



Université de Poitiers Année Universitaire 2023- 2024

U.F.R. Sciences Fondamentales et Appliquées
Master 2 TDSI, parcours Objets Connectés

Rapport de stage

Ajustement spectral avec un réseau de neurones informé par la physique pour la prédiction de la sclérose en plaques

Stage effectué entre le 02/04/2024 et le 27/09/2024

Réalisé par :

RAMCHI Wissal

Tuteur de stage :

Mme. GUILLEVIN Carole, Labcom I3M

Mme. BONNIN Landoline, Laboratoire XLIM

Enseignant référent :

Mr. RICHARD Noel, Laboratoire XLIM

Remerciements

Je tiens à exprimer ma profonde gratitude à toutes les personnes qui ont contribué à la réalisation de ce travail.

Tout d'abord, je souhaite remercier ma maman, dont l'amour inconditionnel et les encouragements constants ont été essentiels tout au long de ce parcours. Sa confiance en moi et sa présence à mes côtés, malgré la distance, ont été une source inestimable de motivation. Merci pour tous les sacrifices que tu as consentis et pour avoir toujours cru en mes capacités.

J'adresse des remerciements tout particuliers à ma tante, dont la force et le courage face au cancer m'inspirent chaque jour. C'est en pensant à toi et à ton combat que j'ai trouvé la motivation de m'engager dans cette voie, avec l'espoir de contribuer un jour à l'amélioration des soins et à la lutte contre cette maladie. Un grand merci également à toute ma famille pour leur soutien inconditionnel.

Je souhaite également exprimer ma gratitude au laboratoire I3M pour leur accueil chaleureux et leur soutien tout au long de cette expérience. Leur environnement stimulant était une source de motivation précieuse.

Je remercie sincèrement mes encadrantes de stage, Madame Carole Guillevin et Madame Landoline Bonnin, pour leur encadrement précieux, leurs conseils avisés, et leur soutien constant tout au long de ce projet. Leur expertise et disponibilité ont été des atouts indispensables à l'aboutissement de ce travail.

Enfin, je souhaite exprimer ma reconnaissance à tous les membres de l'équipe pédagogique des masters "Objets Connectés" de l'Université de Poitiers et "Intelligence Artificielle et Réalité Virtuelle" de l'Université Ibn Tofail, qui ont rendu cette expérience possible grâce à leur partenariat. Je remercie particulièrement Madame Raja Touahni pour son soutien et Monsieur Clency Perrine pour son accueil chaleureux.

Merci à tous.

TABLE DES MATIÈRES

Remerciements

| | |
|--------|---|
| Résumé | 1 |
|--------|---|

| | |
|----------|---|
| Abstract | 2 |
|----------|---|

| | |
|-----------------------|---|
| Introduction générale | 3 |
|-----------------------|---|

| | |
|------------------------------|---|
| 1 Contexte général du projet | 5 |
|------------------------------|---|

| | |
|----------------------------|---|
| 1.1 Introduction | 5 |
|----------------------------|---|

| | |
|---|---|
| 1.2 Présentation de la structure d'accueil et partenaires | 5 |
|---|---|

| | |
|-------------------------------------|---|
| 1.2.1 Structure d'Accueil | 5 |
|-------------------------------------|---|

| | |
|---------------------------------------|---|
| 1.2.2 Partenaires du Labcom | 6 |
|---------------------------------------|---|

| | |
|------------------------------------|---|
| 1.3 Personnes impliquées | 6 |
|------------------------------------|---|

| | |
|-----------------------------|---|
| 1.4 Problématique | 7 |
|-----------------------------|---|

| | |
|------------------------|---|
| 1.5 Objectif | 8 |
|------------------------|---|

| | |
|------------------------|---|
| 1.6 Missions | 8 |
|------------------------|---|

| | |
|--------------------------|---|
| 1.7 Conclusion | 9 |
|--------------------------|---|

| | |
|-----------------|----|
| 2 Etat de l'art | 10 |
|-----------------|----|

| | |
|----------------------------|----|
| 2.1 Introduction | 10 |
|----------------------------|----|

| | |
|--------------------------------------|----|
| 2.2 La sclérose en plaques | 10 |
|--------------------------------------|----|

| | |
|---|----|
| 2.2.1 Dérèglement immunologique | 10 |
|---|----|

| | |
|----------------------------|----|
| 2.2.2 Phénotypes | 11 |
|----------------------------|----|

| | |
|--|----|
| 2.2.3 Symptômes et handicaps | 12 |
|--|----|

| | | |
|----------------------------|--|-----------|
| 2.2.4 | Diagnostic et pronostic | 12 |
| 2.3 | La spectroscopie par résonance magnétique | 13 |
| 2.3.1 | La résolution spectrale | 13 |
| 2.3.2 | Les métabolites et la sclérose en plaques | 14 |
| 2.3.3 | Ajustement spectral | 15 |
| 2.4 | Les réseaux de neurones informés par la physique | 16 |
| 2.5 | Conclusion | 17 |
| 3 | Méthodologie | 18 |
| 3.1 | Introduction | 18 |
| 3.2 | Prise de recul | 18 |
| 3.3 | Etablissement du cahier des charges | 19 |
| 3.4 | Livrables | 19 |
| 3.5 | Critères de validation | 20 |
| 3.6 | Gestion des ressources | 21 |
| 3.6.1 | Ressources temporelles | 22 |
| 3.6.2 | Ressources matérielles | 24 |
| 3.6.3 | Ressources humaines | 24 |
| 3.6.4 | Ressources budgétaires | 25 |
| 3.7 | Analyse stratégique | 26 |
| 3.8 | Gestion des risques | 28 |
| 3.9 | Conclusion | 30 |
| 4 | Approches et résultats | 31 |
| 4.1 | Introduction | 31 |
| 4.2 | Approche médical et motivations | 31 |
| 4.2.1 | Motivation clinique | 31 |
| 4.2.2 | Acquisition des données | 32 |
| 4.3 | Applications | 32 |
| 4.3.1 | MRS-PINN | 33 |
| 4.3.2 | MRS-PINN V2 | 39 |
| 4.3.3 | MRS-PINN Colloc | 46 |
| 4.4 | Conclusion | 48 |
| Conclusion générale | | 49 |

| | |
|-----------------|-----------|
| Annexe | 51 |
| 4.5 Annexe 1 | 51 |
| 4.6 Annexe 2 | 52 |
| 4.7 Annexe 3 | 53 |

TABLE DES FIGURES

| | | |
|------|---|----|
| 2.1 | Altération de la gaine de la myéline | 11 |
| 2.2 | Spectre ^1H -SRM à 3t et spectre ^1H -SRM à 7T | 14 |
| 2.3 | Imagerie et spectre d'un sujet atteint de SEP et d'un autre et sain | 15 |
| 3.1 | Diagramme de GANTT | 23 |
| 3.2 | Répartition des coûts du projet en pourcentage | 26 |
| 3.3 | Analyse SWOT | 27 |
| 3.4 | Analyse des risques et contraintes dans le projet | 29 |
| 3.5 | Matrice de risque | 30 |
| 4.1 | Positionnement d'un voxel de spectroscopie | 32 |
| 4.2 | Annotation des données ^1H -SRM | 33 |
| 4.3 | Architecture du modèle MRS-PINN | 34 |
| 4.4 | Comparaison entre les résultats de MRS-PINN et LCModel | 38 |
| 4.5 | Annotation des données ^1H -SRM | 40 |
| 4.6 | Interface graphique | 41 |
| 4.7 | Architecture proposée pour le MRS-PINN V2 | 42 |
| 4.8 | Comparaison entre les résultats du MRS-PINN V2 et LCModel | 45 |
| 4.9 | Cycle en V | 52 |
| 4.10 | La méthode agile | 53 |

LISTE DES TABLEAUX

| | | |
|-----|---|----|
| 4.1 | Analyse des performances du MRS-PINN | 36 |
| 4.2 | Suite de l'analyse | 37 |
| 4.3 | Comparaison des performances avec LCModel | 38 |
| 4.4 | Etude d'ablation du MRS-PINN | 39 |
| 4.5 | Analyse des performances du MRS-PINN V2 | 44 |
| 4.6 | Comparaison des résultats du MRS-PINN V2 avec LCModel | 44 |
| 4.7 | Etude d'ablation du MRS-PINN V2 | 45 |

Résumé

L'utilisation de la Spectroscopie par Résonance Magnétique du proton ($^1\text{H-SRM}$) en milieu clinique est courante pour le diagnostic des tumeurs, car elle fournit des informations métaboliques importantes en prolongeant la séance d'Imagerie par Résonance Magnétique (IRM) de seulement 10 minutes. Cette technique a également montré des perspectives prometteuses pour le pronostic des maladies neurodégénératives. Vu son potentiel, nous explorons actuellement son application pour la Sclérose En Plaques (SEP). Un ajustement spectral précis est essentiel pour cette application, surtout pour les métabolites en faible concentration ; cependant, les solutions actuelles telles que LCModel ne répondent pas à ce besoin. Par conséquent, l'amélioration des méthodes d'ajustement spectral est importante.

Ce projet, réalisé durant mon stage au Centre Hospitalier Universitaire (CHU) de Poitiers, se concentre sur l'exploration d'une technologie de pointe appelée PINN (Physics-Informed Neural Networks). En combinant l'Intelligence Artificielle (IA) avec l'information mathématique/physique, le PINN nous a permis de surmonter les limites des méthodes existantes. Nous avons développé trois modèles : MRS-PINN, MRS-PINN v2, et MRS-PINN Colloc, chacun intégrant les connaissances physiques de manière différente.

Le modèle MRS-PINN offre une performance modérée. Alors que le modèle MRS-PINN v2, montre une haute précision avec une vitesse de 0,03 secondes. En ce qui concerne le dernier modèle, le MRS-PINN Colloc est en cours de développement et vise à améliorer la précision.

Abstract

The use of proton Magnetic Resonance Spectroscopy (^1H -MRS) in clinical settings is common for diagnosing tumors, as it provides important metabolic information by extending the Magnetic Resonance Imaging (MRI) session by just 10 minutes. This technique has also shown promise in the prognosis of neurodegenerative diseases. Given its potential, we are currently exploring its application for Multiple Sclerosis (MS). Accurate spectral adjustment is essential for this application, especially for low-concentration metabolites ; however, current solutions such as LCModel do not fully address this challenge. Therefore, improving spectral adjustment methods is crucial.

To tackle this issue, this project, realized during my internship at the university hospital of Poitiers, focuses on exploring a cutting-edge technology called Physics-Informed Neural Networks (PINN). By combining Artificial Intelligence (AI) with underlying mathematical and physical information, PINN enables us to overcome the limitations of existing methods. We have developed three models : MRS-PINN, MRS-PINN v2, and MRS-PINN Colloc, each integrating physical knowledge in a different way.

The MRS-PINN model provides moderate performance. For the MRS-PINN v2 model, it demonstrates high performance with a speed of 0.03 seconds. Meanwhile, the MRS-PINN Colloc model is under development and aims improve enhance the accuracy.

Introduction générale

La SEP est une maladie neurodégénérative qui affecte 120 000 personnes aujourd’hui en France. Elle se caractérise en imagerie par l’apparition de lésions dans le cerveau sous forme de plaques et provoque un handicap. Par conséquent un pronostic précis et une prédition fiable de l’apparition de ces lésions sont essentiels pour personnaliser les traitements, améliorer l’état des patients et prévenir l’apparition du handicap. Aujourd’hui aucune méthode ne permet de prévenir en amont la progression de la maladie.

Actuellement, le diagnostic et pronostic de cette maladie se repose sur l’IRM qui fournit des informations anatomiques, offrant ainsi une visualisation détaillée des structures cérébrales et des lésions visibles. Cependant, elle ne permet pas d’accéder à des informations métaboliques ou biochimiques, qui pourraient permettre un diagnostic précoce et la personnalisation du traitement chez les patients atteints de SEP afin de prévenir le handicap. Il existe une technique de IRM, la ^1H -SRM, qui permet d'accéder à ces informations.

Cette technique, bien qu’elle offre des perspectives prometteuses dans la littérature scientifique, n’est pas actuellement recommandée comme un outil dans le diagnostic de la SEP. L’utilisation de la ^1H -SRM pour les maladies neurodégénératives nécessite une quantification des metabolites présent en faible quantité et jouant un rôle déterminant la physiopathologie. Cependant, les logiciels actuels tel que LCModel présentent des difficultés à quantifier de tels métabolites notamment à champ magnétique élevé.

Afin de quantifier ces métabolites, un ajustement spectral est requis. À ce jour, il n’existe pas de solution automatique performante pour accomplir cette tâche. Dans la littérature scientifique, plusieurs chercheurs ont exploré des méthodes d’IA traditionnelle et d’apprentissage profond pour répondre à ce besoin, mais ils rencontrent encore des limites. Pour surmonter ces obstacles, une nouvelle méthode d’IA, nommée PINN, a récemment démontré son potentiel pour résoudre ce défi.

Ce rapport scientifique détaille les méthodes et approches suivies pour obtenir ces résultats. Le premier chapitre se concentre sur le contexte général du projet, fournissant des informations détaillées sur la structure d'accueil, les objectifs et les missions confiés. Le deuxième chapitre offre un état de l'art complet, couvrant les récentes avancées dans

le domaine de la SEP, de la SRM et des techniques d'ajustement spectral. Le troisième chapitre décrit la méthodologie choisie et le processus de planification. Enfin, le dernier chapitre présente les approches sur lesquelles nous nous sommes appuyés pour construire notre hypothèse ainsi que les résultats obtenus, accompagnés d'une analyse critique et d'une comparaison avec l'état de l'art actuel dans le domaine.

CHAPITRE 1

CONTEXTE GÉNÉRAL DU PROJET

1.1 Introduction

Dans ce chapitre, nous commencerons par présenter la structure d'accueil dans laquelle le stage a eu lieu, ses partenaires et les membres impliqués dans le projet, en mettant en lumière leurs rôles et contributions. Ensuite, nous aborderons la problématique, l'objectif, et nous définirons les missions par lesquelles nous passerons afin de mener le projet à bien.

1.2 Présentation de la structure d'accueil et partenaires

1.2.1 Structure d'Accueil

Les travaux de ce projet de stage de Master 2 Objets Connectés de l'Université de Poitiers ont été réalisés sur la plateforme ultra haut champ 7 Tesla (T) du CHU de Poitiers, plus spécifiquement au Laboratoire commun Imagerie Métabolique Multinoyaux Multiorganes (I3M) avec l'équipe DACTIM-MIS (Data Analysis and Computation Through Imaging Modeling-Mathématiques Imagerie Santé). Il a duré six mois, du 2 avril 2024 au 27 septembre 2024, avec un mois de congé du 26 juillet au 20 août 2024 inclus. Il a été financé par l'Université de Poitiers et le laboratoire XLIM.

Le laboratoire commun I3M est un centre de recherche transdisciplinaire basé sur l'imagerie métabolique des organes. Son objectif est de mettre en oeuvre des techniques innovantes de l'IA pour le traitement et l'analyse automatique des images IRM 7T multi-modalités, pour l'aide au diagnostic et au suivi thérapeutique dans les pathologies.

1.2.2 Partenaires du Labcom

Les partenaires du laboratoire sont issus des domaines de la santé et de la recherche technologique. Le **CHU de Poitiers** est un partenaire clé, doté d'une IRM 3T depuis 2011 et d'une IRM 7T depuis octobre 2019. Il est le premier établissement Français à utiliser un tel équipement à des fins cliniques. **Siemens Healthineers**, un leader mondial dans les technologies médicales, collabore également avec le laboratoire. Cette entreprise offre une gamme étendue de solutions d'imagerie médicale, y compris les IRM et les scanners. L'équipe **DACTIM-MIS**, qui fait partie du Laboratoire de Mathématique et Application (LMA) du Centre National de Recherche Scientifique (CNRS) 7348 de l'université de Poitiers, est également un partenaire. Elle réunit des radiologues, informaticiens, statisticiens et mathématiciens, concentrant ses recherches sur l'imagerie cérébrale, la modélisation. Enfin, le laboratoire **XLIM**, CNRS 7252, est aussi un partenaire. Il est un laboratoire de recherche reconnu pour son expertise dans divers domaines, notamment l'électronique, les micro-ondes, l'optique, la photonique, les mathématiques et le traitement d'images, avec des applications dans la biotechnologie, la santé, et l'économie d'énergie.

1.3 Personnes impliquées

Le projet a été réalisé sous la supervision de plusieurs personnes appartenant au laboratoire.

Madame Carole Guillemin, ingénierie de recherche, et co-encadrante de thèse de Madame Bonnin, elle est chargée de la supervision et l'orientation du projet avec un accent sur l'aspect clinique, notamment la caractérisation métabolique et la validation clinique de l'application de l'IA pour prédire l'évolution de la SEP.

Monsieur Pascal Bourdon, maître de conférences à XLIM, et encadrant de thèse de Madame Bonnin, supervise et oriente le projet en se concentrant sur l'aspect IA.

Madame Landoline Bonnin, doctorante en 2^{ème} année. Elle travaille sur le sujet de la prédiction de l'évolution de la SEP à partir des données ¹H-SRM avec l'IA.

Ces personnes ont contribué à l'orientation et à la validation des travaux tout au long du projet, jouant un rôle essentiel dans son succès.

1.4 Problématique

La SEP est aujourd’hui l’une des principales causes de handicap chez les jeunes adultes. Pourtant, ce handicap pourrait être en grande partie évité grâce à une médecine personnalisée, permettant de surveiller les lésions et de prédire leur apparition et leur évolution. Malheureusement, il n’existe pas encore d’outil suffisamment efficace pour répondre pleinement à ce besoin.

Contrairement à l’IRM, la SRM permet de quantifier divers métabolites cérébraux qui peuvent fournir des informations précieuses sur la progression de la maladie. Ces informations métaboliques pourraient ainsi améliorer significativement la précision du diagnostic et la prédition de l’évolution de la SEP. Malgré ses possibilités, l’utilisation de la SRM dans le cadre clinique pour la SEP n’est pas encore courante. De plus, plusieurs défis liés à l’ajustement spectral sont à résoudre. Celle-ci est une étape cruciale pour interpréter correctement les données ^1H -SRM.

Actuellement, certains logiciels comme LCModel sont disponibles afin d’automatiser l’ajustement spectral [1]. Bien qu’ils soient efficaces et rapides, ces outils sont contraints par leur dépendance à des bases de spectres simulés qui varient entre les hôpitaux, ce qui pose des problèmes lors de la comparaison des résultats entre différents centres. De plus, certains hôpitaux n’ont pas les moyens humains et matériels de créer ces spectres simulés. Pour finir, ces logiciels ont des limitations dans l’ajustement précis des métabolites de faible quantité et ayant une importance dans la compréhension des pathologies, en particulier à des intensités de champ magnétique élevées. Ceci limite la fiabilité de la ^1H -SRM en milieu clinique [2–4]. Tout ceci fait que la ^1H -SRM n’est pas encore reconnue comme un outil de suivi de la SEP [5]. Cependant, plusieurs études scientifiques ont démontré son intérêt pour le diagnostic des et pronostic des maladies [6–9]. Il est donc crucial d’obtenir un ajustement précis des métabolites notamment pour les metabolites de faible quantité permettant à la ^1H -SRM d’être un outil performant pour le diagnostic et pronostic de la SEP. Grâce à l’IA, la création d’un outil automatique d’ajustement spectral rapide et efficace est possible aujourd’hui. Ce projet vise à améliorer les techniques d’ajustement spectral pour exploiter ces données dans la prédition de la SEP à l’aide des techniques d’IA.

1.5 Objectif

L'objectif principal de ce projet de stage était de résoudre la problématique précédemment décrite en développant un algorithme capable d'ajuster les données $^1\text{H-SRM}$ 7T en utilisant une approche combinant apprentissage profond et des connaissances physiques, à savoir le PINN. Ce modèle devrait intégrer des contraintes physiques tels que des équations ou des principes liés à la spectroscopie dans le réseau de neurone. En respectant ces contraintes, l'algorithme vise à améliorer à la fois la performance et la rapidité des ajustements spectraux, offrant ainsi un ajustement plus précis et une prédiction plus fine de l'évolution de la SEP. Pour cela le projet est décomposé en quatre grandes missions qui doivent être validées l'une après l'autre.

1.6 Missions

Mission 1 : Développement d'une interface graphique d'ajustement spectral

Cette mission consiste à concevoir une interface graphique pour l'ajustement spectral des données en $^1\text{H-SRM}$ 7T afin de concevoir par la suite une base de données annotée manuellement. Les principales fonctionnalités incluent :

Importation de données spectrales. L'utilisateur doit pouvoir importer les données en 1 Dimension (D) après un traitement spectral intégré. Les fichiers pourront être au format *Digital Imaging and COMmunications* (DICOM) ou NumPy (NPy).

Visualisation et manipulation en 2D. Une fois les données importées, l'utilisateur doit pouvoir les visualiser et les manipuler en 2D.

Barre de contrôle. L'utilisateur doit pouvoir ajuster les courbes des métabolites avec deux sliders l'un pour modifier l'Amplitude A et l'autre pour la largeur à mi-hauteur (FWHM), notée F des raies, choisir le Type T du pic et contrôler sa visibilité.

Exportation des résultats. À la fin du processus d'ajustement, l'utilisateur doit pouvoir exporter les valeurs de A , F,T et la fréquence de résonnance P dans des formats *eXtensible Markup Language* (XML) ou *JavaScript Object Notation* (JSON). Ceci permettra la sauvegarde de l'ajustement des données $^1\text{H-SRM}$ afin de les exploiter dans la mission suivante.

Mission 2 : Annotation des données $^1\text{H-SRM}$ à l'aide de l'interface graphique

Une fois l'interface graphique développée, cette mission a pour but d'utiliser celle-ci pour ajuster manuellement 200 spectres multipathologiques collectés au sein du CHU de Poitiers afin d'obtenir les paramètres d'ajustement précis pour chaque metabolite. Ces paramètres serviront ensuite comme référence pour l'entraînement du PINN développé.

Mission 3 : Algorithme d'ajustement spectral

Cette mission vise à créer un algorithme d'ajustement spectral capable d'ajuster avec précision les métabolites cérébraux d'un spectre ^1H -SRM 7T. Pour cela, nous utiliserons un modèle avancé de réseau neuronal, le PINN. Cet algorithme sera entraîné sur les données annotées dans une étape précédente du projet. Pour finir, la dernière mission aura pour but d'évaluer et analyser les résultats avec une comparaison à la littérature scientifique.

Mission 4 : Comparaison et analyse des résultats

Cette section est dédiée à l'analyse approfondie des résultats de notre PINN par rapport à la littérature scientifique. Tout d'abord, nous examinerons un modèle fondateur en PINN appliqué aux données ^1H -SRM à 3T pour comprendre les principes sur lesquels se fond ce modèle. Ensuite, nous évaluerons ses performances en termes de rapidité et de précision, en mettant en lumière ses avantages et ses limites. Enfin, nous procéderons à une comparaison entre notre modèle et ce modèle. Cette comparaison se focalisera sur trois aspects cruciaux : l'architecture du modèle, la rapidité et la précision de l'ajustement des spectres.

1.7 Conclusion

Dans cette partie, nous avons bien défini le contexte de notre projet en abordant les parties prenantes, la problématique, les objectifs et les missions. Le prochain chapitre portera sur la réalisation d'une bibliographie complète sur les thèmes pertinents du projet, afin d'acquérir une base solide de connaissances et de mieux assimiler les enjeux du projet.

CHAPITRE 2

ETAT DE L'ART

2.1 Introduction

Ce chapitre détaille les bases théoriques sur lesquelles se repose notre projet. Il synthétise l'état de l'art sur lequel nous nous sommes appuyés pour mettre en évidence l'importance de la ^1H -SRM dans le diagnostic et le pronostic de la SEP et l'intérêt de l'amélioration de la précision de l'ajustement spectral.

2.2 La sclérose en plaques

La SEP est une maladie inflammatoire et dégénérative du Système Nerveux Central (SNC). Elle touche majoritairement les femmes entre 20 et 40 ans [10]. La cause de cette pathologie est actuellement inconnue, les hypothèses scientifiques vacillent entre une cause virale, génétique ou environnementale [11]. Elle est caractérisée par une dégradation progressive de la myéline, une gaine protectrice des axones neuronaux, essentielle pour la transmission rapide et efficace des signaux nerveux. Ce dérèglement immunologique qui sera détaillé par la suite, conduit par conséquent à des symptômes neurologiques qui affectent la qualité de vie du patient.

2.2.1 Dérèglement immunologique

La cause primitive connue est la une rupture de la Barrière Hemato-Encéphalique (BHE), une structure qui protège le SNC des agents pathogènes et des cellules immunitaires circulantes [12]. Suite à cette rupture, une perturbation du système immunitaire est

engagée et conduit ensuite à une activation anormale des lymphocytes T. Ces lymphocytes T sont censés réguler les réponses immunitaires dans des conditions saines. Or, dans les conditions de la SEP, ils ne sont pas correctement contrôlés, circulent dans le SNC et activent les cellules immunitaires locales, comme les microglies et les macrophages, qui vont ensuite détruire la myéline axonal (Fig. 2.1) [10].

La myéline joue un rôle crucial dans le bon fonctionnement du système nerveux en isolant les fibres nerveuses, permettant ainsi une conduction rapide des impulsions électriques le long des nerfs. Par conséquent, lorsque la myéline est endommagée, les signaux nerveux sont ralentis ou bloqués, entraînant plusieurs symptômes neurologiques.

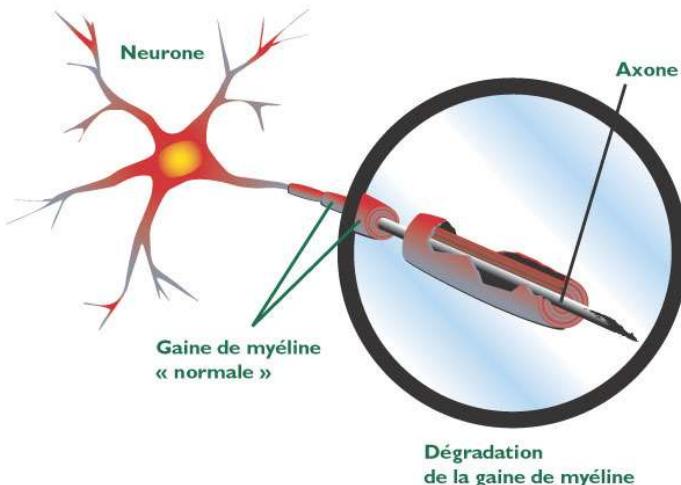


FIGURE 2.1 – Altération de la gaine de la myéline par la SEP

Cette cascade d'événement se manifeste par deux formes cliniques que nous aborderons dans la partie suivante.

2.2.2 Phénotypes

Les formes cliniques de la SEP peuvent être séparées en deux phénotypes : phénotype Rémittentes-Récurrentes (RR) et phénotype Progressives (P).

Dans la forme RR, les patients connaissent des périodes de poussées caractérisées par l'apparition de lésions et donc de symptômes neurologiques, suivies de phases de rémission où ces symptômes peuvent disparaître en partie ou complètement avec la disparition de la lésion. Pendant ces phases de rémission, le handicap ne progresse pas. En revanche, en ce qui concerne le phénotype P, la progression de la maladie se manifeste de manière continue. Elle peut comporter des périodes d'amélioration minimale mais la progression du handicap reste constante [13].

2.2.3 Symptômes et handicaps

La SEP peut avoir un impact important sur la qualité de vie des patients selon la sévérité des handicaps. Étant donné que toutes les parties du SNC peuvent être affectées, une grande variété de symptômes et/ou handicaps divers peut survenir selon le type de neurone [14]. En effet, trois types de neurones sont particulièrement concernés : les neurones cholinergiques, gabaergiques et glutamatergiques.

En cas de dysfonctionnement du neurone cholinergique, les patients peuvent présenter des symptômes affectant le système moteur et/ou sensoriel. Cela peut se traduire par des problèmes tels que la faiblesse musculaire, la paralysie partielle ou complète, ainsi que des sensations altérées comme des engourdissements, des picotements ou des brûlures dans différentes parties du corps [13].

En ce qui concerne l'altération du neurone gabaergique, cela peut entraîner des troubles cognitifs chez les patients. Ces troubles peuvent inclure des difficultés de concentration, des problèmes de mémoire, des altérations du raisonnement, de la logique et d'autres fonctions cognitives [10].

Pour finir, lorsque le fonctionnement du neurone glutamatergique est perturbé, cela peut conduire à des problèmes de mémoire. Les patients peuvent rencontrer des difficultés à se souvenir d'informations récentes, à récupérer des souvenirs à long terme ou à traiter de nouvelles informations de manière efficace [14].

2.2.4 Diagnostic et pronostic

Selon les recommandation de la Haute Autorité de Santé (AHS) le diagnostic et le pronostic de la SEP repose en partie sur l'IRM, permettant de visualiser et quantifier les lésions dans le cerveau (Fig. 2.3). En effet, chaque 6 mois, le patient réalise un examen IRM afin de surveiller l'évolution de ces lésions. Cependant, cet examen ne permet pas d'anticiper leur apparition et par conséquent prévenir le handicap.

2.3 La spectroscopie par résonance magnétique

La ^1H -SRM est une technique non invasive d'IRM permettant d'accéder aux concentrations de divers métabolites d'un échantillon 3D, nommé voxel. Cette technique ne permet pas d'obtenir une image mais un spectre. Chaque pic du spectre correspond à un groupe de protons présents dans une molécule spécifique. Cela permet d'identifier et de quantifier les différents composés moléculaires présents dans l'échantillon. Ces molécules, appelées métabolites, fournissent des informations sur le tissu étudié comme par exemple les interactions entre chaque molécule et la sécretion de celles-ci.

Si le tissu est cérébral, métabolites peuvent être utilisés pour prédire l'état et l'évolution de diverses pathologies cérébrales, telles que la SEP. la qualité des données acquise est généralement influencé par deux facteurs que nous détaillerons dans la section suivante.

2.3.1 La résolution spectrale

La qualité des spectres est principalement attribuable à deux facteurs : la puissance du champ magnétique et le rapport signal-bruit [15]. En effet, la qualité des données obtenues par la ^1H -SRM dépend fortement de la résolution spectrale. Un champ magnétique élevé permet une meilleure résolution spectrale et donc la séparation des raies dans les spectres ^1H -SRM 7T contrairement à la ^1H -SRM 3T, et par conséquent une meilleure distinction et quantification des métabolites cérébraux. En effet, nous pouvons distinguer 18 métabolites avec champ magnétique 7T par rapport à 8 avec champ magnétique 3T.

De plus, un rapport signal-bruit élevé est également crucial pour une quantification précise des métabolites. En effet, un bon rapport signal-bruit indique que le signal des métabolites est plus fort que le bruit de fond, assurant ainsi une mesure plus fiable et une meilleure quantification des concentrations de métabolites dans le cerveau.

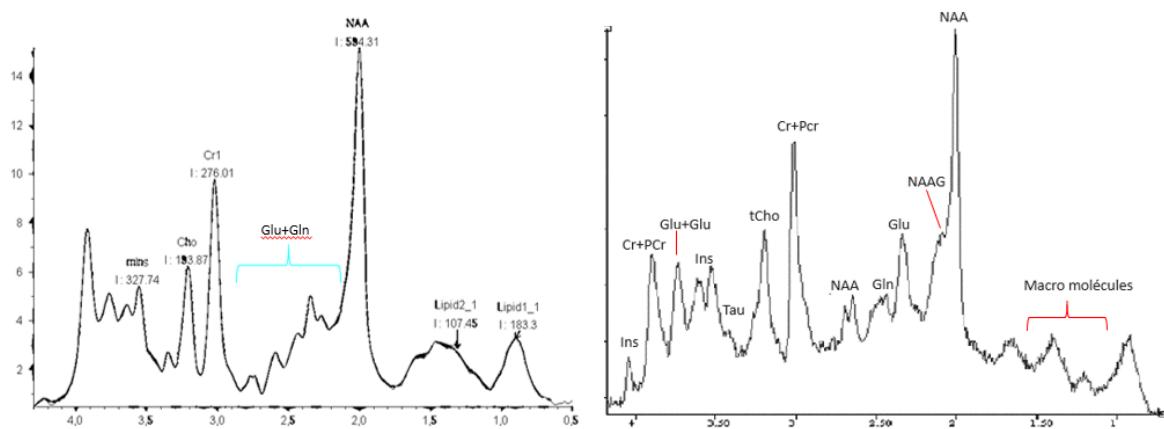


FIGURE 2.2 – A gauche : un spectre ^1H -SRM 3T annoté. **A droite :** un spectre ^1H -SRM 7T annotés. **En abscisse :** les valeurs en PPM. **En ordonnées :** les valeurs de fréquence

Ainsi ces deux facteurs nous permettent d'acquérir des données fiables afin de les exploiter dans l'étude des maladies comme la SEP.

2.3.2 Les métabolites et la sclérose en plaques

La recherche sur la SEP met en évidence le rôle crucial des métabolites dans la démyélinisation et la remyélinisation. Par exemple, le Glutamate (Glu) est un neurotransmetteur qui assure la transmission du message nerveux. Un excès de Glu entraîne une excitotoxicité et endommage les neurones et les oligodendrocytes, aggravant la démyélinisation (Fig 2.2). De même, une diminution du neurotransmetteur Gamma-AminoButyrate Acid (GABA), intensifie les effets nocifs du Glu. En ce qui concerne la remyélinisation, le Lactate (Lac) fournit de l'énergie aux cellules nerveuses, ce qui est également crucial car des niveaux faibles limitent la capacité des cellules à réparer les lésions [16].

Pour conclure, la compréhension des variations métaboliques chez un patient permettrait à travers la ^1H -SRM d'observer des phénomènes non visibles à l'IRM, prédire l'évolution de la maladie et personnaliser son traitement, évitant ainsi une aggravation de l'état du patient.

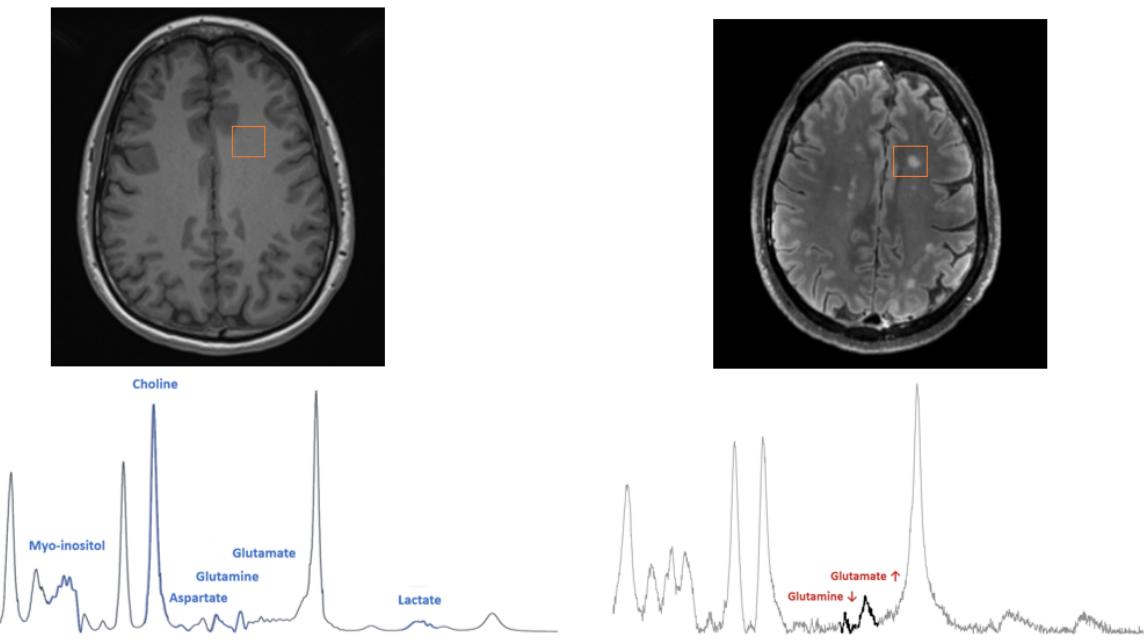


FIGURE 2.3 – **A gauche** : imagerie IRM d'un sujet sain avec le voxel de spectroscopie **en orange** placé sur un tissu sain et le spectre qui présente les concentrations normales des metabolites **A droite** : imagerie IRM du patient atteint de la SEP avec le voxel de spectroscopie **en orange** placé sur une lésion et le spectre qui présente une augmentation du glutamate et diminution de la glutamine.

Afin d'exploiter ces metabolites dans la prediction de l'évolution de la SEP, une étape de prétraitement, nommée ajustement spectrale est essentielle.

2.3.3 Ajustement spectral

Pour quantifier efficacement les spectres ^1H -SRM, une étape de post-traitement nommée ajustement spectral est nécessaire.

L'ajustement spectral souvent connu sous le terme anglais "*fitting*", est une technique d'ajustement des paramètres de fonctions mathématiques afin d'obtenir la meilleure correspondance possible entre un spectre ^1H -SRM mesuré, et établi.

Un spectre ^1H -SRM est composé d'un ensemble de pics de forme gaussienne ou lorentzienne. Ces pics sont également caractérisés par (P), (A) et F [17]. L'ajustement du spectre se fait en ajustant ces paramètres de la fonction gaussienne ou lorentzienne de manière itérative jusqu'à ce que l'écart entre les données spectrales y_i et l'ajustement \hat{y}_i soit minimisé.

Les pics de forme lorentziennes peuvent être décrites dans cette équation :

$$F_L = \frac{A}{1 + \left(\frac{f-P}{F/2}\right)^2} \quad (2.1)$$

Tandis que les raies gaussiennes peuvent être décrites par l'équation (2.2) :

$$F_G = A \times e^{-4 \ln(2) \left(\frac{f-P}{F}\right)^2} \quad (2.2)$$

Ainsi, un spectre $^1\text{H-SRM}$ peut être exprimé comme une somme de N fonctions lorentziennes F_L ou gaussiennes F_G et noté S selon le Type T de la raie :

$$S(f) = \sum_{i=1}^N S_i(A; F, T) \quad (2.3)$$

Actuellement, deux logiciels offrent cette possibilité d'ajustement et sont fréquemment utilisés dans l'application clinique : LCModel et jMRUI [1]. LCModel analyse les signaux dans le domaine fréquentiel, tandis que jMRUI fonctionne dans le domaine temporel. Ces logiciels sont rapides et performants sur des spectres $^1\text{H-SRM}$ 3T. Cependant, ils sont moins performants sur des spectres $^1\text{H-SRM}$ 7T, en particulier sur les métabolites en faible quantité.

2.4 Les réseaux de neurones informés par la physique

Les chercheurs ont montré un grand intérêt à résoudre la problématique d'ajustement des données $^1\text{H-SRM}$ 7T. Des réseaux neuronaux traditionnels ont été explorés à cet effet [18]. Cependant, ils présentent des limitations dans le cadre clinique : i) ils nécessitent une grande base de données d'entraînement ii) ils ne prennent pas en considération la réalité physique [19–21].

Dans ce contexte, le PINN a émergé comme une solution prometteuse pour des tâches complexes contraintes par des aspects physiques. Contrairement aux approches traditionnelles basées sur les données, les PINN intègrent l'information physique/mathématique de la donnée. Cette approche a été initiée par Dissanayake et al. en 1980 et a démontré son succès dans l'article de Raissi et al. en 2019 [22, 23].

Récemment, divers domaines ont montré un intérêt pour les applications des PINN. En biologie et en médecine, par exemple, ils ont été utilisés pour la prédiction de l'hémo-

dynamique en 4D et pour déterminer les distributions de température dans les tissus mammaires afin d'identifier la présence potentielle de tumeurs [24, 25]. Un modèle PINN a également été employé pour prédire la propagation des maladies inflammatoires chroniques de l'intestin, telles que la maladie de Crohn et la colite ulcéreuse [26].

De plus, Shamaei et al. a exploré son application pour ajuster les données ^1H -SRM 3T afin de quantifier les métabolites [27]. Leur étude a montré un grand potentiel pour relever ce défi. Cependant, des difficultés persistent, en particulier dans l'ajustement précis des données ^1H -SRM de faible quantité et des données pathologiques.

Il existe une limitation significative à ce travail. D'une part, l'utilisation des données simulées et de sujets sains à 3T a abouti à un modèle performant dans des conditions idéalisées, mais sujet à des erreurs lors de la prédiction sur des données avec des variations pathologiques. En revanche, ces variations sont plus distinguables sur un spectre 7T, d'où la nécessité d'explorer davantage des modèles capables de mieux gérer les données provenant d'un champ magnétique plus élevé.

2.5 Conclusion

La réalisation de cet synthèse bibliographique détaillée, qui regroupe tous les thèmes pertinents du projet, nous a permis de maîtriser le sujet et de clarifier les exigences pour atteindre nos objectifs. La prochaine étape consistera donc à élaborer une planification méthodologique que nous suivrons tout au long du projet.

CHAPITRE 3

MÉTHODOLOGIE

3.1 Introduction

Dans ce chapitre, nous allons détailler les approches méthodologiques adoptées pour structurer et conduire le projet de manière rigoureuse et organisée. Nous aborderons les outils et les techniques de gestion de projet utilisés pour planifier, exécuter, surveiller et contrôler les différentes missions du projet.

3.2 Prise de recul

Le projet a été décortiqué en quatre missions importantes précédemment citées, chacune ayant des exigences spécifiques et des objectifs clairement définis. Les attendus sont bien définis et les missions clairement énoncées, il est donc préférable d'adopter une méthodologie classique, structurée et stricte, avec une planification rigoureuse. Cette approche exige que chaque tâche soit finalisée et validée avant de passer à la suivante, ce qui garantit le respect des échéances et la qualité du travail. C'est pour cela que nous avons suivi l'approche en cycle V (Annexe 1).

Cependant, étant donné que ce projet de recherche a un caractère exploratoire, qui s'inscrit dans le domaine de la recherche, il est également important de faire preuve de flexibilité en intégrant des éléments itératifs, comme le propose la méthode agile (Annexe 2). Cela permettrait d'ajuster le travail en fonction des découvertes et des évolutions du projet, tout en maintenant une approche méthodique. Par conséquent, la combinaison des

deux méthodologies semble être la solution la plus adaptée, offrant à la fois la flexibilité nécessaire pour répondre aux besoins changeants du projet et une structure rigoureuse pour assurer une gestion efficace des missions. Suite à cette analyse, un cahier des charges a été élaboré accompagné une planification des délais et des jalons définis pour valider chaque mission avant de passer à la suivante. Il est important de noter que cette planification représente une estimation du temps requis, que nous nous efforcerons de respecter, malgré les contraintes inattendues et les éventuels changements au cours du projet.

3.3 Etablissement du cahier des charges

La première étape de la planification du projet était la création du cahier des charges. Après avoir réalisé une bibliographie détaillée et assimilé les principaux axes du projet, l'étape suivante consiste à élaborer un cahier des charges détaillé. Ce document avait pour objectif de décrire minutieusement le projet, ses objectifs, les missions, les critères de validation, les risques et obstacles potentiels à chaque mission. L'objectif de celui-ci était de comprendre le projet en le décomposant des plus grandes missions à ceux plus petites pour assurer sa faisabilité. Après avoir réalisé cette étape, nous passerons à la programmation des livrables.

3.4 Livrables

Afin d'assurer la réussite du projet, nous avons planifié la livraison progressive de ces éléments :

Livrable du cahier des charges prévu pour le 19 avril 2024, ce document détaillera les exigences, les objectifs, les fonctionnalités attendues, les contraintes et les délais du projet.

Livrable de l'interface graphique attendu pour le 26 avril 2024, ce livrable comprendra la conception et le développement de l'interface graphique d'ajustement spectral.

Livrable de l'annotation des données prévu pour le 03 mai 2024, ce livrable consiste d'annoter une base de données de 200 spectres $^1\text{H-SRM}$ 7T à l'aide de l'interface graphique.

Livrable du modèle prévu pour le 07 juin 2024. Il consiste à développer un modèle PINN pour l'ajustement spectral.

Livrable de la comparaison entre la littérature scientifique et notre modèle

PINN attendu pour le 19 juillet 2024, ce livrable comprendra une étude détaillée d'un modèle fondateur en PINN ainsi qu'une comparaison avec notre modèle.

Livrable du rapport technique qui sera une documentation du projet et de ses réalisations à livrer du retour des vacances, le 21 août 2024.

Livrable de l'article prévu pour le 27 septembre 2024 après la fin du projet, sera un article de recherche scientifique détaillant le projet et la contribution apportée à la communauté scientifique soumis à une conférence.

Livrable du script PINN représentant l'implémentation de l'algorithme d'ajustement spectral à partir des données ^1H -SRM 7T, sera fourni avant la fin du stage. Ce livrable inclura des instructions d'utilisation et de déploiement. De plus, le code sera géré avec des sauvegardes via GitHub facilitant la collaboration.

Pour conclure, nous avons réussi à livrer avec succès tous les livrables détaillés ci-dessus, en respectant les délais, à l'exception de celui concernant l'interface graphique. En effet, ce dernier a été retardé de deux semaines en raison d'imprévus que nous aborderons dans la section réservée à la gestion des risques. Pour rattraper ce retard et faire face à ce problème, nous avons dû paralléliser certaines tâches, ce qui nous a permis de gagner du temps et de respecter les délais des autres missions. Suite à cette étape, nous passerons à la définition des critères de validation des livrables.

3.5 Critères de validation

Les livrables attendus, décrits précédemment, ne peuvent être finalisés sans des critères de validation appropriés. De plus notre projet suit la méthodologie en V, par conséquent ces critères sont importants. La validation est essentielle avant de passer à la mission suivante du projet. Pour cela, nous avons défini deux types de critères de validation : des critères de performance, et des critères de documentation.

Critères de performance

Prédiction précise. Le PINN doit garantir un ajustement spectral précis et de chaque métabolite, assurant des prédictions proches aux valeurs réelles des concentrations des métabolites. Pour cela nous avions choisi des métriques pour comparer et analyser les résultats tel que l'erreur quadratique moyenne (*Mean Square Error MSE*) et l'erreur absolue moyenne en pourcentage (*Mean Absolute Percentage Error MAPE*) (Annexe 3).

Vitesse de traitement. Le PINN doit permettre un ajustement rapide des données ^1H -SRM.

Stabilité et robustesse. Notre modèle PINN doit avoir la capacité de reproduire un ajustement spectral même en présence de variations métaboliques pathologiques importante.

Facilité d'utilisation. L'interface graphique doit être conviviale et facile à manipuler. De plus son utilisation doit permettre une annotation rapide et simple.

Critères de documentation

Structuration du script. Le script doit être organisé de manière logique en respectant les conventions de codage, pour faciliter sa compréhension et maintenance.

Commentaires explicatifs. Chaque script doit être accompagné de commentaires clairs et concis.

Documentation technique. Nous allons élaborer une documentation technique complète qui décrit l'architecture de l'algorithme, les méthodes utilisées, les paramètres et leur signification, ainsi que toute autre information technique nécessaire à la compréhension et à la modification de celui-ci avec l'outil Sphynx en format *HyperText Markup Language* (HTML).

Références et citations. Nous allons mettre en place une liste de références bibliographiques avec Zotero, pour toutes les publications, recherches ou méthodes que nous avons utilisés dans le développement du PINN, pour garantir la justification et la transparence de l'approche utilisée.

Une fois les livrables et leurs critères de validation définis, nous passerons à la gestion des ressources.

3.6 Gestion des ressources

Pour garantir la réalisation des livrables correctement et par suite le succès du projet, il est essentiel de mettre en place plusieurs types de ressources : temporelles, matérielles, humaines et budgétaires. Dans cette section, nous décrirons notre planification pour la gestion de ces ressources.

3.6.1 Ressources temporelles

Diagramme de Gantt

Pour assurer un suivi rigoureux des diverses missions du projet tout au long du stage, nous avons utilisé un diagramme de Gantt. Tout d'abord, nous avons débuté le projet en estimant le temps nécessaire pour chaque tâche en fonction de sa complexité. Tout au long de la réalisation, nous avons tenté de respecter ces délais et d'arriver à nos résultats dans les temps, en prenant en compte les contraintes rencontrées. Bien que certaines missions aient été sous-estimées, entraînant des retards, nous avons réussi à nous rattraper et à atteindre tous les objectifs fixés. Ensuite, au fur et à mesure, de nouvelles contraintes imprévues sont apparues, ce qui nous a conduit à fixer de nouveaux objectifs. Malgré ces imprévus, nous avons étudié le projet de manière plus approfondie que prévu.

Le diagramme présenté ci-dessous offre une vue d'ensemble des différentes missions et de leurs durées sur la période de 6 mois de projet.

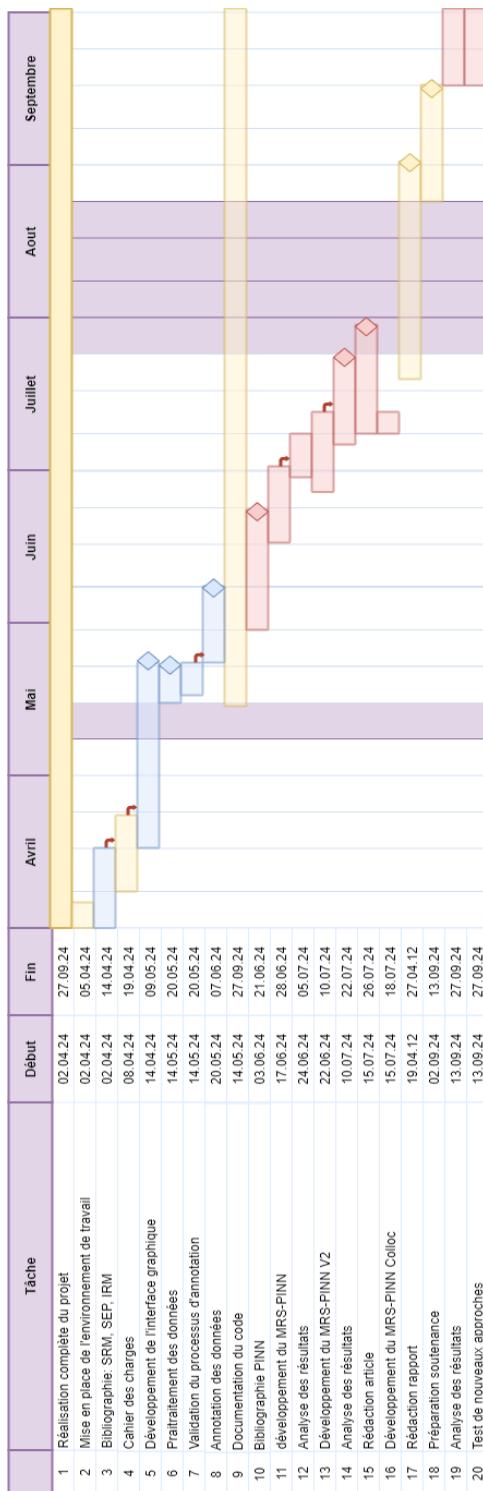


FIGURE 3.1 – Diagramme de GANTT présentant la répartition réelle des missions du projet sur la durée de stage. En **jaune** : tâche influençant l'avancement global du projet. En **Rose** : tâche liée à la partie PINN. En **bleu** : tâche liée à la partie interface graphique et préparation des données

Réunions

Afin de maintenir une communication efficace au sein de l'équipe, nous avons mis en place des réunions régulières. Chaque semaine, nous nous réunissions pour discuter des objectifs atteints, des défis rencontrés et des tâches à venir. Ces réunions hebdomadaires nous permettaient de rester alignés sur les progrès du projet et de partager des retours immédiats.

En complément, nous avons également organisé des réunions mensuelles avec nos encadrants. Ces réunions étaient importantes pour faire le point sur l'avancement global du projet. Elle permettait également de s'assurer de la validité technique et clinique des approches et obtenir des orientations pour les phases suivantes du projet.

En plus de ces réunions programmées, nous avons maintenu une communication quotidienne flexible et réactive. Selon les besoins, des réunions improvisées et des échanges fréquents ont eu lieu, permettant de traiter rapidement les imprévus et d'apporter les ajustements nécessaires pour garantir la bonne progression du projet.

3.6.2 Ressources matérielles

Afin de réaliser ce projet, plusieurs matériaux ont été mis à disposition :

Machine IRM 7T Terra Siemens Healthineers. Cet appareil permet d'obtenir des images cérébrales de haute résolution, essentielles pour la ^1H -SRM. En effet, le champ magnétique ultra-haut de 7T offre une haute résolution et par suite une précision accrue dans l'acquisition des données ^1H -SRM.

Trois ordinateurs. L'ordinateur est doté d'une configuration puissante avec une carte graphique **NVIDIA Quadro RTX 4000**, comprenant 8 Go de RAM GPU GDDR. Cette configuration est idéale pour les tâches de calcul intensif, telles que l'apprentissage profond utilisé dans le développement le PINN d'ajustement spectral.

3.6.3 Ressources humaines

Le projet comprend l'effort combiné de plusieurs professionnels. L'équipe est composée de :

Un radiologue qui est responsable de l'interprétation des résultats d'imagerie mé-

dicale, assurant la précision et la fiabilité des diagnostics. **Deux manipulateurs en électroradiologie médicale** qui sont en charge de la mise en place du patient, préparation et de la réalisation des examens IRM. De plus de **deux ingénieurs de recherche en IA et/ou imagerie médicale** qui sont en charge de la validité clinique et technique des approches utilisées et la supervision du projet. **Une doctorante** qui participe activement à la recherche et au développement d'un modèle pour la prédiction de la SEP avec les données ¹H-SRM. Et enfin **une stagiaire** qui soutient les activités de recherche pour l'amélioration de l'ajustement spectral sous la supervision des ingénieurs et de la doctorante.

Pour conclure, cette équipe est pluridisciplinaire et garantit une approche collaborative, ce qui est essentiel pour le succès du projet.

3.6.4 Ressources budgétaires

La gestion des coûts est un aspect essentiel dans ce projet, car il permet de s'assurer de son bon déroulement tout en respectant les contraintes budgétaires. Les coûts du projet se répartissent principalement en trois catégories : les coûts des examens, des équipements et salariaux.

Coûts des examens

Le projet nécessite la réalisation de 200 examens ¹H-SRM. Chaque examen coûte 270 euros, pour un coût total de 54 000 euros.

Coûts des équipements

Le projet nécessite trois ordinateurs, chacun d'un coût de 4,500 euros. Par conséquent, le coût total pour l'équipement s'élève donc à 13 500 euros.

Coûts salariaux

Les coûts salariaux pour les membres de l'équipe sur la durée totale du projet (6 mois) sont les suivants :

Une stagiaire avec un salaire de 609 euros par mois, soit un total de 3 654 euros.

Une doctorante avec un salaire de 1,500 euros par mois, soit un total de 9 000 euros.

Une ingénierie avec un salaire de 2,800 euros par mois, soit un total de 16 800 euros.

Coût total du projet

En additionnant les coûts des examens, des équipements et des salaires, nous obtenons le coût total du projet :

$$\text{Coût total du projet} = 54,000 + 13,500 + 29,454 = 96,954 \text{ euros} \quad (3.1)$$

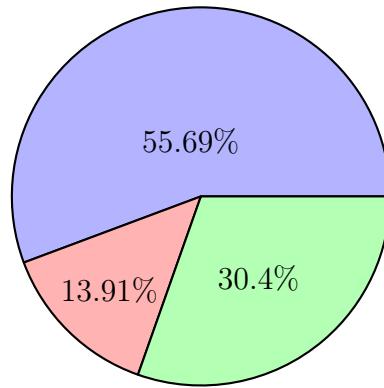


FIGURE 3.2 – Répartition des coûts du projet. **Bleu** : Coût d'examens, **Rose** : Coût des équipements, **Vert** : Coût des Salaires

Ce montant réparti sur la figure 3.2 couvre l'ensemble des aspects critiques du projet, assurant que les ressources nécessaires sont disponibles pour atteindre les objectifs fixés Eq. (3.1). Une fois les ressources et besoins dont spécifiés, nous passerons à une analyse stratégique du projet.

3.7 Analyse stratégique

Une autre étape clé pour le succès du projet est l'analyse stratégique. Cette analyse se concentre sur l'étude des facteurs internes et externes qui influencent la performance du projet. Elle peut être réalisée à l'aide d'une analyse SWOT. SWOT est l'acronyme des termes anglais *strengths*, *weaknesses*, *opportunities* et *threats*. Elle permet d'identifier les forces, faiblesses, opportunités et menaces qui entourent le projet (Fig 3.3). Cette méthode consiste à évaluer la faisabilité du projet en élaborant un plan stratégique capable de répondre aux changements et de s'adapter à l'environnement du projet.



FIGURE 3.3 – Analyse SWOT du projet mettant en évidence les **Forces**, **Faiblesses**, **Opportunités** et **Menaces** associés.

Les forces de ce projet résident dans l'utilisation de la technique innovante appelée PINN. Cette technique permet de pallier à deux contraintes majeures auxquelles d'autres solutions existantes sont confrontées : le respect de la réalité physique des données ^1H -SRM et l'indépendance aux spectres simulés. Une autre force réside dans l'utilisation des données ^1H -SRM 7T, un domaine encore peu exploré dans la littérature. De plus, ces données sont multipathologiques, contrairement à ceux utilisés dans les travaux antérieurs, qui sont souvent des spectres simulés ou proviennent de sujets sains.

En ce qui concerne les opportunités du projet, il se distingue premièrement par sa facilité d'utilisation, sa reproductibilité et ses faibles exigences en termes de ressources. Deuxièmement, il peut être généralisé et appliqué à d'autres maladies, bien qu'il soit conçu pour l'ajustement des données ^1H -SRM pour l'étude de la SEP.

Cependant, le projet est menacé par la sensibilité et la confidentialité des données, étant donné qu'il s'agit de données de patients pour lesquelles une autorisation d'utilisation

est nécessaire. De plus, les travaux actuels utilisent que des spectres simulés 3T ce qui complique la comparaison des résultats. Enfin, la qualité des résultats dépend fortement de l'exactitude de l'annotation. Par conséquent, une annotation incorrecte peut entraîner des résultats incohérents et non validés cliniquement.

Pour les limites, ce projet est limité par le manque d'informations et de publications sur la méthode PINN d'une part, étant donné qu'il s'agit d'une approche récemment explorée. D'une autre part, la complexité des spectres rend l'annotation difficile, ce qui nous limite à 13 métabolites et groupes de métabolites, en privilégiant les plus dominants. Finalement, notre base de données comprend des spectres multipathologiques. Cependant nous avons observé un schéma récurrent : la plupart des spectres suivent une forme spécifique, ce qui pourrait également limiter la performance du modèle face à des spectres plus variés.

Enfin, nous passerons à la dernière étape dans la gestion du projet. Nous allons procéder à une analyse des risques et contraintes.

3.8 Gestion des risques

Pour mener un projet à bien, il est essentiel de réaliser une analyse approfondie des contraintes et des risques. Cette analyse permet d'établir un plan d'action approprié visant à atténuer ces risques, évitant ainsi tout blocage potentiel dans la réalisation du projet. Ces risques peuvent surgir dans différentes catégories du projet, qu'elles soient financières, managériales, stratégiques ou techniques.

Citant un exemple, lors de la réalisation de la première mission du projet, qui consistait au développement de l'interface graphique, nous avons été confrontés à un risque technique majeur : l'annotation inexacte ou incorrecte des données $^1\text{H-SRM}$. Ce risque s'est révélé critique, car il pouvait compromettre la fiabilité des résultats du PINN. En effet, il n'existe actuellement aucun outil disponible dans la littérature pour l'annotation manuelle des données $^1\text{H-SRM}$. Par conséquent, nous avons dû concevoir et développer notre propre approche d'annotation.

Ce processus s'est avéré particulièrement délicat en raison de la complexité du spectre. Par conséquent, nous avons exploré plusieurs approches, notamment l'annotation de l'en-

semble des métabolites en fonction de leur fréquence de résonance, ou bien en se basant sur un spectre modèle simulé. Finalement, nous avons décidé d'utiliser une approche centrée sur des métabolites dominants, qui sont les plus pertinents pour le diagnostic des maladies du SNC 4.5. Ensuite, nous avons sollicité l'expertise d'un clinicien pour valider cette hypothèse et assurer que les résultats soient cohérents avec la réalité clinique. Ce choix a été le résultat d'un processus itératif, dans lequel chaque approche a été testée et affinée.

Cette partie du projet a été complexe et son impact était significatif, car d'une part, la performance du modèle dépend de la qualité d'annotation des données, et d'autre part, les résultats obtenus doivent être cohérents avec la réalité physique et biologique des données, et validés cliniquement.

Le tableau suivant présente les risques identifiés dans ce projet, accompagnés de leur analyse et du plan d'action proposé pour les gérer (Tab 3.4). La fréquence correspond au nombre de fois que ce risque peut survenir au cours d'une période de projet. La sévérité est mesurée par son impact sur le déroulement et l'avancement du projet.

Analyse des risques et contraintes

| N | Catégorie | Hypothèse ou contrainte | Fréquence | Sévérité | Plan d'action |
|---|------------|--|-----------|----------|---|
| 1 | Management | Capacité de calcul des machines insuffisante | 2 | 3 | Mettre le serveur du laboratoire et des machines puissantes en disposition |
| 2 | Technique | Non respect de la sensibilité et sécurité des données | 4 | 4 | Annonimation des données et partage des informations avec une clé USB ou une adresse email du laboratoire sécurisée |
| 3 | Technique | Manque de documentation et bibliographie | 3 | 3 | Recours à des professionnels pour avoir des orientations et des validations |
| 4 | Technique | Processus d'annotation des données incorrect ou inexact | 1 | 4 | Validation des hypothèses auprès des experts dans le domaine avant l'application |
| 5 | Technique | Difficulté de maintenance ou d'opérabilité sur d'autres systèmes | 1 | 3 | Documentation et structuration des codes en respectant les normes |

FIGURE 3.4 – Analyse des risques et contraintes identifiés dans le projet, triés par la catégorie. Chaque risque est identifié par un numéro et évalué en termes de fréquence et de sévérité sur une échelle de 4, avec des hypothèses ou contraintes spécifiques. Ensuite, un plan d'action est proposé pour chaque risque afin de réduire son impact potentiel sur le projet.

Chaque risque peut être classé par sa priorité en considérant deux facteurs : sa fréquence et sa sévérité. En combinant ces deux facteurs, une matrice des risques peut être établie, représentant le positionnement de chaque risque par rapport aux autres et identifiant également sa priorité (Fig. 3.5).

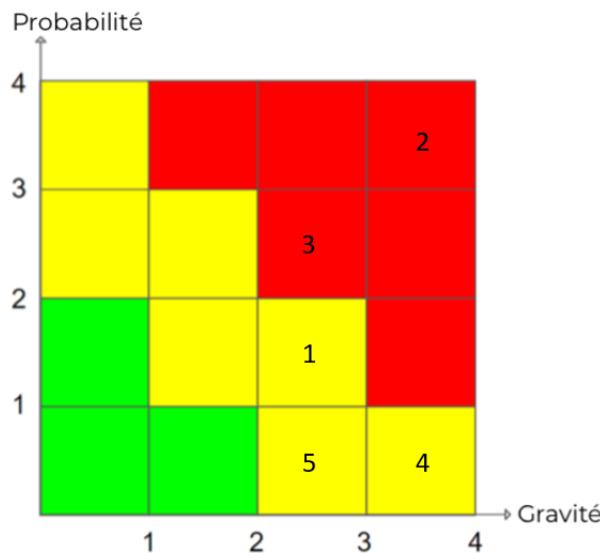


FIGURE 3.5 – Matrice de risque représentant de chaque risque identifié précédemment. les chiffres correspondent aux numéros des tâches dans le tableau précédent (Tab 3.4). **Rouge** : Priorité haute, **Jaune** : Priorité moyenne, **Vert** : Priorité basse. En **abscisse** : la gravité et en **ordonnées** : la probabilité

3.9 Conclusion

En conclusion, ce chapitre fournit une vue d'ensemble complète de la gestion du projet, mettant en lumière les stratégies que nous avons mis en place pour garantir le succès du projet et obtenir des résultats satisfaisants. Suite à cette planification, une phase de conception, développement, réalisation et d'analyse suivra. Elle sera détaillée dans le chapitre suivant.

CHAPITRE 4

APPROCHES ET RÉSULTATS

4.1 Introduction

Ce dernier chapitre présente en détail les approches médicales et d'IA adoptées au cours de ce projet, ainsi que les résultats obtenus. Le chapitre met également en lumière les contributions apportées dans le domaine de la ^1H -SRM. Chaque résultat est accompagné d'une analyse approfondie et d'une comparaison avec les travaux existants dans la littérature scientifique.

4.2 Approche médical et motivations

4.2.1 Motivation clinique

Le profil métabolique diffère entre les régions du cerveau : Matière Blanche (MB) et Matière Grise (MG) et entre les lobes cérébraux (frontal, pariétal, occipital et temporal), que ce soit dans un état sain ou pathologique [28]. Par conséquent, la position du voxel de la ^1H -SRM dans les différentes régions et lobes cérébraux de cerveaux humains sains et pathologiques ce qui fournit des profils spectraux variés [29] (Fig. 4.1).

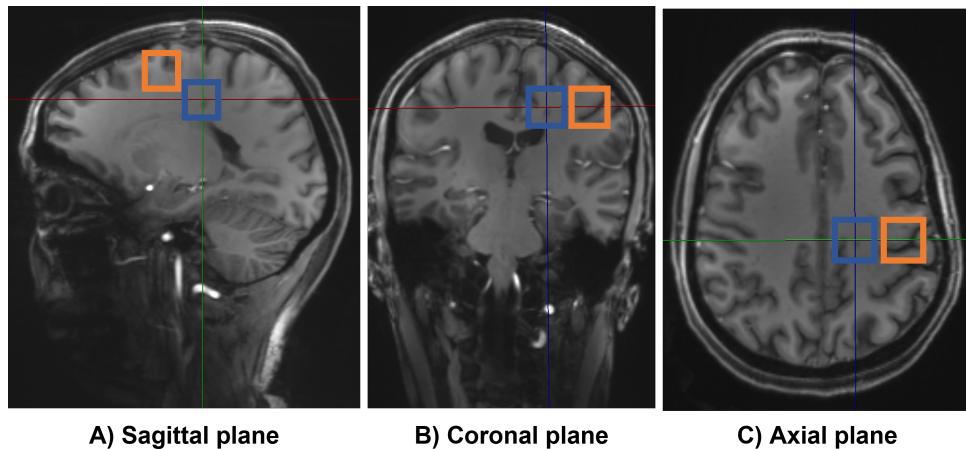


FIGURE 4.1 – Les trois plans spatiaux d'une IRM cérébrale d'un sujet sain (A, B et C). Le voxel bleu indique un exemple de position de la région MB et le voxel orange indique un exemple de position de la région MG. Les voxels bleus et rouges sont situés dans le lobe frontal.

4.2.2 Acquisition des données

Les 200 spectres ^1H -SRM ont été acquis avec une séquence de monovoxel semi-LASER, en utilisant un $\text{TR} = 5000 \text{ ms}$, $\text{TE} = 33 \text{ ms}$ avec une taille de voxel de $20 \times 20 \times 20 \text{ mm}^3$. Pour faciliter le placement du voxel, des images IRM haute résolution ont été acquises avec une séquence T1 pondérée par Écho de Gradient Rapide Préparé par Magnétisation (MPRAGE) avec les paramètres suivants : $\text{TR} = 2300 \text{ ms}$, $\text{TE} = 2.82 \text{ ms}$, $\text{TI} = 1050 \text{ ms}$ et une résolution spatiale de 1 mm^3 . Ensuite nous avons appliqué une Transformée de Fourier avec un décalage pour passer du domaine temporel au domaine fréquentiel et aligner les spectres. Sur la base de nos connaissances dans le domaine médical, nous n'utilisons que la partie réelle du signal spectral [1].

4.3 Applications

Suite à ces choix et acquisitions, nous avons pu ensuite réaliser trois modèles différents que nous avons nommés respectivement MRS-PINN, MRS-PINN V2 et MRS-PINN coloc. Chacun des modèles intègre l'information physique dans le réseau d'une manière différente.

4.3.1 MRS-PINN

Création de la base de données

La première étape de la réalisation du modèle consiste en la création de la base de données. Les données utilisées dans ce travail contiennent 90 spectres ^1H -SRM 7T pour la discrimination de 20 metabolites représentés sur la figure 4.2. Nous avons utilisé uniquement la partie réelle en ne conservant qu'une plage de 3082 fréquences après le pic de l'eau, au lieu d'utiliser l'integratif des 8192 fréquences. Cette réduction de la plage fréquentielle a été effectuée dans le but de focaliser l'analyse sur les régions spectrales les plus pertinentes pour la quantification des métabolites clinique d'intérêt. Les spectres de notre base ont été ensuite traités avec le package SPectroscopy ANalysis Tools (SPANT). Pour finir, SPANT résulte pour chaque spectre un fichier de format Npy contenant un spectre posttraité et un fichier *Comma-separated values* (CSV) comprenant la liste des metabolites d'intérêt sélectionnés avec leurs paramètres : F, P, A.

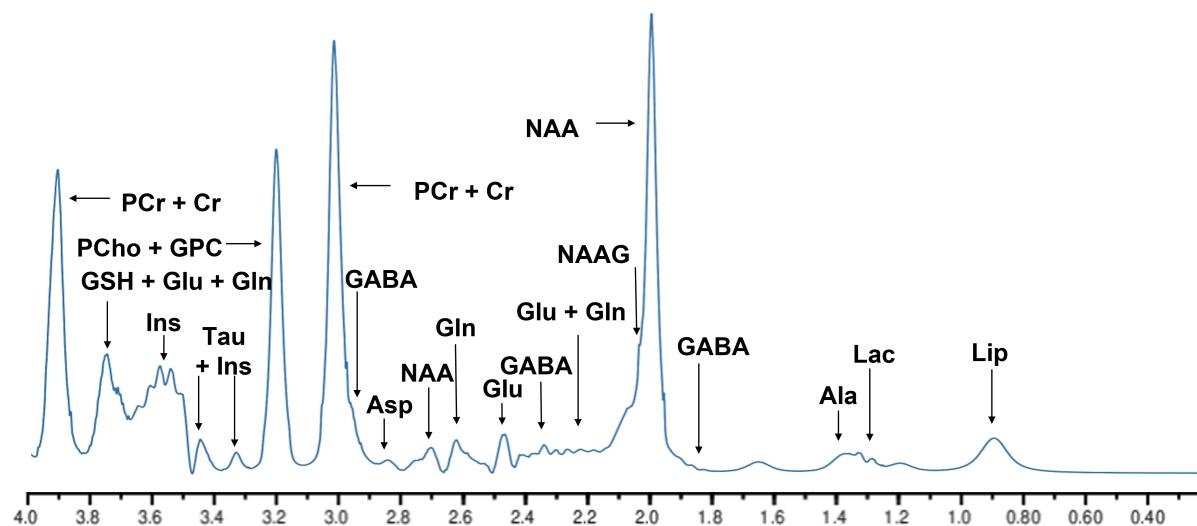


FIGURE 4.2 – Spectre ^1H -SRM 7T acquis dans un cerveau humain sain en utilisant un IRM TERRA (Siemens Healthcare, Erlangen, Allemagne). Les métabolites présents dans les spectres sont : PhosphoCréatine (PCr), Créatine (Cr), PhosphorylCholine (PCh), GlycérophosphorylCholine (GPC), Glutathion (GSH), Glutamate (Glu), Glutamine (Gln), Myo-Inositol (Myo Ins), Taurine (Tau), Scyllo-Inositol (Scyllo Ins), Acide Gamma-Aminobutyrique (GABA), Aspartate (Asp), NAA 2, Gln 2, Glu 2, GABA 2, N-AcetylAspartylGlutamate (NAAG), N-AcetylAspartate (NAA), Alanine (Ala) et Lactate (Lac). L'axe des abscisses représente les PPM du spectre de 0 à 4.

Ces paramètres extraits seront utilisés ensuite pour l'entraînement du modèle.

Création du modèle

Architecture

Ensuite pour l'architecture, le modèle "MRS-PINN" dont l'architecture est sous forme d'un encodeur-décodeur basé sur un réseau de neurones convolutifs 1D (1D-CNN) de huit couches permettra d'ajuster les données $^1\text{H-SRM}$ 7T. Le modèle reconstruit X_R à partir du spectre d'entrée X de 3082 fréquences. Les valeurs discriminées par l'encodeur sont injectées dans deux réseaux neuronaux linéaires parallèles, qui prédisent chacun deux paramètres spectraux, A_P et F_P , pour les 20 métabolites dans le spectre. Chaque réseau contient quatre couches linéaires.

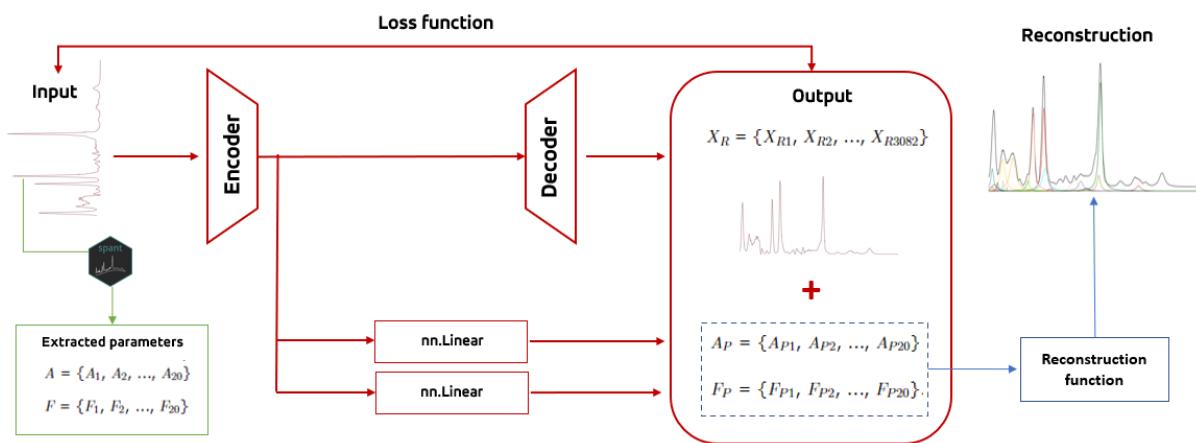


FIGURE 4.3 – Illustration de la méthode d'apprentissage pour notre modèle MRS-PINN : **En vert** : Extraction des paramètres métaboliques réels (A et F). **En rouge** : Réseau de neurones convolutifs 1D encodeur-décodeur + Réseaux prédictifs (A_P et F_P) + Construction et rétropropagation de la fonction de *Loss*. **En bleu** : Reconstruction de l'ajustement $^1\text{H-SRM}$ 7T avec les paramètres prédits (A_P et F_P).

Les paramètres prédits par le modèle serviront ensuite à implémenter la perte avec l'information physique.

Intégration de l'Information Physique

Enfin, nous intégrons l'information physique dans notre modèle en combinant trois fonctions de perte, prenant en compte la réalité physique des données $^1\text{H-SRM}$ 7T grâce aux paramètres réels A et F et prédit A_P et F_P . La Fig. 4.3 illustre l'intégration de l'information physique dans le réseau. Nous entraînons notre modèle en cherchant à minimiser le MSE entre le vecteur des paramètres réels x_i et ceux prédits y_i (Eq. 4.1). Pour cela, nous minimisons la MSE entre les composants suivants : le spectre réel X et le spectre X_R généré par le décodeur, les paramètres spectraux réels, A et F , et les para-

mètres spectraux prédits, A_P et F_P . Pour ajuster les poids du modèle, les trois MSE sont additionnées et rétropagés (Eq. 4.3).

$$MSE_{(x,y)} = \frac{1}{I} \sum_{i=1}^I (x_i - y_i)^2 \quad (4.1)$$

Où l'indice i varie de 1 à I et I représente la longueur de x et y.

$$Loss = \lambda_1, MSE_{(A,A_P)} + \lambda_2, MSE_{(F,F_P)} + \lambda_3, MSE_{(X,X_R)} \quad (4.2)$$

Où les facteurs de pondération pour chaque MSE sont λ_1 , λ_2 et λ_3 .

Ensuite, le model sera entrainé avec une configuration expérimentale de notre choix.

Configuration expérimentale

Le modèle a été implémenté à l'aide de la bibliothèque Pytorch. Il a été entraîné sur une NVIDIA Quadro RTX4000 avec 8 Go de RAM GPU GDDR6 et a pris environ 20 minutes pour 500 époques. La taille du *batch* est 50 avec 80 % d'échantillons aléatoires pour l'entraînement, 10% pour la validation et 10% pour le test. La *Perte* est minimisée par un optimiseur Adam. Suite à une étude pour déterminer les facteurs de pondération optimaux, nous utilisons les valeurs : $\lambda_1 = 1$, $\lambda_2 = 0.7$ et $\lambda_3 = 0.2$ dans la *Perte* (Eq. 4.3). Une fois l'entraînement fini, nous passerons à la reconstruction (Fig. 4.3).

Reconstruction

Pour chaque métabolite M , l'ajustement est reconstruit en utilisant leurs P, les paramètres spectraux prédits A_P et F_P , et les fonctions Gaussienne (Eq. 2.2) ou Lorentzienne (Eq.2.1). Ces fonctions définissent la réalité physique des données $^1\text{H-SRM}$ 7T. La figure 4.3 illustre la reconstruction de l'ajustement. Une fois le résultat obtenu, une étape d'analyse, évaluation et de comparaison est entamée.

Résultats et discussion

Les ensembles de données publiques $^1\text{H-SRM}$ 7T sont rares et nos données internes sont soumises à des restrictions strictes de confidentialité qui empêchent leur partage. Pour surmonter cette limitation, nous avons utilisé plusieurs métriques tel que le MSE, MAPE et le Coefficient de Détermination (R^2). Cette évaluation nous a permis de valider notre modèle de manière transparente et scientifique, malgré les limitations liées à la disponibilité des données. Cette comparaison nous a permis de conclure trois points

importants.

Tout d'abord, pour les métabolites tels que Asp, Cr, GABA, Glu, Glu 2 et GPC, le modèle présente une faible MSE, indiquant que les prédictions pour ces métabolites sont proches des valeurs réelles. Leurs valeurs de R^2 élevées indiquent que le modèle a une bonne capacité à expliquer la variance observée. De plus, leurs MAPE sont faibles, confirmant à nouveau l'exactitude des prédictions en pourcentage du modèle (Tab. 4.1).

Ensuite, pour des métabolites tels que Ala, GSH, Gln, Gln 2 et NAAG, les résultats montrent une MSE élevée malheureusement, indiquant des différences significatives entre les prédictions et les valeurs réelles. Leur R^2 est également faible ou négatif, indiquant que le modèle a des difficultés à comprendre la variance observée. Enfin, leurs MAPE sont élevés, ce qui indique que les prédictions du modèle sont moins précises en termes de pourcentage que les valeurs réelles (Tab. 4.1) (Tab. 4.2).

Enfin, le modèle montre des performances variables pour certains métabolites tels que GABA 2, Tau, Myo Ins, NAA et NAA 2, avec des MAPE et R^2 variables. Bien que certains aient des R^2 élevés, leurs MAPE peuvent varier largement, indiquant des prédictions absolues précises mais avec des erreurs en pourcentage élevées (Tab. 4.1) (Tab. 4.2).

| Entraînement MRS-PINN | | | | | | | | | | |
|-----------------------|--------------|--------------|-------------|--------------|-------|-------------|---------|-------------|-------|--------------|
| | Ala | Asp | Cr | GABA | GABA2 | Gln | Gln2 | Glu | Glu2 | GPC |
| MSE | 1.05 | 0.58 | 0.01 | 0.49 | 1.00 | 0.27 | 0.26 | 0.19 | 0.16 | 0.05 |
| R^2 | -0.05 | 0.41 | 0.98 | 0.50 | -0.02 | 0.72 | 0.73 | 0.80 | 0.83 | 0.94 |
| MAPE % | 199.04 | 67.64 | 17.55 | 47.37 | 74.74 | 1525.20 | 2196.64 | 34.16 | 32.44 | 15.55 |
| Validation MRS-PINN | | | | | | | | | | |
| MSE | 1.07 | 0.75 | 0.12 | 0.71 | 1.46 | 3.85 | 11.06 | 0.70 | 0.58 | 0.37 |
| R^2 | -0.07 | 0.24 | 0.88 | 0.28 | -0.46 | -2.85 | -10.06 | 0.29 | 0.41 | 0.62 |
| MAPE % | 63.85 | 44.44 | 10.11 | 14.22 | 20.40 | 80.07 | 121.19 | 50.69 | 46.87 | 27.07 |

^aNon Analysé.

TABLE 4.1 – Analyse de la performance d'ajustement entre notre modèle MRS-PINN. L'évaluation est réalisée sur 20 métabolites, sur les ensembles de données d'entraînement et de validation, et en utilisant les trois métriques suivantes : MSE, R^2 et MAPE.

| Entraînement MRS-PINN | | | | | | | | | | | | |
|-----------------------|-------------|-------------|-------------|-------------|-------|-------------|-------------|-------------|-------------|--------|-----|-----|
| | GSH | Myo | Ins | Lac | NAA | NAA2 | NAAG | PCh | PCr | Scyllo | Ins | Tau |
| MSE | 0.45 | 0.21 | 0.19 | 0.01 | 0.01 | 0.33 | 0.22 | 0.13 | 0.87 | 0.69 | | |
| R ² | 0.54 | 0.78 | 0.80 | 0.98 | 0.99 | 0.66 | 0.77 | 0.86 | 0.12 | 0.30 | | |
| MAPE % | 79.60 | 24.03 | 94.84 | 33.40 | 24.97 | 66.51 | 39.42 | 42.01 | 484.15 | 105.08 | | |
| Validation MRS-PINN | | | | | | | | | | | | |
| MSE | 1.61 | 0.81 | 2.55 | 0.03 | 0.03 | 0.35 | 1.17 | 0.67 | 1.25 | 0.71 | | |
| R ² | -0.61 | 0.18 | -1.55 | 0.96 | 0.96 | 0.64 | -0.17 | 0.32 | -0.25 | 0.28 | | |
| MAPE % | 79.47 | 21.41 | 72.12 | 8.62 | 10.05 | 78.35 | 53.75 | 27.39 | 617.79 | 23.57 | | |

^aNon Analysé.

TABLE 4.2 – Suite de la Table I (Tab. 4.1).

En conclusion, l'analyse des métriques MSE, MAPE et R² fournit un aperçu de la robustesse du modèle. Comme dans l'étude de Shamaei *et al.*, nous constatons que le R² est plus pertinent que le MAPE pour évaluer MRS-PINN [27] [30]. Notre analyse se concentre donc peu sur le MAPE. Les résultats de notre modèle MRS-PINN sont moyennement satisfaisant, nous avons pu avoir une bonne précision de 14 métabolites sur un total de 20 (Tab. 4.1) (Tab. 4.2). Cependant, certains métabolites nécessitent encore des améliorations. Grâce à notre méthode et à l'utilisation des spectres ¹H-SRM 7T, nous avons accès à de nouveaux métabolites, tels que Glu 2, que nous pouvons bien quantifier.

Enfin, les résultats de notre modèle MRS-PINN ont été comparés avec ceux de LCModel. Dans l'ensemble, MRS-PINN se performe mieux. En effet, sur quatre échantillons, par exemple, notre modèle présente un MSE plus faible et un R² plus élevé, indiquant une meilleure précision. Les valeurs de MAPE sont également comparables, démontrant la robustesse de notre approche proposée. Les résultats mettent en évidence l'efficacité de notre méthode par rapport aux outils existants et son potentiel pour une analyse plus précise et fiable des spectres ¹H-SRM 7T (Tab. 4.3) (Fig.4.4). Pour déterminer le principal facteur qui influence les résultats du modèle et a conduit à cette performance, nous allons réaliser une étude par ablation.

| Modèle/Logiciel | Métrique | Échantillon 0 | Échantillon 1 | Échantillon 2 | Échantillon 3 |
|-----------------|----------------|---------------|---------------|---------------|---------------|
| MRS-PINN | MSE | 0.37 | 0.30 | 0.41 | 1.46 |
| | R ² | 0.62 | 0.67 | 0.59 | -0.46 |
| | MAPE | 79.56 | 81.95 | 85.97 | 96.20 |
| LCModel | MSE | 0.41 | 0.50 | 0.45 | 23.99 |
| | R ² | 0.59 | 0.50 | 0.5447 | -22.99 |
| | MAPE | 79.77 | 83.53 | 79.39 | 206.80 |

^aAjustement de LCModel

TABLE 4.3 – Comparaison des performances d'ajustement entre notre modèle MRS-PINN et un exemple de logiciel d'ajustement de référence, LCModel. L'évaluation est réalisée sur quatre échantillons de test en utilisant trois métriques : MSE, R² et MAPE.

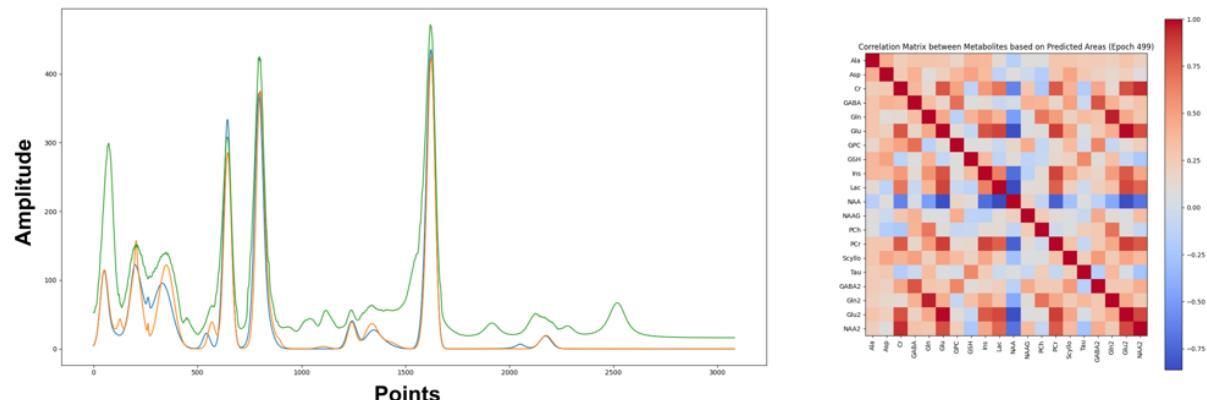


FIGURE 4.4 – **À gauche** : Prédiction de l'ajustement de l'échantillon 14 par notre modèle MRS-PINN. L'axe des x est en points du spectre de 0 à 3082. L'axe des y représente l'intensité. **À droite** : Matrice de corrélation de la prédition de concentration des 20 métabolites par notre modèle MRS-PINN à la dernière époque. L'axe des y représente la concentration réelle des 20 métabolites extraite par notre interface graphique et l'axe des x représente la concentration des 20 métabolites prédicta par notre modèle MRS-PINN. La barre de couleur représente le degré de corrélation : **En rouge** : corrélation positive, **En blanc** : corrélation nulle et **En bleu** : corrélation négative

Pour déterminer le principal facteur qui influence les résultats du modèle et a conduit à cette performance, nous allons réaliser une étude par ablation.

Étude d'ablation

Nous avons également réalisé une étude d'ablation pour examiner les performances de MRS-PINN en supprimant certains composants : les facteurs de pondération, la fonction de perte $MSE_{(A,A_P)}$ et la fonction de perte $MSE_{(F,F_P)}$. Cette étude d'ablation montre

les contributions de ces composants à notre modèle MRS-PINN (Tab. 4.4).

| Perte | MSE | R ² | MAPE |
|--------------------------------------|-------------|----------------|--------------|
| <i>Perte</i> | 0.37 | 0.62 | 79.56 |
| sans $\lambda_{1 to 3}$ | 0.37 | 0.62 | 79.05 |
| sans $MSE_{(A,A_P)}$ | 1.76 | -0.76 | 99.93 |
| sans $MSE_{(F,F_P)}$ | 1.76 | -0.76 | 99.94 |
| sans $MSE_{(A,A_P)}$ $MSE_{(F,F_P)}$ | 1.76 | -0.76 | 100.00 |

TABLE 4.4 – Évaluation des performances de notre modèle MRS-PINN avec l’ablation des facteurs de pondération, de la perte de A, de la perte de F et des deux. L’évaluation est basée sur trois métriques : MSE, R² et MAPE.

Cette étude évaluée par le MSE, R² et MAPE, met en évidence deux conclusions importantes :

- Les facteurs de pondération ont un impact minimal sur les performances du modèle.
- Les fonctions de perte $MSE_{(A,A_P)}$ et $MSE_{(F,F_P)}$ sont essentielles pour maintenir les performances du modèle.

En somme, le modèle est moyennement performant, avec des résultats nettement supérieurs à ceux de LCModel. Cependant, il rencontre des difficultés d’ajustement pour certains métabolites. Ceci est une conséquence de l’annotation imprécise des données issue de LCModel. Suite à cette conclusion, nous passerons à une deuxième étape du projet, où nous essaierons d’améliorer notre modèle avec des données annotées plus précisément.

4.3.2 MRS-PINN V2

Création de la base de données

Dans le cadre de cette deuxième phase de travail, nous avons enrichi notre base de données en l’augmentant à 200 spectres avec des spectres collectés récemment dans notre centre hospitalier. Cette expansion permet de couvrir une plus grande variété de cas cliniques afin d’améliorer la robustesse de notre modèle d’ajustement spectral. Ce modèle sera dédié à la discrimination précise de 13 métabolites et groupes de métabolites. La figure 4.5 représente les différentes entités que notre algorithme analysera.

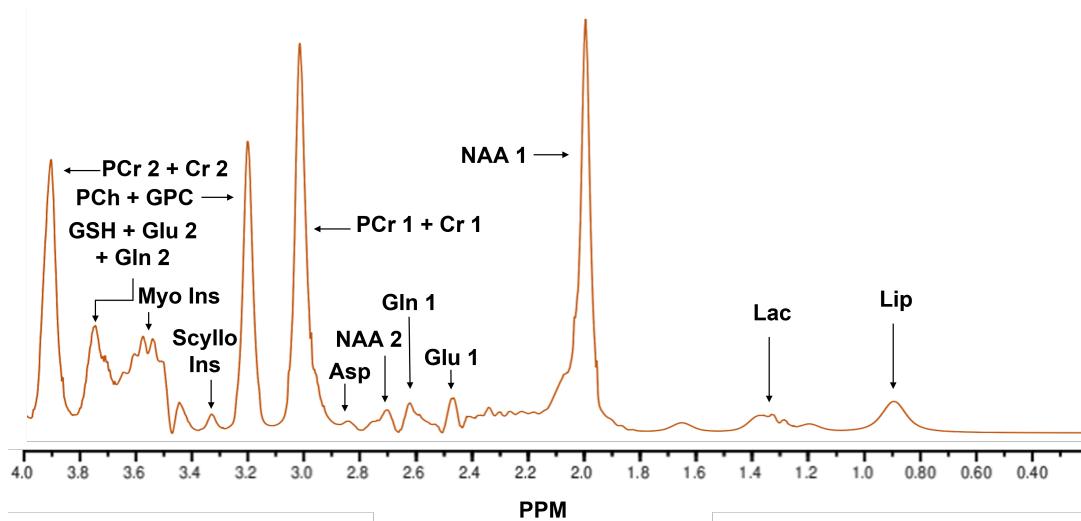


FIGURE 4.5 – ^1H -SRM 7T spectra acquises dans un cerveau humain sain à l'aide du scanner TERRA MRI (Siemens Healthcare, Erlangen, Allemagne). Les métabolites ou groupes de métabolites présents dans les spectres sont : PhosphoCréatine (PCr) 2 + Créatine (Cr) 2, Choline Phosphorylée (PCh) + Glycérophosphorylcholine (GPC), Glutamate Sulphyrule (GSH) + Glutamate (Glu) 2 + Glutamine (Gln) 2, Myo-Inositol (Myo Ins), Scyllo-Inositol (Scyllo Ins), PCr 1 + Cr 1, Aspartate (Asp), N-Acétylaspartate (NAA) 2, Gln 1, Glu 1, NAA 1, Lactate (Lac) et Lipides (Lip). L'axe des x représente les Parts Par Million (PPM) du spectre de 0 à 4.

Ensuite, afin de construire notre base d'apprentissage, nous avons procédé premièrement à l'annotation des données d'une manière différente cette fois ci (voir Fig. 3). Alors que d'autres chercheurs ont concentré leurs efforts sur le développement des interfaces avec des fonctionnalités plus avancées, incluant le prétraitement et l'ajustement automatique, notre objectif était de créer une interface simple et "conviviale", axée sur la modélisation des spectres et l'extraction des paramètres avec PYQT5 [31].

Notre outil repose sur le principe de modélisation d'un spectre comme une somme de fonctions gaussiennes et lorentziennes. Il intègre les 13 métabolites cérébraux pertinents pour l'étude des maladies du SNC. Chacun est ajusté en manipulant les paramètres spectraux de la gaussienne ou bien lorentzienne afin de se rapprocher au spectre. Ce processus itératif conduit à la minimisation le résidu spectral et par conséquent un ajustement précis.



FIGURE 4.6 – L’interface graphique développée pour l’extraction des paramètres spectraux (A, F et T). Les différents composants sont : **Partie 1** : la barre de menu, qui comprend plusieurs fonctions : importation d’un spectre "Importer les spectres" au format Numpy ou Digital Imaging and Communications in Medicine (DICOM) + importation "Importer les ajustements" ou exportation "Exporter" d'un ajustement et des paramètres spectraux au format JavaScript Object Notation (JSON) ou eXtensible Markup Language (XML) + zoom. **Partie 2** : La barre de contrôle, qui a plusieurs fonctions : liste des métabolites et de leurs paramètres à définir : A, F et T avec une barre de défilement ou une boîte. **Partie 3** : Affichage du spectre ^1H -SRM 7T et des ajustements des différents métabolites définis. L’axe des x représente les Parts Par Million (PPM) du spectre de 0 à 4. L’axe des y représente l’intensité. **Partie 4** : Exemple de fichier JSON exporté lors de l’ajustement du métabolite Glu 1.

Une fois les données sont annotées, nous passerons à la création de notre architecture PINN.

Création du modèle

Pour la réalisation du modèle, nous nous sommes inspirés de la première architecture avec des améliorations pour arriver au modèle MRS-PINN V2. Ce deuxième modèle est également d’une architecture encodeur-décodeur CNN 1D à huit couches, comprenant un encodeur, un décodeur et trois prédicteurs linéaires (Fig. 4).

Les 512 fréquences à la sortie de l’encodeur sont aplatis en une dimension. Ce tenseur est ensuite injecté dans trois réseaux neuronaux linéaires distincts pour prédire A_P , F_P et T_P .

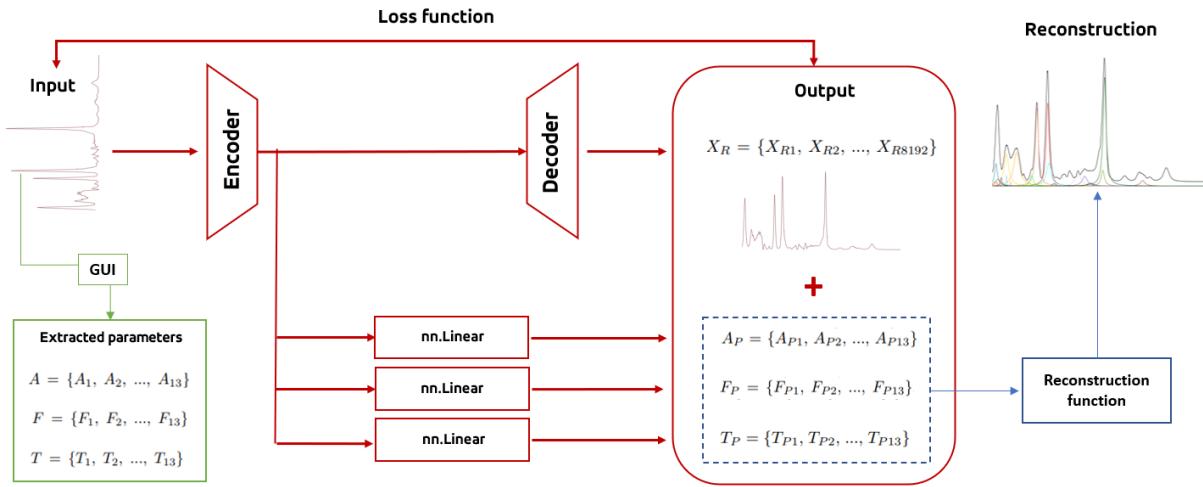


FIGURE 4.7 – Illustration de la méthode d’apprentissage pour notre modèle MRS-PINN V2 : **En vert** : Extraction des paramètres métaboliques réels (A , F et T). **En rouge** : Réseau de neurones convolutifs 1D encodeur-décodeur + Réseaux prédictifs (A_P , F_P et T_P) + Construction et rétropropagation de la fonction de *Loss*. **En bleu** : Reconstruction de l’ajustement $^1\text{H-SRM}$ 7T avec les paramètres prédits (A_P , F_P et T_P).

Les paramètres spectraux seront intégrés dans le calcul de l’information physique du modèle que nous détaillerons par la suite.

Intégration des informations physiques

De la même manière que le modèle précédent, il intègre les paramètres spectraux du spectre $^1\text{H-SRM}$ 7T dans la fonction de perte. Pendant l’entraînement, il minimise le MSE entre les valeurs observées (X : spectre d’entrée, A , F , T : paramètres spectraux) et les valeurs prédites (X_R : spectre généré, A_P , F_P , T_P : paramètres prédits).

$$\text{Perte} = \lambda_1 \text{MSE}_{(A, A_P)} + \lambda_2 \text{MSE}_{(F, F_P)} + \lambda_3 \text{MSE}_{(T, T_P)} + \lambda_4 \text{MSE}_{(X, X_R)} \quad (4.3)$$

La somme pondérée de ces MSE est rétropropagée avec des facteurs λ_1 à λ_4 pour chaque MSE (Eq. 4.3). Ensuite le modèle sera entraîné avec une configuration de notre choix, que nous présenterons dans la prochaine partie.

Configuration expérimentale

Le modèle a été implémenté à l’aide de la bibliothèque Pytorch. Il a été entraîné sur une NVIDIA Quadro RTX4000 avec 8 Go de RAM GPU GDDR6 et a pris environ 15 minutes pour 300 époques. La taille du *batch* est 50 avec 80 % d’échantillons aléatoires pour l’entraînement, 20% pour la validation et 10% de test. La *Perte* est minimisée par

un optimiseur Adam. Suite à une étude pour déterminer les facteurs de pondération optimaux, nous utilisons les facteurs de pondération : $\lambda_1 = 0.45$, $\lambda_2 = 0.25$, $\lambda_3 = 0.2$ et $\lambda_4 = 0.1$ dans la *Perte* (Eq. 4.3). Une fois l'entraînement est fini, nous passerons à l'étape de reconstruction de résultats.

Reconstruction

Les trois réseaux neuronaux linéaires séparés prédisent les paramètres (A_P , F_P et T_P) pour les 13 métabolites M dans le spectre X . Avec la valeur connue en littérature de la position P , ils sont injectés dans la fonction gaussienne ou lorentzienne en fonction de la valeur de T_P (Eq. 2.2) (Eq. 2.1) [1]. La partie bleue de la figure 4.7 illustre la construction de l'ajustement. Cette reconstruction obtenue sera ensuite utilisée afin d'évaluer le modèle et se comparer au travaux de la littérature scientifique.

Résultats et discussion

Pour la concentration des métabolites, le modèle MRS-PINN V2 fournit de bons résultats avec une MSE très faible et un R^2 élevé pour 13 des 10 métabolites. Par exemple, PCr + Cr 2 montre un ajustement très précis, tandis que PCh + GPC présente des ajustements moins satisfaisants (Tab. 4.5) (Fig. 4.8).

De plus, notre modèle donne également de bons résultats pour les métabolites en faible quantité, tels que Scyllo Ins, qui est actuellement difficile à ajuster avec les logiciels existants (Tab. 4.5). Pour les échantillons ajustés, notre modèle surpasse le logiciel LC MODEL en termes de performance globale. Le MRS-PINN V2 présente une MSE plus faible, un R^2 plus élevé, des résidus et KLPD normalisés inférieurs par rapport à LC MODEL (Tab. 4.6) (Fig. 4.8). Par exemple, pour les échantillons 0 et 33, il affiche des MSE et R^2 significativement meilleurs (Tab. 4.6). En résumé, MRS-PINN V2 offre un ajustement plus précis et réduit les erreurs comparées à LC MODEL.

| Métabolite | Asp | NAA-2 | Lac | Glu-1 | NAA-1 | Gln-1 | Lip |
|----------------|----------|------------|---------|----------|---------|---------------|------|
| MSE | 0.20 | 0.39 | 0.66 | 0.18 | 0.22 | 0.44 | 0.67 |
| R ² | 0.80 | 0.61 | 0.34 | 0.82 | 0.78 | 0.56 | 0.33 |
| Métabolite | PCr2+Cr2 | Scyllo-Ins | PCh+GPC | PCr1+Cr1 | Myo-Ins | GSH+Glu2+Gln2 | |
| MSE | 0.05 | 0.20 | 0.73 | 0.04 | 0.18 | 0.19 | |
| R ² | 0.95 | 0.80 | 0.27 | 0.96 | 0.82 | 0.81 | |

TABLE 4.5 – Scores de MSE et R² pour la concentration des métabolites à la dernière époque d’entraînement.

| Modèle/Logiciel | Métrique | Échantillon 0 | Échantillon 14 | Échantillon 24 | Échantillon 33 | Échantillon 84 |
|--------------------|------------------|---------------|----------------|----------------|----------------|----------------|
| MRS-PINN V2 | MSE | 0.37 | 0.68 | 0.80 | 0.39 | 0.35 |
| | R ² | 0.63 | 0.32 | 0.20 | 0.61 | 0.65 |
| | Résidu normalisé | 0.10 | 0.03 | 0.03 | 0.11 | 0.07 |
| | KLPD normalisé | 0.19 | 0.20 | 0.22 | 0.21 | 0.19 |
| LC MODEL | MSE | 0.52 | 0.68 | 0.62 | 0.65 | 0.45 |
| | R ² | 0.47 | 0.31 | 0.37 | 0.34 | 0.54 |
| | Résidu normalisé | 0.24 | 0.23 | 0.17 | 0.18 | 0.21 |
| | KLPD normalisé | 0.27 | 0.36 | 0.37 | 0.36 | 0.33 |

TABLE 4.6 – Scores de MSE, R², résidu normalisé et KLPD pour cinq échantillons ajustés.

Pour déterminer le principal facteur qui influence les résultats du modèle et a conduit à cette performance, nous allons réaliser une étude par ablation.

Étude d’ablation

Nous avons également réalisé une étude d’ablation pour examiner la performance du modèle MRS-PINN V2 en supprimant certains composants : le $MSE_{(A,A_P)}$, $MSE_{(F,F_P)}$, $MSE_{(T,T_P)}$, $MSE_{(X,X_R)}$ et la combinaison des fonctions de perte $MSE_{(A,A_P)} (F,F_P) (T,T_P)$. Cette étude d’ablation montre les contributions de ces composants à notre modèle MRS-PINN V2 (Tab. 4.7). Évaluée par MSE et R², l’étude met en évidence trois conclusions importantes.

TABLE 4.7 – Étude d’ablation de la performance de MRS-PINN V2.

| Métrique | Ajustement | sans A | sans F | sans T | sans X | sans AFT |
|----------------|------------|--------|---------|--------|--------|----------|
| MSE | 0.23 | 1.29 | 371.51 | 0.34 | 0.31 | 1.06 |
| R ² | 0.77 | -0.29 | -370.51 | 0.66 | 0.69 | -0.06 |

Tout d’abord, la fonction de perte $MSE_{(F,F_P)}$ est la plus critique pour notre modèle MRS-PINN V2 car sa suppression entraîne une augmentation énorme de la MSE et une forte diminution du R².

Ensuite, la suppression des fonctions de perte (T,T_P) (X,X_R) a également un effet négatif sur la performance de notre modèle MRS-PINN V2, mais l’effet est très faible et la performance reste bonne.

Enfin, la suppression des fonctions de perte $MSE_{(A,A_P)}$ ou la combinaison des fonctions de perte $MSE_{(A,A_P)}$ (F,F_P) (T,T_P) dégrade la performance de notre modèle MRS-PINN V2, mais pas autant que la suppression de $MSE_{(F,F_P)}$ seule.

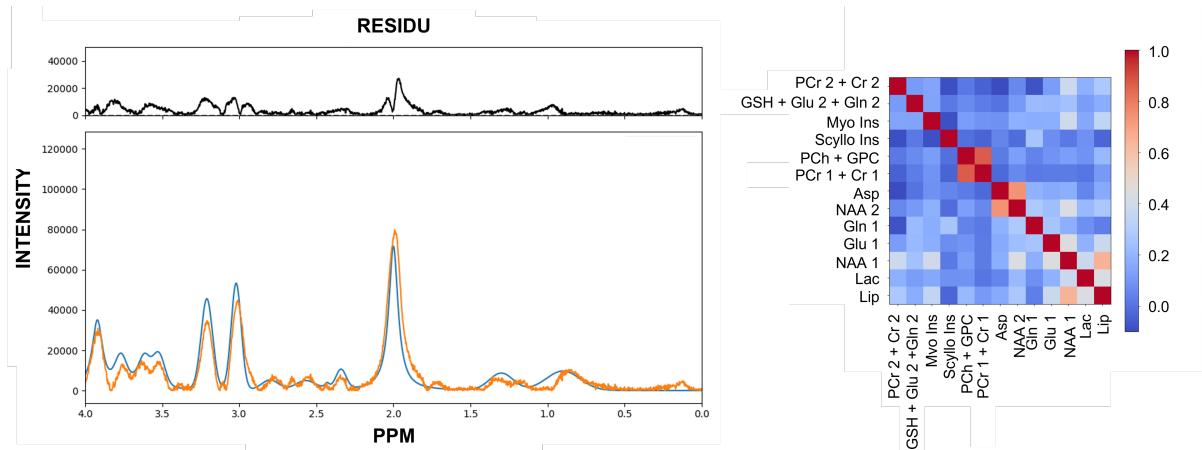


FIGURE 4.8 – **À gauche :** Prédiction de l’ajustement de l’échantillon 14 par notre modèle MRS-PINN V2 avec le résidu ci-dessus. L’axe des x représente le PPM du spectre de 0 à 4. L’axe des y représente l’intensité. **À droite :** Matrice de corrélation de la prédiction de concentration des 13 métabolites par notre modèle MRS-PINN V2 à la dernière époque. L’axe des y représente la concentration réelle des 13 métabolites extraite par notre interface graphique et l’axe des x représente la concentration des 13 métabolites prédictive par notre modèle MRS-PINN V2. La barre de couleur représente le degré de corrélation : **En rouge** : corrélation positive, **En blanc** : corrélation nulle et **En bleu** : corrélation négative.

Ce travail représente un avancement significatif dans la quantification des métabolites et ouvre la voie pour une amélioration. Dans la prochaine étape, nous allons nous concentrer sur l'amélioration de l'ajustement avec une autre technique de PINN.

4.3.3 MRS-PINN Colloc

Création des données

Pour ce travail, nous avons utilisé les mêmes données annotées de l'interface graphique que celles employées dans la section précédente du modèle MRS-PINN V2.

Création du modèle

Pour le dernier volet de projet nous avons créé le modèle nommé MRS-PINN Colloc, dans lequel nous avons réutilisé l'architecture du MRS-PINN V2, tout en apportant une modification importante dans le calcul de la perte. Ceci vise à améliorer l'efficacité et la précision du modèle tout en le contraignant plus à respecter la réalité physique des données en intégrant des équations et des points de collocations.

Points de collocation

Dans notre travail précédent, nous avons constaté que le MRS-PINN V2 offrait des performances satisfaisantes pour 10 des 13 métabolites étudiés. Afin d'améliorer ces résultats, nous avons mis en place une nouvelle approche intégrant des points de collocation.

Le principe de cette méthode consiste à renforcer l'entraînement du modèle en sélectionnant des segments de données de manière aléatoire ou stratégique, et en ajoutant des points d'entraînement supplémentaires, appelés points de collocation. Ces points permettent de mieux contraindre le modèle en minimisant l'erreur de reconstruction du spectre sur ces points.

L'intégration des équations physiques et mathématiques depuis les points de collocations permet également de généraliser le modèle à des scénarios dans lesquels les données sont rares ou bruitées, en s'appuyant sur les lois physiques pour guider l'apprentissage, au-delà des simples données d'entraînement. Ceci sera plus détaillé dans la partie suivante.

Intégration des Informations Physiques

Dans le modèle MRS-PINN Coloc, l'intégration des informations physiques a été réalisée de manière différente par rapport aux approches précédentes. Nous avons combiné trois types de fonctions de perte : la perte des paramètres spectraux, la perte de reconstruction contrainte par des équations physiques, et la perte associée aux points de collocations.

Perte des Paramètres Spectraux : La régularisation des paramètres spectraux est maintenue de manière proche à celle du modèle précédent en gardant la perte des trois paramètres, sans la perte du décodeur et sans facteurs de pondération. Cette régularisation permet d'assurer que les paramètres spectraux prédits sont conformes aux attentes théoriques (4.4).

$$\text{Perte}_{\text{reg}} = \text{MSE}_{(A, A_P)} + \text{MSE}_{(F, F_P)} + \text{MSE}_{(T, T_P)} \quad (4.4)$$

Perte de Reconstruction Physique : La perte de reconstruction dans ce modèle ne se base pas uniquement sur la sortie du décodeur. Elle est calculée à partir de la reconstruction du spectre, en utilisant les paramètres prédits injectés dans des équations gaussiennes et lorentziennes. Celles-ci sont conçues pour modéliser la réalité physique des spectres, afin d'obtenir une reconstruction plus précise et réaliste (4.5).

$$\text{Perte}_{\text{phys}} = \text{MSE}(X, S(A, F, T)) \quad (4.5)$$

Perte des Points de collocation : En complément, la perte associée aux points de collocation est utilisée pour améliorer l'entraînement du modèle. Ces points sont sélectionnés de manière stratégique sur le spectre reconstruit et placés dans les intervalles des trois métabolites pour lesquels la prédiction du modèle précédent est moins précise (4.6).

$$\text{Perte}_{\text{coloc}} = \sum_{i \in \mathcal{I}} \text{MSE}(X[i], S(A, F, T)[i]) \quad (4.6)$$

$\sum_{i \in \mathcal{I}}$ est calculée sur un ensemble d'indices i appartenant à un ensemble d'intervalles \mathcal{I} . Cela signifie que la perte totale est obtenue en additionnant les erreurs sur ces intervalles.

En combinant ces trois types de perte, le modèle est non seulement contraint par les données observées, mais aussi par les lois physiques qui les sous-tendent (4.7).

$$\text{Perte} = \text{MSE}_{\text{phys}} + \text{MSE}_{\text{reg}} + \text{MSE}_{\text{coloc}} \quad (4.7)$$

Ce travail est encore en cours de réalisation, les résultats et l'analyse sont en cours de finalisation et représentent la dernière étape prévue pour ce projet.

4.4 Conclusion

En conclusion, ce chapitre a mis en évidence les résultats obtenus tout au long du projet. Nous avons démontré l'efficacité des PINN pour l'ajustement spectral, réussissant à surmonter des défis techniques complexes et à résoudre les limitations des approches actuelles. Ces avancées confirment la pertinence de notre approche et posent les bases pour de futures recherches et applications.

Conclusion générale

Conclusion du projet

Ce travail a démontré l'efficacité des PINN pour améliorer l'ajustement spectral des données ^1H -SRM à 7T, apportant ainsi une contribution significative au domaine de la quantification des métabolites.

Le premier modèle, basé sur les données générées par LCModel, a permis de surmonter l'une des principales limitations de l'ajustement spectral classique, en éliminant la dépendance aux spectres simulés.

Ensuite, une approche avec identification manuelle à l'aide d'une interface graphique a permis de pallier une autre contrainte, à savoir l'amélioration de l'ajustement des métabolites de faible quantité dans des spectres de sujets sains et pathologiques. Le modèle montre d'excellentes performances dans la prédiction de 10 métabolites et groupes de métabolites sur un total de 13. Cependant, il a également révélé des domaines où des améliorations sont possibles pour certains autres métabolites.

Le troisième volet de notre travail visait à perfectionner les résultats obtenus et explorer une nouvelle approche combinant des fonctions de perte, des équations physiques, et des points de collocation.

En conséquence, ce travail a abouti à la publication d'un article, marquant une contribution significative à la communauté scientifique. De plus, il ouvre la voie à de nouvelles perspectives, telles que l'application de cet outil à d'autres maladies neurologiques et psychiatriques. Une autre avenue prometteuse réside dans l'extension de la méthode PINN à d'autres types de noyaux et de séquences SRM, notamment la SRM au phosphore-31 (^{31}P -SRM), qui est particulièrement pertinente pour l'étude du métabolisme énergétique. Enfin, pour élargir son utilité médicale, il sera nécessaire d'adapter la méthode à des données acquises à 3T ou bien à 1.5T, qui sont fréquemment utilisées.

En conclusion, ces avancées offrent des opportunités passionnantes pour le développement futur des PINN dans l'ajustement spectral. Cette technologie promet de transformer les approches diagnostiques et thérapeutiques, améliorant ainsi de manière significative

les soins aux patients.

Conclusion personnelle

La réalisation de ce travail m'a permis d'acquérir des compétences précieuses sur le plan professionnel aussi que personnel. J'ai approfondi mes connaissances sur des sujets complexes tels que la SEP, l'IRM, la SRM et les PINN. J'ai également exploité les compétences acquises au cours de mon parcours universitaire, notamment en traitement du signal, intelligence artificielle, mathématiques et développement web, pour concevoir une solution innovante. De plus, j'ai collaboré à l'écriture de mon premier article scientifique, qui a été publié avec succès [32]. Cette expérience a renforcé mon intérêt pour la recherche et m'a motivé à poursuivre en thèse.

Annexe

4.5 Annexe 1

La méthode en "V" est une approche systématique de gestion de projet, particulièrement utilisée pour les projets où chaque phase dépend de la réalisation de la phase précédente. Cette méthode se distingue par une progression étape par étape, garantissant que chaque phase est complétée avant de passer à la suivante.

La **descente du "V"** représente la phase d'analyse et de conception. Pendant cette phase :

- **Identification des besoins** : Les besoins du projet sont identifiés et documentés.
- **Établissement des spécifications** : Les spécifications fonctionnelles et techniques sont définies.
- **Conception générale et détaillée** : Les conceptions générale et détaillée sont réalisées, posant ainsi les fondations du projet et définissant clairement ce qui doit être réalisé.

Cette phase est cruciale car elle établit les bases sur lesquelles le projet sera construit et assure que toutes les exigences sont bien comprises et documentées.

La **phase de développement** se situe au croisement des deux branches du "V". C'est ici que le travail de mise en œuvre du système est effectué. Cette phase implique :

- **Mise en œuvre du système** : Le système est développé conformément aux conceptions établies lors de la phase descendante.
- **Tests et validations** : Les tests nécessaires sont réalisés pour assurer la qualité du produit final.

La **montée du "V"** correspond à la phase de validation. Pendant cette phase :

- **Tests unitaires** : Les différentes composantes du système sont testées individuellement.
- **Tests d'intégration** : Les composantes sont intégrées et testées ensemble pour vérifier leur interopérabilité.
- **Validation et recette** : Le produit final est testé pour s'assurer qu'il répond aux spécifications établies et qu'il est conforme aux attentes du client.

Cette phase garantit que le produit final est conforme aux attentes et aux exigences définies lors de la phase descendante.

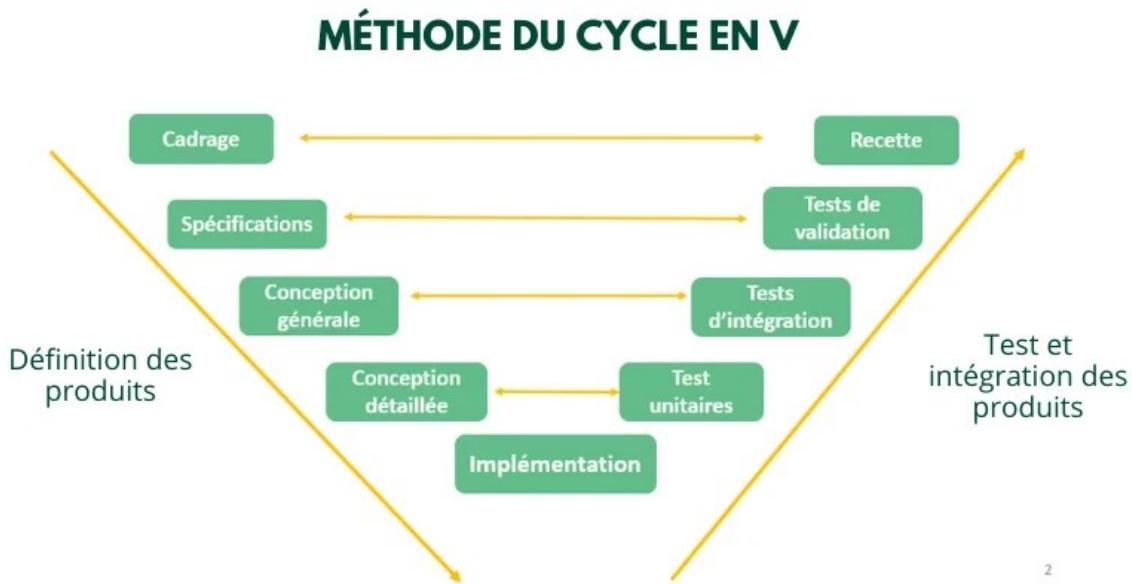


FIGURE 4.9 – Schéma modélisant les différentes phases d'un projet suivant la méthodologie cycle en V.

4.6 Annexe 2

L'Agile est un cadre de gestion de projet qui privilégie la flexibilité et l'adaptabilité. Le projet est divisé en phases courtes et dynamiques appelées sprints, chacune vise à livrer des fonctionnalités spécifiques dans un délai. Cette approche itérative permet d'ajuster le développement en continu, en fonction des retours du client. L'un des aspects clés de l'Agile est la communication continue et l'engagement de l'équipe. Les réunions et les échanges fréquents, comme les *daily stand-ups* et les rétrospectives, assurent une communication constante au sein de l'équipe, facilitant la détection rapide des problèmes et la mise en place de solutions.

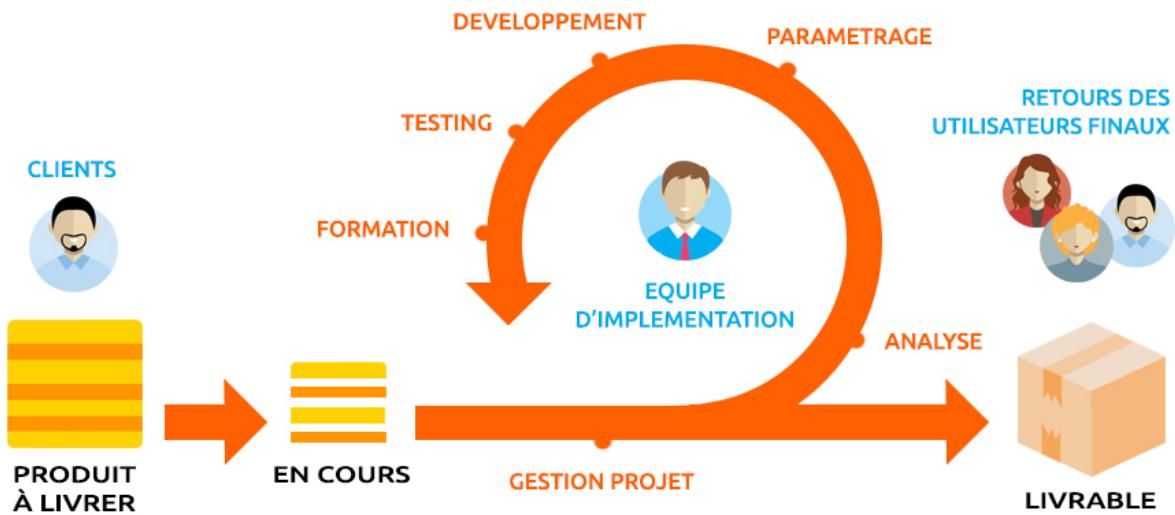


FIGURE 4.10 – Schéma modélisant la méthode agile.

4.7 Annexe 3

L'erreur quadratique moyenne

L'*erreur quadratique moyenne (MSE)* est une métrique qui mesure la moyenne des carrés des écarts entre les valeurs réelles et les valeurs prédites par le modèle. Plus le MSE est faible, plus le modèle est précis. La formule du MSE est donnée par :

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

où :

- n est le nombre total de données,
- y_i est la valeur réelle de la donnée i ,
- \hat{y}_i est la valeur prédite par le modèle pour la donnée i .

Le coefficient de détermination

Le *Coefficient de Détermination (R^2)* est une métrique qui indique la proportion de la variance des données réelles expliquée par le modèle. Un R^2 proche de 1 signifie que le modèle explique bien la variance des données, tandis qu'un R^2 proche de 0 signifie le contraire. La formule du R^2 est :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

où :

- y_i est la valeur réelle de la donnée i ,
- \hat{y}_i est la valeur prédite par le modèle pour la donnée i ,
- \bar{y} est la moyenne des valeurs réelles.

L'erreur moyenne absolue en pourcentage

L'erreur moyenne absolue en pourcentage(MAPE) mesure l'erreur moyenne en pourcentage entre les valeurs réelles et les valeurs prédites. Il s'agit d'une métrique interprétable qui exprime l'erreur en termes relatifs, ce qui la rend utile pour comparer la précision des prédictions entre différents modèles. La formule du MAPE est :

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100$$

où :

- n est le nombre total de données,
- y_i est la valeur réelle de la donnée i ,
- \hat{y}_i est la valeur prédite par le modèle pour la donnée i .

La pseudo divergence de Kullback-Leibler

La divergence de Kullback-Leibler (Divergence KL) est une mesure de la différence entre deux distributions de probabilité : la distribution réelle P et la distribution prédite Q . Elle quantifie la perte d'information lors de l'utilisation de Q pour approximer P . Une divergence KL plus faible indique que les distributions sont plus similaires. La formule de la divergence KL est :

$$D_{KL}(P||Q) = \sum_{i=1}^n P(x_i) \log \left(\frac{P(x_i)}{Q(x_i)} \right)$$

où :

- $P(x_i)$ est la probabilité réelle associée à l'observation x_i ,
- $Q(x_i)$ est la probabilité prédite pour l'observation x_i ,

— La somme est effectuée sur l'ensemble des points de données n .

Dans le contexte de l'imagerie spectrale, où les spectres ne sont pas des fonctions de probabilité, nous devons normaliser les spectres pour appliquer correctement la divergence KL. Pour ce faire, nous définissons le spectre normalisé S à partir du spectre initial s comme suit :

$$S(\lambda) = \frac{s(\lambda)}{k}, \quad \forall \lambda \in [\lambda_{\min}, \lambda_{\max}] \quad (4.8)$$

où k est la constante de normalisation donnée par :

$$k = \int_{\lambda_{\min}}^{\lambda_{\max}} s(\lambda) d\lambda \quad (4.9)$$

La pseudo-divergence KL spectrale, est une adaptation de la divergence KL appliquée aux données spectrales. Cette mesure est composée de deux parties : la première évalue les différences de forme entre les spectres (en utilisant les spectres normalisés), et la seconde évalue les différences d'intensité (en utilisant les spectres initiaux) [33]. Elle est exprimée par la formule suivante :

$$\text{div}_{KL'}(S_1, S_2) = k_1 \cdot KL(S_1, S_2) + k_2 \cdot KL(S_2, S_1) + (k_1 - k_2) \log \left(\frac{k_2}{k_1} \right) \quad (4.10)$$

Cette mesure est utile pour comparer les spectres même s'ils ne peuvent pas être considérés comme des distributions de probabilité. Elle est composée de deux parties : la première évalue les différences de forme entre les spectres, tandis que la seconde évalue les variations d'intensité globales. Ainsi, la pseudo-divergence KL permet de quantifier la similarité entre les spectres en tenant compte à la fois des différences de forme et des variations d'intensité.

BIBLIOGRAPHIE

- [1] P. A. Bottomley and J. R. Griffiths, Eds., *Handbook of magnetic resonance spectroscopy in vivo : MRS theory, practice and applications*, ser. NMR-MRI - biology, chemistry, medicine, physics. Chichester, West Sussex : Wiley, 2016.
- [2] G. Hangel, E. Niess, P. Lazen, P. Bednarik, W. Bogner, and B. Strasser, “Emerging methods and applications of ultra-high field MR spectroscopic imaging in the human brain,” *Analytical Biochemistry*, vol. 638, p. 114479, Feb. 2022. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0003269721003808>
- [3] R. Bartha, D. Drost, R. Menon, and P. Williamson, “Comparison of the quantification precision of human short echo time ^1H spectroscopy at 1.5 and 4.0 Tesla,” *Magnetic Resonance in Medicine*, vol. 44, no. 2, pp. 185–192, Aug. 2000. [Online]. Available : [https://onlinelibrary.wiley.com/doi/10.1002/1522-2594\(200008\)44:2<185::AID-MRM4>3.0.CO;2-V](https://onlinelibrary.wiley.com/doi/10.1002/1522-2594(200008)44:2<185::AID-MRM4>3.0.CO;2-V)
- [4] I. Tkáč, P. Andersen, G. Adriany, H. Merkle, K. Uğurbil, and R. Gruetter, “In vivo ^1H NMR spectroscopy of the human brain at 7 T,” *Magnetic Resonance in Medicine*, vol. 46, no. 3, pp. 451–456, Sep. 2001. [Online]. Available : <https://onlinelibrary.wiley.com/doi/10.1002/mrm.1213>
- [5] HAS, Jul. 2006. [Online]. Available : https://www.has-sante.fr/jcms/c_272001/fr/la-sclerose-en-plaques
- [6] C. Guillemin, P. Agius, M. Naudin, G. Herpe, S. Ragot, N. Maubeuge, J. Philippe Neau, and R. Guillemin, “ ^1H - ^{31}P magnetic resonance spectroscopy : effect of biotin in multiple sclerosis,” *Annals of Clinical and Translational Neurology*, vol. 6, no. 7, pp. 1332–1337, Jul. 2019. [Online]. Available : <https://onlinelibrary.wiley.com/doi/10.1002/acn3.50825>

- [7] C. Menuel, R. Guillevin, R. Costalat, M. Perrin, M. Sahli-Amor, N. Martin-Duverneuil, and J. Chiras, “Spectroscopie du phosphore 31 par résonance magnétique : applications en pathologies cérébrales,” *Journal of Neuroradiology*, vol. 37, no. 2, pp. 73–82, May 2010. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0150986109001175>
- [8] M. Naudin, M. Lapert, P. Fayolle, C. Guillevin, and R. Guillevin, “La spectroscopie clinique à ultra haut champ au chu de poitiers,” *Journal of Neuroradiology*, vol. 50, no. 2, pp. 142–143, Mar. 2023. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0150986123000688>
- [9] R. Guillevin, C. Menuel, H. Duffau, M. Kujas, L. Capelle, A. Aubert, S. Taillibert, A. Idbaih, J. Pallud, G. Demarco, R. Costalat, K. Hoang-Xuan, J. Chiras, and J.-N. Vallée, “Proton magnetic resonance spectroscopy predicts proliferative activity in diffuse low-grade gliomas,” *Journal of Neuro-Oncology*, vol. 87, no. 2, pp. 181–187, Apr. 2008. [Online]. Available : <http://link.springer.com/10.1007/s11060-007-9508-y>
- [10] G.-J. Déborah, “PREDICTION D’UNE ATTEINTE COGNITIVE DANS LA SCLE-ROSE EN PLAQUES,” Ph.D. dissertation, UNIVERSITE JOSEPH FOURIER FACULTE DE MEDECINE DE GRENOBLE, GRENOBLE, 2014.
- [11] J.-C. Ouallet and B. Brochet, “Aspects cliniques, physiopathologiques, et thérapeutiques de la sclérose en plaques,” Ph.D. dissertation, Oct. 2004. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S1762423104000515>
- [12] P.-O. Bertho, M. Carpentier, L. Carpentier, A. Hay-Lombardie, and E. Bigot-Corbel, “Marqueurs biologiques utilisés dans le diagnostic de la sclérose en plaques,” *Revue Francophone des Laboratoires*, vol. 2021, no. 534, pp. 66–73, Jul. 2021. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S1773035X21001969>
- [13] a. leprieur, “Sclérose en plaques : la maladie et les avancées thérapeutiques,” Ph.D. dissertation, ’Université de Caen Normandie, Caen Normandie, 2019.
- [14] P. Gallien and S. Robineau, “Sensory-motor and genito-sphincter dysfunctions in multiple sclerosis,” *Biomedicine & Pharmacotherapy*, vol. 53, no. 8, pp. 380–385, Sep. 1999. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0753332299801097>
- [15] R. Mekle, V. Mlynárik, G. Gambarota, M. Hergt, G. Krueger, and R. Gruetter, “MR spectroscopy of the human brain with enhanced signal intensity at ultrashort echo times on a clinical platform at 3T and 7T,” *Magnetic Resonance*

- in Medicine*, vol. 61, no. 6, pp. 1279–1285, Jun. 2009. [Online]. Available : <https://onlinelibrary.wiley.com/doi/10.1002/mrm.21961>
- [16] C. Lubetzki, “Physiopathologie de la sclérose en plaques : actualités,” *Bulletin de l’Académie Nationale de Médecine*, vol. 206, no. 6, pp. 721–726, Jun. 2022. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0001407922001017>
- [17] G. Yu, “Représentations parcimonieuses pour la séparation des signaux : quantification en IRM spectroscopique pour l’étude du métabolisme,” Ph.D. dissertation, Université de REIMS CHAMPAGNE-ARDENNE, 2011.
- [18] R. Rizzo, M. Dziadosz, S. P. Kyathanahally, A. Shamaei, and R. Kreis, “Quantification of MR spectra by deep learning in an idealized setting : Investigation of forms of input, network architectures, optimization by ensembles of networks, and training bias,” *Magnetic Resonance in Medicine*, vol. 89, no. 5, pp. 1707–1727, May 2023. [Online]. Available : <https://onlinelibrary.wiley.com/doi/10.1002/mrm.29561>
- [19] Y.-L. Huang, Y.-R. Lin, and S.-Y. Tsai, “Comparison of Convolutional Neural Networks Based Method and LCmodel on the Quantification of in Vivo Magnetic Resonance Spectroscopy,” Jan. 2022. [Online]. Available : <https://www.researchsquare.com/article/rs-1190469/v1>
- [20] N. Hatami, M. Sdika, and H. Ratiney, “Magnetic Resonance Spectroscopy Quantification using Deep Learning,” Jun. 2018, arXiv :1806.07237 [cs]. [Online]. Available : <http://arxiv.org/abs/1806.07237>
- [21] H. H. Lee and H. Kim, “Deep learning-based target metabolite isolation and big data-driven measurement uncertainty estimation in proton magnetic resonance spectroscopy of the brain,” *Magnetic Resonance in Medicine*, vol. 84, no. 4, pp. 1689–1706, Oct. 2020. [Online]. Available : <https://onlinelibrary.wiley.com/doi/10.1002/mrm.28234>
- [22] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, “Scientific Machine Learning Through Physics-Informed Neural Networks : Where we are and What’s Next,” *Journal of Scientific Computing*, vol. 92, no. 3, p. 88, Sep. 2022. [Online]. Available : <https://link.springer.com/10.1007/s10915-022-01939-z>
- [23] M. Raissi, P. Perdikaris, and G. Karniadakis, “Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0021999118307125>

- [24] X. Zhang, B. Mao, Y. Che, J. Kang, M. Luo, A. Qiao, Y. Liu, H. Anzai, M. Ohta, Y. Guo, and G. Li, “Physics-informed neural networks (PINNs) for 4D hemodynamics prediction : An investigation of optimal framework based on vascular morphology,” *Computers in Biology and Medicine*, vol. 164, p. 107287, Sep. 2023. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0010482523007527>
- [25] O. Mukhmetov, Y. Zhao, A. Mashekova, V. Zarikas, E. Y. K. Ng, and N. Aidossov, “Physics-informed neural network for fast prediction of temperature distributions in cancerous breasts as a potential efficient portable AI-based diagnostic tool,” *Computer Methods and Programs in Biomedicine*, vol. 242, p. 107834, Dec. 2023. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S016926072300500X>
- [26] A. Rebai, L. Boukhris, R. Toujani, A. Gueddiche, F. A. Banna, F. Souissi, A. Lasram, E. B. Rayana, and H. Zaag, “Unsupervised physics-informed neural network in reaction-diffusion biology models (Ulcerative colitis and Crohn’s disease cases) A preliminary study,” 2023, version Number : 1. [Online]. Available : <https://arxiv.org/abs/2302.07405>
- [27] A. Shamaei, J. Starcukova, and Z. Starcuk, “Physics-informed deep learning approach to quantification of human brain metabolites from magnetic resonance spectroscopy data,” *Computers in Biology and Medicine*, vol. 158, p. 106837, May 2023. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S0010482523003025>
- [28] Y. Zhang and J. Shen, “Regional and tissue-specific differences in brain glutamate concentration measured by in vivo single voxel MRS,” *Journal of Neuroscience Methods*, vol. 239, pp. 94–99, Jan. 2015. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S016502701400346X>
- [29] S. Köhler, U. Winkler, and J. Hirrlinger, “Heterogeneity of Astrocytes in Grey and White Matter,” *Neurochemical Research*, vol. 46, no. 1, pp. 3–14, Jan. 2021. [Online]. Available : <http://link.springer.com/10.1007/s11064-019-02926-x>
- [30] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” *PeerJ Computer Science*, vol. 7, p. e623, Jul. 2021. [Online]. Available : <https://peerj.com/articles/cs-623>

- [31] S. Van Meerten, W. Franssen, and A. Kentgens, “ssNake : A cross-platform open-source NMR data processing and fitting application,” *Journal of Magnetic Resonance*, vol. 301, pp. 56–66, Apr. 2019. [Online]. Available : <https://linkinghub.elsevier.com/retrieve/pii/S1090780719300308>
- [32] L. Bonnin, W. Ramchi, P. Bourdon, C. Guillevin, R. Guillevin, C. Giraud, and C. Fernandez-Maloigne, “Mr spectroscopy spectrum fitting using a physically informed neural network,” *Conference Proceedings of MICAD*, vol. X, no. Y, pp. Z–Z, 2024.
- [33] N. Richard, D. Helbert, C. Olivier, and M. Tamisier, “Pseudo-Divergence and Bidimensional Histogram of Spectral Differences for Hyperspectral Image Processing,” *Journal of Imaging Science and Technology*, vol. 60, no. 5, pp. 050 402–1–050 402–13, Sep. 2016. [Online]. Available : <https://library.imaging.org/jist/articles/60/5/jist0168>