles acquisitions annotées. Il serait aussi intéressant de créer des bases de données annotées d'autres structures cérébrales, telles que le corps calleux, les noyaux caudés, la substance blanche, ... Cela permettrait de se rapprocher des segmentations multi-structures réalisées en IRM. Lors de la réalisation de ces nouvelles segmentations, l'algorithme développé permettra d'estimer au préalable l'espacement entre les annotations permettant d'obtenir des segmentations 3D de bonne qualité dans un temps optimal.

CHAPITRE VI

SEGMENTATION AUTOMATIQUE DU SYSTÈME VENTRICULAIRE CÉRÉBRAL ET DES THALAMI PAR APPRENTISSAGE SUPERVISÉ

Contents

VI.1 Bases de données	140
VI.1.1 Pré-traitement des données	140
VI.1.2 Base de données du SVC	142
VI.1.3 Base de données des thalami	144
VI.2 Réseaux de neurones convolutifs : définition, optimisation et comparaison	145
VI.2.1 Apprentissage profond et réseaux de neurones convolutifs	145
Apprentissage automatique et apprentissage profond	145
Réseaux de neurones convolutifs	146
Réseaux de neurones convolutifs dédiés à la segmentation d'images médicales	148
VI.2.2 Optimisation, précision et comparaison de réseaux de neurones convolutifs .	150
Optimisation d'un réseau de neurones convolutifs	150
Segmentation du jeu de test et métrique d'évaluation de la précision	151
Comparaison de la précision obtenue par deux CNNs.	152
VI.3 Fonction de coût et taille des patches d'entraînement	153
VI.3.1 Définition des fonctions de coût	153
VI.3.2 Optimisation par rapport à l'entropie croisée	154
VI.3.3 Optimisation par rapport à l'entropie croisée puis le softDice	154
VI.3.4 Comparaison de différentes fonctions de coût	156
VI.3.5 Définition de la taille des patches d'entraînement	156
VI.3.6 Conclusion	158
VI.4 Optimisation du nombre de couches de U-net et V-net	159
VI.4.1 Méthode	159
VI.4.2 Expérience et résultats	160
VI.4.3 Volumes mesurés après segmentation du SVC par V-net.	163
VI.4.4 Conclusion	163

VI.5 Apprentissage de la localisation du SVC	164
VI.5.1 Motivations	164
Réseau générateur de formes par compositions.	164
VI.5.2 Architecture du CPPN et intégration à U-net et V-net	164
Architecture du CPPN	164
Intégration du CPPN à U-net et à V-net	166
VI.5.3 Expériences et résultats	166
VI.5.4 Discussion	169
VI.5.5 Conclusion	170
VI.6 Taille adéquate du jeu d'entraînement et augmentation artificielle de données	171
VI.6.1 Taille adéquate de l'ensemble d'entraînement	171
Méthode	171
Résultats et discussion	171
Conclusion	173
VI.6.2 Augmentation artificielle de données	173
Intérêt de l'augmentation de données artificielle	173
Méthodes	173
Résultats	174
Entraînement de V-net avec augmentation artificielle de données.	175
Conclusion	176
VI.7 Segmentation de données atypiques	177
VI.7.1 Présentation des données atypiques	177
VI.7.2 Expériences et résultats	178
VI.7.3 Conclusion	181
VI.8 Segmentation des thalami	182
VI.8.1 Expérience et résultats	182
VI.8.2 Volume thalamique mesuré après segmentation par V-net.	184
VI.8.3 Conclusion	185
VI.9 Conclusion	186
VI.9.1 Résumé du travail effectué	186
VI.9.2 Résumé des contributions	186
VI.9.3 Perspectives	187

CHAPITRE VI

C E chapitre est dédié à la segmentation du SVC et des thalami dans les données échographiques 3D créées en utilisant l'algorithme détaillé dans le Chapitre.IV. Pour résoudre ce problème, nous avons mis en œuvre des méthodes d'apprentissages profonds et plus particulièrement des réseaux de neurones convolutifs. Ceux-ci ont été entraînés à produire des segmentations à partir des annotations 3D de référence obtenues avec l'algorithme présenté dans le Chapitre.V.

La majeure partie de ce chapitre concerne la segmentation du SVC. La raison est que, comme cela a été évoqué dans le Chapitre.I, la segmentation de cette structure a un intérêt clinique immédiat pour la prise en charge des prématurés et un intérêt pour le pronostic des troubles du neuro-développement. Le choix et l'optimisation des réseaux de neurones convolutifs ont été réalisés pour segmenter au mieux cette structure. La méthode retenue a ensuite été utilisée pour segmenter les thalami, ce qui a permis de savoir si cette méthode pourrait permettre de segmenter d'autres structures du cerveau avec précision.

Les contributions majeures présentées dans ce chapitre sont :

- La première segmentation automatique du SVC, par des réseaux de neurones convolutifs, dans des données échographiques 3D et dans un temps clinique.
- La première segmentation automatique des thalami dans des données échographiques 3D, dans un temps clinique.
- La modification des architectures de CNNs qui permet à ces réseaux d'encoder la position anatomique du SVC.
- L'estimation du nombre de volumes d'apprentissage nécessaires pour segmenter le SVC avec précision ($\text{Dice} > 0.8$) et la comparaison de plusieurs stratégies d'augmentation artificielle de données.

Ce chapitre est divisé en huit parties. La première présente les bases de données qui ont été utilisées. La seconde partie est dédiée au fonctionnement des CNNs. La troisième détaille comment la fonction de coût et la taille des patches mis en entrée des réseaux ont été définies. Les performances obtenues par des CNNs 2D et 3D pour segmenter le SVC sont ensuite comparées. La cinquième partie présente les modifications effectuées dans les architectures des CNNs pour encoder la position anatomique du SVC. La partie suivante est dédiée à l'estimation du nombre de volumes nécessaires pour obtenir une bonne précision après apprentissage et aux types d'augmentation artificielle de données qui peuvent être utilisés pour cette application. La septième partie discute de la segmentation de données atypiques du SVC. Enfin, la dernière partie est dédiée à la segmentation des thalami.

VI.1 Bases de données

Cette première partie présente les bases de données qui ont été utilisées pour travailler sur les problèmes de segmentation du SVC et des thalami.

VI.1.1 Pré-traitement des données

Pour cette étude, 25 volumes échographiques positionnés dans le repère commun (Figure.IV.5) ont été utilisés. Ceux-ci ont été créés à partir d'ETFs 2D réalisées avec une sonde multi-D 9L4 en utilisant l'algorithme de création de volumes échographiques présenté dans le Chapitre.IV. Les segmentations 3D associées à ces volumes ont été créées avec l'algorithme présenté dans le Chapitre.V. Un exemple de volume échographique et les segmentations du SVC et des thalami qui lui sont associées sont visibles sur la Figure.VI.1.

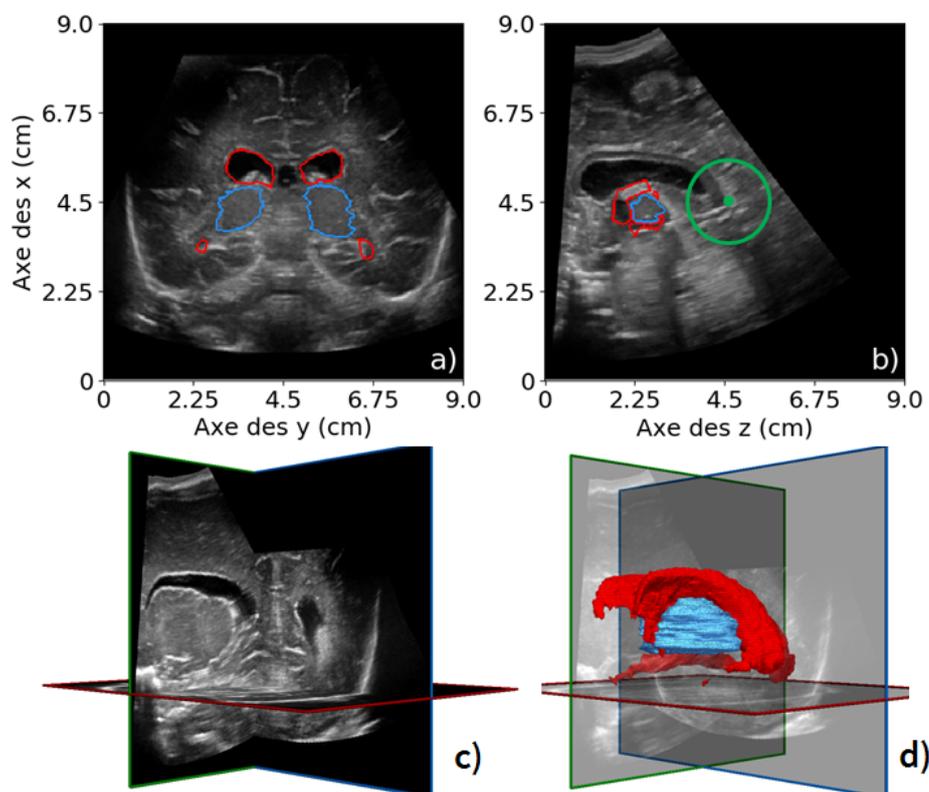


FIGURE VI.1 – Coupe coronale (a) et sagittale (b) de données échographiques 3D positionnées dans le repère commun (c) et segmentation 3D interpolée du SVC (en rouge) et des thalami (en bleu) associées (d), l'origine du repère commun est visible en vert sur (b)¹.

Avant d'être utilisées pour entraîner les CNNs, les données échographiques 3D et les segmentations 3D ont été pré-traitées. Ce pré-traitement avait pour objectifs de réduire les besoins en mémoire et de créer des volumes de même dimension. Dans un premier temps, des sous-volumes de taille $640 \times 640 \times 640$ (pixels) centrés sur l'origine du repère commun ont été extraits, en prenant soin

1. Images issues de notre base de données.

de ne pas perdre une partie des segmentations. Les dimensions des données ont ensuite été divisées par deux en utilisant une interpolation bicubique pour les données échographiques et sur le proche voisin pour les segmentations de référence, ce qui a permis d'obtenir des données de taille $320 \times 320 \times 320$. Les données échographiques ont été ensuite centrées et réduites : dans chaque coupe coronale dans le cas des réseaux 2D et par rapport à tout le volume dans le cas des réseaux 3D. Les segmentations du SVC et des thalami ont respectivement été représentées par les entiers 1 et 2 dans les volumes contenant les segmentations. Lorsqu'un même pixel avait deux étiquettes (les segmentations 3D sont interpolées), nous avons considéré qu'il appartenait aux thalami car les segmentations manuelles de cette structure avaient été réalisées dans les volumes 3D, contrairement à celles du SVC.

Dans le cas de la segmentation du SVC, les données X et Y utilisées pour entraîner les CNNs sont visibles en coupe coronale sur les Figure.VI.2.a et Figure.VI.2.b.

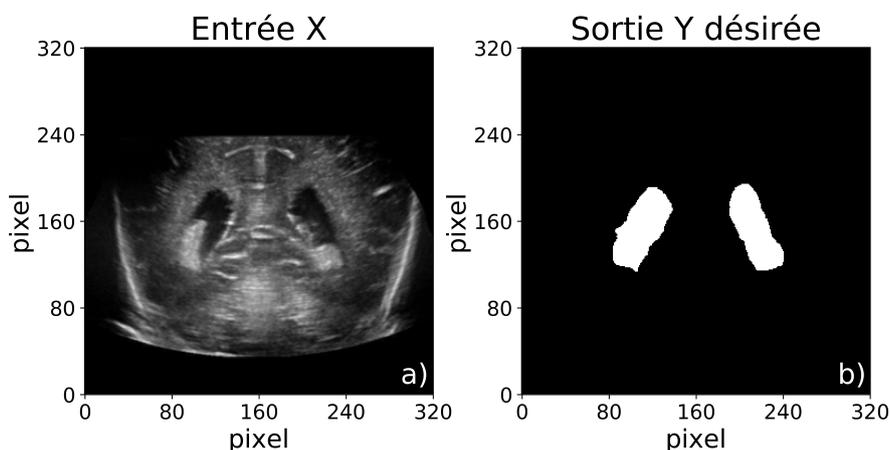


FIGURE VI.2 – Données (X, Y) dans le cas de la segmentation du SVC.

On voit que l'arrière-plan est représenté par 0 et le SVC par 1. Dans le cas de la segmentation du SVC et des thalami, les données X et Y sont visibles en coupe coronale sur la Figure.VI.3.a et la Figure.VI.3.b.

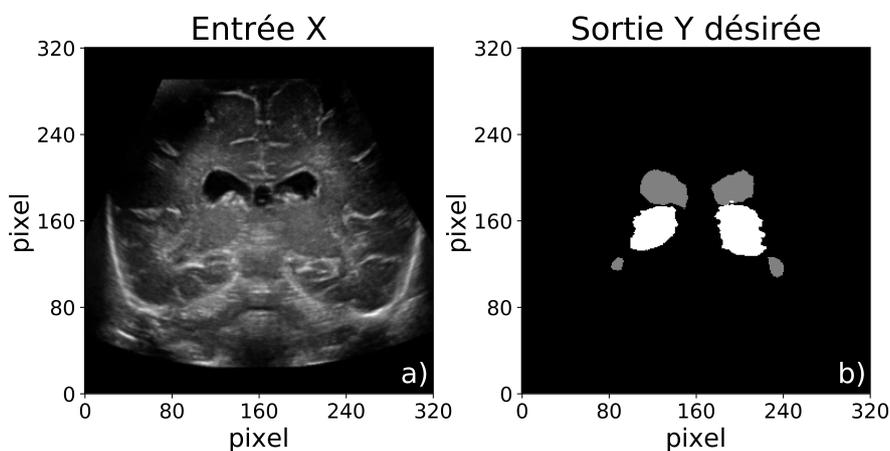


FIGURE VI.3 – Données (X, Y) dans le cas de la segmentation du SVC et des thalami.

Dans ce cas, l'arrière-plan est représenté par 0, le SVC par 1 et les thalami par 2.

VI.1.2 Base de données du SVC

La base de données du SVC a été créée à partir des 25 volumes échographiques disponibles. Ces volumes correspondaient à 21 patients dont l'âge moyen au moment des acquisitions était de $35,81 \pm 1,59$ SA. Cette base a été divisée en un jeu d'entraînement, de validation et de test contenant respectivement 13, 5 et 7 volumes qui incluaient chacun 4, 2 et 2 volumes où le SVC était dilaté. Étant donné que nous avons 25 volumes pour 21 patients, nous avons pris soin de placer les volumes correspondant au même patient dans le jeu d'entraînement pour ne pas biaiser les résultats obtenus. La répartition et le VV moyen dans chaque jeu de données sont présentés dans la Table VI.1.

TABLE VI.1 – Base de données du SVC et répartition entre les jeux d'entraînement, de validation et de test.

Base de données	Nombre de volumes	VV moyen (cm ³)	Nombre de volumes (SVC dilaté)	VV moyen (cm ³)	Nombre de volumes (SVC normal)	VV moyen (cm ³)
Complète	25	4.8 ± 3.2	8	9.1 ± 1.5	17	3.2 ± 2.1
Entraînement	13	4.8 ± 3.1	4	9.4 ± 0.6	9	2.7 ± 0.8
Validation	5	5.2 ± 3.1	2	8.4 ± 2.6	3	3.1 ± 0.9
Test	7	4.5 ± 3.2	2	9.4 ± 1.1	5	2.6 ± 0.5

Les jeux d'entraînement, de validation et de test ont été équilibrés de manière à avoir des distributions proches. Ainsi, environ un tiers des SVC contenus dans chaque jeu de données étaient dilatés. On constate que le VV moyen de tous les SVC, des SVC dilatés et des SVC normaux sont à peu près les mêmes dans chaque jeu de données.

Dans le cas où les SVC étaient dilatés, les segmentations de référence contenues dans l'ensemble de test sont visibles sur la Figure VI.4.

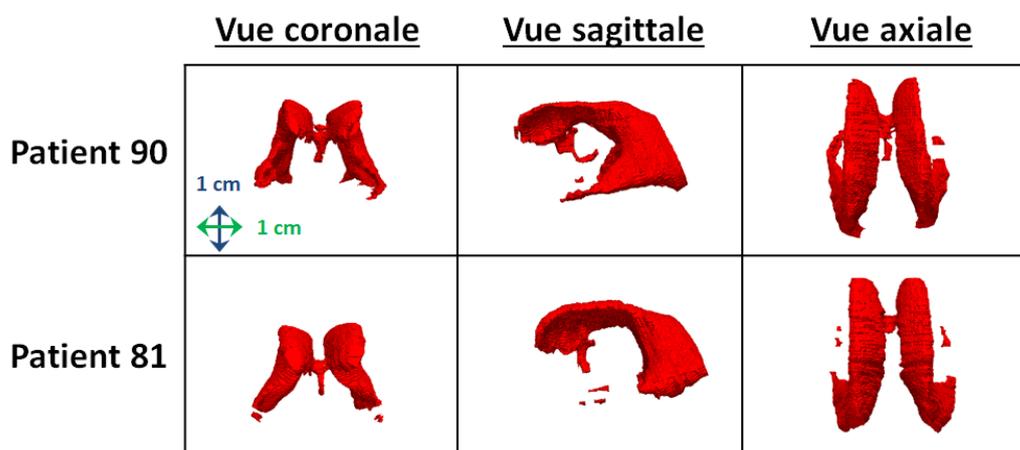


FIGURE VI.4 – SVC dilatés contenus dans le jeu de test.

On voit sur ces images que les segmentations de référence sont très peu fractionnées. Il y a des discontinuités au niveau des cornes temporales qui peuvent s'expliquer par le fait que cette structure est extrêmement fine et qu'elle n'est pas toujours visible dans les images échographiques, ce qui la rend difficile à segmenter. Néanmoins l'aspect global des segmentations est très cohérent avec l'anatomie de cette structure. Les segmentations de référence du jeu de test qui correspondent à des SVC normaux sont visibles sur la Figure.VI.5.

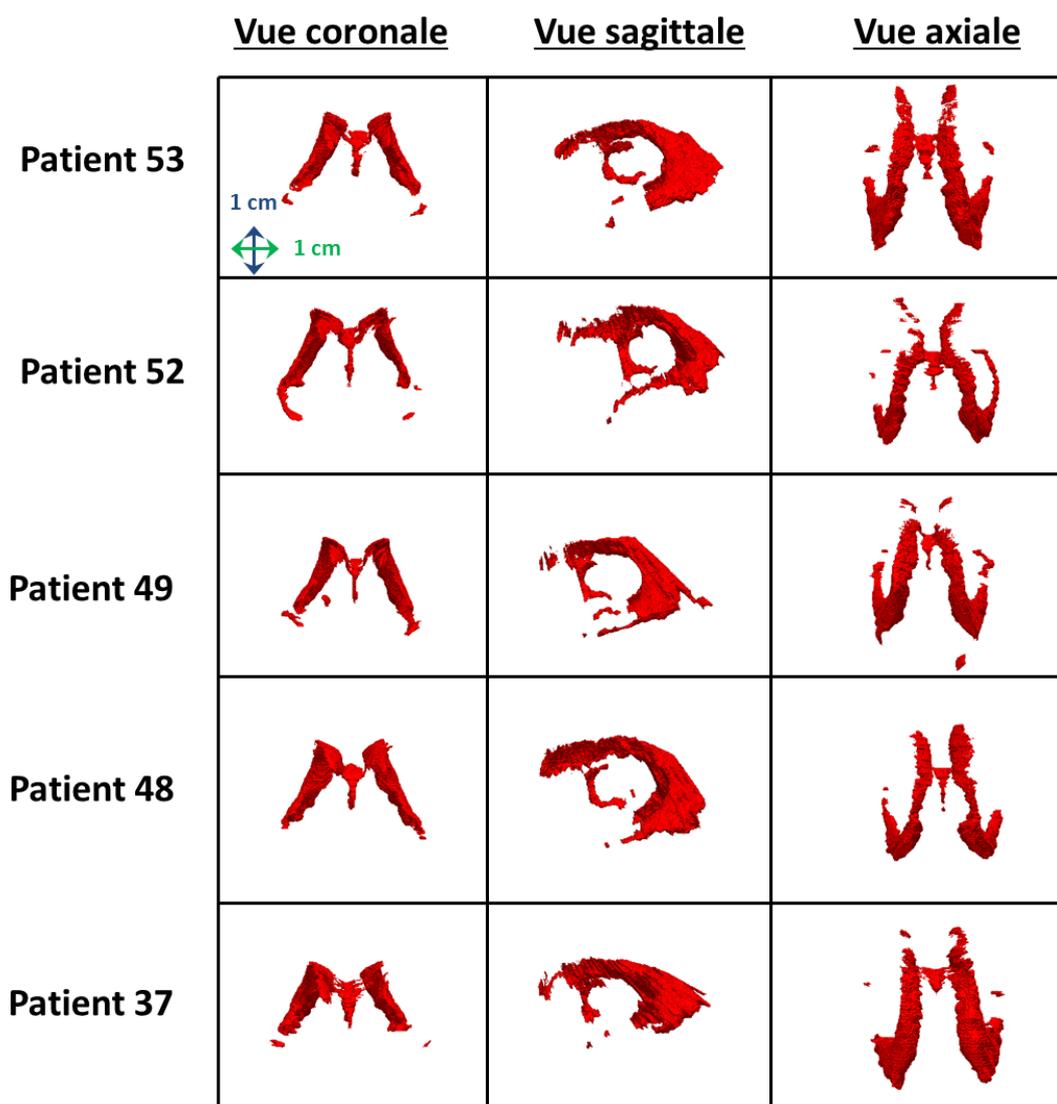


FIGURE VI.5 – SVC normaux contenus dans le jeu de test.

Ces segmentations de référence paraissent cohérentes vis à vis de l'anatomie du SVC. On remarque néanmoins qu'elles peuvent être fractionnées au niveau des cornes frontales et des cornes temporales. Comme pour les SVCs dilatés, cela peut s'expliquer par le fait que ces structures sont extrêmement fines et donc peu visibles dans les données échographiques, ce qui les rend difficiles à segmenter. Ces segmentations restent d'excellente qualité compte tenu de la difficulté que constitue le problème de segmentation manuelle du SVC.

VI.1.3 Base de données des thalami

La base de données des thalami a été créée à partir de 16 volumes échographiques (inclus dans les 25 volumes dont nous disposons). Ces volumes correspondaient à 15 patients dont l'âge moyen au moment des acquisitions était de $35,4 \pm 1,05$ SA. Le nombre de volumes dans les jeux d'entraînement, de validation et de test a été respectivement fixé à 8, 3 et 5. Cette répartition, le VV et le volume thalamique moyen sont donnés dans la Table.VI.2.

TABLE VI.2 – Base de données des thalami et répartition entre les ensembles d'entraînement, de validation et de test

Base de données	Nombre de volumes	VV moyen (cm ³)	Volume thalamique moyen (cm ³)
Complète	16	5.3 ± 3.3	5.3 ± 0.8
Entraînement	8	5.4 ± 3.3	4.8 ± 0.4
Validation	3	5.7 ± 3.7	6.1 ± 0.7
Test	5	4.9 ± 3.0	5.5 ± 0.6

On constate que le VV et le volume thalamique sont proches dans tous les jeux de données. On remarque également que les écart-types associés aux volumes thalamiques moyen sont bien plus faibles que ceux associés au VV. Ce résultat est cohérent car cette structure n'est pas sujette à des variations de taille aussi importante que le SVC. Les segmentations de référence des thalami, contenues dans le jeu de test, sont visibles sur la Figure.VI.6.

La forme de ces segmentations de référence est très cohérente par rapport à l'anatomie de cette structure. Des variations anatomiques peuvent être observées d'une segmentation à l'autre mais celles-ci restent faibles. Cet ensemble de test a 4 patients en commun avec l'ensemble de test défini pour la base de données du SVC.

Cette première partie a détaillé le pré-traitement réalisé sur les données échographiques 3D et les segmentations de référence associées. Elle a également présenté les bases de données du SVC et des thalami qui ont été utilisées ainsi que la répartition des volumes dans les jeux d'entraînement, de validation et de test. La partie suivante est dédiée à l'apprentissage profond et aux réseaux de neurones convolutifs que nous avons utilisés pour segmenter ces deux structures.

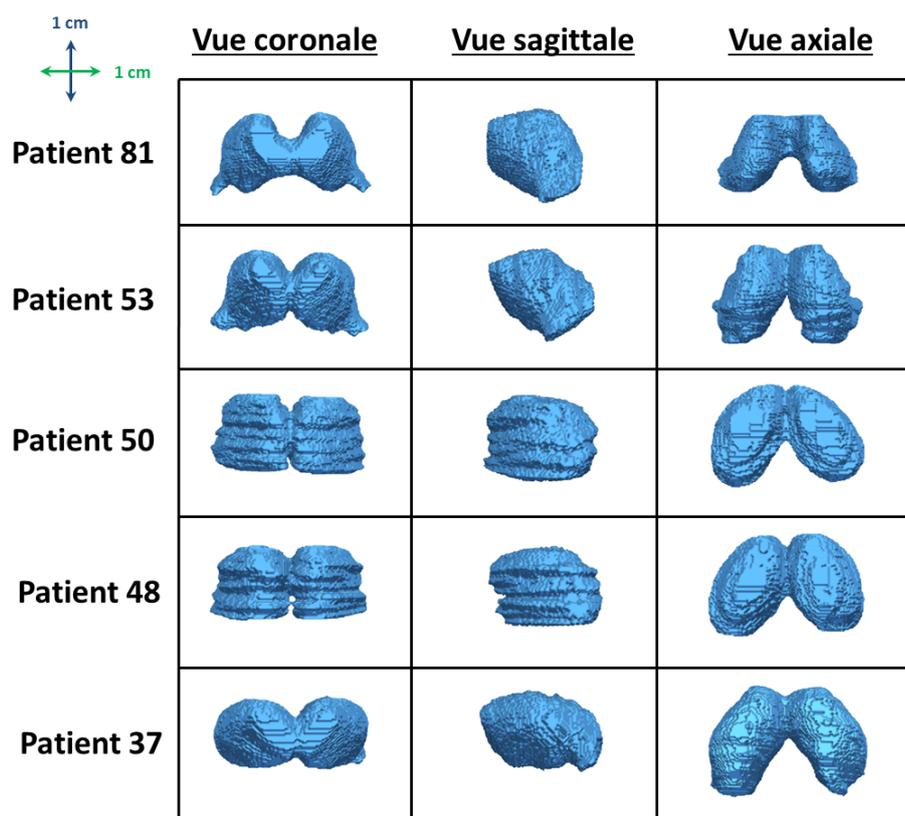


FIGURE VI.6 – Thalami contenus dans le jeu de test.

VI.2 Réseaux de neurones convolutifs : définition, optimisation et comparaison

VI.2.1 Apprentissage profond et réseaux de neurones convolutifs

Apprentissage automatique et apprentissage profond

Un algorithme d'apprentissage automatique est défini comme un algorithme capable d'apprendre à partir de données. Une définition plus précise a été donnée par [Mitchell et Learning \(1997\)](#) : "On dit qu'un programme informatique apprend de l'expérience E par rapport à une catégorie de tâches T et à une mesure de performance P , si sa performance à l'exécution des tâches T , telle que mesurée par P , s'améliore avec la pratique de l'expérience E ."

Dans notre cas, l'expérience E correspond à des couples (X, Y) où $X \in \mathcal{X}$, l'ensemble des données échographiques 3D, et $Y \in \mathcal{Y}$, l'ensemble des segmentations 3D associées à ces volumes. Dans les données Y chaque structure est représentée par un entier naturel (étiquette) : 0 pour l'arrière plan, 1 pour le SVC et 2 pour les thalami. La tâche T que l'on souhaite faire apprendre à l'algorithme est de produire une segmentation \hat{Y} à partir de données d'entrée X , c'est à dire un volume étiqueté. Ce problème est donc un problème de classification supervisée. Les mesures P utilisées pour mesurer la performance des réseaux seront définies dans la Partie.VI.2.2 et la Partie.VI.3.

En apprentissage supervisé, les algorithmes d'apprentissage automatique ont pour objectif d'ap-

prendre la distribution de probabilité conditionnelle $P(Y/X)$. L'approche classique pour résoudre les problèmes de classification consiste à définir et extraire des caractéristiques de X puis à attribuer une étiquette à chacun de ses pixels à partir de celles-ci. Néanmoins, il peut être extrêmement long et difficile de définir ces caractéristiques de façon pertinente. Pour faciliter cette étape, il est possible d'utiliser des méthodes d'apprentissage de représentation pour apprendre à représenter automatiquement les données X . Les méthodes d'apprentissage profond font partie de ces méthodes, elles ont la particularité de permettre l'apprentissage de représentations très abstraites, voire de concepts. La hiérarchie de ces classes de méthodes est représentée par un diagramme de Venn sur la Figure.VI.7.

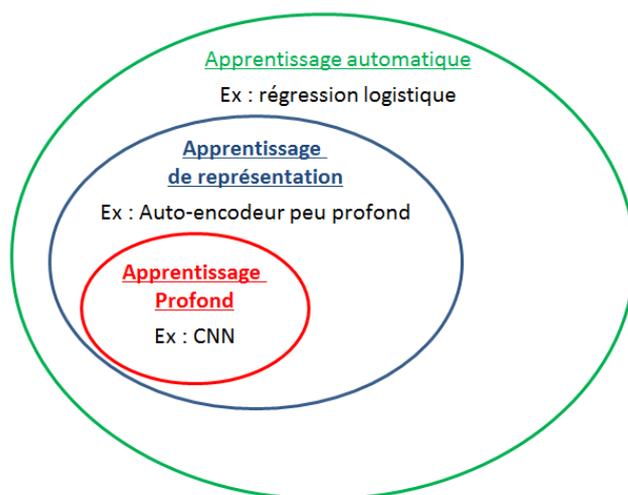


FIGURE VI.7 – Classes des méthodes d'apprentissage automatique.

Les réseaux de neurones convolutifs font partie des méthodes d'apprentissage profond, ils sont particulièrement adaptés aux problèmes de segmentation d'images.

Réseaux de neurones convolutifs

Par leur capacité à apprendre des représentations complexes et abstraites, les CNNs sont devenus récemment l'état de l'art sur divers problèmes de segmentation d'images (Long *et al.* (2015), Badrinarayanan *et al.* (2017), Isola *et al.* (2017)). Ils ont obtenu une excellente précision en quelques secondes sur ceux-ci, ce qui rend leur utilisation pertinente pour les problèmes de segmentation d'images médicales. Un CNN est défini comme un réseau de neurones qui utilise une opération de convolution à la place de l'opération de multiplication matricielle dans au moins une de ces couches de neurones (Goodfellow *et al.* (2016)). Ces réseaux sont généralement composés de plusieurs couches convolutives qui comprennent au moins deux opérations :

- une convolution
- une fonction d'activation non linéaire

et parfois d'autres, telles que :

- une normalisation du batch
- un pooling

Une couche convolutive est représentée sur la Figure.VI.8.

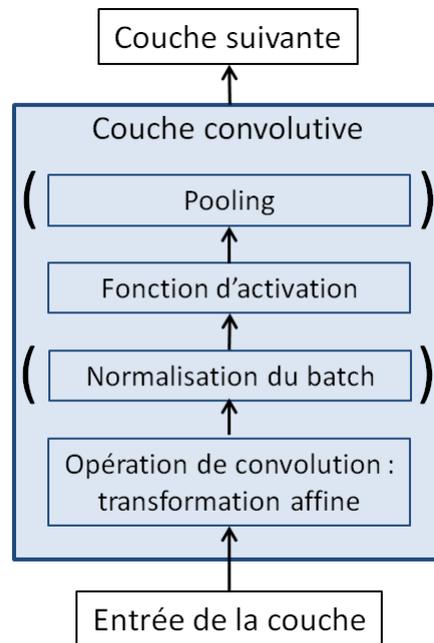


FIGURE VI.8 – Représentation d'une couche convolutive.

L'opération de convolution, notée $*$, réalise une transformation affine d'une image d'entrée I à l'aide d'un noyau de convolution K . Elle est définie par l'Équation.VI.1

$$(I * K)(i, j) = b_K + \sum_m \sum_n I(i - m, j - n) K(m, n) \quad (\text{VI.1})$$

En fonction des réseaux et des images traitées, le noyau de convolution K peut être de dimension 2 (réseau 2D) ou 3 (réseau 3D), ces éléments sont appelés poids du réseau. Les réels b_K correspondent aux biais du réseau.

Les fonctions d'activation utilisées dans les CNNs sont non-linéaires. Pour notre application, nous avons utilisé la fonction ReLU (Nair *et* Hinton (2010)) définie par l'Équation.VI.2 qui est classiquement utilisée dans les CNNs.

$$\begin{aligned} \text{ReLU}: \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \max(0, x) \end{aligned} \quad (\text{VI.2})$$

Cette opération permet d'augmenter la capacité des réseaux en autorisant les modèles appris à être non-linéaires.

Les couches convolutives peuvent intégrer une opération de pooling. Celle-ci permet de résumer l'information contenue dans un certain voisinage. Les réseaux avec lesquels nous avons travaillé utilisent l'opération de max-pooling (Zhou *et* Chellappa (1988)) qui extrait la valeur maximale dans un certain voisinage. Appliquée avec un pas de 2, cette opération permet de diviser la dimension des

images par 2.

Un batch est un ensemble d'images données en entrée du réseau. L'opération de normalisation du batch (Ioffe *et Szegedy* (2015)), notée BN, est souvent utilisée dans les CNNs car elle permet d'améliorer la convergence lors de l'apprentissage du réseau (Santurkar *et al.* (2018)).

Une succession de couches convolutives donne une capacité importante à un CNN pour apprendre des représentations abstraites de données d'entrée X . Néanmoins, l'architecture (taille des convolutions, connexions entre les couches, nombre de couches, ...) doit toujours être optimisée en fonction de l'application sur laquelle on travaille (Wolpert *et al.* (1997)). La partie suivante présente des CNNs 2D et 3D classiquement utilisés pour la segmentation d'images médicales.

Réseaux de neurones convolutifs dédiés à la segmentation d'images médicales

Pour les problèmes de segmentation d'images médicales, des CNNs 2D (Ronneberger *et al.* (2015)) et 3D (Milletari *et al.* (2016)) ont récemment vu le jour. Ces deux architectures sont représentées respectivement sur la Figure.VI.9 et la Figure.VI.10, elles vont être décrites dans les deux paragraphes suivants.

U-net U-net est un CNN 2D, il se compose d'une première partie dont le rôle est d'extraire des représentations abstraites des données d'entrée X et d'une deuxième partie qui utilise ces représentations pour calculer la probabilité qu'a chaque pixel d'appartenir à chaque classe (\hat{Y}).

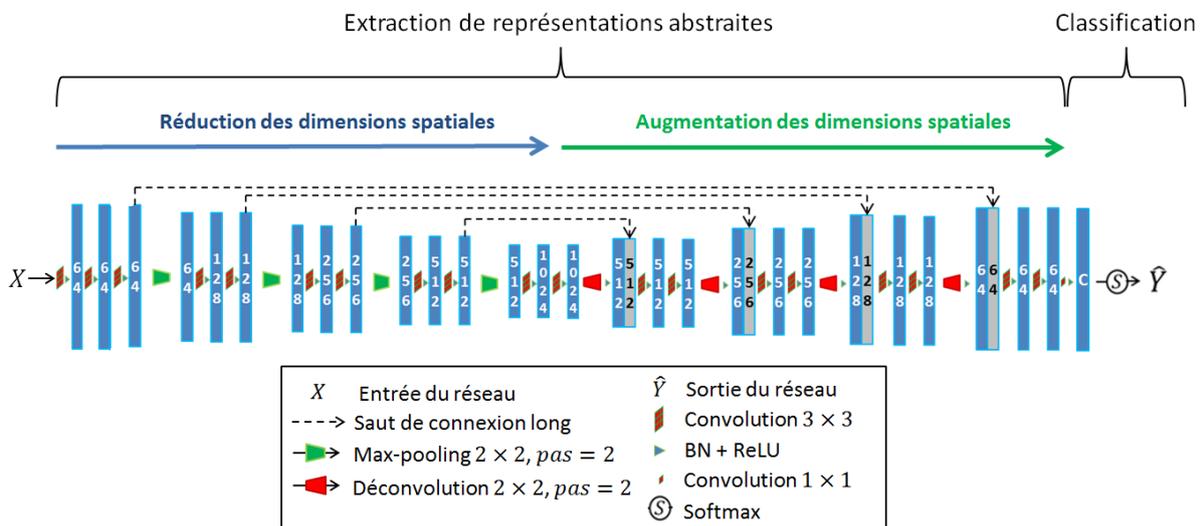


FIGURE VI.9 – Représentation du réseau de neurones convolutifs U-net. Les tenseurs sont représentés par des rectangles bleus ou gris, le nombre de canaux qu'ils contiennent est donné à l'intérieur de ceux-ci.

La première partie du réseau est composée d'une phase où les dimensions spatiales sont réduites et d'une phase où les dimensions spatiales sont augmentées.

La réduction des dimensions spatiales se fait sur quatre étages qui sont chacun composés de deux ou trois couches convolutives et d'une opération de max-pooling. Cette partie permet d'extraire des caractéristiques des images d'entrée à différentes échelles. On remarque aussi que le nombre

de carte de caractéristiques (nombre de canaux) augmente avec la réduction des dimensions spatiales, ce qui permet de générer de plus en plus de cartes de caractéristiques.

De manière analogue, l'augmentation des dimensions spatiales des cartes de caractéristiques obtenues en bas du U est réalisée avec 4 étages qui se composent chacun de deux couches convolutives et d'une couche déconvolutive. Après chaque opération de déconvolution, un saut de connexion long fait le lien entre les étages de même niveau des parties de contraction et d'expansion du réseau. Cela permet d'utiliser les cartes de caractéristiques générées avant réduction des dimensions de manière à améliorer la précision des cartes de caractéristiques obtenues lors de l'augmentation des dimensions spatiales. Cette méthode permet en outre d'améliorer la convergence lors de la phase d'apprentissage de ces réseaux (Drozdal *et al.* (2016)).

La deuxième partie du réseau est composée d'une couche convolutive, sans BN, qui réalise des opérations de convolution de taille 1×1 . Cette couche joue le rôle d'un classifieur linéaire, on obtient à sa sortie un tenseur Z composé d'autant de canaux que de classes C considérées dans le problème de segmentation. Cette couche est suivie de la fonction softmax définie par l'Équation.VI.3 qui calcule un tenseur \hat{Y} ayant C canaux.

$$\forall j \in [1, C], \hat{y}_j = \frac{\exp(z_j)}{\sum_{c=1}^C \exp(z_c)} \quad (\text{VI.3})$$

Où \hat{y}_j et z_j désignent respectivement la carte de caractéristique contenue dans le j -ème canal de \hat{Y} et Z . Cette fonction permet d'attribuer à chaque pixel de l'image d'entrée une probabilité d'appartenir à chaque classe. La couche convolutive de dimension 1, proposée par Long *et al.* (2015), permet d'attribuer simultanément cette probabilité à tous les pixels, ce qui permet de segmenter des volumes entiers en quelques secondes.

V-net Le réseau V-net a une architecture analogue au réseau U-net mais utilise des convolutions 3D. Nous avons fixé la taille de celles-ci à $3 \times 3 \times 3$ pour permettre des comparaisons avec le réseau U-net.

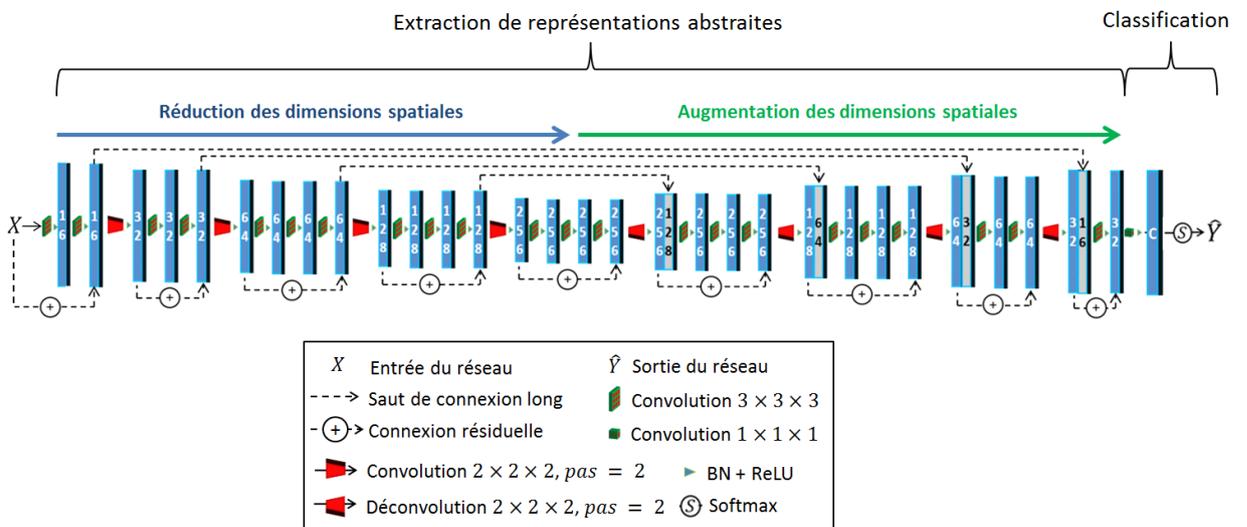


FIGURE VI.10 – Représentation du réseau de neurones convolutifs V-net.

Il se distingue principalement de U-net par la présence d'une connexion résiduelle (He *et al.* (2016)) à chaque étage. Cette connexion permet d'améliorer la vitesse de convergence lors de la phase d'optimisation et d'optimiser des réseaux plus profonds (Drozdzal *et al.* (2016)). De plus, les opérations de max-pooling sont remplacées par des opérations de convolutions $2 \times 2 \times 2$ avec un $pas = 2$.

Pour résoudre les problèmes de segmentation du SVC et des thalami, nous avons travaillé avec ces deux architectures. Cela nous a permis de déterminer si une architecture 2D ou 3D était plus adaptée.

VI.2.2 Optimisation, précision et comparaison de réseaux de neurones convolutifs

Optimisation d'un réseau de neurones convolutifs

Après avoir créé une base de données et défini un réseau de neurones convolutif, il faut entraîner ce dernier à produire des segmentations des structures qui nous intéressent. Cette section décrit la méthode et les paramètres qui ont été utilisés pour entraîner U-net et V-net.

Les poids des réseaux ont été initialisés selon une distribution uniforme avec l'initialiseur de Xavier (Glorot *et Bengio* (2010)) et les biais ont été initialisés à 0.01. Les paramètres des réseaux ont ensuite été optimisés de la manière suivante :

1. Un ensemble de patch (images échographiques) extrait aléatoirement du jeu d'entraînement est donné en entrée du réseau.
2. Le réseau produit une sortie \hat{Y} .
3. \hat{Y} est comparée à la segmentation de référence Y avec une fonction de coût (\mathcal{L}). Le gradient de l'erreur est ensuite rétro-propagé dans les différentes couches du réseau pour ajuster ses paramètres de manière à minimiser \mathcal{L} .

L'itération de ces trois étapes permet d'optimiser les paramètres du réseau par rapport à la fonction de coût \mathcal{L} . La taille et le nombre de patches contenus dans chaque batch donné en entrée de U-net et V-net se trouvent dans la Table.VI.3.

TABLE VI.3 – Taille et nombre de patches donnés en entrée des réseaux lors de la phase d'entraînement.

Réseau	Taille des patches (pixel)	Taille des batchs (patch)
U-net	128×128	10
V-net	$128 \times 128 \times 64$	5

Le choix de la taille des patches sera justifié dans la partie.VI.3.

L'optimisation des paramètres du réseau a été réalisée en utilisant l'algorithme de rétropropagation du gradient et l'optimiseur Adam (Kingma *et Ba* (2014)) avec $\beta_1 = 0.9$, $\beta_2 = 0.999$ et $\epsilon = 1 \times 10^{-8}$. Le

pas de temps a été réglé à 1×10^{-4} pendant les 10000 premières itérations, 2×10^{-5} pendant les 10000 itérations suivantes et à 5×10^{-6} jusqu'à la fin de l'entraînement.

Entraîner un CNNs ne correspond pas exactement à résoudre un problème d'optimisation. On cherche en effet à optimiser au mieux les paramètres du réseau en minimisant une fonction de coût mais on souhaite également que le réseau soit capable de réaliser des segmentations précises sur des données qui n'ont pas servi à son optimisation. Pour éviter le phénomène de sur-apprentissage l'erreur donnée par la fonction de coût a été calculée sur le jeu de validation toutes les 1000 itérations. Les paramètres finalement conservés sont ceux qui avaient permis d'obtenir la plus petite erreur sur l'ensemble de validation.

Le processus d'optimisation a été arrêté avec un critère d'arrêt précoce : plus de 15000 itérations réalisées et erreur calculée sur le jeu de validation qui n'a pas diminuée depuis 5000 itérations.

Segmentation du jeu de test et métrique d'évaluation de la précision

Segmentation du jeu de test. Après qu'un réseau ait été optimisé à partir des jeux d'entraînement et de validation, ses performances ont été évaluées sur le jeu de test.

Dans le cas 2D, la segmentation de chaque volume (taille de $320 \times 320 \times 320$ pixels) a été réalisée en segmentant une par une ses images coronales (taille de 320×320 pixels).

Dans le cas 3D, les volumes ont été segmentés par sous-volumes de taille $320 \times 320 \times 64$ avec un recouvrement de 75 %. Les probabilités d'appartenir à chacune des classes, obtenues en chaque voxel, ont ensuite été sommées. La segmentation finale a été obtenue en attribuant à chaque voxel l'étiquette qui correspondait à la plus haute valeur des sommes de probabilités.

Métrique d'évaluation de la précision d'une segmentation. Après optimisation, la précision des réseaux a été évaluée sur le jeu de test avec deux mesures de la performance (P) ou métriques : le Dice (Équation.VI.4) et la MAD (Équation.VI.5).

$$Dice = 2 \times \frac{|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \quad (VI.4)$$

$$MAD = \frac{1}{2} \left(\frac{\sum_{x \in \partial Y} \min d(x, \partial \hat{Y})}{|\partial Y|} + \frac{\sum_{x \in \partial \hat{Y}} \min d(x, \partial Y)}{|\partial \hat{Y}|} \right) \quad (VI.5)$$

Le Dice est compris entre 0 et 1 (segmentation parfaite), il permet d'estimer la superposition des segmentations de références et des segmentations automatiques. Il s'intéresse donc à la segmentation des structures dans leur totalité, il est intéressant d'utiliser cette métrique car notre objectif final est de mesurer des volumes.

La MAD correspond à la distance moyenne entre les bords des segmentations automatiques et de référence. Elle permet donc de savoir si les structures sont contourées avec précision.

L'erreur volumique entre une segmentation de référence et une segmentations produite par un réseau a été calculée avec deux métriques : l'erreur volumique absolue (ΔV_a définie par l'Équation.VI.6) et l'erreur volumique relative (ΔV_r définie par l'Équation.VI.7).

$$\Delta V_a = \|V_{ref} - V_{seg}\|_1 \quad (VI.6)$$

$$\Delta V_r = \frac{\|V_{ref} - V_{seg}\|_1}{V_{seg}} \quad (VI.7)$$

Où V_{ref} désigne le volume de la segmentation de référence et V_{seg} le volume de la segmentation produite par un CNN. Ces deux métriques ne permettent pas d'estimer la précision des segmentations mais ont été utilisées pour avoir une idée de l'erreur commise sur le volume des structures cérébrales qui est un paramètre clinique qui intéresse les médecins.

Comparaison de la précision obtenue par deux CNNs.

Lorsque l'on optimise un CNN, on résout un problème d'optimisation non convexe. La solution trouvée a donc de fortes chances de correspondre à un minimum local. Ainsi, étant donné que les paramètres d'un réseau sont initialisés aléatoirement et qu'une descente de gradient stochastique par batch est utilisée lors de l'optimisation, il y a peu de chance que la même solution soit trouvée plusieurs fois si le processus d'optimisation est répété. Les performances obtenues sur le jeu de test varient donc d'un entraînement à un autre, même si les jeux d'entraînement et de validation restent inchangés. Les Dices moyens obtenus sur le jeu de test pour 30 optimisations de U-net sur les données du SVC sont visibles sur l'histogramme de la Figure.VI.11.

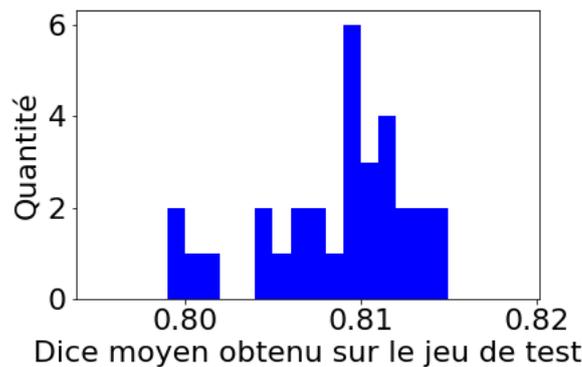


FIGURE VI.11 – Histogramme du Dice moyen obtenu sur le jeu de test après avoir optimisé 30 U-net.

On peut voir que le Dice moyen obtenu sur le jeu de test se situe entre 0.8 et 0.815 et dépend des entraînements. Ainsi, lorsqu'on cherche à mettre en évidence une différence significative entre deux architectures dont les performances sont proches, il est nécessaire d'optimiser plusieurs fois (disons N fois) chaque réseau. Cela permet d'obtenir une distribution de N valeurs du Dice et de la MAD pour chaque architecture. Dans notre cas, des t-test ont été réalisés pour comparer ces distributions et une différence a été considérée comme significative lorsque la p-valeur, notée p , était inférieure à 0.05. On considère par exemple le cas où les Dice moyens obtenus sur le jeu de test

après avoir optimisé cinq fois U-net et V-net sont respectivement 0.808,0.809,0.811,0.800,0.810 et 0.821,0.823,0.820,0.819,0.820. Alors, la p-valeur obtenu en réalisant un t-test entre ces deux distribution est inférieur à 0.05. On conclut donc que V-net est plus précis (par rapport au Dice) que U-net si on cherche à segmenter le SVC.

VI.3 Fonction de coût et taille des patches d'entraînement

Cette partie détaille le choix de la fonction de coût et de la taille des patches d'entraînement.

VI.3.1 Définition des fonctions de coût

Une fonction de coût sert comme mesure de la performance P , elle est utilisée pour optimiser les paramètres d'un réseau de manière à le rendre plus performant.

L'optimisation d'un CNN pour le type de problème auquel on s'intéresse, la classification multi-labels, se fait classiquement par rapport à l'entropie croisée (EC) qui est définie par l'Équation.VI.8.

$$EC(Y, \hat{Y}) = -\sum_{c=1}^C \sum_{i=1}^N y_{i,c} \ln(\hat{y}_{i,c}) \quad (\text{VI.8})$$

Où $y_{i,c} \in \{0, 1\}$ vaut 1 si le i -ème pixel de X appartient à la classe c et $\hat{y}_{i,c} \in [0, 1]$ est la probabilité, calculée par le réseau, qu'a le i -ème pixel de X d'appartenir à la classe c .

Dans le cas où il existe une forte disparité entre les fréquences d'apparition de chaque classe dans Y , l'entropie croisée pondérée (ECP) est généralement utilisée. Celle-ci est définie par l'Équation.VI.9.

$$ECP(Y, \hat{Y}) = -\sum_{c=1}^C \frac{1}{f_c} \sum_{i=1}^N y_{i,c} \ln(\hat{y}_{i,c}) \quad (\text{VI.9})$$

Où f_c désigne la fréquence normalisée qu'à la classe c dans Y . La fréquence d'apparition du SVC était par exemple de $\frac{1}{170}$ dans notre cas.

Le softDice, qui est une généralisation de l'indice de Dice à une distribution de labels (au sens probabiliste), est aussi parfois utilisé (Milletari *et al.* (2016)). Il est défini par l'Équation.VI.10.

$$\text{softDice}(Y, \hat{Y}) = 1 - \frac{1}{C} \sum_{c=1}^C 2 \times \frac{\sum_{i=1}^N y_{i,c} \hat{y}_{i,c}}{\delta + \sum_{i=1}^N y_{i,c} + \hat{y}_{i,c}} \quad (\text{VI.10})$$

Le paramètre δ a été fixé à 10^{-10} , il permet d'éviter de diviser par 0 dans le cas où Y (qui est en pratique un batch) ne contient pas une des classes. Avec cette fonction de coût, chaque classe a le même poids, indépendamment de sa fréquence dans Y , dans le calcul de l'erreur. Le calcul du SoftDice a donc été réalisé tel que défini par l'Équation.VI.11 :

$$\text{softDice}(Y, \hat{Y}) = 1 - \frac{1}{C-1} \sum_{c=2}^C 2 \times \frac{\sum_{i=1}^N y_{i,c} \hat{y}_{i,c}}{\delta + \sum_{i=1}^N y_{i,c} + \hat{y}_{i,c}} \quad (\text{VI.11})$$

C'est à dire en excluant la classe $c = 1$ qui correspondait à l'arrière-plan.

VI.3.2 Optimisation par rapport à l'entropie croisée

Dans un premier temps, l'EC a été utilisée pour optimiser U-net dans le cas du problème de segmentation du SVC. Les courbes d'apprentissage obtenues, moyennée sur cinq optimisations, sont visibles sur la Figure.VI.12.a. L'évolution de la valeur du softDice, qui n'a pas été utilisé comme fonction de coût mais évalué lors de l'optimisation, est visible sur la Figure.VI.12.b.

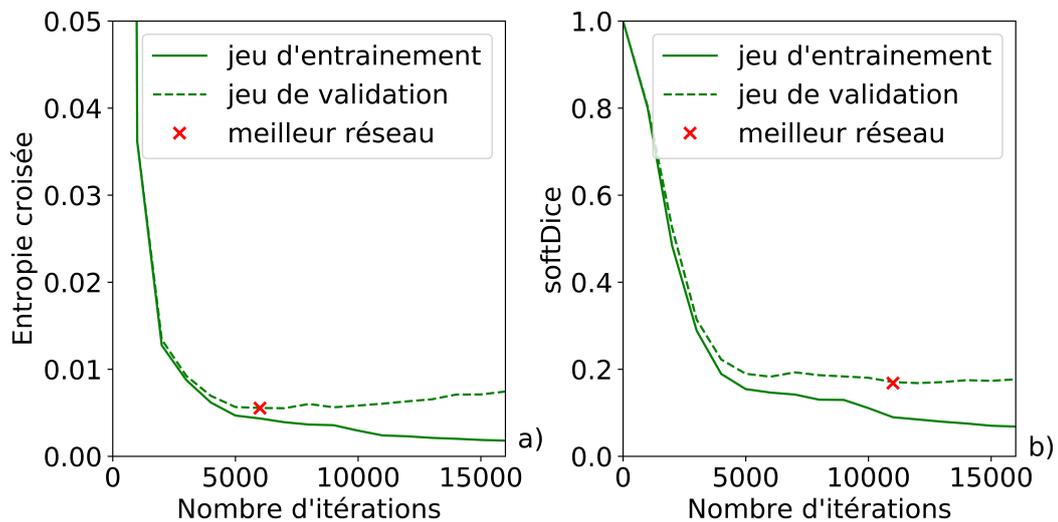


FIGURE VI.12 – Courbes d'apprentissages moyennes pour 5 optimisations du réseau U-net par rapport à l'entropie croisée (a) et valeur du softDice au cours de ces optimisations (b).

On voit sur la Figure.VI.12.a que la valeur de l'EC décroît continuellement sur l'ensemble d'entraînement tandis qu'elle diminue puis augmente sur l'ensemble de validation. Ce comportement signifie qu'à partir d'un certain nombre d'itérations, le modèle sur-apprend le jeu d'entraînement et perd de sa capacité à généraliser au jeu de validation. Cela est caractéristique du sur-apprentissage. C'est pour cela qu'on conserve le réseau à l'itération où la valeur de la fonction de coût était la plus faible sur le jeu de validation (6000 itérations ici) comme expliqué dans la partie.VI.3.

Sur les courbes de la Figure.VI.12, on constate que le softDice calculé sur le jeu de validation n'augmente pas, ou très peu, même après l'itération où le minimum de l'EC a été obtenu sur le jeu de validation. De plus la valeur minimale du softDice sur le jeu de validation a été obtenue après 11000 itérations. Or, la précision des segmentations est évaluée par rapport au Dice qui est une métrique proche du softDice. Choisir le réseau obtenu après 11000 itérations plutôt que celui obtenu après 6000 itérations pourrait donc permettre d'obtenir de meilleures performances.

VI.3.3 Optimisation par rapport à l'entropie croisée puis le softDice

Pour tirer profit de l'observation faite dans la partie précédente, la fonction de coût a été modifiée au cours des entraînements. U-net a par exemple été optimisé par rapport à l'EC lors des 5000 premières itérations, pour prévenir des instabilités lors d'une optimisation uniquement avec la fonction SoftDice, et ensuite par rapport au softDice. Pour ce cas, les courbes d'apprentissages moyennes pour 5 optimisations sont visibles sur la Figure.VI.13.

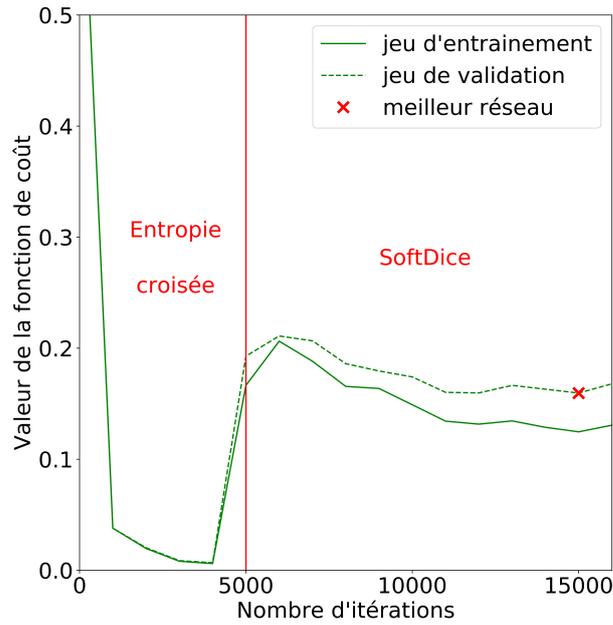


FIGURE VI.13 – Courbes d’apprentissages moyennes pour 5 optimisations du réseau U-net par rapport à l’entropie croisée puis le softDice.

Le changement de fonction de coût après 5000 itérations peut être observé sur ces courbes. On constate que la valeur du softDice calculée sur le jeu de validation diminue bien après le changement de fonction de coût.

La Figure.VI.14 montre l’évolution de la valeur du softDice au cours de l’optimisation de U-net en utilisant l’EC ou l’EC puis le softDice comme fonction de coût.

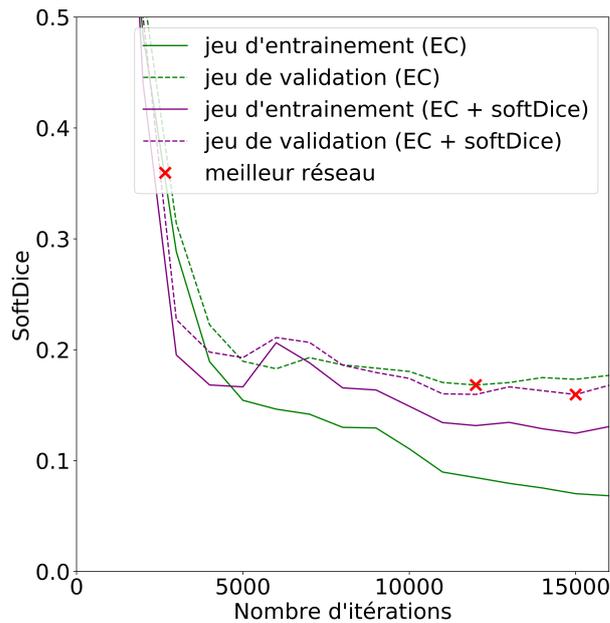


FIGURE VI.14 – Courbes d’apprentissages moyennes pour 5 optimisations du réseau U-net par rapport à l’entropie croisée et par rapport à l’entropie croisée puis au softDice.

On voit sur ces courbes que la valeur minimum obtenue pour le softDice sur le jeu de validation est plus faible lorsqu'on utilise l'EC puis le softDice comme fonction de coût.

On remarque aussi que l'écart entre la valeur du softDice obtenue sur le jeu d'entraînement et le jeu de validation est plus faible en fin d'entraînement lorsqu'on optimise sur l'EC puis le softDice. Cela montre que le réseau sur-apprend moins avec cette fonction de coût.

VI.3.4 Comparaison de différentes fonctions de coût

Pour choisir la fonction de coût qui serait la plus adaptée à notre problème, U-net a été optimisé avec cinq fonctions de coût différentes : l'EC (Équation.VI.8), l'EC pondérée (Équation.VI.9), l'EC puis le softDice (Équation.VI.10), l'EC pondérée puis le softDice et le softDice. Ces optimisations ont été réalisées en utilisant des batchs contenant 2 patchs de taille 320×320 . Les Dice moyens obtenus sur les jeux d'entraînement et de validation pour cinq optimisations du réseau U-ne sont donnés dans la Table.VI.4.

TABLE VI.4 – Dice moyens obtenus sur les jeux d'entraînement et de validation pour différentes fonctions de coût.

Fonction de coût	Dice moyen (entraînement)	Dice moyen (validation)	p-valeur
Entropie croisée	0.816 ± 0.01	0.812 ± 0.01	4.96e-03
Entropie croisée pondérée	0.832 ± 0.02	0.813 ± 0.012	1.87e-02
Entropie croisée + softDice	0.835 ± 0.01	0.827 ± 0.004	5.72e-02
Entropie croisée pondérée + softDice	0.858 ± 0.01	0.832 ± 0.003	1
softDice	0.847 ± 0.006	0.828 ± 0.004	1.29e-01

Le meilleur Dice sur l'ensemble de validation a été obtenu en optimisant U-net par rapport à l'EC pondérée puis le softDice. En utilisant ce résultat comme référence, les p-valeurs montrent que l'écart de performance obtenu est significatif par rapport au cas où l'optimisation a été réalisée par rapport à l'EC ou l'EC pondérée. En revanche, l'écart n'est pas significatif par rapport aux optimisations réalisées avec l'EC puis le softDice ($p = 5.72 \times 10^{-2}$) ou le softDice ($p = 1.29 \times 10^{-1}$).

D'après les résultats obtenus dans cette partie, utiliser le softDice durant l'optimisation permet d'obtenir de meilleurs résultats, au regard du Dice, sur l'ensemble de validation. La prochaine partie détaille comment la taille des patchs d'entraînement a été optimisée et comment le choix entre les trois fonctions de coût utilisant le softDice a été réalisé.

VI.3.5 Définition de la taille des patchs d'entraînement

Pour les trois fonctions de coût retenues, U-net a été optimisé cinq fois en faisant varier la taille des patchs d'entraînement. Cette expérience avait pour objectifs de déterminer quelle fonction de coût et quelle taille de patch seraient les plus adaptées au problème de segmentation du SVC. Le réseau a été optimisé avec chacune des tailles de patch suivantes : 320×320 , 288×288 , 208×208 ,

128 × 128 et 64 × 64 avec des tailles de batch respectif de 2, 5, 5, 10 et 10 de manière à avoir une quantité de données à peu près constante entre les expériences. Le Dice moyen obtenu sur le jeu de validation est donné dans la Table.VI.5 pour chaque configuration.

TABLE VI.5 – Dice moyens obtenus sur l’ensemble de validation en fonction de la fonction de coût utilisée et de la taille des patches d’entraînement.

Taille de patch	Entropie croisée + softDice	Entropie croisée pondérée + softDice	softDice
320 × 320	0.827 ± 0.004	0.832 ± 0.003	0.828 ± 0.004
288 × 288	0.838 ± 0.002	0.832 ± 0.003	0.84 ± 0.002
208 × 208	0.837 ± 0.002	0.835 ± 0.002	0.84 ± 0.002
128 × 128	0.842 ± 0.002	0.841 ± 0.003	0.841 ± 0.001
64 × 64	0.732 ± 0.016	0.745 ± 0.023	0.728 ± 0.018

On constate qu’optimiser U-net avec des patches d’entraînements dont la taille est plus petite que la dimension des images permet d’améliorer les performances obtenues sur l’ensemble de validation, sauf dans les cas où cette taille est de 64 × 64.

Cette amélioration peut s’expliquer par le fait que l’opération de max-pooling n’est pas équivariante par translation. Le fait d’utiliser des sous-images permet de créer artificiellement des translations dans les données d’entrées et donc d’apprendre au réseau à produire les mêmes segmentations quelque soit la position du SVC dans les images.

La perte de performance constatée lorsque les patches d’entrée avaient une taille de 64 × 64 peut s’expliquer par le fait que les dimensions spatiales sont trop réduites en bas du U (4×4) dans cette configuration. Ce qui ne permet pas de restaurer précisément les cartes de caractéristiques dans la partie d’expansion du réseau U-net.

La meilleure performance a été obtenue dans le cas où le réseau a été optimisé avec des patches de taille 128 × 128 avec l’EC puis le softDice. En utilisant ce résultat comme référence, les p-valeurs sont données dans la Table.VI.6.

TABLE VI.6 – p-valeurs obtenues pour différentes fonctions de coût et tailles de patch.

Taille de patch	Entropie croisée + softDice	Entropie croisée pondérée + softDice	softDice
320 × 320	2.13e-05	1.08e-04	6.92e-05
288 × 288	9.78e-03	4.63e-04	9.66e-02
208 × 208	4.62e-03	8.62e-04	7.37e-02
128 × 128	1	2.97e-01	3.32e-01
64 × 64	1.31e-07	5.69e-06	5.69e-06

Ces résultats montrent qu'il n'y a pas de différences significatives dans les cas suivants :

- taille de patch : 128×128 / fonction de coût : EC puis softDice.
- taille de patch 128×128 / fonction de coût : ECP puis softDice.
- taille de patch 128×128 / fonction de coût : softDice.
- taille de patch 208×208 / fonction de coût : softDice.
- taille de patch 288×288 / fonction de coût : softDice.

Chacune de ces cinq configurations semblent donc être optimale pour optimiser le réseau U-net sur le problème de segmentation du SVC. Les résultats obtenus en utilisant uniquement le softDice semblent stables à partir du moment où l'entraînement est réalisé sur des sous-images d'une taille suffisante.

VI.3.6 Conclusion

A partir de ces résultats, nous avons fait le choix d'entraîner les réseaux avec des patches d'une taille de 128×128 en utilisant l'EC puis le softDice comme fonction de coût ($Dice = 0.842 \pm 0.002$). En comparaison, le Dice obtenu sur le jeu de validation en optimisant par rapport à L'EC avec des patches d'entrée de taille 320×320 était de 0.812 ± 0.01 . La méthode d'optimisation choisie permet donc d'améliorer fortement les résultats obtenus sur l'ensemble de validation par rapport à la méthode d'optimisation classique. Nous avons étendu ces résultats au cas 3D en choisissant la même fonction de coût et des patches d'une taille $128 \times 128 \times 64$.

VI.4 Optimisation du nombre de couches de U-net et V-net

Cette partie a pour objectifs d'estimer le niveau d'abstraction optimal de U-net et V-net pour la segmentation du SVC dans nos données échographiques 3D et de déterminer si l'un de ces deux types d'architecture est plus adapté à ce problème.

VI.4.1 Méthode

Pour déterminer le niveau d'abstraction optimal de ces deux architectures, qui dépend de l'application pour laquelle les réseaux sont entraînés, nous avons fait varier leur nombre de couches. Celui-ci a progressivement été augmenté en partant d'un U-net 9 couches et d'un V-net 7 couches. Dans le cas 2D, cette augmentation a été réalisée en ajoutant dans le bas du U :

- un max-pooling (taille : 2×2) et deux couches convolutives (taille : 3×3).
- une couche convolutive (taille : 2×2 , $pas = 2$) et deux couches convolutives (taille : 3×3).

Cette méthode est représentée sur la Figure.VI.15 où l'on passe d'un U-net 17 couches à un U-net 24 couches.

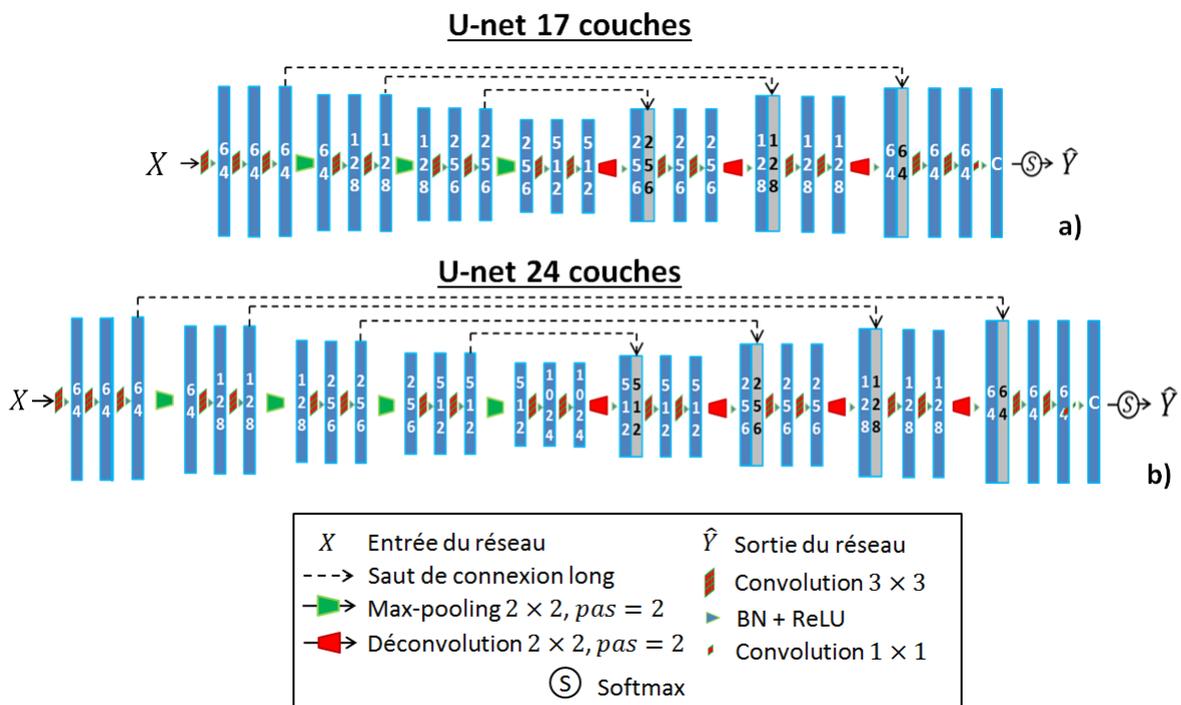


FIGURE VI.15 – Augmentation du nombre de couches (17 (a) à 24 (b)) de l'architecture U-net.

Dans le cas 3D, le nombre de couches de V-net a été augmenté en ajoutant dans le bas du V :

- une couche de convolution (taille : $2 \times 2 \times 2$) avec un $pas = 2$ et trois couches de convolutions (taille : $3 \times 3 \times 3$).
- une couche de déconvolution (taille : $2 \times 2 \times 2$) avec un $pas = 2$ et trois couches de convolutions (taille : $3 \times 3 \times 3$).

Cette méthode est illustrée sur la Figure.VI.16 où l'on passe d'un V-net 23 couches à un V-net 31 couches.

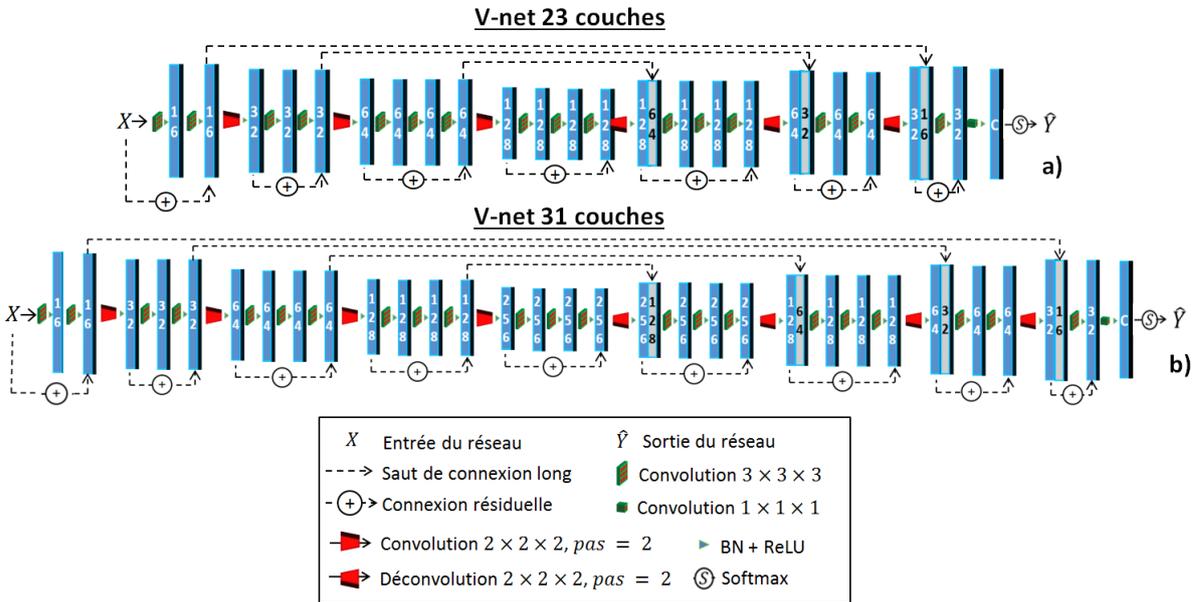


FIGURE VI.16 – Augmentation du nombre de couches (23 (a) à 31 (b)) de l'architecture V-net.

Le nombre de couches de V-net et U-net n'a pas été augmenté au delà de 31 et 23 couches car la taille des patches d'entraînement (128×128) avait été optimisée avec un réseau U-net 24 couches qui comprenait quatre opérations réduisant les dimensions spatiales (comme le V-net 31 couches).

VI.4.2 Expérience et résultats

Ces réseaux ont chacun été optimisés 5 fois avec la méthode décrite dans la partie.VI.2.2 et la fonction de coût et la taille de patch définies dans la partie.VI.3. Les résultats quantitatifs obtenus sont donnés dans la Table.VI.7. Les valeurs du Dice et de la MAD ont été calculés sur le jeu de test après optimisation. La troisième colonne donne le Dice moyen obtenu sur l'ensemble de test par le réseau, parmi les cinq réseaux optimisés, qui avait obtenu le meilleur Dice sur l'ensemble de validation (MV). Les colonnes 4 et 5 donnent le Dice et la MAD qui ont été obtenus en moyenne sur le jeu de test par les 5 réseaux optimisés.

TABLE VI.7 – Résultats quantitatifs obtenus par U-net et V-net avec différents nombres de couches.

Réseau	Couches	Dice (MV)	Dice	MAD (mm)	Temps de segmentation (s)	Nombre de paramètres
U-net	24	0.81 ± 0.062	0.808 ± 0.004	0.51 ± 0.02	10.1 ± 0.3	31 081 410
	19	0.8 ± 0.062	0.795 ± 0.005	0.55 ± 0.02	8.4 ± 0.2	7 739 842
	14	0.761 ± 0.07	0.762 ± 0.005	1.06 ± 0.08	7.6 ± 0.2	1 901 762
	9	0.612 ± 0.1	0.605 ± 0.008	3.62 ± 0.17	6.7 ± 0.3	440 898
V-net	31	0.82 ± 0.055	0.821 ± 0.001	0.51 ± 0.03	4.3 ± 0.2	16 317 042
	23	0.818 ± 0.049	0.819 ± 0.001	0.55 ± 0.02	4.0 ± 0.3	3 891 826
	15	0.795 ± 0.06	0.796 ± 0.001	0.86 ± 0.03	4.1 ± 0.2	783 986
	7	0.524 ± 0.076	0.516 ± 0.006	4.98 ± 0.17	4.0 ± 0.2	117 042

Les p-valeurs obtenues en comparant les résultats obtenus par les réseaux U-net 24 couches et V-net 31 couches avec les mêmes architectures composées de moins de couches sont données dans la Table.VI.8.

TABLE VI.8 – p-valeurs obtenues en comparant U-net et V-net avec des architectures analogues composées de moins de couches.

U-net			V-net		
Couche	p-valeur (Dice)	p-valeur (MAD)	Couche	p-valeur (Dice)	p-valeur (MAD)
24	1	1	31	1	1
19	3.89e-03	7.65e-03	23	9.51e-02	4.03e-02
14	3.57e-07	9.62e-07	15	7.02e-09	1.83e-07
9	5.68e-11	4.34e-10	7	1.08e-13	1.83e-11

Dans le cas 2D, les performances des réseaux ont augmentées jusqu'à l'optimisation d'un U-net 24 couches. Dans ce cas, un Dice moyen de 0.808 ± 0.004 et une MAD moyenne de 0.51 ± 0.02 mm ont été obtenus. Les p-valeurs données dans la Table.VI.8 montrent que ces résultats sont significativement meilleurs que ceux obtenus par les U-net ayant moins de couches. Dans le cas 3D, les performances des réseaux ont augmentées jusqu'à l'optimisation d'un V-net 31 couches. En revanche, les résultats obtenus pour le Dice avec ce réseau ne sont pas significativement meilleurs que ceux obtenus avec un V-net 23 couches ($p = 9.51 \times 10^{-2}$).

Ces résultats suggèrent que le niveau d'abstraction optimal est de 24 couches pour U-net et 31 couches pour V-net, même si ce résultat peut être débattu dans le cas de V-net car la p-valeur est proche de 0.05.

Tous ces réseaux ont permis de segmenter ces volumes 3D en quelques secondes en utilisant des GPUs NVIDIA Tesla V100. Les meilleures performances ont été obtenues par le réseau V-net 23 couches qui a mis 4.0 ± 0.3 secondes en moyenne pour segmenter des volumes d'une taille de $320 \times 320 \times 320$ pixels.

Pour déterminer si une architecture 2D ou 3D était plus adaptée à ce problème, les réseaux ont été comparés au regard de leur précision et de leur nombre de paramètres. Si on compare le U-net 24 couches et le V-net 31 couches qui est composé d'environ deux fois moins de paramètres, le V-net 31 couches est significativement meilleur par rapport au Dice ($p = 2.19 \times 10^{-4}$) mais pas par rapport à la MAD ($p = 8.78 \times 10^{-1}$). De plus, le réseau V-net 23 couches, qui a environ 8 fois moins de paramètres que le réseau U-net 24 couches est aussi significativement meilleur par rapport au Dice ($p = 5.29 \times 10^{-4}$). En revanche, il est significativement moins bon par rapport à la MAD ($p = 1.04 \times 10^{-2}$). D'après ces résultats, on peut dire qu'une architecture V-net (3D) est plus adaptée à ce problème qu'une architecture U-net (2D).

De manière à comprendre cette différence de performances, la Figure.VI.17 montre les résultats qualitatifs obtenus avec les meilleurs réseaux 2D et 3D (meilleur Dice sur l'ensemble de validation) dans le cas des données ayant respectivement obtenues le Dice le plus élevé (Figure.VI.17.a) et le Dice le moins élevé (Figure.VI.17.b).

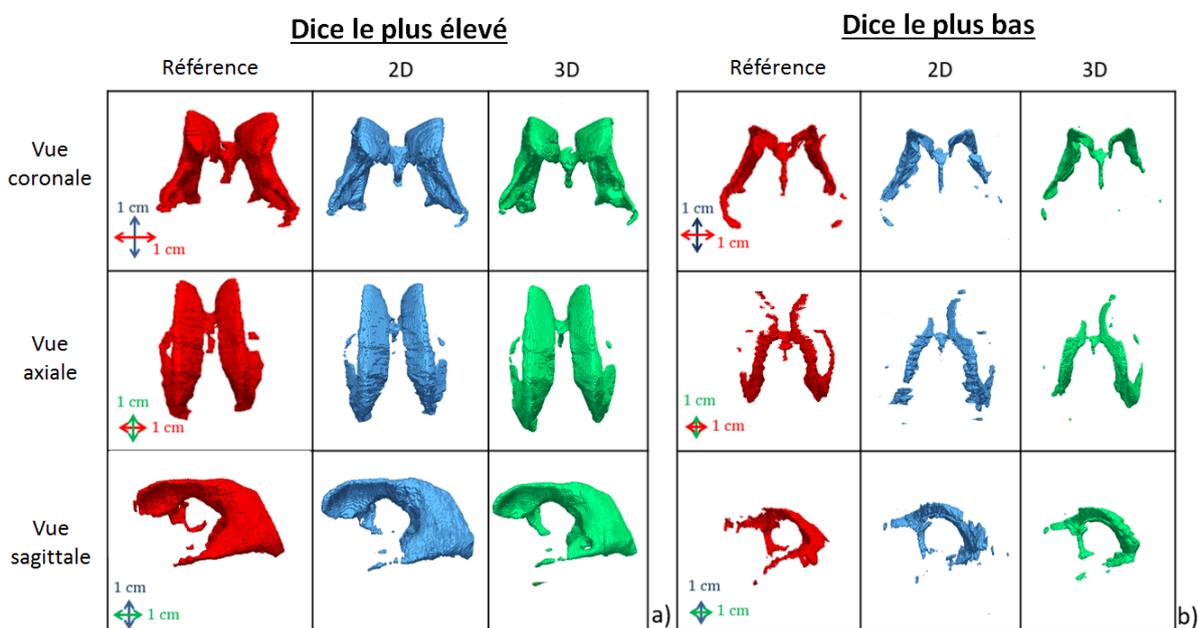


FIGURE VI.17 – Résultats qualitatifs obtenus pour la segmentation du SVC par les réseaux U-net et V-net dans le cas du Dice le plus élevé (a) et le moins élevé (b).

Dans le cas ayant obtenu le Dice le plus élevé, qui correspond à un SVC dilaté, on peut voir que les architectures 2D et 3D produisent toutes les deux de bonnes segmentations. Néanmoins, l'architecture 3D a été légèrement meilleure pour segmenter le troisième ventricule, même si les deux architectures ont échoué à segmenter sa partie la plus fine. Dans le cas ayant obtenu le Dice le moins élevé, qui correspond à un SVC normal, le même comportement est observé. De manière

générale, les performances obtenues par ces deux réseaux sur les autres parties du SVC semblent équivalentes. Néanmoins on observe une différence majeure au niveau de la régularité des contours qui sont beaucoup plus lisses dans le cas des segmentations effectuées avec le V-net.

VI.4.3 Volumes mesurés après segmentation du SVC par V-net.

La paramètre que l'on cherche à estimer en clinique pour évaluer la gravité d'une dilatation ventriculaire cérébrale et suivre son évolution est le VV. Pour avoir une idée de l'erreur volumique commise lors des segmentations automatiques, le jeu de test a été segmenté par le réseau V-net 31 couches qui avait obtenu le meilleur Dice sur le jeu de validation. Les erreurs volumique absolues (ΔV_a) et relatives (ΔV_r) obtenues en moyenne sur le jeu de test sont données dans la Table.VI.9 dans le cas des SVC normaux et dilatés.

TABLE VI.9 – Erreurs volumiques commises par V-net lors de la segmentation des SVC normaux et dilatés.

Etat du SVC	Nombre de patients	VV (cm ³)	ΔV_a (cm ³)	ΔV_r (%)
Normal	5	2.54 ± 0.54	0.36 ± 0.3	18.7 ± 16.7
Dilaté	2	9.39 ± 1.07	0.34 ± 0.31	4.4 ± 4.1

Dans le cas des SVC normaux, le VV moyen des segmentations de référence était de 2.54 ± 0.54 cm³. L'erreur volumique absolue moyenne commise est de 0.36 ± 0.3 cm³, ce qui est faible. En revanche, l'erreur volumique relative (%) est importante et avec un fort écart-type : 18.7 ± 16.7 %. Ce dernier résultat s'explique par le faible volume des ventricules normaux et par la difficulté que peut rencontrer le réseau à segmenter certaines structures fines dans ce cas : une partie non segmentée du SVC augmentera fortement ΔV_r . Néanmoins, cette erreur n'aurait pas un impact clinique important car les médecins cherchent principalement à mesurer le VV des SVC dilatés.

Dans le cas des SVC dilatés, le VV moyen des segmentations de référence était de 9.39 ± 1.07 cm³. L'erreur volumique absolue commise est faible et proche de celle obtenue dans le cas des SVC normaux ($\Delta V_a = 0.34 \pm 0.31$ cm³). L'erreur volumique relative est aussi faible 4.4 ± 4.1 %, ce qui montre que dans le cas où le SVC est dilaté le réseau est capable de produire une segmentation permettant de mesurer précisément le VV. Ce résultat devra néanmoins être vérifié sur plus de cas et des études cliniques devront être déterminées si l'erreur commise est admissible ou non.

VI.4.4 Conclusion

Dans cette partie, U-net et V-net ont été optimisés en faisant varier leur nombre de couches. Cette expérience avait pour objectifs de déterminer le niveau d'abstraction optimal de ces deux structures et de savoir si l'une d'elles était plus adaptée au problème de segmentation du SVC.

Les résultats obtenus ont conclu que le niveau d'abstraction optimal pour U-net et V-net dans notre configuration étaient respectivement de 24 et 31 couches. Le meilleur réseau a permis d'obtenir un excellent Dice de 0.82 ± 0.055 sur le jeu de test. Ces segmentations ont été obtenues en un temps moyen de 4.1 ± 0.2 s qui est compatible avec les contraintes cliniques.

VI.5 Apprentissage de la localisation du SVC

Dans cette partie, une modification des architectures U-net et V-net est proposée pour permettre l'apprentissage de la position anatomique du SVC. Cette modification a été réalisée de telle sorte que cette apprentissage se fasse en même temps que l'optimisation de ces CNNs.

VI.5.1 Motivations

L'anatomie du cerveau est sujette à des variations interindividuelles mais les structures cérébrales et leur localisation sont bien définies. Par exemple, on sait que le SVC se situe au centre du cerveau, comme les thalami. Les CNNs n'apprennent pas la position absolue d'une structure dans une image mais sont capables d'extraire des caractéristiques indépendamment de leur position dans une image grâce aux bonnes propriétés des noyaux de convolution. Alors qu'il est facile pour un observateur humain de différencier des structures d'aspect similaire, comme les plexus choroïde et la boîte crânienne, en se basant sur leur position au sein des volumes, les CNNs n'utilisent que le contexte autour d'un pixel pour lui attribuer une étiquette. Néanmoins, donner accès au réseau au positionnement spatial à grande échelle peut demander beaucoup de ressources : augmentation du nombre de couches ou noyau de convolution d'une taille plus importante. Ainsi, il serait intéressant de déterminer une méthode pour que les CNNs puissent apprendre la position anatomique des structures cérébrales, de manière à réduire les besoins en contexte des CNNs. Pour répondre à cet objectif, nous proposons d'utiliser un réseau générateur de compositions de formes.

Réseau générateur de formes par compositions.

Les réseaux générateurs de formes par compositions (CPPN) ont été introduit par [Stanley \(2007\)](#) dans le domaine de la génétique dans le but de modéliser les changements de phénotype. Ce type de réseau utilise des sommes et des compositions de fonctions mathématiques basiques qui permettent de produire des formes géométriques à partir d'entrées comme des cartes de coordonnées. En combinant ces formes avec les cartes de caractéristiques de bas niveau produites par U-net et V-net, il est possible d'apporter une information sur la position absolue de ces caractéristiques dans le cerveau. Cette approche a l'avantage de permettre l'apprentissage de la position anatomique directement à partir des données d'entrée lors de l'optimisation d'un CNN. En d'autres termes, les CPPNs permettent aux CNNs d'apprendre leur propre dictionnaire de formes. L'objectif de ces formes est d'aider le réseau à déterminer où se situe le CVS. Par exemple, certaines formes pourraient correspondre à des régions où se situent le SVC et d'autres à des régions où le SVC ne se trouve pas. Cependant, un encodage plus complexe de la position du SVC pourrait être appris.

VI.5.2 Architecture du CPPN et intégration à U-net et V-net

Architecture du CPPN

Le CPPN que nous avons défini prend des cartes de coordonnées normalisées, notées X_c , Y_c et Z_c , en entrée. Les coordonnées normalisées (x_n, y_n, z_n) sont définies (Équation.VI.12) de telles sortes que pour un point de coordonnées (x, y, z) :

$$(x_n, y_n, z_n) = (X_c(x, y, z), Y_c(x, y, z), Z_c(x, y, z)) \quad (\text{VI.12})$$

Où $x_n, z_n \in [-1, 1]$ et $y_n \in [0, 1]$. Les coordonnées normalisées de y_n sont comprises entre 0 et 1 de manière à prendre en compte la symétrie entre les deux hémisphères du cerveau. Cela permet en pratique de générer des formes symétriques par rapport au plan sagittal. Ces données d'entrées sont montrées sur la Figure.VI.18.a à une certaine profondeur z_n . L'architecture du CPPN est décrite sur la Figure.VI.18.b.

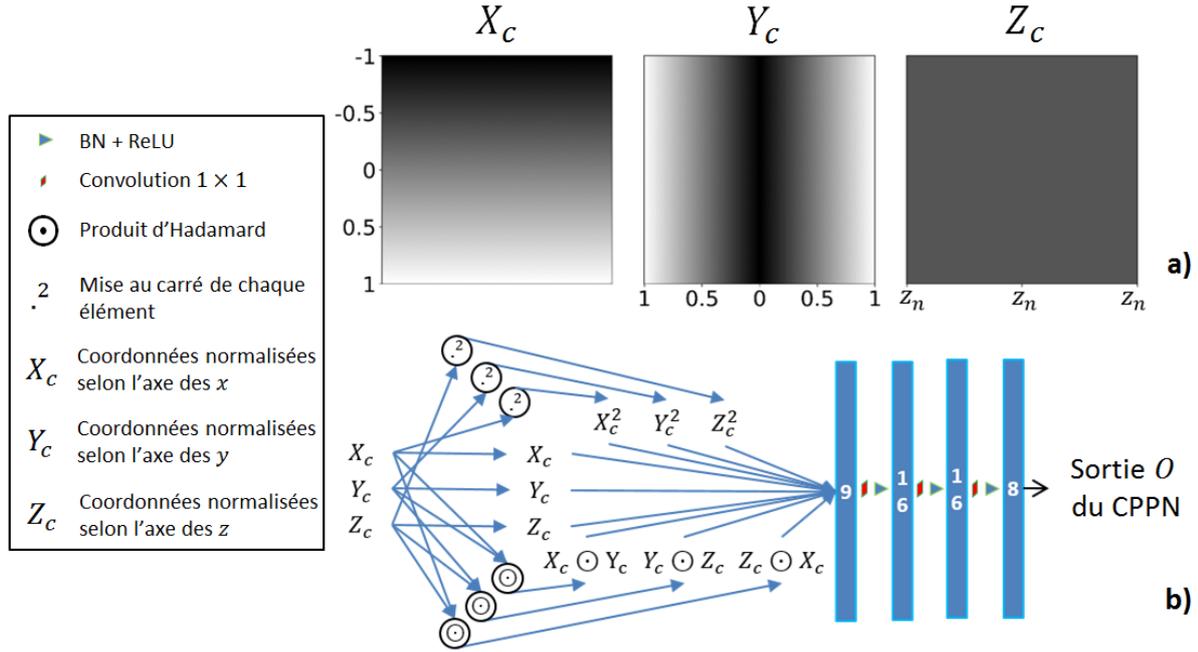


FIGURE VI.18 – Données d'entrées à une profondeur z_n fixée (a) et architecture du CPPN (b).

La concaténation des cartes de coordonnées normalisées, leur carré et leur produit terme à terme (produit d'Hadamard) a été utilisée comme entrée du CPPN. Ce dernier est composé de trois couches convolutives qui sont chacune composées d'une opération de convolution $1 \times 1 (\times 1)$, d'une normalisation du batch et de la fonction d'activation ReLU. Dans la première couche convolutive, l'opération de convolution permet de générer des quadriques (ellipsoïde, hyperboloïde, ...). Par exemple, la i -ème sortie de la première couche $O_{1,i}$, avant la BN et la fonction ReLU, est donnée par l'Équation.VI.13.

$$O_{1,i} = w_{i,1}X_c^2 + w_{i,2}Y_c^2 + w_{i,3}Z_c^2 + w_{i,4}X_c \odot Y_c + w_{i,5}X_c \odot Z_c + w_{i,6}Y_c \odot Z_c + w_{i,7}X_c + w_{i,8}Y_c + w_{i,9}Z_c + b_i \quad (\text{VI.13})$$

Cette équation correspond exactement à une équation de quadrique où $w_{i,k}$ et b_i sont le k -ème poids et le biais du i -ème neurones de cette première couche. Les sorties $O_{1,i}$ sont ensuite combinées les une avec les autres pour créer des formes plus complexes qui correspondent à des combinaisons non-linéaires de quadrique. Cette architecture a l'avantage d'utiliser uniquement des couches convolutives pour générer des formes, ce qui permet de conserver les propriétés des réseaux complètement convolutifs (Long *et al.* (2015)).

Intégration du CPPN à U-net et à V-net

Pour permettre à U-net et V-net d'être entraînés en même temps que le CPPN, les formes sont apprises en parallèle des cartes de caractéristiques de bas niveau avant d'être concaténées au réseau principal (Figure.VI.19). Cela a pour conséquence de permettre l'extraction des cartes de caractéristiques de bas niveau en étant invariant par translation.

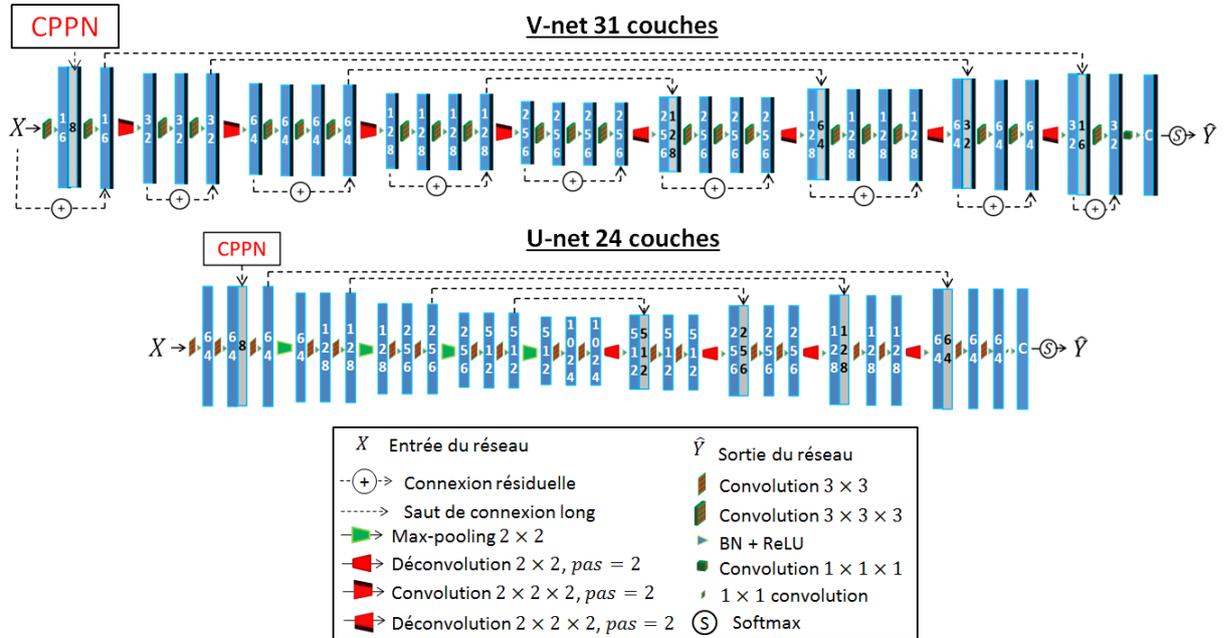


FIGURE VI.19 – Intégration du CPPN à U-net et V-net.

VI.5.3 Expériences et résultats

Tous les réseaux étudiés dans la partie.VI.4 ont été ré-entraînés avec le CPPN pour évaluer l'impact de celui-ci. Les 8 canaux de sortie du CPPN sont visibles à la profondeur $z = 80$ et $z = 160$ (Figure.VI.20.a) sur la Figure.VI.20.b et Figure.VI.20.c. Les données X mises en entrée de U-net sont également visibles.

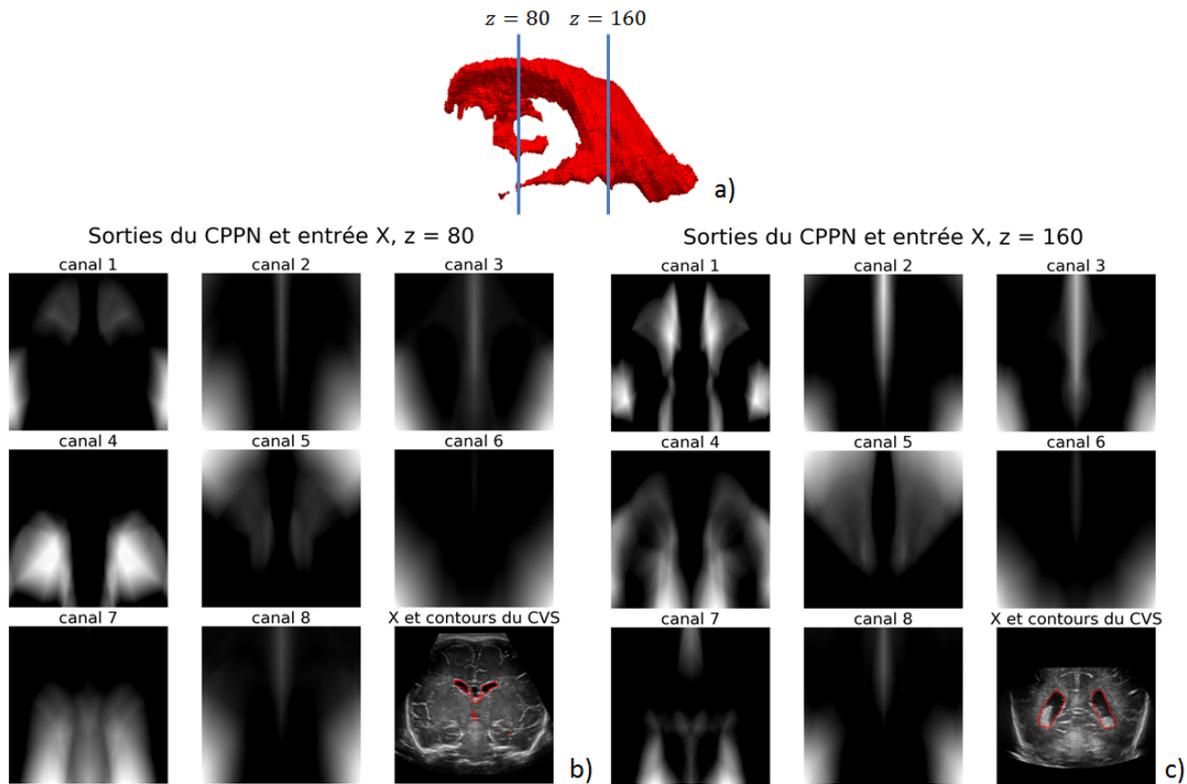


FIGURE VI.20 – Images des formes produites par le CPPN du réseau ayant obtenu le meilleur Dice sur l'ensemble de validation.

On voit que certaines formes, comme celles du canal 5, recouvrent la zone où se situe le SVC, ce qui est intéressant car cela pourrait aider à le localiser. Les autres formes semblent partitionner l'image en plusieurs régions, ce qui pourrait permettre de les utiliser pour déterminer la position des caractéristiques dans les images. Les résultats quantitatifs obtenus après les entraînements de U-net et V-net avec CPPN sont donnés dans la Table VI.10.

TABLE VI.10 – Résultats quantitatifs obtenus en utilisant un CPPN avec U-net et V-net pour différents nombres de couches.

Réseau	Couches	Dice (MV)	Dice	MAD (mm)	Temps de segmentation (s)	Nombre de paramètres
U-net	24	0.817 ± 0.05	0.812 ± 0.003	0.51 ± 0.01	10.1 ± 0.5	31 086 666
	19	0.808 ± 0.055	0.807 ± 0.001	0.5 ± 0.02	8.8 ± 0.2	7 745 098
	14	0.804 ± 0.056	0.801 ± 0.002	0.54 ± 0.02	7.9 ± 0.2	1 907 018
	9	0.765 ± 0.074	0.757 ± 0.008	0.82 ± 0.06	6.2 ± 0.4	446 154
V-net	31	0.824 ± 0.046	0.823 ± 0.001	0.53 ± 0.03	3.7 ± 0.2	16 321 146
	23	0.82 ± 0.05	0.82 ± 0.002	0.66 ± 0.2	4.4 ± 0.3	3 895 930
	15	0.811 ± 0.059	0.807 ± 0.005	0.64 ± 0.02	4.0 ± 0.2	788 090
	7	0.724 ± 0.107	0.714 ± 0.005	1.31 ± 0.04	4.3 ± 0.2	121 146

On peut voir que le meilleur Dice a été obtenu par le V-net 31 couches et la meilleure MAD par le U-net 19 couches. Dans le cas des réseaux ayant peu de couches, les performances obtenues sont bien meilleures quand le CPPN est intégré au réseau. Les p-valeurs calculées à partir des résultats obtenus par chaque réseau avec et sans CPPN sont données dans la Table.VI.11.

TABLE VI.11 – p-valeurs comparant les performances de chaque réseau avec et sans CPPN.

U-net			V-net		
layers	p-valeur (Dice)	p-valeur (MAD)	layers	p-valeur (Dice)	p-valeur (MAD)
24	1.53e-01	8.69e-01	31	3.31e-02	3.74e-01
19	1.54e-03	2.91e-03	23	4.95e-01	3.04e-01
14	3.17e-07	1.60e-06	15	1.44e-03	1.24e-06
9	3.39e-09	1.42e-09	7	3.05e-11	9.99e-11

Dans le cas 2D, le CPPN a permis d'améliorer significativement les performances obtenues par les U-net 9 couches, 14 couches et 19 couches, au regard du Dice et de la MAD.

Dans le cas 3D, le CPPN a permis d'améliorer significativement le Dice obtenu par les V-net 7 couches, 15 couches et 31 couches. Une différence significative, en faveur de l'utilisation du CPPN, a également été obtenue pour la MAD dans le cas des V-net 7 couches et 15 couches. On peut dire que les performances obtenues avec le CPPN sont meilleures que celles obtenues sans CPPN lorsque que le nombre de couches des réseaux est faible et que ces performances deviennent semblables lorsque que le nombre de couches dépassent un certain seuil. Ce résultat s'observe plus aisément sur la Figure.VI.21 où les évolutions du Dice et de la MAD en fonction du nombre de couches sont tracées pour U-net (Figure.VI.21.a) et V-net (Figure.VI.21.b) avec et sans CPPN.

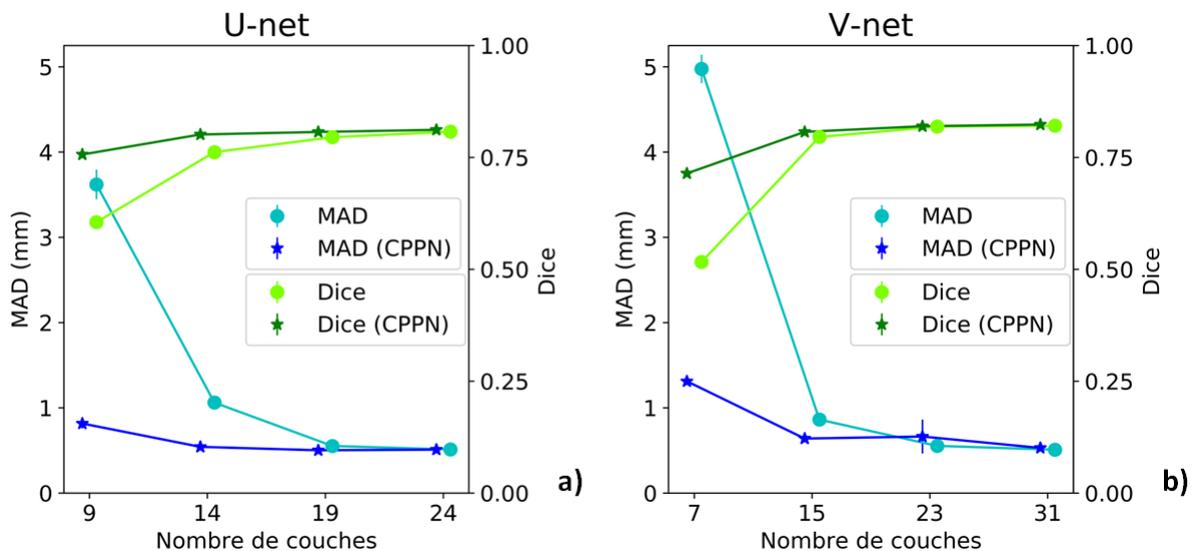


FIGURE VI.21 – Évolution du Dice et de la MAD en fonction du nombre de couches pour U-net (a) et V-net (b) avec et sans CPPN.

Des résultats qualitatifs obtenus dans le cas des réseaux U-net 24 couches et V-net 31 couches avec et sans CPPN sont visibles sur la Figure.VI.22.a. Ces résultats correspondent aux patients ayant obtenus le Dice le plus élevé pour le réseau qui a obtenu le meilleur Dice sur l'ensemble de validation. Des résultats analogues sont présentés dans le cas des U-net 9 couches et V-net 7 couches (Figure.VI.22.b).

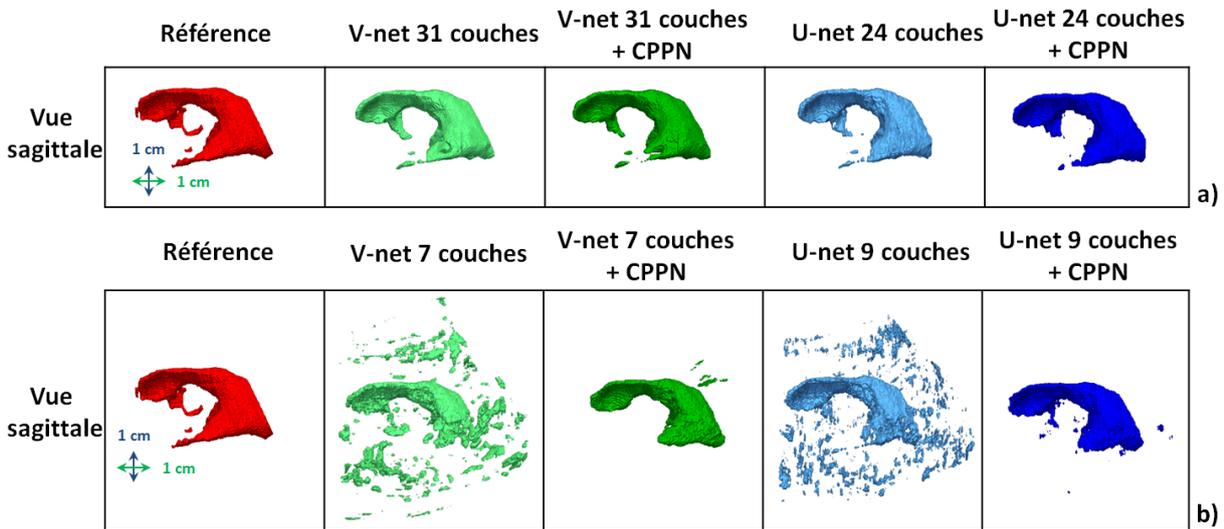


FIGURE VI.22 – Évolution du Dice et de la MAD en fonction du nombre de couches pour U-net (a) et V-net (b) avec et sans CPPN.

Ces images montrent qu'il n'y a pas de différences significatives entre les segmentations réalisées avec ou sans CPPN dans le cas des réseaux U-net 24 couches et V-net 31 couches. En revanche, on constate une réduction importante des faux positifs, pour les réseaux U-net 9 couches et V-net 7 couches lorsque le CPPN est utilisé.

VI.5.4 Discussion

Le CPPN a été utilisé pour générer des formes (Figure.VI.20) qui pourrait fournir des informations sur la position anatomique du SVC.

Une diminution du nombre de faux-positifs correspondant à des structures semblables à certaine partie du SVC était attendue. Cette hypothèse a été validée par les résultats visibles sur la Figure.VI.22.b où de nombreux faux-positifs sont évités lorsque le CPPN est utilisé. Des exemples de faux-positifs obtenus avec le réseau V-net à 7 couches sont visibles sur la Figure.VI.23.b, celle-ci montre que la boîte crânienne et le CSP, qui ressemblent respectivement aux plexus choroïde et au LCS ont été segmentés par le réseau. Dans le cas où le CPPN est utilisé avec ce même V-net, la boîte crânienne n'est plus segmentée et seul un morceau du bord du CSP l'est, ce dernier étant une des parties du CSP la plus proche du CVS.

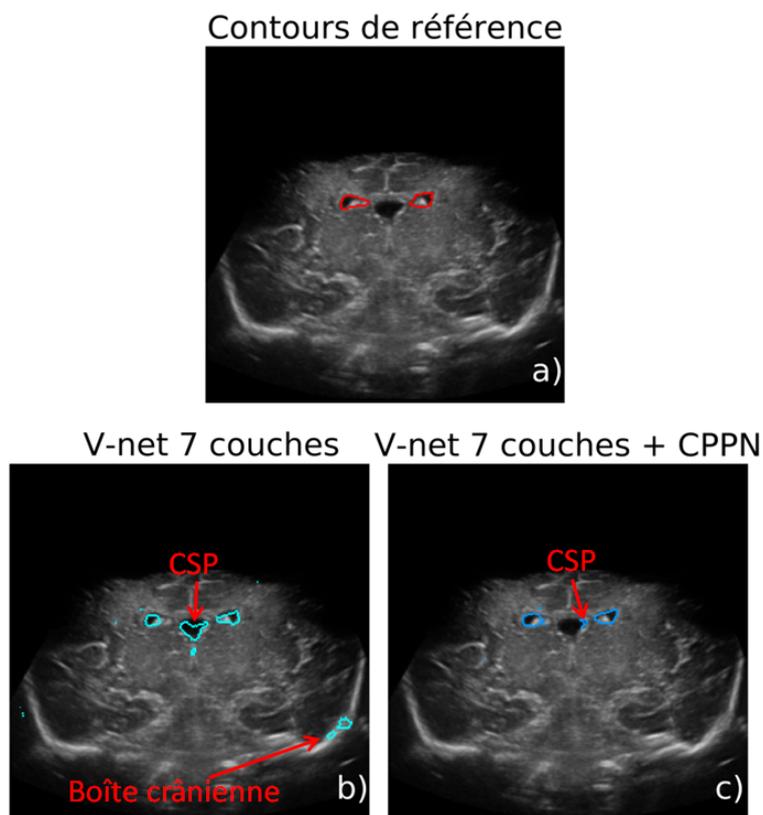


FIGURE VI.23 – Contours de référence (a) et segmentation produit par un V-net 7 couches sans (b) et avec CPPN (c).

VI.5.5 Conclusion

Cette partie a présenté comment les architectures de U-net et V-net ont été modifiées, par l'ajout d'un CPPN, de manière à être en mesure d'apprendre la position anatomique du SVC à partir de cartes de coordonnées. Ces formes ont permis d'améliorer significativement les performances de U-net et V-net quand ces architectures avaient peu de couches et n'étaient donc pas en mesure d'utiliser beaucoup de contexte. Notamment en réduisant le nombre de faux-positifs qui ressemblaient à des parties du SVC mais qui n'étaient pas situés dans la même partie du cerveau que celui-ci.

Nous avons choisi d'utiliser des cartes de coordonnées en entrée du CPPN de manière à créer des combinaison de quadrique. Néanmoins, une étude plus approfondie pourrait être réalisée pour déterminer quelles entrées et fonctions mathématiques seraient les plus adaptées pour générer des formes qui encodent la position anatomique du SVC.

VI.6 Taille adéquate du jeu d'entraînement et augmentation artificielle de données

VI.6.1 Taille adéquate de l'ensemble d'entraînement

L'objectif de cette partie est de répondre aux questions suivantes :

1. Combien de volumes d'entraînement sont nécessaires pour que U-net (24 couches) obtienne une bonne précision ($Dice > 0.8$) sur l'ensemble d'entraînement.
2. Quelle est la meilleure stratégie pour améliorer la précision des segmentations : augmenter le nombre de volumes d'entraînement ou définir un réseau avec une architecture plus adaptée au problème.

Méthode

Créer une base de données annotées nécessite beaucoup d'efforts et de temps, il est donc important d'avoir une idée du nombre de volumes d'entraînement nécessaire pour obtenir des segmentations précises et de savoir quelle stratégie utiliser pour améliorer la performance des réseaux. Pour évaluer cela, nous avons fixé un ensemble de validation et un ensemble de test (définis dans la partie.VI.1.2) puis nous avons entraîné un réseau U-net (24 couches) en faisant varier le nombre de volumes dans la base d'entraînement. Lors de chaque entraînement, ce nombre a été fixé à 1, 3, 5, 7, 9, 11 ou 13 (base d'entraînement complète). Pour chacune de ces tailles, les volumes ont été tirés de manière aléatoire dans le jeu d'entraînement, dans chaque cas, l'expérience a été reproduite 10 fois. Les paramètres utilisés pour optimiser U-net ont été décrits dans la partie.VI.3.

Résultats et discussion

L'évolution du Dice moyen (sur les 10 optimisations) en fonction du nombre de volumes dans la base d'entraînement est montré sur la Figure.VI.24.

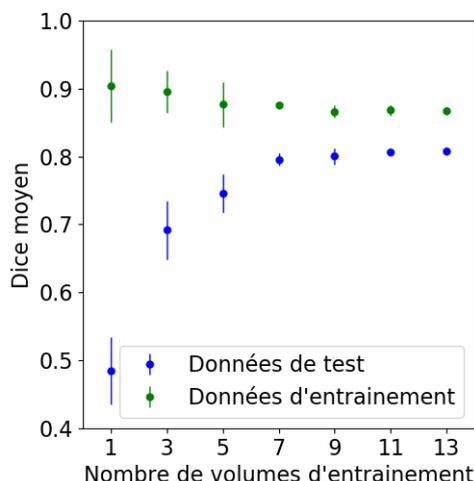


FIGURE VI.24 – Évolution du Dice moyen obtenu sur les jeux d'entraînement et de test en fonction de la taille du jeu d'entraînement.

D'après ces courbes, les Dice moyens, calculés sur l'ensemble d'entraînement et l'ensemble de test, diminuent et augmentent respectivement en fonction du nombre de volumes dans le jeu d'entraînement. Ce comportement est celui-ci qui est attendu. Il s'explique par le fait que, plus il y a de volumes dans le jeu d'entraînement, plus celui-ci contient de variabilité, ce qui le rend plus difficile à apprendre mais permet également une meilleure généralisation au jeu de test.

Une fois que la taille du jeu d'entraînement a atteint 7 volumes, on voit que les Dice moyens calculés sur les jeux d'entraînement et de test évoluent peu et avec un écart-type très faible.

Les valeurs quantitatives obtenues lors de cette expérience sont données dans la Table.VI.12. Les résultats obtenus avec un jeu d'entraînement contenant 13 volumes ont été utilisés comme référence. Des p-valeurs ont ensuite été calculées entre les résultats de référence pour 13 volumes et les résultats obtenus avec un jeu d'entraînement d'une taille plus faible. Les valeurs obtenues sont également données dans la Table.VI.12.

TABLE VI.12 – Dice moyen obtenus sur les jeux d'entraînement et de test en fonction du nombre de volume dans le jeu d'entraînement.

Nombre de volumes d'entraînement	Dice moyen (entraînement)	Dice moyen (test)	p-valeur
1	0.904 ± 0.054	0.485 ± 0.049	1.22e-13
3	0.895 ± 0.031	0.692 ± 0.043	1.97e-07
5	0.877 ± 0.033	0.746 ± 0.028	4.81e-06
7	0.875 ± 0.006	0.796 ± 0.009	2.08e-03
9	0.866 ± 0.009	0.8 ± 0.012	9.10e-02
11	0.868 ± 0.008	0.807 ± 0.004	6.85e-01
13	0.867 ± 0.006	0.808 ± 0.005	1

On constate qu'une bonne précision est obtenue avec seulement 9 patients dans l'ensemble d'entraînement ($Dice = 0.8 \pm 0.012$). Ainsi, construire une base de données d'entraînement qui permette d'obtenir des segmentations de bonne qualité reste coûteux en temps (9 volumes d'entraînement + 5 de validation) mais réaliste.

En ce qui concerne les p-valeurs, on voit qu'il existe une différence significative entre la référence et les résultats obtenus avec 1, 3, 5 et 7 patients. En revanche, une différence significative n'a pas été obtenue avec 9 volumes ($p = 0.091$) ni avec 11 volumes ($p = 0.685$). Pour 9 volumes d'entraînement, la p-valeur obtenue est néanmoins proche de 0.05. Ces résultats suggèrent qu'il faudrait ajouter beaucoup de patients ou des patients atypiques à la base de données avant d'améliorer la précision de U-net. Ce qui risquerait d'être coûteux en temps et en énergie. Une meilleure solution serait donc de définir une architecture plus adaptée au problème. Ce résultat a été confirmé en utilisant le réseau V-net dans la partie précédente.

Conclusion

Dans cette partie, les performances de U-net pour segmenter le SVC ont été évaluées en faisant varier le nombre de volumes dans le jeu d'entraînement. Cette expérience avait pour objectif d'estimer le nombre de volumes nécessaires pour obtenir des segmentations précises et de déterminer qu'elle serait la meilleure stratégie à adopter pour améliorer la précision des segmentations du SVC.

Les résultats ont montré qu'entraîner U-net avec un jeu d'entraînement de 9 volumes et un jeu de validation de 5 volumes permettait d'obtenir des segmentations précises ($Dice \geq 0.8$) sur le jeu de test. De plus, cette expérience a montré qu'il serait préférable d'utiliser une architecture plus adaptée à ce problème plutôt que d'augmenter le nombre de volumes dans le jeu d'entraînement.

Il serait néanmoins intéressant d'utiliser des stratégies d'augmentation artificielle de données pour enrichir la base de données d'entraînement à moindre coût. Cette solution sera étudiée dans la partie suivante.

VI.6.2 Augmentation artificielle de données

Cette partie présente les stratégies d'augmentation artificielle de données que nous avons utilisées dans le but d'améliorer les performances de U-net (24 couches) et de V-net (31 couches) pour la segmentation du SVC.

Intérêt de l'augmentation de données artificielle

Comme nous l'avons vu dans la partie précédente, il risquerait d'être très coûteux d'augmenter la taille du jeu d'entraînement afin d'améliorer les performances de U-net de manière significative. Ainsi, plutôt que d'annoter manuellement de nouveaux volumes, nous avons cherché à déterminer des stratégies d'augmentation artificielle de données qui seraient adaptées à ce problème.

L'augmentation artificielle de données consiste à enrichir le jeu d'entraînement en appliquant des transformations aux données (X, Y) qu'il contient : rotations, symétries, Cette méthode a pour objectif de permettre l'apprentissage d'une distribution $P(Y/X)$ sur le jeu d'entraînement qui soit plus représentative de celle du jeu de test. Les transformations utilisées doivent donc permettre d'augmenter la variabilité contenu dans l'ensemble d'entraînement tout en générant de nouvelles données qui soient réalistes. Cette méthode permet d'enrichir le jeu d'entraînement à moindre coût, il faut néanmoins déterminer les transformations qui sont adaptées au problème que l'on souhaite résoudre.

Méthodes

Pour augmenter artificiellement la variabilité contenu dans le jeu d'entraînement, différentes transformations ont été appliquées aux données échographiques et à leurs labels associés :

- Des symétries horizontales (par rapport au plan sagittale).
- Des rotations d'un angle compris entre -10 et 10 degrés.
- Des zooms d'un facteur compris entre 1 et 3.
- L'ajout d'un bruit blanc gaussien ($\sigma^2 = 0.0625$).

- Des transformations affines des coordonnées des pixels, définies par une matrice d'Équation.VI.14 :

$$\begin{pmatrix} 1 + m_{00} & m_{01} \\ m_{10} & 1 + m_{11} \end{pmatrix} \quad (\text{VI.14})$$

avec $m_{00}, m_{01}, m_{10}, m_{11} \in [-0.15, 0.15]$

- Des transformations non-linéaire des images par un champ de déplacement aléatoire Gaussien.

Un exemple de patch, avant et après transformations, est visible sur la Figure.VI.25.

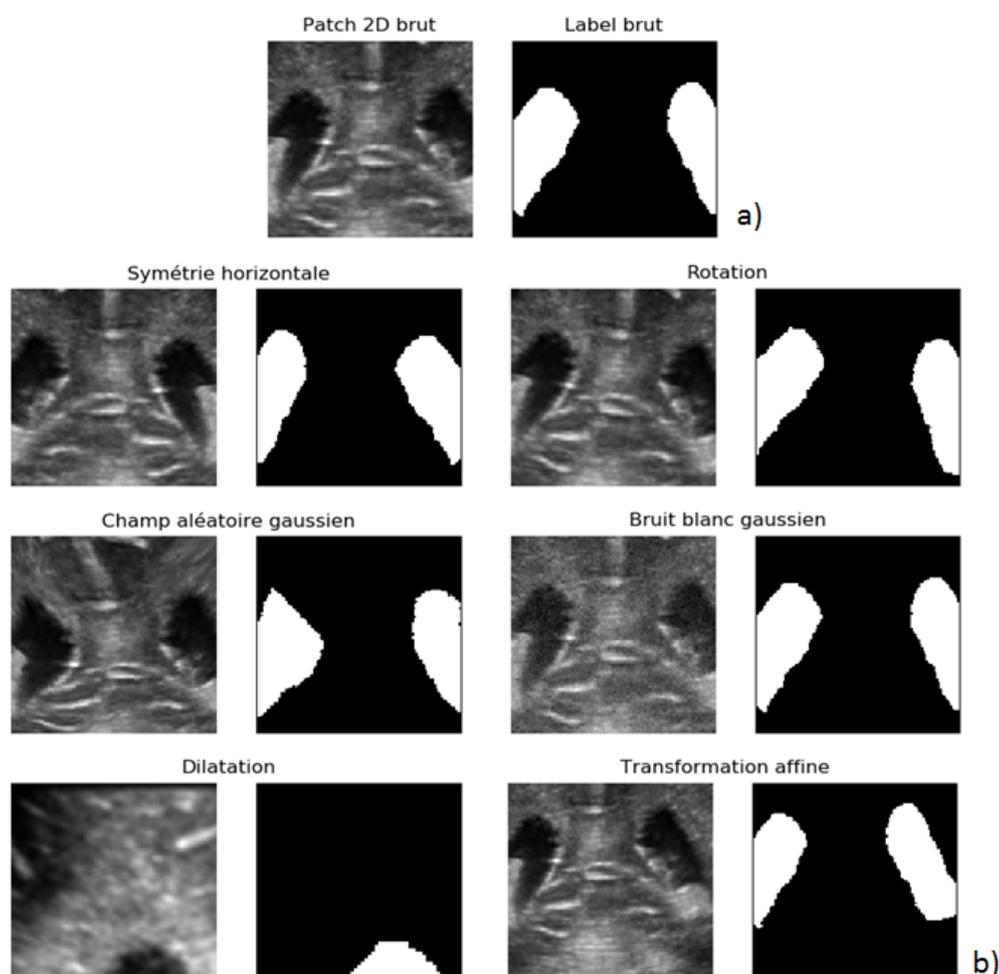


FIGURE VI.25 – Patch avant (a) et après application d'une transformation utilisée pour faire de l'augmentation artificielle de données (b).

Pour déterminer si ces transformations permettaient d'améliorer significativement les performances de U-net, celui-ci a été entraîné 10 fois pour chacune d'elles. Lors du tirage aléatoire d'un patch pendant la phase d'entraînement, les transformations ont été réalisées avec une probabilité de 0.2.

Résultats

Les résultats obtenus après avoir optimisé U-net pour chaque type d'augmentation de données sont visibles dans la Table.VI.25. Les valeurs du Dice et de la MAD sont celles obtenues en moyenne sur le jeu de test (pour les 10 optimisations).

TABLE VI.13 – Précisions des segmentations après entraînement de U-net avec augmentation de données.

Type daugmentation	Dice	p-valeur	MAD	p-valeur
Sans augmentation	0.809 ± 0.003	1.00e+00	0.52 ± 0.03	1
Symétrie horizontale	0.817 ± 0.003	3.42e-05	0.5 ± 0.02	1.41e-01
Transformation affine	0.811 ± 0.006	5.06e-01	0.56 ± 0.06	1.00e-01
Champ aléatoire gaussien	0.811 ± 0.005	3.52e-01	0.54 ± 0.04	1.71e-01
Bruit blanc gaussien	0.806 ± 0.003	1.12e-01	0.53 ± 0.03	6.15e-01
Rotation	0.811 ± 0.005	3.62e-01	0.54 ± 0.04	3.17e-01
Dilatation	0.805 ± 0.009	2.06e-01	0.6 ± 0.07	7.72e-03

Parmi les transformations utilisées, seule la symétrie horizontale a permis d'améliorer significativement le Dice : $p = 3.42 \times 10^{-5} < 0.05$. La MAD est, en moyenne, meilleure lorsqu'on utilise ce type de transformation mais l'écart n'est pas significatif : $p = 1.41 \times 10^{-1}$. Ce résultat paraît cohérent car la réalisation d'une symétrie horizontale permet d'inverser les hémisphères droit et gauche du cerveau et ne modifie pas la statistique des images (pas d'interpolation nécessaire pour obtenir les nouvelles images). Cette transformation permet donc d'augmenter artificiellement la variabilité contenue dans le jeu d'entraînement en créant de nouvelles données réalistes.

La réalisation de dilatation a réduit significativement la MAD ($p = 4.65e - 03 < 0.05$). Ce résultat peut paraître surprenant car la taille du SVC peut beaucoup varier d'un individu à l'autre, on aurait donc pu imaginer que cette transformation permettrait de simuler cette variabilité. Ce résultat peut s'expliquer par le fait que les données obtenues après transformations n'étaient pas réalistes : statistique du speckle différente, forme du SVC non réaliste, ...

Les autres transformations n'ont pas permis d'améliorer significativement les performances de U-net. Un meilleur Dice a néanmoins été obtenu lors de l'application aux images : de rotations, de transformations affines ou de champs de déplacements aléatoires gaussiens.

Entraînement de V-net avec augmentation artificielle de données.

Dans la section précédente, nous avons vu que les symétries horizontales avaient permis d'améliorer significativement les performances de U-net. Cette transformation a également été utilisée lors de l'entraînement du réseau V-net pour savoir si ce résultat était aussi valable pour cette architecture.

Pour cette expérience, V-net a été optimisé 5 fois avec et sans la réalisation de symétries horizontales sur les données d'entrées, le processus d'optimisation utilisé était le même que pour le réseau

U-net. Les résultats obtenus sur le jeu de test, en moyenne (5 optimisations) et dans le cas du réseau ayant obtenu le meilleur Dice sur le jeu de validation (MV), sont donnés dans la Table.VI.14.

TABLE VI.14 – Précisions des segmentations après entraînement de V-net avec augmentation de données.

Augmentation	Dice (MV)	Dice	p-valeur	MAD (mm) (MV)	MAD (mm)	p-valeur
Sans Augmentation	0.82 ± 0.055	0.821 ± 0.001	1	0.5 ± 0.13	0.51 ± 0.03	1
Symétrie horizontale	0.828 ± 0.044	0.829 ± 0.002	$8.11e-05$	0.49 ± 0.13	0.49 ± 0.03	4.58e-01

Comme pour le réseau U-net, on constate que réaliser des symétries horizontales sur les données d'entraînement permet d'améliorer significativement le Dice ($p = 8.11 \times 10^{-5}$) mais pas la MAD ($p = 0.458$).

Conclusion

Dans cette partie, différentes stratégies d'augmentation artificielle de données ont été mises en œuvre pour améliorer la précision des segmentations du SVC obtenues avec U-net et V-net. Dans le cas de U-net, les stratégies d'augmentation de données qui ont été utilisées sont : les symétries horizontales, les rotations, les dilatations, les transformations affines, l'ajout d'un bruit blanc gaussien et les transformations par un champ de déplacement aléatoire gaussien. Les résultats obtenus ont permis de conclure que seule la réalisation de symétries horizontales permettait d'améliorer significativement les performances de U-net. Ce dernier résultat a également été confirmé avec V-net.

VI.7 Segmentation de données atypiques

L'objectif de cette partie est d'évaluer la précision des réseaux U-net et V-net, avec et sans CPPN, pour la segmentation de données atypiques. On entend par données atypiques des cas qui n'ont pas été vus dans le jeu d'entraînement.

VI.7.1 Présentation des données atypiques

En plus de la base de données annotées du SVC, nous disposons de deux cas atypiques :

- un patient atteint d'hydrocéphalie dont le VV était de 46.73 cm^3 . En comparaison, le VV des SVC dilatés était de $9.4 \pm 0.6 \text{ cm}^3$ dans le jeu d'entraînement.
- un volume échographique reconstruit à partir d'une acquisition réalisée avec une autre sonde : Philips PureWave eL18-4. Le patient associé présentait de plus une dilatation atypique au niveau de la corne occipitale droite.

La Figure.VI.26 montre des données échographiques et leurs segmentations 3D du SVC dans un cas normal (Figure.VI.26.a), un cas dilaté (Figure.VI.26.b) et le cas atypique où le SVC est extrêmement dilaté (Figure.VI.26.c).

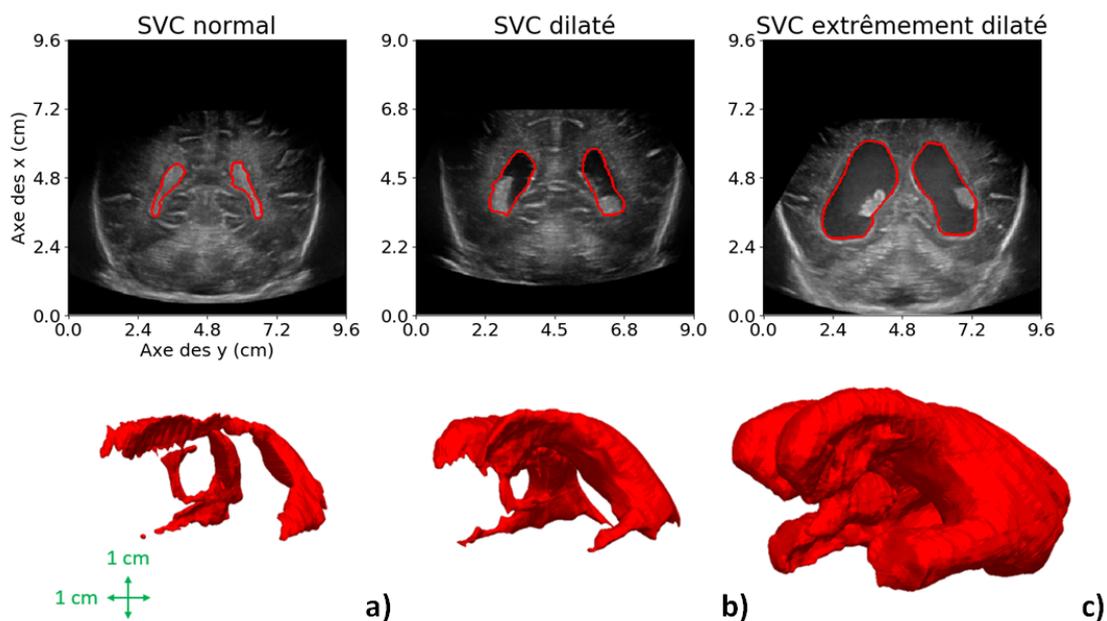


FIGURE VI.26 – Données échographiques en coupe coronale et segmentations 3D associées de référence du SVC dans un cas où le SVC est normal (a), dilaté (b) ou extrêmement dilaté (c).

Les images US 2D et les segmentations 3D montrent clairement la dilatation anormale du SVC. Sur les images échographiques, cela se traduit par une région hypoéchogène de taille très importante par rapport au cas où le SVC est normal ou dilaté.

Sur la Figure.VI.27.b, on peut voir les données échographiques acquises avec la sonde Philips PureWave et la segmentation associée du SVC.

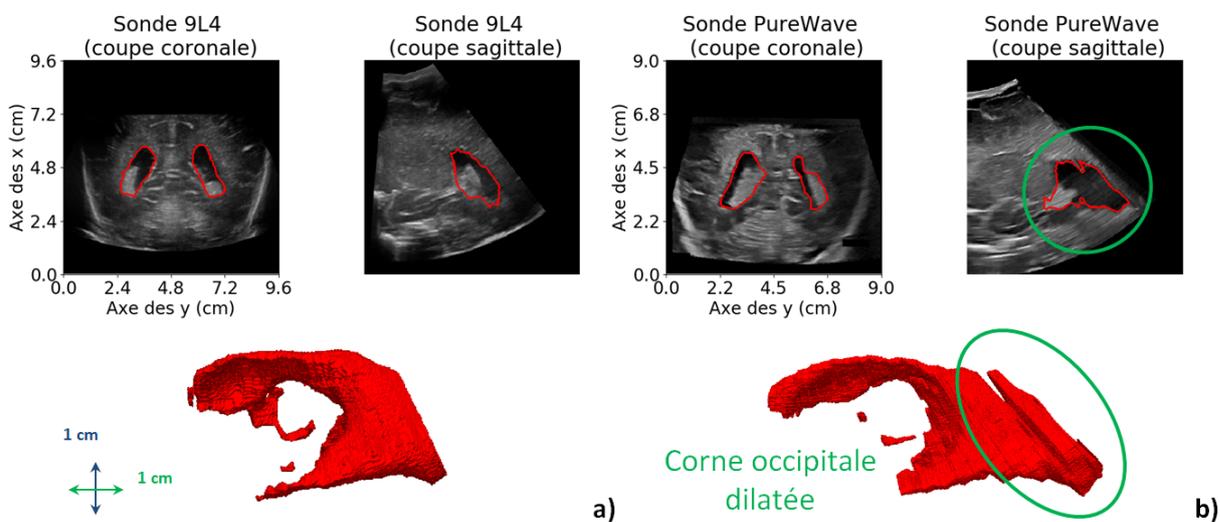


FIGURE VI.27 – Données échographiques et segmentation 3D de référence associées de SVC dilatés dans le cas d’une acquisition avec la sonde 9L4 (a) et la sonde Philips PureWave (b).

On peut voir que la corne occipitale de ce ventricule est anormalement dilatée, la qualité des données US est, de plus, inférieure à celle obtenue avec la sonde 9L4. A titre de comparaisons, une acquisition réalisée sur un patient avec un SVC dilaté avec la sonde 9L4 est montrée sur la Figure.VI.27.a.

VI.7.2 Expériences et résultats

Ces données atypiques ont été segmentées automatiquement avec les réseaux U-net 24 couches et V-net 31 couches, avec et sans CPPN. Dans chacun de ces quatre cas, le réseau qui avait obtenu le meilleur Dice sur le jeu de validation a été utilisé. Les Dice obtenus sur ces données atypiques et en moyenne sur le jeu de test sont donnés dans la Table.VI.15.

TABLE VI.15 – Dice obtenus par U-net et V-net, avec et sans CPPN, sur le jeu de test et sur des données atypiques.

Données	U-net	U-net + CPPN	V-net	V-net + CPPN
Jeu de test	0.81 ± 0.062	0.817 ± 0.05	0.82 ± 0.055	0.824 ± 0.046
Sonde PureWave	0.688	0.738	0.612	0.707
Cas d’hydrocéphalie	0.304	0.25	0.111	0.068

Dans le cas de la segmentation du volume acquis avec la sonde Philips PureWave, on constate que les Dice sont moins bons que ceux obtenus en moyenne sur le jeu de test. De plus, la précision obtenue est largement meilleure lorsque que les réseaux intègrent le CPPN. Dans ce dernier cas, les Dice obtenus sont corrects (> 0.7) mais restent en dessous des performances moyennes obtenues sur le jeu de test.

Dans le cas de la segmentation du volume correspondant à un cas d'hydrocéphalie, les Dice obtenus ne sont pas bons. La segmentation la plus précise ($Dice = 0.304$) a été réalisée par U-net. De plus, on constate que la précision obtenue par ces réseaux est moins bonne lorsqu'ils intègrent le CPPN.

Pour comprendre ces résultats quantitatifs, les segmentations produites par ces quatre réseaux ainsi que la segmentation de référence associée sont visibles sur la Figure.VI.28 dans le cas d'hydrocéphalie et sur la Figure.VI.29 dans le cas de l'acquisition réalisée avec la sonde PureWave.

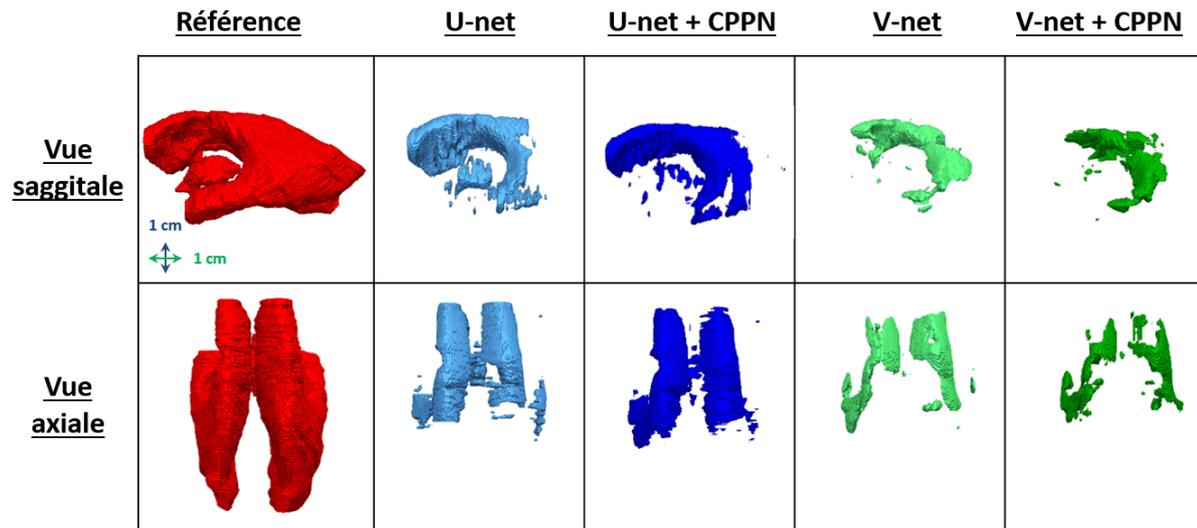


FIGURE VI.28 – Segmentations 3D du SVC produites par U-net et V-net, avec et sans CPPN, dans le cas du patient atteint d'hydrocéphalie.

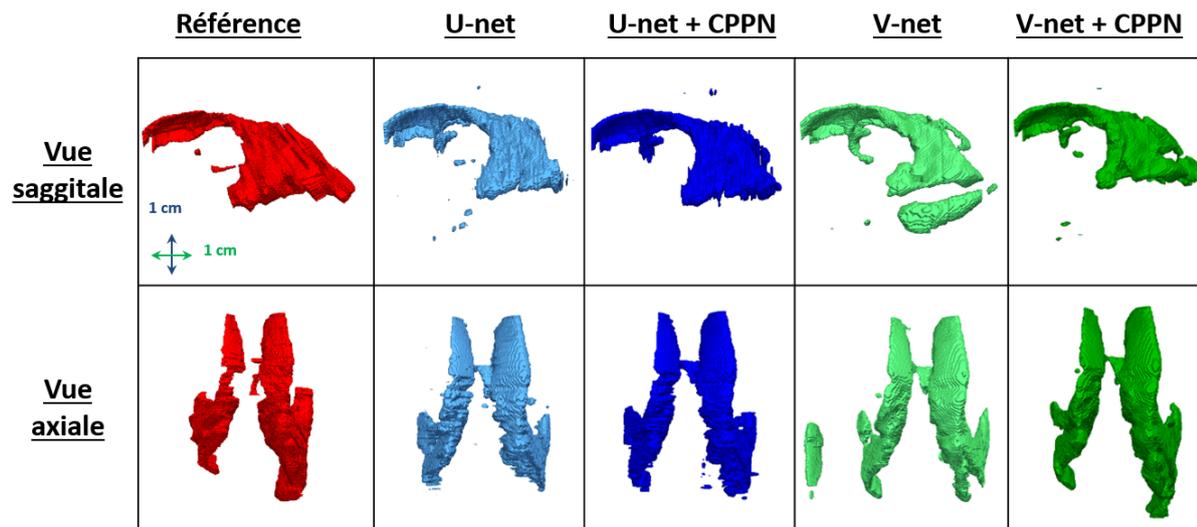


FIGURE VI.29 – Segmentations 3D du SVC produites par U-net et V-net, avec et sans CPPN, dans le cas de l'acquisition réalisée avec la sonde Philips PureWave.

Dans le cas de l'acquisition réalisée avec la sonde Philips PureWave (Figure.VI.29), on peut voir que les cornes frontales, la partie centrale et les cornes temporales sont correctement segmentées

par les quatre réseaux. En revanche, la corne occipitale dilatée n'est vraiment segmentée que par le réseau V-net qui inclut le CPPN. On peut aussi voir que V-net produit un gros faux-positif sur la gauche du SVC, juste en dehors de la boîte crânienne. Il est possible que celui-ci soit dû à une zone correspondant à un champ DICOM, lui aussi inhabituel, qui avait été éliminé en mettant tous les pixels de la zone où il se trouvait à zéro (Figure.VI.30).

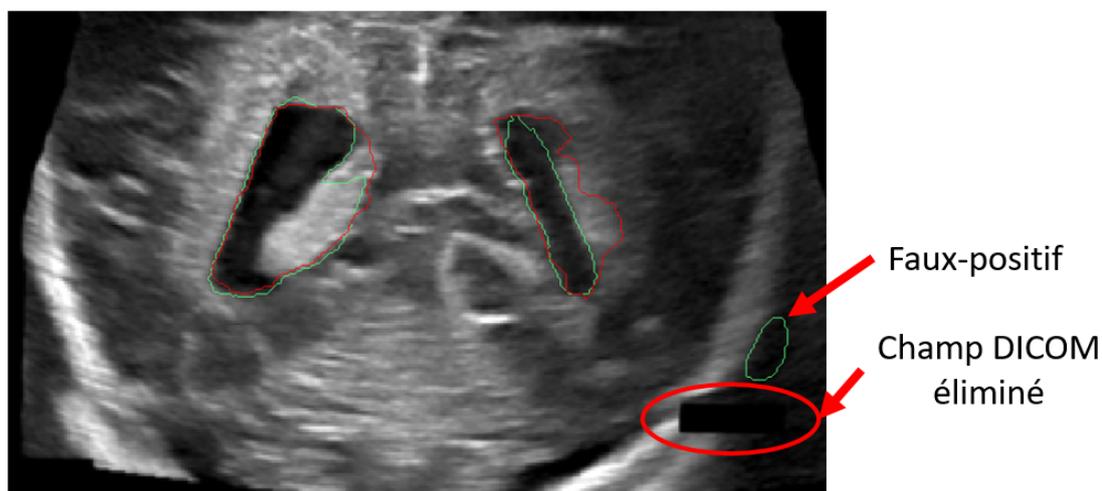


FIGURE VI.30 – Faux-positif proche d'un champ DICOM éliminé sur l'acquisition réalisée avec la sonde Pure-Wave. La segmentation de référence est en rouge et la segmentation produite par le réseau en vert.

Cette expérience montre trois choses :

- Les réseaux sont capables de segmenter des données acquises avec une autre sonde, lorsque la topologie du SVC est semblable aux cas observés dans le jeu d'entraînement.
- Les réseaux ne parviennent pas à segmenter les cas atypiques de dilatation des cornes occipitales (aucun volume semblable dans le jeu d'entraînement).
- Les réseaux ne sont pas robustes en cas de modification arbitraire dans une image, comme par exemple, la mise à zéro d'un groupe de pixel (élimination d'un champ DICOM).

Le premier point devra être vérifié par la segmentation d'autres SVCs, ayant des topologies semblables à celles observées dans le jeu d'entraînement, dans des volumes acquis avec d'autres sondes. Il semble que le réseau ne soit pas capable de segmenter le SVC si sa topologie est éloignée de ce qui a été observé dans le jeu d'entraînement.

Ce dernier point est totalement vérifié si on regarde le cas d'hydrocéphalie (Figure.VI.28) qui est très mal segmenté par tous les réseaux. Dans les cas 2D, on observe que seules les cornes occipitales sont à peu près trouvées ainsi que la partie antérieure de la partie centrale du SVC. Dans les cas 3D, seuls les bords du SVC au niveau des cornes occipitales et du début de la partie antérieure de la partie centrale du SVC sont segmentés. Ces résultats peuvent s'expliquer par le fait que la topologie de ce SVC est trop éloignée de ce qui a été appris par le réseau dans le jeu d'entraînement. En particulier au niveau des cornes temporales, de la partie centrale et des cornes occipitales.

VI.7.3 Conclusion

Dans cette partie, deux cas atypiques ont été segmentés par les réseaux U-net et V-net avec et sans CPPN. Le premier cas était un cas d'hydrocéphalie et le second une acquisition réalisée avec une sonde Philips PureWave sur un patient dont l'une des corne occipitale était anormalement dilatée. Dans le premier cas, les performances des quatre réseaux ont été moins bonnes qu'en moyenne sur le jeu de test et la corne occipitale dilatée n'a pas été segmentée correctement. Des performances correctes ($Dice > 0.7$) ont néanmoins été obtenues pour certains réseaux. Dans le second cas, aucun des quatre réseaux n'a pu segmenter le SVC avec précision, le meilleur Dice étant de 0.3.

Ces résultats montrent que les réseaux n'ont pas la capacité de segmenter des SVC qui n'ont pas une topologie semblable à celles observées dans le jeu d'entraînement. En revanche, ils semblent être en mesure de segmenter correctement des données acquises avec une autre sonde. Ce dernier point devra néanmoins être vérifié sur d'autres acquisitions, mais la segmentation de données provenant d'une sonde différente semble être envisageable. À l'avenir, il sera nécessaire d'introduire des cas où la topologie du SVC est atypique dans le jeu d'entraînement.

VI.8 Segmentation des thalami

Dans cette partie, le travail effectué sur la segmentation du SVC est étendu aux thalami. Comme cela a été évoqué dans le Chapitre I, cette structure cérébrale contraste très faiblement avec certaines structures adjacentes, ce qui rend ce problème de segmentation difficile.

VI.8.1 Expérience et résultats

Dans les parties précédentes, nous avons vu que les CNNs U-net et V-net permettaient de segmenter le SVC avec précision. Dans l'avenir, ce projet a pour objectif la réalisation d'une segmentation multi-structure de données échographiques 3D. Disposer d'une architecture unique qui permette de segmenter celles-ci simultanément, avec précision et dans un temps clinique serait donc idéal. A titre de faisabilité, les réseaux U-net 24 couches et V-net 31 couches ont été entraînés à segmenter le SVC et les thalami en utilisant la base de données décrite dans la Partie.VI.1.3 et la procédure d'optimisation décrite dans la Partie.VI.3. Chaque réseau a été optimisé 5 fois. Le Dice et la MAD obtenus en moyenne (sur les 5 optimisations) par U-net et V-net sur le jeu de test sont donnés dans la Table.VI.16. Dans le cas des thalami, le Dice obtenu sur le jeu de test par le réseau ayant obtenu le meilleur Dice sur le jeu de validation (MV) est également donné. Enfin, le temps de segmentation moyen du SVC et des thalami est donné dans la dernière colonne.

TABLE VI.16 – Résultats quantitatifs obtenus par U-net et V-net pour la segmentation du SVC et des thalami : Dice, MAD et p-valeurs.

Réseau	SVC		Thalami			Temps (s)
	Dice	MAD (mm)	Dice (MV)	Dice	MAD (mm)	
U-net	0.823 ± 0.004	0.47 ± 0.01	0.883 ± 0.042	0.881 ± 0.003	0.6 ± 0.01	6.3 ± 0.0
V-net	0.83 ± 0.002	0.5 ± 0.02	0.891 ± 0.016	0.889 ± 0.002	0.58 ± 0.01	5.6 ± 0.1
p-valeurs	1.50e-02	2.47e-02		5.79e-03	2.22e-02	

En ce qui concerne le SVC, le Dice et la MAD obtenus en moyenne par U-net et V-net sur le jeu de test sont très bons et légèrement supérieurs aux résultats donnés dans la Table.VI.7. Cela peut s'expliquer par le fait que le jeu de test de la base de données du SVC et de la base de données du SVC et thalami (Partie.VI.1.3) ne sont pas exactement les mêmes, tout comme les jeux d'entraînement et de validation. De plus, on trouve que V-net est significativement meilleur que U-net au regard du Dice ($p = 1.50 \times 10^{-2}$) mais pas de la MAD ($p = 2.47 \times 10^{-2}$).

Les segmentations produites pour les thalami par les deux réseaux sont d'une excellente précision. Les meilleurs résultats ont été obtenus par V-net : $Dice = 0.889 \pm 0.02$ et une $MAD = 0.58 \pm 0.01$. Ces résultats sont significativement meilleurs que ceux obtenus par U-net avec des p-valeurs respectives de $p = 5.79 \times 10^{-3}$ et $p = 2.22 \times 10^{-2}$. Le Dice obtenu par V-net sur le jeu de test dans le cas du réseau ayant obtenu le meilleur Dice sur le jeu de validation est de 0.91 ± 0.016 , ce qui est excellent compte tenu de la difficulté du problème.

La segmentation la plus rapide de ces deux structures a été réalisée par V-net en 5.6 secondes. Cette valeur est supérieur de 1,5 seconde au temps obtenu dans le cas de la segmentation du SVC seul par V-net mais reste un temps de segmentation compatible avec un usage clinique. Les volumes segmentés avaient une taille de $320 \times 320 \times 320$ et les segmentations ont été réalisées avec un GPU NVIDIA Tesla V100.

De manière à interpréter ces résultats, des segmentations des thalami contenus dans le jeu de test ont été réalisées avec les réseaux U-net et V-net ayant obtenus le meilleur Dice (pour les thalami) sur le jeu de validation. La Figure.VI.31 montre les segmentations correspondant au meilleur Dice (Figure.VI.31.a) et au moins bon Dice (Figure.VI.31.b) obtenus sur le jeu de test.

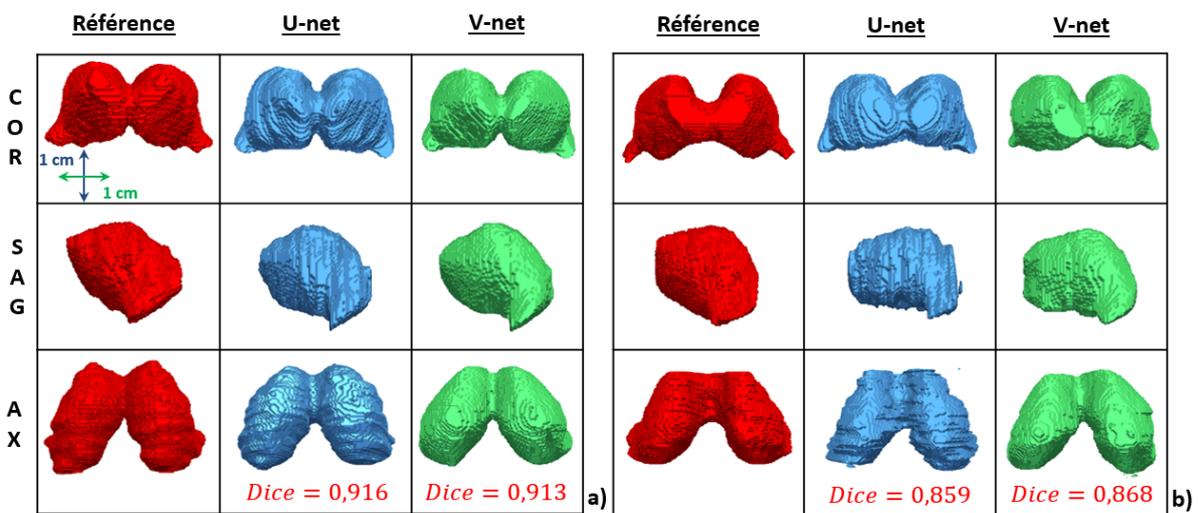


FIGURE VI.31 – Segmentation des thalami produite par U-net et V-net : Dice le plus élevé (a) et le plus bas (b).

Les segmentations produites par U-net et V-net sont extrêmement cohérentes avec ce qu'on observe sur la référence dans les deux cas. Les bords des segmentations réalisées avec V-net sont néanmoins plus réguliers que ceux des segmentations produites avec U-net, ce qui peut s'expliquer par l'utilisation de convolutions 3D dans le cas de V-net. Pour ces deux patients, les segmentations du SVC et des thalami produites par V-net sont visibles sur la Figure.VI.32.

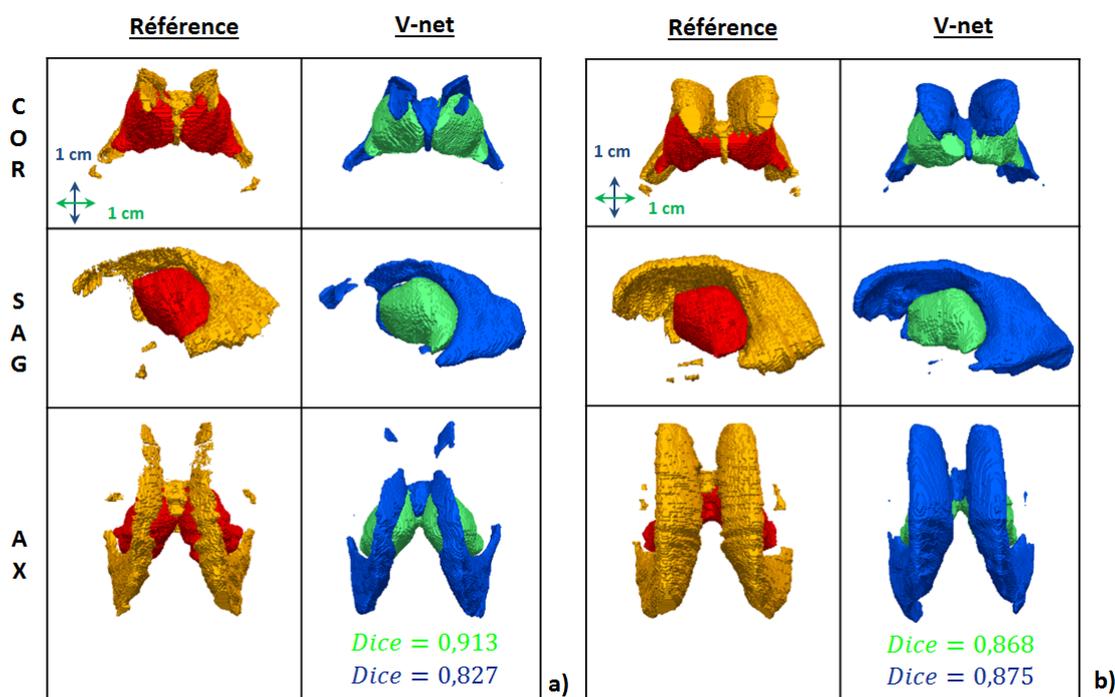


FIGURE VI.32 – Segmentation du SVC et des thalami produites par V-net : Dice le plus élevé (a) et le plus bas (b) obtenus pour les thalami.

Dans les deux cas, on peut voir que les segmentations automatiques obtenues sont très bonnes. En ce qui concerne les segmentations du SVC par V-net, on peut faire les mêmes constatations que dans la Partie.VI.4.

VI.8.2 Volume thalamique mesuré après segmentation par V-net.

Dans l'avenir on va s'intéresser au volume thalamique (VT). Celui-ci est en effet corrélé à la présence de PWMLs (Tusor *et al.* (2017)), il pourrait donc constituer un biomarqueur qui, dans le cas de l'échographie, viendrait remplacer le volume de la SB qui est extrêmement difficile à mesurer. Les erreurs volumiques commises sur le jeu de test par le réseau V-net ayant obtenu le meilleur Dice (pour les thalami) sur le jeu de validation sont données dans la Table.VI.17.

TABLE VI.17 – Erreurs volumiques commises par V-net lors de la segmentation des thalami.

Nombre de patients	VT moyen (cm ³)	ΔV_a (cm ³)	ΔV_r (%)
5	5.51 ± 0.65	0.59 ± 0.35	9.8 ± 5.6

Les erreurs volumiques absolues et relatives obtenues sont correctes : $\Delta V_a = 0.59 \pm 0.35$ cm³ et $\Delta V_r = 9.8 \pm 5.6$ cm³. Des études cliniques plus approfondies devront être déterminées si ces valeurs sont admissibles ou non.

VI.8.3 Conclusion

Dans cette partie, U-net et V-net ont été optimisés pour segmenter le SVC et les thalami. Les résultats obtenus ont montrés que V-net était capable de segmenter les thalami avec une excellente précision : $Dice = 0.891 \pm 0.016$ dans un temps compatible avec la clinique (5.6 ± 0.1 secondes). Les expériences réalisées ont de plus montrés que cette architecture était significativement meilleure que U-net pour ce problème. V-net semble donc être une structure adéquate pour segmenter les thalami et le SVC dans ces données échographiques 3D.

Compte tenu de la difficulté de ces deux problèmes de segmentation et des bonnes performances obtenues par V-net, on peut s'attendre à ce que cette architecture permette d'obtenir de bon résultats pour d'autres structures cérébrales.

VI.9 Conclusion

VI.9.1 Résumé du travail effectué

Ce chapitre a traité des problèmes de segmentation du SVC et des thalami dans des données échographiques 3D. La majorité de ce travail a été effectué sur le SVC avant d'être étendu aux thalami. Les premières parties ont présenté les bases de données que nous avons construites pour travailler sur ces problèmes ainsi que les architectures des CNNs utilisés et la manière dont ces réseaux ont été optimisés.

Les expériences réalisées ont permis de conclure que V-net (CNN 3D) était une meilleure architecture que U-net (CNN 2D) pour résoudre ces deux problèmes, même si U-net permet également d'obtenir de très bonnes segmentations. Dans le cas de la segmentation du SVC, V-net a permis d'obtenir un $Dice = 0.828 \pm 0.044$ pour un temps de segmentation de 4.1 ± 0.2 secondes, et dans le cas de la segmentation des thalami un $Dice = 0.891 \pm 0.016$ pour un temps de segmentation de 5.6 ± 0.1 secondes. Cette architecture est donc capable de segmenter ces deux structures avec une excellente précision dans un temps compatible avec la clinique.

Les expériences menées sur la taille du jeu d'entraînement et sur l'augmentation artificielle de données dans le cas du SVC ont montré deux choses : le nombre de volumes dans le jeu d'entraînement est suffisant pour obtenir une bonne précision sur le jeu de test ($Dice > 0.8$ à partir de 9 volumes d'entraînement avec U-net) et réaliser des symétries horizontales pendant la phase d'optimisation permet d'obtenir de meilleures performances sur le jeu de test ($Dice = 0.828 \pm 0.044$ avec V-net).

Les architectures de U-net et V-net ont également été modifiées, par l'ajout d'un CPPN, pour que ces réseaux apprennent la position anatomique du SVC et puisse tirer profit de cette information. Les expériences menées ont montré une amélioration significative des performances lorsque le nombre de couches de ces réseaux est faible et qu'ils ne sont donc pas en mesure d'utiliser beaucoup de contexte.

Enfin, les performances de U-net et V-net ont été évaluées pour la segmentation du SVC dans deux cas atypiques : un volume acquis avec une autre sonde que la 9L4 et un cas d'hydrocéphalie. Les résultats obtenus suggèrent que ces réseaux sont capables de segmenter des volumes acquis avec une autre sonde mais échouent à segmenter un SVC dont la topologie est trop éloignée de celle observée dans le jeu d'entraînement.

VI.9.2 Résumé des contributions

Les contributions majeures apportées par ce chapitre sont :

1. La première segmentation automatique du SVC, par des réseaux de neurones convolutifs, dans des données échographiques 3D et dans un temps clinique (Martin *et al.* (2018)).

2. La première segmentation automatique des thalami dans des données échographiques 3D, dans un temps clinique.
3. La modification des architectures de U-net et V-net qui permet à ces réseaux d'encoder la position anatomique du SVC.
4. L'estimation du nombre de volumes d'apprentissage nécessaires pour segmenter le SVC avec précision (Dice > 0.8) et la comparaison de plusieurs stratégies d'augmentation artificielle de données (Martin *et al.* (2019)).

Les points 1 et 4 ont été valorisés lors de présentations orales au congrès IUS (Martin *et al.* (2018), Martin *et al.* (2019)). Les points 1 et 3 ont été valorisés par un article soumis dans le journal MEDIA.

VI.9.3 Perspectives

Dans le cas de la segmentation du SVC, plusieurs points pourraient être étudiés :

- Enrichir la base de données avec des cas d'hydrocéphalie pour apprendre aux réseaux à traiter ces cas.
- Vérifier que des réseaux U-net et V-net, entraînés sur nos données, sont capable de segmenter des données échographiques 3D obtenues à partir d'acquisitions réalisées avec d'autres sondes.

En ce qui concerne les thalami, les points qui pourraient être étudiés sont :

- Évaluer l'influence du CPPN, comme cela a été fait pour le SVC.
- Déterminer si certaines stratégies d'augmentation artificielle de données permettent d'améliorer la précision des segmentations.
- Déterminer si il est possible d'améliorer significativement la précision des segmentations en augmentant le nombre de volumes dans le jeu d'entraînement.

Enfin, la base de données pourraient être enrichie avec des segmentations de nouvelles structures, ce qui permettrait de vérifier que cette architecture est adéquate pour segmenter l'ensemble des structures cérébrales chez le prématuré. De plus, il serait intéressant d'optimiser ou de modifier celle-ci pour qu'elle soit dédiée à ce problème et soit la moins demandeuse en ressources (de manière à être utilisable sur la plupart des GPUs). Pour cela, intégrer le CPPN aux architectures est un bon point de départ car il paraît important d'encoder la position anatomique des structures cérébrales pour faciliter leur segmentation.

CONCLUSION

Cette dernière partie résume le travail qui a été réalisé pendant cette thèse, les contributions apportées et les perspectives pour la suite du projet.

Reconstruction 3D de haute qualité de données échographiques 2D du cerveau du prématuré

Concernant la reconstruction 3D de haute qualité de données échographiques 2D du cerveau du prématuré deux objectifs avaient été évoqués dans le Chapitre.II :

1. A court terme : reconstruire en 3D les ETFs 2D réalisées au CH d'Avignon avec une sonde multi-D 9L4 en conservant une excellente qualité d'images et dans un temps raisonnable.
2. A long terme : identifier et étudier une solution qui permette d'acquérir des volumes 3D de haute qualité en routine clinique.

Pour répondre au premier objectif, une chaîne de reconstruction 3D, détaillée dans le Chapitre.IV a été développée. Cet algorithme a été validé in-vivo sur le corps calleux où une MAD de 0.69 ± 0.014 mm a été obtenue. Cette précision est nettement moins bonne que celle qui serait obtenue avec une acquisition IRM ou avec une sonde matricielle mais est excellente dans le cadre de la reconstruction 3D d'acquisitions main-libres. Sans compter les temps d'interaction avec l'utilisateur, cette chaîne a permis de reconstruire des volumes en 6 minutes 45 secondes (± 1 min 42 s), ce qui est raisonnable pour un algorithme dédié à la recherche. Au total, des reconstructions 3D de haute qualité de 25 ETF 2D ont été réalisées pour constituer la base d'exams médicaux qui ont été utilisés pour travailler sur le problème de segmentation des structures cérébrales. Cette chaîne répond donc parfaitement au premier objectif que nous nous étions fixés.

Plusieurs points pourraient être approfondis dans l'avenir. Il faudrait notamment valider cette chaîne, et plus particulièrement les modèles d'acquisition, sur l'ensemble des structures du cerveau à partir du fantôme de tête de nouveau-né ou d'acquisitions IRMs. De plus, pour envisager un usage clinique de cet algorithme, toutes ses étapes, et notamment celle de positionnement des repères anatomiques, devront être automatisées.

En ce qui concerne le deuxième objectif, si on se base sur l'étude bibliographique réalisée dans le Chapitre.II, la solution qui nous semble la plus pertinente est d'utiliser un capteur enregistrant le déplacement de la sonde au cours de l'acquisition. La solution proposée par Morgan *et al.* (2018), qui utilise un accéléromètre et un gyromètre nous paraît être la plus adaptée. Elle est peu coûteuse, adaptable à n'importe quelle sonde et permet de réaliser l'acquisition dans les conditions habituelles. Une étude de faisabilité sera réalisée en partenariat avec l'université de Duke, aux Etats-Unis,

dans la suite de ce projet. Ceux-ci nous ont déjà fait parvenir un de leur capteur qui sera testé pour reconstruire des acquisitions réalisées sur le fantôme de tête de prématuré. Pour aller plus loin, cette méthode permettrait de constituer une base de données pour entraîner un CNN à reconstruire automatiquement en 3D les ETFs 2D en utilisant uniquement les images acquises.

Annotations 3D des structures cérébrales dans des données échographiques 3D

Pour étudier un problème de segmentation, il est nécessaire d'avoir des segmentations de référence. Pour obtenir celles-ci dans nos données échographiques, la segmentation manuelle de chacune de leurs images n'était pas réaliste : charge de travail et temps requis trop importants. Un algorithme permettant d'interpoler des segmentations 3D à partir de segmentations manuelles 2D réalisées de manière discontinues a donc été développé (Chapitre.V). Celui-ci nous a permis d'obtenir des segmentations 3D de référence du SVC et des thalami dans respectivement 25 et 16 volumes, cette dernière base étant unique au monde à notre connaissance. D'après nos estimations, le temps de segmentation manuel a été réduit par 5 dans le cas du SVC, 6 dans le cas des thalami annotés en coupe coronale et 4 dans le cas des thalami annotés en coupe axiale, ce qui est un gain extrêmement important et précieux. La précision des segmentations interpolées a été estimée à partir de modèles CAO de référence, l'écart estimé entre une segmentation manuelle et interpolée est donné dans la Table.VI.18.

TABLE VI.18 – Précision estimée des reconstruction 3D du SVC et des thalami.

Structure	SVC	Thalami (coronale)	Thalami (axiale)
Espacement utilisé (images)	4.1 ± 2.1	5.4 ± 1.8	3.2 ± 0.4
Dice	0.969 ± 0.006	0.984 ± 0.0	0.982 ± 0.001
MAD (mm)	0.16 ± 0.04	0.13 ± 0.03	0.16 ± 0.02

Les segmentations 3D ont donc été interpolées avec une très bonne précision. Les valeurs du Dice et de la MAD obtenues sont des valeurs que l'algorithme de segmentation automatique ne pourra théoriquement pas dépasser. Si on considère de plus la variabilité intra-opérateur et le fait que le SVC est souvent beaucoup plu fin que le modèle CAO (base de données Chapitre.VI) la précision maximale pouvant être atteinte pas un algorithme de segmentation automatique est probablement bien inférieure à ces valeurs. Cette précision maximale est difficile à évaluer, sur nos données elle se situe probablement entre un Dice de 0.85 et 0.90 et une MAD de 0.25 mm et 0.35 mm dans le cas du SVC et entre un Dice de 0.90 et 0.95 et une MAD de 0.2 mm et 0.3 mm dans le cas des thalami.

Dans l'avenir, l'algorithme développé pourra être utilisé pour obtenir des segmentations 3D de référence d'autres structures cérébrales : corps calleux, cervelet, noyaux caudés, ... L'espacement optimal (compromis entre précision des segmentions 3D interpolées et temps d'annotation manuel) entre les segmentations 2D pourra alors être déterminé avant d'annoter les images. Une base de référence 3D des lésions ponctuées (PWMLs) est déjà en cours de réalisation et contient aujourd'hui environ 20 patients, celle-ci est également unique au monde à notre connaissance.

Segmentation automatique des structures cérébrales du cerveau du prématuré dans des données échographiques 3D

Les algorithmes présentés dans le Chapitre.IV et le Chapitre.V ont permis de créer des bases de données annotées du SVC et des thalami dans des données échographiques 3D. Des méthodes d'apprentissage supervisé, choisies pour travailler sur les problèmes de segmentation de ces deux structures (Chapitre.III), ont alors pu être optimisées. En particulier, des CNNs ont été entraînés, parmi les architectures avec lesquelles nous avons travaillées, V-net (CNN 3D) est celle qui a permis d'obtenir les meilleurs résultats pour résoudre ces deux problèmes, même si U-net a aussi permis d'obtenir des segmentations très précises. Dans le cas du SVC, V-net entraîné avec des symétries horizontales a segmenté cette structure avec un $Dice = 0.828 \pm 0.044$ et une $MAD = 0.49 \pm 0.13$ mm pour un temps de segmentation de 4.4 secondes. Dans le cas des thalami, sans faire d'augmentation de données artificielle, un $Dice = 0.891 \pm 0.016$ et une $MAD = 0.53 \pm 0.1$ mm pour un temps de segmentation de 5.6 secondes (SVC et thalami) ont été obtenus. Cette architecture est donc capable de segmenter ces deux structures avec une excellente précision dans un temps compatible avec la clinique. Si on considère les valeurs limites du Dice et de la MAD qui ont été données dans la partie précédente, cette précision pourraient encore être légèrement améliorées.

Pour cela, la meilleure option serait de déterminer une architecture de CNN qui soit dédiée à la segmentation des structures cérébrales du prématuré. L'intégration d'un CPPN à l'architecture de U-net et de V-net, pour que ces réseaux apprennent la position anatomique du SVC et puisse tirer profit de cette information, va dans ce sens. Les expériences qui ont été menées ont montré une amélioration significative des performances lorsque le nombre de couches des architectures était faible : dans le cas du réseau V-net à 7 couches, le Dice a ainsi augmenté de 0.524 ± 0.076 à 0.724 ± 0.107 . Ce qui est une importante amélioration, même si les performances deviennent équivalentes à celles obtenues sans CPPN quand le nombre de couches augmente. L'intégration d'un CPPN est un bon point de départ pour créer une architecture dédiée avec peu de paramètres.

L'évaluation des performances de U-net et V-net sur des cas atypiques a montré que ces réseaux semblaient capables de segmenter des données acquises avec une sonde différente de la 9L4. Cette faculté devra néanmoins être vérifiée sur des acquisitions réalisées avec d'autres modèles de sondes. Cette expérience a de plus permis de conclure que ces réseaux n'étaient pas capables de segmenter des SVCs dont la topologie est trop éloignée de celle observée dans le jeu d'entraînement. La base de données du SVC devra donc être enrichie avec des cas de dilatations atypiques (hydrocéphalie, dilatation inhabituelle de certaines parties du SVC, ...) pour que les CNNs soient capables de segmenter ces cas.

Conclusion générale

Cette thèse montre qu'il est possible de segmenter automatiquement, avec précision et dans un temps clinique, des structures cérébrales de l'enfant prématuré dans des données échographiques 3D de haute qualité. Cela montre que l'échographie 3D, sous réserve d'être de haute qualité, pourrait être utilisée en routine clinique pour quantifier le volume des structures cérébrales et ouvre la

Conclusion

voie aux études d'évaluation de son bénéfice pour les patients. La suite du projet devra permettre l'acquisition en ligne de volume échographique de haute qualité du cerveau du prématuré, la segmentation de structures cérébrales autre que les thalami et le SVC et la détection des PWMLs. Cela passe par la création de nouvelles bases de données annotées avec les outils développés durant cette thèse.

BIBLIOGRAPHIE

- Allen, M. C., 2008. Neurodevelopmental outcomes of preterm infants. *Current opinion in neurology* 21 (2), 123–128.
- Altaye, M., Holland, S. K., Wilke, M., Gaser, C., 2008. Infant brain probability templates for mri segmentation and normalization. *Neuroimage* 43 (4), 721–730.
- Anbeek, P., Išgum, I., van Kooij, B. J., Mol, C. P., Kersbergen, K. J., Groenendaal, F., Viergever, M. A., de Vries, L. S., Benders, M. J., 2013. Automatic segmentation of eight tissue classes in neonatal brain mri. *PLoS One* 8 (12), e81895.
- Anbeek, P., Vincken, K. L., Groenendaal, F., Koeman, A., Van Osch, M. J., Van der Grond, J., 2008. Probabilistic brain tissue segmentation in neonatal magnetic resonance imaging. *Pediatric research* 63 (2), 158.
- Ancel, P.-Y., Goffinet, F., 2014. Epipage 2 : a preterm birth cohort in france in 2011. *BMC pediatrics* 14 (1), 97.
- Ancel, P.-Y., Veronique Pierrat, V., MDe, S. N. G., MDg, A. B., 2006. Cerebral palsy among very preterm children in relation to gestational age and neonatal ultrasound abnormalities : the epipage cohort study. In : *peds*. Vol. 2005. p. 0091.
- Arnaud, C., Daubisse-Marliac, L., White-Koning, M., Pierrat, V., Larroque, B., Grandjean, H., Alberge, C., Marret, S., Burguet, A., Ancel, P.-Y., *et al.*, 2007. Prevalence and associated factors of minor neuromotor dysfunctions at age 5 years in prematurely born children : the epipage study. *Archives of pediatrics & adolescent medicine* 161 (11), 1053–1061.
- Arpino, C., Compagnone, E., Montanaro, M. L., Cacciatore, D., De Luca, A., Cerulli, A., Di Girolamo, S., Curatolo, P., 2010. Preterm birth and neurodevelopmental outcome : a review. *Child's nervous system* 26 (9), 1139–1149.
- Artaechevarria, X., Munoz-Barrutia, A., Ortiz-de Solórzano, C., 2009. Combination strategies in multi-atlas image segmentation : application to brain mr data. *IEEE transactions on medical imaging* 28 (8), 1266–1277.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2015. Segnet : A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv :1511.00561*.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet : A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence* 39 (12), 2481–2495.

-
- Barry, C., Allott, C., John, N., Mellor, P., Arundel, P., Thomson, D., Waterton, J., 1997. Three-dimensional freehand ultrasound : image reconstruction and volume analysis. *Ultrasound in medicine & biology* 23 (8), 1209–1224.
- Bax, J., Cool, D., Gardi, L., Knight, K., Smith, D., Montreuil, J., Sherebrin, S., Romagnoli, C., Fenster, A., 2008. Mechanically assisted 3d ultrasound guided prostate biopsy system. *Medical physics* 35 (12), 5397–5410.
- Beaino, G., Khoshnood, B., Kaminski, M., Pierrat, V., Marret, S., Matis, J., LedÉsert, B., Thiriez, G., Fresson, J., ROZÉ, J.-C., *et al.*, 2010. Predictors of cerebral palsy in very preterm infants : the epi-page prospective population-based cohort study. *Developmental Medicine & Child Neurology* 52 (6), e119–e125.
- Bentley, J. L., 1975. Multidimensional binary search trees used for associative searching. *Communications of the ACM* 18 (9), 509–517.
- Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A.-B., Narwal, R., Adler, A., Garcia, C. V., Rohde, S., Say, L., *et al.*, 2012. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries : a systematic analysis and implications. *The lancet* 379 (9832), 2162–2172.
- Boucher, M.-A., Lippé, S., Damphousse, A., El-Jalbout, R., Kadoury, S., 2018a. Dilatation of lateral ventricles with brain volumes in infants with 3d transfontanelle us. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 557–565.
- Boucher, M.-A., Lippé, S., Dupont, C., Knoth, I. S., Lopez, G., Shams, R., El-Jalbout, R., Damphousse, A., Kadoury, S., 2018b. Computer-aided lateral ventricular and brain volume measurements in 3d ultrasound for assessing growth trajectories in newborns and neonates. *Physics in Medicine & Biology* 63 (22), 225012.
- Brouwer, M. J., de Vries, L. S., Groenendaal, F., Koopman, C., Pistorius, L. R., Mulder, E. J., Benders, M. J., 2012. New reference values for the neonatal cerebral ventricles. *Radiology* 262 (1), 224–233.
- Burstein, J., Papile, L., Burstein, R., 1979. Intraventricular hemorrhage and hydrocephalus in premature newborns : a prospective study with ct. *American journal of roentgenology* 132 (4), 631–635.
- Cardoso, M. J., Melbourne, A., Kendall, G. S., Modat, M., Robertson, N. J., Marlow, N., Ourselin, S., 2013. Adapt : an adaptive preterm segmentation algorithm for neonatal brain mri. *NeuroImage* 65, 97–108.
- Chen, H., Dou, Q., Yu, L., Qin, J., Heng, P.-A., 2018. Voxresnet : Deep voxelwise residual networks for brain segmentation from 3d mr images. *NeuroImage* 170, 446–455.
- Chen, J.-F., Fowlkes, J. B., Carson, P. L., Rubin, J. M., 1997. Determination of scan-plane motion using speckle decorrelation : Theoretical considerations and initial test. *International Journal of Imaging Systems and Technology* 8 (1), 38–44.

-
- Cheong, J. L., Thompson, D. K., Spittle, A. J., Potter, C. R., Walsh, J. M., Burnett, A. C., Lee, K. J., Chen, J., Beare, R., Matthews, L. G., *et al.*, 2016. Brain volumes at term-equivalent age are associated with 2-year neurodevelopment in moderate and late preterm children. *The Journal of pediatrics* 174, 91–97.
- Coleman, T. E., Li, Y., 1994. On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds. *Mathematical programming* 67 (1-3), 189–224.
- Coleman, T. E., Li, Y., 1996. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on optimization* 6 (2), 418–445.
- de Brebisson, A., Montana, G., 2015. Deep neural networks for anatomical brain segmentation. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. pp. 20–28.
- Delobel-Ayoub, M., Kaminski, M., Marret, S., Burguet, A., Marchand, L., Sylvie, N., Matis, J., Thiriez, G., Fresson, J., Arnaud, C., *et al.*, 2006. Behavioral outcome at 3 years of age in very preterm infants : the epipage study. *Pediatrics* 117 (6), 1996–2005.
- Destrempe, F., Cloutier, G., 2010. A critical review and uniformized representation of statistical distributions modeling the ultrasound echo envelope. *Ultrasound in medicine & biology* 36 (7), 1037–1051.
- Devi, C. N., Chandrasekharan, A., Sundararaman, V., Alex, Z. C., 2015. Neonatal brain mri segmentation : a review. *Computers in biology and medicine* 64, 163–178.
- Dewi, D., Wilkinson, M., Mengko, T., Purnama, I., Van Ooijen, P., Veldhuizen, A., Maurits, N., Verkerke, G., 2009. 3d ultrasound reconstruction of spinal images using an improved olympic hole-filling method. In : *International Conference on Instrumentation, Communication, Information Technology, and Biomedical Engineering 2009*. IEEE, pp. 1–5.
- Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The importance of skip connections in biomedical image segmentation. In : *Deep Learning and Data Labeling for Medical Applications*. Springer, pp. 179–187.
- Fenster, A., Downey, D. B., Cardinal, H. N., 2001. Three-dimensional ultrasound imaging. *Physics in medicine & biology* 46 (5), R67.
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., *et al.*, 2002. Whole brain segmentation : automated labeling of neuroanatomical structures in the human brain. *Neuron* 33 (3), 341–355.
- Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C., Collins, D., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47, S102.
- Fontanarosa, I., 2015. L'intérêt pronostique de l'irm cérébrale chez les prématurés. Ph.D. thesis, Université de Lorraine.

-
- Fox, L. M., Choo, P., Rogerson, S. R., Spittle, A. J., Anderson, P. J., Doyle, L., Cheong, J. L., 2014. The relationship between ventricular size at 1 month and outcome at 2 years in infants less than 30 weeks' gestation. *Archives of Disease in Childhood-Fetal and Neonatal Edition* 99 (3), F209–F214.
- Fuerst, B., Wein, W., Müller, M., Navab, N., 2014. Automatic ultrasound–mri registration for neurosurgery using the 2d and 3d lc2 metric. *Medical image analysis* 18 (8), 1312–1319.
- Gilmore, J. H., Gerig, G., Specter, B., Charles, H. C., Wilber, J. S., Hertzberg, B. S., Kliewer, M. A., 2001. Infant cerebral ventricle volume : a comparison of 3-d ultrasound and magnetic resonance imaging. *Ultrasound in medicine & biology* 27 (8), 1143–1146.
- Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In : *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. pp. 249–256.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep learning*. MIT press.
- Gousias, I. S., Edwards, A. D., Rutherford, M. A., Counsell, S. J., Hajnal, J. V., Rueckert, D., Hammers, A., 2012. Magnetic resonance imaging of the newborn brain : manual segmentation of labelled atlases in term-born and preterm infants. *Neuroimage* 62 (3), 1499–1509.
- Gousias, I. S., Hammers, A., Counsell, S. J., Srinivasan, L., Rutherford, M. A., Heckemann, R. A., Hajnal, J. V., Rueckert, D., Edwards, A. D., 2013. Magnetic resonance imaging of the newborn brain : automatic segmentation of brain images into 50 anatomical regions. *PloS one* 8 (4), e59990.
- Haiden, N., Klebermass, K., Rücklinger, E., Berger, A., Prusa, A. R., Rohrmeister, K., Wandl-Vergesslich, K., Kohlhauser-Vollmuth, C., 2005. 3-d ultrasonographic imaging of the cerebral ventricular system in very low birth weight infants. *Ultrasound in medicine & biology* 31 (1), 7–14.
- Hassenpflug, P., Prager, R. W., Treece, G. M., Gee, A. H., 2005. Speckle classification for sensorless freehand 3-d ultrasound. *Ultrasound in medicine & biology* 31 (11), 1499–1508.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 770–778.
- Housden, R., Gee, A. H., Prager, R. W., Treece, G. M., 2008. Rotational motion in sensorless freehand three-dimensional ultrasound. *Ultrasonics* 48 (5), 412–422.
- Housden, R. J., Gee, A. H., Treece, G. M., Prager, R. W., 2006. Subsample interpolation strategies for sensorless freehand 3d ultrasound. *Ultrasound in medicine & biology* 32 (12), 1897–1904.
- Housden, R. J., Gee, A. H., Treece, G. M., Prager, R. W., 2007. Sensorless reconstruction of unconstrained freehand 3d ultrasound data. *Ultrasound in medicine & biology* 33 (3), 408–419.
- Huang, C.-S., Yang, Y.-W., Chen, R.-T., Lo, C.-M., Lo, C., Cheng, C.-F., Lee, C.-S., Chang, R.-F., 2017a. Whole-breast ultrasound for breast screening and archiving. *Ultrasound in medicine & biology* 43 (5), 926–933.

-
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q., 2017b. Densely connected convolutional networks. In : Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708.
- Huang, Q., Huang, Y., Hu, W., Li, X., 2014. Bezier interpolation for 3-d freehand ultrasound. IEEE Transactions on Human-Machine Systems 45 (3), 385–392.
- Ioffe, S., Szegedy, C., 2015. Batch normalization : Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv :1502.03167.
- Isola, P., Zhu, J.-Y., Zhou, T., Efros, A. A., 2017. Image-to-image translation with conditional adversarial networks. In : Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134.
- Jeon, T. Y., Kim, J. H., Yoo, S.-Y., Eo, H., Kwon, J.-Y., Lee, J., Lee, M., Chang, Y. S., Park, W. S., 2012. Neurodevelopmental outcomes in preterm infants : comparison of infants with and without diffuse excessive high signal intensity on mr images at near-term-equivalent age. Radiology 263 (2), 518–526.
- Keunen, K., Išgum, I., van Kooij, B. J., Anbeek, P., van Haastert, I. C., Koopman-Esseboom, C., Fieret-van Stam, P. C., Nivelstein, R. A., Viergever, M. A., de Vries, L. S., *et al.*, 2016. Brain volumes at term-equivalent age in preterm infants : imaging biomarkers for neurodevelopmental outcome through early school age. The journal of Pediatrics 172, 88–95.
- Kim, H., Lepage, C., Evans, A. C., Barkovich, A. J., Xu, D., 2015. Neocivet : Extraction of cortical surface and analysis of neonatal gyrification using a modified civet pipeline. In : International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 571–579.
- Kingma, D. P., Ba, J., 2014. Adam : A method for stochastic optimization. arXiv preprint arXiv :1412.6980.
- Kishimoto, J., de Ribaupierre, S., Lee, D., Mehta, R., St Lawrence, K., Fenster, A., 2013. 3d ultrasound system to investigate intraventricular hemorrhage in preterm neonates. Physics in Medicine & Biology 58 (21), 7513.
- Kishimoto, J., de Ribaupierre, S., Salehi, E., Romano, W., Lee, D. S., Fenster, A., 2016. Preterm neonatal lateral ventricle volume from three-dimensional ultrasound is not strongly correlated to two-dimensional ultrasound measurements. Journal of Medical Imaging 3 (4), 046003.
- Kittler, J., 1998. Combining classifiers : A theoretical framework. Pattern analysis and Applications 1 (1), 18–27.
- Kovesi, P., *et al.*, 1997. Symmetry and asymmetry from local phase. In : Tenth Australian joint conference on artificial intelligence. Vol. 190. Citeseer, pp. 2–4.
- Kurian, J., Sotardi, S., Liszewski, M. C., Gomes, W. A., Hoffman, T., Taragin, B. H., 2017. Three-dimensional ultrasound of the neonatal brain : technical approach and spectrum of disease. Pediatric radiology 47 (5), 613–627.

-
- Laporte, C., Arbel, T., 2011. Learning to estimate out-of-plane motion in ultrasound imagery of real tissue. *Medical image analysis* 15 (2), 202–213.
- Larroque, B., Marret, S., Ancel, P.-Y., Arnaud, C., Marpeau, L., Supernant, K., Pierrat, V., Rozé, J.-C., Matis, J., Cambonie, G., *et al.*, 2003. White matter damage and intraventricular hemorrhage in very preterm infants : the epipage study. *The Journal of pediatrics* 143 (4), 477–483.
- Leviton, A., Gilles, F., 1996. Ventriculomegaly, delayed myelination, white matter hypoplasia, and “periventricular” leukomalacia : how are they related? *Pediatric neurology* 15 (2), 127–136.
- Lind, A., Parkkola, R., Lehtonen, L., Munck, P., Maunu, J., Lapinleimu, H., Haataja, L., Group, P. S., *et al.*, 2011. Associations between regional brain volumes at term-equivalent age and development at 2 years of age in preterm children. *Pediatric radiology* 41 (8), 953–961.
- Liu, L., Oza, S., Hogan, D., Chu, Y., Perin, J., Zhu, J., Lawn, J. E., Cousens, S., Mathers, C., Black, R. E., 2016a. Global, regional, and national causes of under-5 mortality in 2000–15 : an updated systematic analysis with implications for the sustainable development goals. *The Lancet* 388 (10063), 3027–3035.
- Liu, M., Kitsch, A., Miller, S., Chau, V., Poskitt, K., Rousseau, F., Shaw, D., Studholme, C., 2016b. Patch-based augmentation of expectation–maximization for brain mri tissue segmentation at arbitrary age after premature birth. *NeuroImage* 127, 387–408.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 3431–3440.
- Makropoulos, A., Counsell, S. J., Rueckert, D., 2018. A review on automatic fetal and neonatal brain mri segmentation. *NeuroImage* 170, 231–248.
- Makropoulos, A., Gousias, I. S., Ledig, C., Aljabar, P., Serag, A., Hajnal, J. V., Edwards, A. D., Counsell, S. J., Rueckert, D., 2014. Automatic whole brain mri segmentation of the developing neonatal brain. *IEEE transactions on medical imaging* 33 (9), 1818–1831.
- Martin, M., Sciolla, B., Sdika, M., Quétin, P., Delachartre, P., 2019. Segmentation of neonates cerebral ventricles with 2d cnn in 3d us data : suitable training-set size and data augmentation strategies. In : *2019 IEEE International Ultrasonics Symposium (IUS)*. IEEE, pp. 1–4.
- Martin, M., Sciolla, B., Sdika, M., Wang, X., Quetin, P., Delachartre, P., 2018. Automatic segmentation of the cerebral ventricle in neonates using deep learning with 3d reconstructed freehand ultrasound imaging. In : *2018 IEEE International Ultrasonics Symposium (IUS)*. IEEE, pp. 1–4.
- Maunu, J., Lehtonen, L., Lapinleimu, H., Matomäki, J., Munck, P., Rikalainen, H., Parkkola, R., Haataja, L., Group, P. S., 2011. Ventricular dilatation in relation to outcome at 2 years of age in very preterm infants : a prospective finnish cohort study. *Developmental Medicine & Child Neurology* 53 (1), 48–54.
- Mazzola, C. A., Choudhri, A. F., Auguste, K. I., Limbrick, D. D., Rogido, M., Mitchell, L., Flannery, A. M., 2014. Pediatric hydrocephalus : systematic literature review and evidence-based guidelines. part

-
- 2 : management of posthemorrhagic hydrocephalus in premature infants. *Journal of Neurosurgery : Pediatrics* 14 (Supplement_1), 8–23.
- Melbourne, A., Cardoso, M. J., Kendall, G. S., Robertson, N. J., Marlow, N., Ourselin, S., 2012. Neo-brains12 challenge : adaptive neonatal mri brain segmentation with myelinated white matter class and automated extraction of ventricles i-iv. *MICCAI grand challenge : neonatal brain segmentation (NeoBrainSI2)*, 16–21.
- Melhem, E. R., Hoon Jr, A. H., Ferrucci Jr, J. T., Quinn, C. B., Reinhardt, E. M., Demetrides, S. W., Freeman, B. M., Johnston, M. V., 2000. Periventricular leukomalacia : relationship between lateral ventricular volume on brain mr images and severity of cognitive and motor impairment. *Radiology* 214 (1), 199–204.
- Ment, L. R., Vohr, B., Allan, W., Westerveld, M., Katz, K. H., Schneider, K. C., Makuch, R. W., 1999. The etiology and outcome of cerebral ventriculomegaly at term in very low birth weight preterm infants. *Pediatrics* 104 (2), 243–248.
- Miller, D., Lippert, C., Vollmer, F., Bozinov, O., Benes, L., Schulte, D., Sure, U., 2012. Comparison of different reconstruction algorithms for three-dimensional ultrasound imaging in a neurosurgical setting. *The International Journal of Medical Robotics and Computer Assisted Surgery* 8 (3), 348–359.
- Milletari, F., Ahmadi, S.-A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Bötzel, K., *et al.*, 2017. Hough-cnn : deep learning for segmentation of deep brain regions in mri and ultrasound. *Computer Vision and Image Understanding* 164, 92–102.
- Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-net : Fully convolutional neural networks for volumetric medical image segmentation. In : *3D Vision (3DV)*, 2016 Fourth International Conference on. IEEE, pp. 565–571.
- Mitchell, T. M., Learning, M., 1997. *Mcgraw-hill science. Engineering/Math* 1, 27.
- Moeskops, P., Benders, M. J., Chiță, S. M., Kersbergen, K. J., Groenendaal, F., de Vries, L. S., Viergever, M. A., Išgum, I., 2015. Automatic segmentation of mr brain images of preterm infants using supervised classification. *NeuroImage* 118, 628–641.
- Moeskops, P., Viergever, M. A., Mendrik, A. M., de Vries, L. S., Benders, M. J., Išgum, I., 2016. Automatic segmentation of mr brain images with a convolutional neural network. *IEEE transactions on medical imaging* 35 (5), 1252–1261.
- Morgan, M. R., Broder, J. S., Dahl, J. J., Herickhoff, C. D., 2018. Versatile low-cost volumetric 3-d ultrasound platform for existing clinical 2-d systems. *IEEE transactions on medical imaging* 37 (10), 2248–2256.
- Mozaffari, M. H., Lee, W.-S., 2017. Freehand 3-d ultrasound imaging : a systematic review. *Ultrasound in medicine & biology* 43 (10), 2099–2124.

-
- Nagdyman, N., Walka, M. M., Kampmann, W., Stöver, B., Obladen, M., 1999. 3-d ultrasound quantification of neonatal cerebral ventricles in different head positions. *Ultrasound in medicine & biology* 25 (6), 895–900.
- Nair, V., Hinton, G. E., 2010. Rectified linear units improve restricted boltzmann machines. In : *Proceedings of the 27th international conference on machine learning (ICML-10)*. pp. 807–814.
- Nelson, T. R., Pretorius, D. H., 1997. Interactive acquisition, analysis, and visualization of sonographic volume data. *International Journal of Imaging Systems and Technology* 8 (1), 26–37.
- Neshat, H., Cool, D. W., Barker, K., Gardi, L., Kakani, N., Fenster, A., 2013. A 3d ultrasound scanning system for image guided liver interventions. *Medical physics* 40 (11).
- Netter, F. H., 2019. *Atlas d'anatomie humaine*. Elsevier Masson.
- Osher, S., Sethian, J. A., 1988. Fronts propagating with curvature-dependent speed : algorithms based on hamilton-jacobi formulations. *Journal of computational physics* 79 (1), 12–49.
- Pan, S. J., Yang, Q., 2009. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22 (10), 1345–1359.
- Papile, L.-A., Burstein, J., Burstein, R., Koffler, H., 1978. Incidence and evolution of subependymal and intraventricular hemorrhage : a study of infants with birth weights less than 1,500 gm. *The Journal of pediatrics* 92 (4), 529–534.
- Pappas, A., Adams-Chapman, I., Shankaran, S., McDonald, S. A., Stoll, B. J., Laptook, A. R., Carlo, W. A., Van Meurs, K. P., Hintz, S. R., Carlson, M. D., *et al.*, 2018. Neurodevelopmental and behavioral outcomes in extremely premature neonates with ventriculomegaly in the absence of periventricular-intraventricular hemorrhage. *JAMA pediatrics* 172 (1), 32–42.
- Pierrat, V., Marchand-Martin, L., Arnaud, C., Kaminski, M., Resche-Rigon, M., Lebeaux, C., Bodeau-Livinec, F., Morgan, A. S., Goffinet, F., Marret, S., *et al.*, 2017. Neurodevelopmental outcome at 2 years for preterm children born at 22 to 34 weeks' gestation in france in 2011 : Epipage-2 cohort study. *bmj* 358, j3448.
- Poulsen, C., Pedersen, P. C., Szabo, T. L., 2005. An optical registration method for 3d ultrasound free-hand scanning. In : *IEEE Ultrasonics Symposium, 2005. Vol. 2. IEEE*, pp. 1236–1240.
- Prager, R. W., Gee, A. H., Treece, G. M., Cash, C. J., Berman, L. H., 2003. Sensorless freehand 3-d ultrasound using regression of the echo intensity. *Ultrasound in medicine & biology* 29 (3), 437–446.
- Prastawa, M., Gilmore, J. H., Lin, W., Gerig, G., 2005. Automatic segmentation of mr images of the developing newborn brain. *Medical image analysis* 9 (5), 457–466.
- Prevost, R., Salehi, M., Sprung, J., Ladikos, A., Bauer, R., Wein, W., 2017. Deep learning for sensorless 3d freehand ultrasound imaging. In : *International conference on medical image computing and computer-assisted intervention*. Springer, pp. 628–636.

-
- Qiu, W., Chen, Y., Kishimoto, J., de Ribaupierre, S., Chiu, B., Fenster, A., Yuan, J., 2017. Automatic segmentation approach to extracting neonatal cerebral ventricles from 3d ultrasound images. *Medical image analysis* 35, 181–191.
- Qiu, W., Yuan, J., Kishimoto, J., McLeod, J., Chen, Y., de Ribaupierre, S., Fenster, A., 2015a. User-guided segmentation of preterm neonate ventricular system from 3-d ultrasound images using convex optimization. *Ultrasound in medicine & biology* 41 (2), 542–556.
- Qiu, W., Yuan, J., Rajchl, M., Kishimoto, J., Chen, Y., de Ribaupierre, S., Chiu, B., Fenster, A., 2015b. 3d mr ventricle segmentation in pre-term infants with post-hemorrhagic ventricle dilatation (phvd) using multi-phase geodesic level-sets. *NeuroImage* 118, 13–25.
- Roberts, G., Howard, K., Spittle, A. J., Brown, N. C., Anderson, P. J., Doyle, L. W., 2008. Rates of early intervention services in very preterm children with developmental disabilities at age 2 years. *Journal of paediatrics and child health* 44 (5), 276–280.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net : Convolutional networks for biomedical image segmentation. In : *International Conference on Medical image computing and computer-assisted intervention*. Springer, pp. 234–241.
- Salerno, C. C., Pretorius, D. H., Hilton, S. v. W., O’Boyle, M. K., Hull, A. D., James, G. M., Riccabona, M., Mannino, E., Craft, A., Nelson, T. R., 2000. Three-dimensional ultrasonographic imaging of the neonatal brain in high-risk neonates : preliminary study. *Journal of ultrasound in medicine* 19 (8), 549–555.
- San José-Estépar, R., MARTIn-FERNAndez, M., Caballero-Martínez, P. P., Alberola-López, C., Ruiz-Alzola, J., 2003. A theoretical framework to three-dimensional ultrasound reconstruction from irregularly sampled data. *Ultrasound in medicine & biology* 29 (2), 255–269.
- Sanroma, G., Benkarim, O. M., Piella, G., Ballester, M. Á. G., 2016. Building an ensemble of complementary segmentation methods by exploiting probabilistic estimates. In : *International Workshop on Machine Learning in Medical Imaging*. Springer, pp. 27–35.
- Santurkar, S., Tsipras, D., Ilyas, A., Madry, A., 2018. How does batch normalization help optimization? In : *Advances in Neural Information Processing Systems*. pp. 2483–2493.
- Sciolla, B., Martin, M., Delachartre, P., Quetin, P., 2016. Segmentation of the lateral ventricles in 3d ultrasound images of the brain in neonates. In : *2016 IEEE International Ultrasonics Symposium (IUS)*. IEEE, pp. 1–4.
- Sendur, L., Selesnick, I. W., 2002. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on signal processing* 50 (11), 2744–2756.
- Serenius, F., Källén, K., Blennow, M., Ewald, U., Fellman, V., Holmström, G., Lindberg, E., Lundqvist, P., Maršál, K., Norman, M., *et al.*, 2013. Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in sweden. *Jama* 309 (17), 1810–1820.

-
- Solberg, O. V., Lindseth, F., Torp, H., Blake, R. E., Hernes, T. A. N., 2007. Freehand 3d ultrasound reconstruction algorithms—a review. *Ultrasound in medicine & biology* 33 (7), 991–1009.
- Song, Z., Tustison, N., Avants, B., Gee, J. C., 2006. Integrated graph cuts for brain mri segmentation. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 831–838.
- Spittle, A. J., Cheong, J., Doyle, L. W., Roberts, G., Lee, K. J., Lim, J., Hunt, R. W., Inder, T. E., Anderson, P. J., 2011. Neonatal white matter abnormality predicts childhood motor impairment in very preterm children. *Developmental Medicine & Child Neurology* 53 (11), 1000–1006.
- Srhøj-Egekher, V., Benders, M. J., Viergever, M. A., Išgum, I., 2013. Automatic neonatal brain tissue segmentation with mri. In : *Medical Imaging 2013 : Image Processing*. Vol. 8669. International Society for Optics and Photonics, p. 86691K.
- Stanley, K. O., 2007. Compositional pattern producing networks : A novel abstraction of development. *Genetic programming and evolvable machines* 8 (2), 131–162.
- Sun, S.-Y., Gilbertson, M., Anthony, B. W., 2013. 6-dof probe tracking via skin mapping for freehand 3d ultrasound. In : *2013 IEEE 10th International Symposium on Biomedical Imaging*. IEEE, pp. 780–783.
- Sun, S.-Y., Gilbertson, M., Anthony, B. W., 2014. Probe localization for freehand 3d ultrasound by tracking skin features. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 365–372.
- Tabrizi, P. R., Obeid, R., Cerrolaza, J. J., Penn, A., Mansoor, A., Linguraru, M. G., 2018. Automatic segmentation of neonatal ventricles from cranial ultrasound for prediction of intraventricular hemorrhage outcome. In : *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 3136–3139.
- Treece, G. M., Gee, A. H., Prager, R. W., Cash, C. J., Berman, L. H., 2003. High-definition freehand 3-d ultrasound. *Ultrasound in medicine & biology* 29 (4), 529–546.
- Tusor, N., Benders, M. J., Counsell, S. J., Nongena, P., Ederies, M. A., Falconer, S., Chew, A., Gonzalez-Cinca, N., Hajnal, J. V., Gangadharan, S., *et al.*, 2017. Punctate white matter lesions associated with altered brain development and adverse motor outcome in preterm infants. *Scientific reports* 7 (1), 13250.
- Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 1999. Automated model-based tissue classification of mr images of the brain. *IEEE transactions on medical imaging* 18 (10), 897–908.
- Wang, L., Gao, Y., Shi, F., Li, G., Gilmore, J. H., Lin, W., Shen, D., 2015. Links : Learning-based multi-source integration framework for segmentation of infant brain images. *NeuroImage* 108, 160–172.
- Wang, P., Cuccolo, N. G., Tyagi, R., Hacihaliloglu, I., Patel, V. M., 2018. Automatic real-time cnn-based neonatal brain ventricles segmentation. In : *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, pp. 716–719.

-
- Weisenfeld, N. I., Warfield, S. K., 2009. Automatic segmentation of newborn brain mri. *Neuroimage* 47 (2), 564–572.
- Wolpert, D. H., Macready, W. G., *et al.*, 1997. No free lunch theorems for optimization. *IEEE transactions on evolutionary computation* 1 (1), 67–82.
- Wright, R., Kyriakopoulou, V., Ledig, C., Rutherford, M. A., Hajnal, J. V., Rueckert, D., Aljabar, P., 2014. Automatic quantification of normal cortical folding patterns from fetal brain mri. *NeuroImage* 91, 21–32.
- Yan, P., Cheeseborough III, J. C., Chao, K. C., 2012. Automatic shape-based level set segmentation for needle tracking in 3-d trus-guided prostate brachytherapy. *Ultrasound in medicine & biology* 38 (9), 1626–1636.
- Zhang, W. Y., Rohling, R. N., Pai, D. K., 2004. Surface extraction with a three-dimensional freehand ultrasound system. *Ultrasound in medicine & biology* 30 (11), 1461–1473.
- Zhou, Y.-T., Chellappa, R., 1988. Computation of optical flow using a neural network. In : *IEEE International Conference on Neural Networks*. Vol. 1998. pp. 71–78.
- Zubiaurre-Elorza, L., Soria-Pastor, S., Junqué, C., Fernandez-Espejo, D., Segarra, D., Bargalló, N., Romano-Berindoague, C., Macaya, A., 2012. Thalamic changes in a preterm sample with periventricular leukomalacia : correlation with white-matter integrity and cognitive outcome at school age. *Pediatric research* 71 (4-1), 354.



FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : MARTIN

DATE de SOUTENANCE : 03/12/2019

Prénoms : Matthieu, Henri

TITRE : Reconstruction 3D de données échographiques du cerveau du prématuré et segmentation automatique des ventricules cérébraux et thalami par apprentissage supervisé.

NATURE : Doctorat

Numéro d'ordre : 2019LYSEI118

Ecole doctorale : Electronique, Electrotechnique et Automatiques (ED 160)

Spécialité : Traitement du signal et de l'image

RESUME : Environ 15 millions d'enfants naissent prématurément chaque année dans le monde. Ces patients peuvent présenter des anomalies du développement cérébral qui peuvent causer des troubles du neuro-développement : paralysie cérébrale, surdité, cécité, retard du développement intellectuel, ... En pratique clinique, la quantification du volume des structures cérébrales est un bon indicateur qui permet de réduire ces risques et de les pronostiquer pour orienter les patients dans des parcours de soins adaptés pendant l'enfance. Cette thèse a pour objectif de montrer que l'échographie 3D pourrait être une alternative à l'IRM qui permettrait de quantifier le volume des structures cérébrales chez 100 % des prématurés.

Ce travail se focalise plus particulièrement sur la segmentation des ventricules latéraux (VL) et des Thalami, il apporte trois contributions principales : le développement d'un algorithme de création de données échographiques 3D à partir d'échographie transfontanellaire 2D du cerveau du prématuré, la segmentation des ventricules latéraux et des thalami dans un temps clinique et l'apprentissage par des réseaux de neurones convolutionnels (CNN) de la position anatomique des ventricules latéraux. En outre, nous avons créé plusieurs bases de données annotées en partenariat avec le CH d'Avignon. L'algorithme de création de données échographiques 3D a été validé in-vivo où une précision de 0.69 ± 0.14 mm a été obtenue sur le corps calleux. Les VL et les thalami ont été segmentés par apprentissage profond avec l'architecture V-net. Les segmentations ont été réalisées en quelques seconde par ce CNN et des Dice respectifs de 0.828 ± 0.044 et de 0.891 ± 0.016 ont été obtenus. L'apprentissage de la position anatomique des VL a été réalisée via un CPPN (Compositional Pattern Producing Network), elle a permis d'améliorer significativement la précision de V-net lorsqu'il était composé de peu de couches, faisant passer le Dice de 0.524 ± 0.076 à 0.724 ± 0.107 dans le cas d'un réseau V-net à 7 couches.

MOTS-CLÉS : Données échographique 3D, reconstruction 3D de volume échographique, enfants prématurés, segmentation, thalami, système ventriculaire cérébral, apprentissage supervisé, CNN, CPPN

Laboratoire (s) de recherche : CREATIS

Directeur de thèse: Philippe DELACHARTRE

Président de jury : Jean-Philippe THIRAN

Composition du jury :	Delachartre, Philippe	PU	INSA Lyon	Directeur de thèse
	Mateus, Diana	PU	Ecole Centrale de Nantes	Rapporteure
	PetitJean, Caroline	MC-HDR	Université de Rouen	Rapporteure
	Sciolla, Bruno	Docteur	Eco-SI	Examineur
	Thiran, Jean-Philippe	PU	EPFL	EXamineur
	Quétin, Philippe	PH	CH d'Avignon	Examineur