



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Gabriel Santos Martins Dias

**Spectro-ViT: Um modelo Vision Transformer
para Reconstrução em MEGA-PRESS editada
para GABÁ usando Espectrogramas**

Campinas

2024



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Gabriel Santos Martins Dias

**Spectro-ViT: Um modelo Vision Transformer para
Reconstrução em MEGA-PRESS editada para GABA
usando Espectrogramas**

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na Área de Engenharia de Computação.

Orientador: Profa. Dra. Letícia Rittner

Este exemplar corresponde à versão final da tese defendida pelo aluno Gabriel Santos Martins Dias, e orientada pela Profa. Dra. Letícia Rittner

Campinas

2024

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

D543s Dias, Gabriel Santos Martins, 1997-
Spectro-ViT: um modelo vision transformer para reconstrução em MEGA-PRESS editada para GABA usando espectrogramas / Gabriel Santos Martins Dias. – Campinas, SP : [s.n.], 2024.

Orientador: Letícia Rittner.

Dissertação (mestrado) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Elétrica e de Computação.

1. MEGA-PRESS. 2. Espectroscopia de prótons por ressonância magnética. 3. Inteligência artificial. I. Rittner, Letícia, 1972-. II. Universidade Estadual de Campinas (UNICAMP). Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações Complementares

Título em outro idioma: Spectro-ViT: a vision transformer model for GABA-edited MEGA-PRESS reconstruction using spectrograms

Palavras-chave em inglês:

MEGA-PRESS

Magnetic resonance spectroscopy

Artificial intelligence

Área de concentração: Engenharia de Computação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora:

Letícia Rittner [Orientador]

Maria Concepcion Garcia Otaduy

Levy Boccato

Data de defesa: 02-07-2024

Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0009-0003-8422-4692>

- Currículo Lattes do autor: <http://lattes.cnpq.br/0576036812532468>

COMISSÃO JULGADORA - DISSERTAÇÃO DE MESTRADO

Candidato: Gabriel Santos Martins Dias RA: 172441

Data da Defesa: 02 de julho de 2024

Título da Tese: “Spectro-ViT: Um modelo Vision Transformer para Reconstrução em MEGA-PRESS editada para GABA usando Espectrogramas”.

Prof. Dra. Letícia Rittner (Presidente)

Dra. Maria Concepcion Garcia Otaduy

Prof. Dr. Levy Boccato

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

Agradecimentos

Este trabalho só foi possível graças ao apoio, orientação e incentivo de várias pessoas e instituições, às quais sou profundamente grato.

Primeiramente, gostaria de expressar minha sincera gratidão à minha orientadora, Profa. Dra. Leticia Rittner. Sua orientação e apoio incondicional foram fundamentais para o desenvolvimento deste trabalho. Agradeço por ter investido no meu potencial e por ter me auxiliado a utilizá-lo ao máximo. Sua dedicação e conhecimento foram essenciais para meu crescimento acadêmico e pessoal ao longo desta jornada.

Agradeço também às instituições que me concederam bolsas de estudo valorizadas e reconhecidas, possibilitando a realização deste mestrado. O presente trabalho foi realizado com apoio do processo nº 2020/09838-0, da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Além disso, sou extremamente grato pela oportunidade de realizar um intercâmbio de quatro meses na Universidade de Calgary (Canadá), onde tive o privilégio de interagir com pessoas que são referência na minha área de pesquisa.

Não posso deixar de agradecer à minha família, pelo suporte e amor incondicional. Vocês estiveram ao meu lado em todos os momentos, especialmente nos mais desafiadores, quando o cansaço e o esgotamento eram mais intensos. Seu apoio e encorajamento sempre me deram forças para continuar e alcançar meus objetivos.

A todos que contribuíram, direta ou indiretamente, para a realização deste trabalho, meu mais sincero agradecimento. Suas contribuições foram inestimáveis para a conclusão desta dissertação e para minha trajetória acadêmica e profissional.

Resumo

Este estudo investigou o uso de um Vision Transformer (ViT) para reconstruir dados de Espectroscopia por Ressonância Magnética (MRS) editada para GABA a partir de um número reduzido de transientes. Transientes referem-se às amostras coletadas por meio da repetição do experimento durante uma aquisição de MRS, com o objetivo de gerar um sinal de qualidade adequada. Especificamente, foram utilizados 80 transientes em vez dos típicos 320, visando reduzir o tempo de aquisição em quatro vezes. Os 80 transientes foram pré-processados e convertidos em uma representação de espectrograma usando a Transformada de Fourier de Tempo Curto (STFT). Um ViT, denominado Spectro-ViT, foi treinado e testado utilizando dados *in-vivo* da sequência MEGA-PRESS editada para GABA. Seu desempenho foi comparado com outros métodos de reconstrução na literatura, tanto do ponto de vista de qualidade do espectro reconstruído, quanto em relação aos valores estimados de concentração de metabólitos a partir deste espectro, com os escaneamentos típicos de 320 transientes servindo como referência. O modelo Spectro-ViT exibiu as melhores métricas de qualidade geral entre todos os outros métodos avaliados. As concentrações de metabólitos das reconstruções do Spectro-ViT para GABA atingiram o melhor valor médio de R^2 de 0,67 e o melhor valor médio de Erro Percentual Absoluto Médio de 9,68%, sem diferenças estatísticas significativas em comparação com a referência de 320 transientes. O código para reproduzir esta pesquisa está disponível em <https://github.com/MICLab-Unicamp/Spectro-ViT>.

Keywords: MEGA-PRESS editada para GABA; Espectroscopia por Ressonância Magnética; Aprendizado Profundo; Vision Transformer

Abstract

This research explored the application of a Vision Transformer (ViT) for reconstructing GABA-edited Magnetic Resonance Spectroscopy (MRS) data from a smaller number of transients. Transients are the samples obtained during an MRS acquisition by repeating the experiment to produce a signal of adequate quality. In this study, 80 transients were utilized instead of the usual 320 transients, aiming to decrease scan time. The 80 transients were pre-processed and transformed into a spectrogram representation using the Short-Time Fourier Transform (STFT). A ViT, referred to as Spectro-ViT, was trained and subsequently tested on in-vivo GABA-edited MEGA-PRESS data. Its performance was benchmarked against other existing pipelines in the literature having in view the spectral reconstruction and the estimated metabolite concentration values, with the standard 320-transient scans serving as the baseline (reference). The Spectro-ViT model achieved the highest overall quality metrics among all compared pipelines. The metabolite concentrations from Spectro-ViT's reconstructions for GABA obtained the best average R^2 value of 0.67 and the best average Mean Absolute Percentage Error value of 9.68%, with no significant statistical differences when compared to the 320-transient reference. The code for reproducing this research is available at <https://github.com/MICLab-Unicamp/Spectro-ViT>.

Keywords: *GABA-edited MEGA-PRESS; Magnetic Resonance Spectroscopy; Deep Learning; Vision Transformer*

Lista de ilustrações

Figura 2.1 – Magnetização dos núcleos de hidrogênio	25
Figura 2.2 – Aquisição do sinal FID	26
Figura 2.3 – Exemplo de aquisição do sinal e localização do voxel	27
Figura 2.4 – Exemplo de espectro proveniente de aquisição MRS não editada	29
Figura 2.5 – Esquemático da STFT	30
Figura 2.6 – Sinal de MRS no domínio da STFT	31
Figura 2.7 – Obtenção do espectro de diferença em GABA MEGA-PRESS	32
Figura 2.8 – Principais etapas do <i>pipeline</i> convencional GABA MEGA-PRESS	34
Figura 2.9 – Diagrama do objetivo de reconstrução	34
Figura 3.1 – Geração dos dados da Trilha 1 do desafio	38
Figura 3.2 – Intervalo de cálculo do MSE	40
Figura 3.3 – Exemplo de largura à meia largura destacada no pico de GABA	41
Figura 3.4 – Exemplo ilustrativo para a métrica Shape Score	42
Figura 3.5 – Arquitetura do Spectro-ViT	44
Figura 3.6 – Preparação dos dados simulados para treinamento	44
Figura 3.7 – Método da janela deslizante	45
Figura 3.8 – Pré-Processamento para geração dos espectrogramas	46
Figura 3.9 – Esquemático do treinamento do modelo empregado no desafio	47
Figura 3.10–Análise qualitativa da Trilha 1 do desafio	49
Figura 3.11–Distribuição de métricas na Trilha 2 do desafio	51
Figura 3.12–Análise qualitativa da Trilha 2 do desafio	52
Figura 3.13–Distribuição de métricas da Trilha 3 do desafio	53
Figura 3.14–Análise qualitativa da Trilha 3 do desafio	54
Figura 3.15–Distribuição dos erros de ajuste de GABA/Água para Trilhas 2 e 3 do desafio	55
Figura 3.16–Espectros reconstruídos com respectivos erros de ajuste para Trilhas 2 e 3 do desafio	56
Figura 4.1 – Amostras de espectros de diferença geradas por modelos DL compara- dos com Gannet Quarter	69
Figura 4.2 – Amostras de espectros de diferença geradas por modelos DL compara- dos com o espectro EMRT	70
Figura 4.3 – Distribuição das concentrações obtidas pelas pipelines	72
Figura 4.4 – Distribuições de robustez do Spectro-ViT	74

Figura 4.5 – Resultado do cálculo de R^2 em função do número de transientes para a quantificação de GABA/Água 77

Lista de abreviações

ADAM	Adaptative Moment Estimation (Estimação de Momento Adaptativo)
ANU	Amostragem Não Uniforme
CNN	Convolutional Neural Network (Rede Neural Convolutacional)
DL	Deep Learning (Aprendizado Profundo)
EMRT	Espectro Médio de Referência do Treino
FID	Free Induction Decay (Decaimento de Indução livre)
FWHM	Full Width at Half Maximum (Largura à Meia Altura)
GABA	Gamma-AminoButyric Acid (Ácido Gama-Aminobutírico)
GABA MEGA-PRESS	MEGA-PRESS editada para GABA
GE	General Electric
IA	Inteligência Artificial
LR	Learning Rate (Taxa de Aprendizagem)
MAE	Mean Absolute Error (Erro Absoluto Médio)
MAPE	Mean Absolute Percentage Error (Erro Percentual Absoluto Médio)
MEGA-PRESS	MEscher–GARwood Point RESolved Spectroscopy
ML	Machine Learning (Aprendizado de Máquina)
MRS	Magnetic Resonance Spectroscopy (Espectroscopia por RM)
MRSI	MRS Imaging (MRS multi-voxel)
MSE	Mean Squared Error (Erro Quadrático Médio)
ReLU	Rectified Linear Unit (Unidade Linear Retificada)
RM	Ressonância Magnética
SNC	Sistema Nervoso Central
SNR	Signal-to-Noise (Relação Sinal-Ruído)
STFT	Short-time Fourier Transform (TF de Tempo Curto)
TE	Tempo de Eco
TF	Transformada de Fourier
TR	Tempo de Repetição
ViT	Vision Transformer

Sumário

1	Introdução	14
1.1	Motivação	15
1.2	Trabalhos relacionados	16
1.2.1	Aprendizado de máquina em MRS	16
1.2.2	Reconstrução em MRS utilizando DL	17
1.3	Objetivos	21
1.4	Contribuições	22
1.5	Organização do trabalho	23
2	Fundamentação Teórica	24
2.1	Princípios físicos da RM	24
2.2	Sinal de MRS	27
2.3	Representação do sinal MRS em Espectrograma	29
2.4	GABA MEGA-PRESS	31
3	Spectro-ViT: solução vencedora do <i>Edited-MRS Reconstruction Challenge</i>	36
3.1	Conjunto de Dados	37
3.1.1	Dados Simulados (Trilha 1)	37
3.1.2	Dados <i>in-vivo</i> (Trilha 2 e Trilha 3)	39
3.2	Métricas de Avaliação	40
3.3	Modelo Proposto: Spectro-ViT	43
3.3.1	Arquitetura do Modelo	43
3.3.2	Preparação dos Dados	43
3.3.3	Geração dos Espectrogramas	45
3.3.4	Treinamento	46
3.4	Resultados	48
3.4.1	Trilha 1: Dados Simulados	48
3.4.2	Trilha 2: Dados <i>in-vivo</i> com mesmos parâmetros de aquisição	48
3.4.3	Trilha 3: Dados <i>in-vivo</i> com diferentes parâmetros de aquisição	50
3.4.4	Análise Adicional: Erro de ajuste	52
3.5	Discussão	54
3.6	Estudos intermediários para a definição do método final	59
3.6.1	Modelo codificador (<i>encoder</i>)	59
3.6.2	Parâmetros da STFT	60
3.6.3	Normalização	61
3.7	Conclusão do Capítulo	62

4	Spectro-ViT: versão aprimorada	64
4.1	Alterações na metodologia original	64
4.1.1	Conjunto de dados	64
4.1.2	Preparação dos dados	64
4.1.3	Geração dos Espectrogramas	65
4.1.4	Treinamento	66
4.2	Experimentos e Resultados	67
4.2.1	Qualidade da Reconstrução	67
4.2.2	Quantificação	71
4.2.3	Análise de Robustez	73
4.2.4	Redução de Transientes	75
4.3	Discussão	77
4.4	Conclusão do Capítulo	79
5	Conclusão	82
5.1	Premiações	83
5.2	Publicações resultantes	83
5.3	Repositórios de código público	84
5.4	Bolsas de Estudo	85
	Referências	86

1 Introdução

O ácido gama-aminobutírico, mais conhecido pela sigla inglesa GABA (*Gamma-AminoButyric Acid*), é o principal neurotransmissor inibitório no Sistema Nervoso Central (SNC) dos seres humanos. Ele tem um papel fundamental na modulação da atividade cerebral, influenciando tanto funções normais quanto uma variedade de condições patológicas (MCCORMICK, 1989). Devido à sua capacidade de regular os receptores que modulam a excitabilidade neuronal, o GABA é essencial no estudo e no desenvolvimento de medicamentos para tratar distúrbios como insônia e epilepsia (GHIT *et al.*, 2021). Além disso, alterações nesses receptores estão associadas a transtornos de humor, dependência de substâncias e síndromes comportamentais (PAZ *et al.*, 2021).

Avaliar os níveis de GABA no cérebro através da Espectroscopia por Ressonância Magnética (MRS, do inglês *Magnetic Resonance Spectroscopy*) é de grande importância para a comunidade neurocientífica. Atualmente, esta técnica é o único método não invasivo e sem radiação ionizante que permite a detecção e quantificação de metabólitos cerebrais (DHAMALA *et al.*, 2019). A utilização da MRS pode proporcionar uma análise detalhada dos processos bioquímicos cerebrais, fornecendo dados cruciais para o avanço de diagnósticos e tratamentos. A MRS tem sido extensivamente utilizada em pesquisas sobre GABA, envolvendo condições como depressão (RITTER *et al.*, 2022), transtorno do espectro autista (DEMAYO *et al.*, 2021), transtorno de déficit de atenção com hiperatividade (HARRIS *et al.*, 2021), dor crônica (PIGOTT *et al.*, 2023), desenvolvimento neurológico (BELL *et al.*, 2021) e dependência de substâncias (SHYU *et al.*, 2022).

A seleção da sequência de pulsos adequada constitui a primeira etapa de qualquer aquisição por MRS. Essas sequências determinam a implementação precisa e o tempo dos pulsos de radiofrequência e gradientes magnéticos durante o procedimento, gerando dados que elucidam o foco pretendido da investigação. Entre as diversas sequências de MRS (BOTTOMLEY *et al.*, 1987; FRAHM *et al.*, 1987; ORDIDGE *et al.*, 1986; BROWN *et al.*, 1982; STAR-LACK *et al.*, 1998), a mais utilizada para a medição de GABA é conhecida pelo nome em inglês *MEscher-GArwood Point RESolved Spectroscopy* (MEGA-PRESS) (MESCHER *et al.*, 1998). Esta sequência é uma implementação específica da técnica de edição *J-difference editing* (CHOI *et al.*, 2021).

Embora a técnica de MEGA-PRESS editada para GABA (GABA MEGA-PRESS) tenha proporcionado avanços na quantificação desse metabólito, ela enfrenta desafios técnicos, principalmente devido à baixa relação sinal-ruído (SNR, do inglês *Signal-to-Noise Ratio*). Uma limitação crítica dessa metodologia é o tempo extenso necessário

para coletar dados de qualidade suficiente que permitam uma detecção e quantificação confiáveis. Este desafio está diretamente relacionado à técnica de combinação de sinais, conhecida em inglês como *signal averaging*, uma componente essencial da aquisição de GABA MEGA-PRESS. Nesse procedimento, múltiplas medições são repetidamente coletadas ao longo do experimento e, após a aquisição, os sinais obtidos são submetidos a um cálculo de média para melhorar a SNR. A problemática temporal está associada ao fato de que a melhoria da SNR ocorre proporcionalmente à raiz quadrada do número de sinais acumulados para a média, levando a um *trade-off* significativo entre a qualidade da SNR e o tempo de aquisição (GRAAF, 2019).

Na literatura, esses sinais repetidos, coletados para a técnica de *signal averaging*, são denominados transientes (KREIS *et al.*, 2021). Há amplas discussões sobre o número ideal de transientes necessários para garantir uma qualidade de sinal adequada (MIKKELSEN *et al.*, 2018; PEEK *et al.*, 2023). O intervalo de transientes recomendado começa em 126 e chega ao valor mais comumente recomendado de 320. O tempo de aquisição atual estende-se até cerca de 10 minutos, baseando-se na recomendação mais frequente de 320 transientes (PEEK *et al.*, 2023).

Mesmo aderindo às recomendações atuais de consenso sobre o número de transientes para a aquisição de GABA MEGA-PRESS, pode ser necessário aumentar a SNR ao analisar regiões cerebrais menores ou mais complexas. Contudo, este aumento é muitas vezes inviável na maioria das aplicações. Por exemplo, uma aquisição realizada em um volume cerebral de 8 mililitros (ml) exibe apenas cerca de 30% da SNR em comparação a um volume de 27 ml, quando o número de transientes é o mesmo (PEEK *et al.*, 2023). Nesse contexto, contrapondo um volume de 8 ml para 27 ml, seria necessário estender o tempo de varredura em um fator de nove para obter uma SNR comparável. Um volume maior de voxel aumentará a SNR; no entanto, isso também reduz a especificidade regional e aumenta ainda mais os efeitos de volume parcial (PEEK *et al.*, 2023). Adicionalmente, à medida que o tempo de aquisição aumenta, o experimento torna-se mais suscetível aos movimentos do paciente, o que pode gerar artefatos e comprometer tanto a qualidade do sinal medido quanto a confiabilidade da análise (EVANS *et al.*, 2013; HARRIS *et al.*, 2014; CHOI *et al.*, 2021).

1.1 Motivação

A aquisição de MRS é realizada em equipamentos de Ressonância Magnética (RM, do inglês *Magnetic Resonance*), que não emitem radiação iônica. Contudo, durações prolongadas de aquisições limitam a aplicabilidade clínica e de pesquisa devido ao aumento do desconforto do paciente, dos custos e dos artefatos de movimento, além de levar à

rigidez dos protocolos de aquisição.

A duração da aquisição da MRS está diretamente ligada à quantidade de transientes coletados. Uma melhoria significativa da SNR pode resultar em um aumento impraticável do tempo de exame, pois a SNR aumenta apenas com a raiz quadrada do número de transientes. Por exemplo, para obter um aumento de dez vezes na SNR, seria necessário aumentar o tempo de aquisição em cem vezes. Este *trade-off* entre a qualidade desejada de sinal e as limitações de duração do exame torna-se evidente ao lidar com sinais de baixa SNR.

A combinação de transientes é crucial nas aquisições de GABA MEGA-PRESS, apesar de suas limitações, pois possibilita a obtenção de sinais de boa qualidade. Para ampliar a aplicabilidade clínica da MRS, é essencial maximizar a SNR e a qualidade dos dados, mas tentando simultaneamente reduzir os tempos prolongados de exame. O desenvolvimento e a implementação de técnicas que reduzam o tempo de aquisição não apenas aumentariam a viabilidade e a praticidade da MRS em aplicações clínicas, mas também reduziriam custos e melhorariam a eficiência do agendamento dos pacientes. Além disso, permitiriam mais aquisições em uma única sessão, como a avaliação de metabólitos em mais regiões do cérebro ou a realização de outras sequências de RM.

Nesse contexto, a reconstrução de sinais se torna uma abordagem fundamental. Reconstrução é um termo comum na literatura de RM, descrevendo o processo de recuperação de informações perdidas por meio de técnicas de aceleração (MUCKLEY *et al.*, 2021; BEAUFERRIS *et al.*, 2022). Essa área é essencial para o desenvolvimento de métodos que agilizem o processo de aquisição sem comprometer a qualidade dos dados. Recentemente, avanços em técnicas de Inteligência Artificial (IA), especialmente em aprendizado profundo (DL, do inglês *Deep Learning*) (GOODFELLOW *et al.*, 2016), impulsionaram estudos de reconstrução em GABA MEGA-PRESS (BERTO *et al.*, 2023; BERTO *et al.*, 2024). Até o momento, há poucas publicações sobre o tema e nenhuma implementação prática foi efetivamente adotada, indicando um vasto campo de pesquisa pouco explorado que necessita de mais investigações e aperfeiçoamentos.

1.2 Trabalhos relacionados

1.2.1 Aprendizado de máquina em MRS

O potencial de aplicações de aprendizado de máquina (ML, do inglês *Machine Learning*) no meio clínico corresponde a um dos aspectos mais relevantes no campo de pesquisa atual em MRS (SANDE *et al.*, 2023). Tentativas de auxiliar e otimizar o trabalho de especialistas, diminuir o tempo de aquisição e aumentar a robustez e a generalização

das ferramentas de MRS existentes são essenciais para alavancar o uso da técnica. O uso de ML em MRS ainda está em seus estágios iniciais, com os temas mais explorados incluindo controle de qualidade (KYATHANAHALLY *et al.*, 2018b; JANG *et al.*, 2021; HERNÁNDEZ-VILLEGAS *et al.*, 2022), quantificação (SHAMAEI *et al.*, 2023; SCHMID *et al.*, 2023; ALMEIDA *et al.*, 2023) e classificação (DIKAIOS, 2021; ZARINABAD *et al.*, 2018; ZHAO *et al.*, 2022).

Estudos recentes têm se concentrado em modelos profundos, sendo a rede neural convolucional (CNN, do inglês *Convolutional Neural Network*) o tipo mais comumente utilizado. Muitos estudos adotam arquiteturas de modelos semelhantes, com comparações limitadas entre arquiteturas alternativas. De acordo com a revisão recente feita por Sande *et al.* (2023), trabalhos futuros deveriam explorar diferentes e novos tipos de modelos, como os *Transformers* (ISLAM *et al.*, 2023), que demonstraram potencial em outros campos de imagens médicas (LI *et al.*, 2023). À medida que o campo de ML evolui rapidamente, torna-se essencial manter as pesquisas atualizadas com os desenvolvimentos recentes e avaliar seus impactos nos diversos domínios de MRS.

Os principais desafios na aplicação de ML em MRS incluem a escassez de dados *in-vivo* públicos, as dificuldades de generalização do modelo para diferentes grupos populacionais, scanners e parâmetros de aquisição, além da complexidade em realizar simulações e aumento de dados, sem métodos consistentes de geração desenvolvidos até o momento. Apesar desses desafios, os estudos revelam um grande potencial para implementação prática (SANDE *et al.*, 2023).

1.2.2 Reconstrução em MRS utilizando DL

Lee *et al.* (2020) investigaram o uso de CNNs para reconstruir espectros de MRS a partir da subamostragem de pontos, truncando entre 8 e 512 pontos de um total de 1024. A reconstrução a partir de aquisições truncadas pode tornar as sequências de MRS mais flexíveis e eficientes em aplicações clínicas. Três CNNs baseadas na arquitetura U-Net foram avaliadas: FIDCNNFID (tempo-tempo), specCNNspec (frequência-frequência) e FIDCNNspec (tempo-frequência), com a specCNNspec apresentando o melhor desempenho. As CNNs foram treinadas em dados simulados baseados em amostras *in-vivo* de cérebros de ratos, obtidos com um scanner de ressonância magnética de 9.4 Tesla. A specCNNspec apresentou o melhor desempenho no conjunto de teste de dados simulados, recuperando efetivamente os detalhes espectrais mesmo com truncamento significativo. Devido a esse desempenho, a specCNNspec foi selecionada para avaliação *in-vivo*. No entanto, apesar de sua eficácia em dados simulados, seu desempenho *in-vivo* foi significativamente inferior, indicando uma capacidade limitada de generalização.

A queda de desempenho nos dados *in-vivo* destaca a importância do uso de dados de treinamento mais realistas e em maior quantidade. A presença de sinais indesejados, como água residual e outros artefatos espectrais nos dados *in-vivo*, pode ter afetado o desempenho da CNN. Além disso, conjuntos de dados maiores poderiam melhorar a capacidade de generalização da rede. Estratégias adicionais de regularização, que não foram extensivamente exploradas por Lee *et al.* (2020), podem ser necessárias para melhorar essa capacidade. Também é necessário validar os resultados para scanners com intensidade de campo magnético menores do que 9.4 Tesla, onde a resolução e a SNR podem ser consideravelmente menores.

Além da aquisição convencional de MRS unidimensional, existem técnicas de MRS bidimensionais, como a aquisição L-COSY (THOMAS *et al.*, 2001). As aquisições MRS multidimensionais podem fornecer melhor informação estrutural e ajudar a distinguir metabólitos que sofrem sobreposição de sinais mais fortes (SANDE *et al.*, 2023). No entanto, para obter sinais de boa qualidade e resolução em aquisições MRS multidimensionais, é necessário um longo tempo de aquisição, o qual aumenta exponencialmente com o número de dimensões (KAZIMIERCZUK; OREKHOV, 2011). A técnica de subamostragem *Amostragem Não Uniforme* (ANU) tem sido amplamente utilizada para reduzir o tempo de aquisições multidimensionais (SCHMIEDER *et al.*, 1993; HYBERTS *et al.*, 2012; YING *et al.*, 2017). No entanto, a qualidade dos espectros obtidos a partir da técnica ANU depende diretamente de algoritmos de reconstrução apropriados.

Luo *et al.* (2020) propuseram um novo método baseado em DL para a reconstrução de espectros sub-amostrados por ANU, visando otimizar aquisições multidimensionais. A metodologia introduzida utiliza uma rede CNN, denominada EDHRN, projetada para melhorar a qualidade da reconstrução de espectros preservando picos fracos e removendo artefatos com maior eficácia do que CNNs típicas, como a DenseNet (HUANG *et al.*, 2017). Comparado com os dois métodos tradicionais, SMILE (YING *et al.*, 2017) e hmsIST (HYBERTS *et al.*, 2012), o modelo EDHRN de Luo *et al.* (2020) pode fornecer uma qualidade de reconstrução comparável em termos de intensidades de pico e fidelidade da forma dos picos. O tempo de reconstrução dos métodos também é comparável ou mais rápido do que os dois métodos, especialmente para aquisições tridimensionais.

No entanto, a performance da reconstrução espectral do modelo EDHRN depende do nível de ruído, e a reconstrução de espectros com SNR muito baixa pode não ser satisfatória. A distribuição das amostras simuladas utilizadas para treinamento é limitada, o que pode comprometer a generalização dos resultados para outros tipos de distribuição de dados não cobertos pelo conjunto de treinamento. Embora o treinamento seja mais rápido do que em trabalhos anteriores de DL, a necessidade de um pré-processamento específico e a estrutura de redes paralelas aumentam a complexidade computacional do

método. Por fim, a EDHRN pode não ser totalmente eficaz em todas as aquisições de MRS, especialmente quando picos fracos estão próximos de artefatos fortes, o que pode levar a desafios adicionais de generalização dos dados reconstruídos pelo modelo EDHRN.

Iqbal *et al.* (2021) também investigaram um modelo de aprendizado profundo para acelerar as aquisições L-COSY por meio da reconstrução de espectros obtidos a partir da subamostragem ANU. A técnica de DL proposta utiliza uma arquitetura U-Net densamente conectada (D-UNet) (IQBAL *et al.*, 2019) para a reconstrução. Os resultados mostraram que o modelo D-UNet superou significativamente a reconstrução tradicional baseada em detecção comprimida (LUSTIG *et al.*, 2007) em fatores de aceleração mais altos, com erro quadrático médio normalizado inferior a 5%. Além da reconstrução, o modelo D-UNet foi treinado para quantificar espectros de L-COSY totalmente amostrados, identificando e quantificando 17 metabólitos. Mesmo em condições de baixa SNR (25% de ruído comparado ao sinal máximo), o erro quadrático médio normalizado foi inferior a 8%, destacando a eficácia da metodologia. Tal robustez é crucial para a aplicação prática do modelo desenvolvido em ambientes clínicos

As limitações do desenvolvimento do modelo D-UNet no estudo de Iqbal *et al.* (2021) incluem a necessidade de conhecimento prévio dos metabólitos presentes no tecido e a compreensão de como esses sinais são afetados por padrões específicos da subamostragem ANU. Isso pode limitar a aplicabilidade do método em patologias menos compreendidas ou com mudanças químicas desconhecidas. Além disso, a metodologia atual, desenvolvida apenas com dados simulados, não incorporou sinais com presença de água residual e lipídios nos espectros de treinamento, o que pode impactar negativamente a generalização da quantificação dos metabólitos, especialmente em aquisições onde a remoção total desses sinais é desafiadora. O estudo também simplificou a simulação do alargamento de largura espectral ao usar apenas um termo exponencial, embora picos experimentais possam apresentar formas de linha gaussianas, lorentzianas ou voigt (FUCHS *et al.*, 2014). Isso compromete a generalização do modelo, aumentando a complexidade do treinamento e exigindo um conjunto de dados maior. Por fim, diferenças de potencial hidrogeniônico (pH) que podem causar desvios de frequência não foram consideradas, sendo este um fator importante a ser abordado em futuras aplicações.

A sequência de MRS multi-voxel (MRSI, do inglês *Magnetic Resonance Spectroscopy Imaging*) (BROWN *et al.*, 1982) oferece vantagens significativas em comparação com a aquisição de voxel único. Ela proporciona uma resolução espacial aprimorada e uma cobertura anatômica mais abrangente para mapear metabólitos no cérebro. A MRSI permite a aquisição simultânea de múltiplos voxels em diferentes regiões cerebrais, tornando-a especialmente benéfica para investigar patologias cerebrais heterogêneas (MAUDSLEY *et al.*, 2021). No entanto, apesar do grande potencial clínico, a MRSI enfrenta desafios de

vido aos longos tempos de aquisição, o que prejudica severamente sua aplicação. Além disso, a presença de sinais indesejados, como os sinais de lipídios subcutâneos próximos à região do crânio, torna desafiadora a aplicação de técnicas convencionais de aceleração. Diversas técnicas de aceleração têm sido exploradas para MRSI, principalmente através da subamostragem do espaço-k (KREIS *et al.*, 2021) e reconstrução utilizando técnicas de detecção comprimida ou de imagem paralela (PRUESSMANN *et al.*, 1999; GRISWOLD *et al.*, 2002; BREUER *et al.*, 2006). No entanto, ainda são necessários estudos para otimizar essas técnicas, visando uma reconstrução mais eficiente e livre de artefatos (SANDE *et al.*, 2023).

Nassirpour *et al.* (2018) buscaram aprimorar a reconstrução da técnica tradicional GRAPPA (GRISWOLD *et al.*, 2002) estimando os pesos do espaço-k com múltiplas redes neurais de arquitetura simples. Todo o estudo foi desenvolvido com um conjunto de dados *in-vivo* adquiridos de um scanner com intensidade de campo magnético de 9.4 Tesla. Foram treinadas redes neurais especializadas para prever os pontos de dados ausentes de vizinhos cruzados ou de vizinhos adjacentes. Durante a reconstrução, os dois tipos de rede são implementadas sequencialmente de tal maneira que primeiro as redes de vizinho cruzado preencham os valores ausentes do espaço-k e, depois, as redes de vizinho adjacente estimam os pontos restantes. Os resultados de Nassirpour *et al.* (2018) sugerem que o uso das redes neurais para reconstrução GRAPPA reduz artefatos de *aliasing*, impactando positivamente os mapas de concentração de metabólitos e aumentando significativamente o desempenho em comparação com o GRAPPA regular. No entanto, como a intensidade do campo magnético de 9.4 Tesla utilizado no estudo não é comum em ambientes clínicos, é necessário validar a generalização do modelo. A SNR para essa intensidade de aquisição é alta, exigindo avaliação da performance do modelo para valores menores. Também é preciso avaliar o desempenho do modelo para outros parâmetros de aquisição que variam consideravelmente em campos magnéticos de intensidades menores.

No contexto de reconstrução em GABA MEGA-PRESS, a investigação de Berto *et al.* (2023) se destaca como pioneira ao introduzir um modelo DL chamado Rdc-UNET2D. Este modelo visa reconstruir aquisições GABA MEGA-PRESS usando apenas 80 transientes, tendo como referência o total de 320 transientes tipicamente adquiridos. Este fator de aceleração, que utiliza um quarto do total de dados da referência, foi escolhido por ser um valor típico no campo de RM (BEAUFERRIS *et al.*, 2022; MUCKLEY *et al.*, 2021). O Rdc-UNET2D é um modelo puramente convolucional, sendo uma adaptação da arquitetura U-Net com operações de convolução em duas dimensões. O modelo processa uma entrada composta por 40 canais bidimensionais. Cada canal é uma matriz, sendo que uma matriz representa a parte real e outra a parte imaginária dos pontos espectrais. Essas matrizes são geradas pela concatenação dos pontos espectrais ao longo das

linhas. Na saída, o modelo gera o espectro reconstruído

O modelo Rdc-UNET2D demonstrou boas métricas de qualidade espectral nas reconstruções, apresentando altos valores de SNR e baixos erros de ajuste, evidenciando o potencial de se realizar a aquisição GABA MEGA-PRESS com apenas 80 transientes por meio de DL. No entanto, Berto *et al.* (2023) não avaliaram a quantificação dos sinais reconstruídos para verificar se a informação metabólica estava realmente sendo preservada. Um espectro reconstruído pode apresentar boas métricas de qualidade, como SNR e erro de ajuste, mas ainda assim não representar uma boa reconstrução se as concentrações metabólicas divergirem da aquisição convencional. Além disso, a robustez do modelo foi pouco explorada, pois apenas os transientes com menor probabilidade de serem impactados por desvios de fase e frequência foram utilizados na reconstrução.

1.3 Objetivos

Objetivo Geral

O principal objetivo deste projeto é desenvolver um modelo de DL capaz de reconstruir espectros de aquisições GABA MEGA-PRESS utilizando apenas um quarto do número total de transientes convencionalmente adquiridos, resultando em aquisições quatro vezes mais rápidas. O espectro reconstruído deve se assemelhar ao espectro obtido a partir de uma aquisição convencional com a quantidade total de transientes, tanto na qualidade do sinal quanto nos resultados da quantificação de GABA.

Objetivos Específicos

Propor um modelo Vision Transformer treinado a partir de espectrogramas:

Desenvolver um modelo Vision Transformer (ViT) (DOSOVITSKIY *et al.*, 2020) capaz de reconstruir aquisições GABA MEGA-PRESS utilizando a representação dos sinais como espectrogramas, e avaliar o desempenho do modelo treinado em conjuntos de dados públicos *in-vivo* coletados de múltiplos scanners, a fim de replicar cenários do mundo real;

Avaliar a qualidade dos sinais reconstruídos: Comparar a qualidade dos sinais reconstruídos pelo modelo proposto com a qualidade obtida usando o número total de transientes, tomando como referência o software Gannet (EDDEN *et al.*, 2014). Além disso, comparar a qualidade dos sinais reconstruídos pelo modelo proposto com a dos sinais reconstruídos pelo modelo Rdc-UNET2D (BERTO *et al.*, 2023);

Avaliar a quantificação dos sinais reconstruídos: Comparar os resultados de quantificação obtidos a partir dos sinais reconstruídos pelo modelo proposto com aqueles obtidos utilizando o número completo de transientes, tomando como referência o software Gannet (EDDEN *et al.*, 2014). Adicionalmente, realizar uma comparação entre os resultados de quantificação dos sinais reconstruídos pelo modelo proposto e os obtidos pelo modelo Rdc-UNET2D (BERTO *et al.*, 2023).

1.4 Contribuições

As principais contribuições deste trabalho são:

Modelo estado da arte: proposta de um modelo de DL estado da arte para reconstrução que aborda o *trade-off* entre o tempo de aquisição e a quantidade de dados necessários para adquirir sinais de boa qualidade para GABA MEGA-PRESS.

Modelagem no domínio do espectrograma: o uso bem-sucedido de características extraídas simultaneamente dos domínios do tempo e da frequência para treinar um modelo DL de reconstrução, demonstrando uma alternativa promissora para a modelagem em MRS. Enquanto modelos anteriores na literatura escolheram uma representação de domínio de frequência ou tempo como entrada, a abordagem proposta explora a capacidade bidimensional de representação tempo-frequência do espectrograma.

Utilização de modelo baseado em Transformer: emprego da arquitetura ViT para análise em MRS. Uma revisão recente da literatura indica uma carência de modelos do tipo *transformer* nesse campo (SANDE *et al.*, 2023), ressaltando a importância de se explorar diversos tipos de modelos inovadores (ISLAM *et al.*, 2023). Assim, esta dissertação é pioneira na adoção do modelo *Transformer* para MRS, marcando sua estreia em pesquisas na área. Através da exploração das capacidades do ViT, este estudo revela novas possibilidades e percepções na modelagem e análise de MRS, com potencial para melhorar o desempenho das descobertas e fomentar avanços futuros.

Avaliação da Quantificação de metabólitos: condução da primeira análise de quantificação espectral em reconstruções realizadas em GABA MEGA-PRESS, estendendo a avaliação além das métricas de reconstrução tradicionais.

Disponibilização de repositório de código público: compartilhamento em repositório público dos resultados e código¹, visando facilitar a reutilização, aperfeiçoamento

¹ <<https://github.com/MICLab-Unicamp/Spectro-ViT>>

e incentivar pesquisas futuras semelhantes. O *pipeline* de treinamento desenvolvida é robusta e flexível, podendo ser adaptada para diversos problemas em MRS. O aproveitamento desse repositório de código público já se provou fundamental em pesquisas envolvendo quantificação (ALMEIDA *et al.*, 2023; DIAS *et al.*, 2024).

1.5 Organização do trabalho

O restante desta dissertação está organizado em quatro capítulos. O Capítulo 2 apresenta um resumo dos principais conceitos teóricos de MRS. A primeira versão do modelo de reconstrução desenvolvido, denominado Spectro-ViT, que foi vencedor do desafio *Edited-MRS Reconstruction Challenge*, juntamente com os detalhes dos conjuntos de dados utilizados e os métodos de implementação, estão descritos no Capítulo 3. Também fazem parte deste capítulo os resultados obtidos no desafio com relação às métricas de qualidade dos sinais reconstruídos. O Capítulo 4 apresenta a versão aprimorada do modelo Spectro-ViT, apontando as principais modificações em relação à primeira versão e pequenos experimentos que levaram a essas modificações. São apresentados também os experimentos realizados, comparando desta vez o modelo proposto com outros métodos, tanto do ponto de vista de qualidade dos sinais reconstruídos, quanto com relação à quantificação de metabólitos resultante. Por fim, o Capítulo 5 traz as conclusões desta dissertação e aponta ideias para trabalhos futuros.

2 Fundamentação Teórica

Este capítulo apresenta os principais conceitos teóricos sobre MRS que servem de base para o entendimento desse estudo.

2.1 Princípios físicos da RM

A MRS compartilha os mesmos princípios físicos da RM, fundamentando-se nas propriedades magnéticas e nas energias associadas dos núcleos atômicos. Partículas subatômicas, como elétrons, nêutrons e prótons, possuem um momento angular intrínseco conhecido como spin. Essa propriedade permite que elas se alinhem com ou contra um campo magnético externo, absorvam e emitam energia, e, finalmente, produzam os sinais detectados na RM. O núcleo magnético do isótopo 1H , que consiste em um único próton, é frequentemente o alvo da RM, pois é abundante em compostos orgânicos e possui um sinal magnético elevado, caracterizado pela sua alta razão giromagnética. O elemento químico 1H está abundantemente presente na maioria dos tecidos na forma de água ou lipídios. No cérebro humano, um pequeno volume cúbico de $1 \times 1 \times 1$ mm contém cerca de 6×10^{19} spins atômicos de 1H (GRAAF, 2019). Quando os núcleos não estão sob a influência de um campo magnético externo, seus spins magnéticos estão orientados aleatoriamente, resultando em uma magnetização nula de todos os momentos magnéticos individuais.

Se os núcleos forem colocados em um campo magnético externo de alta intensidade \vec{B}_0 , eles se alinham em uma das duas direções paralelas ao campo: na mesma direção de \vec{B}_0 (paralelo) ou na direção oposta (antiparalelo). Núcleos paralelos e antiparalelos têm momentos magnéticos iguais, mas opostos, que se anulam mutuamente. No entanto, há uma predominância de núcleos alinhados paralelamente a \vec{B}_0 , resultando em uma magnetização líquida \vec{M}_0 diferente de zero. Por convenção, o campo magnético \vec{B}_0 está orientado ao longo do eixo z (BLOCH, 1946; PURCELL, 1953).

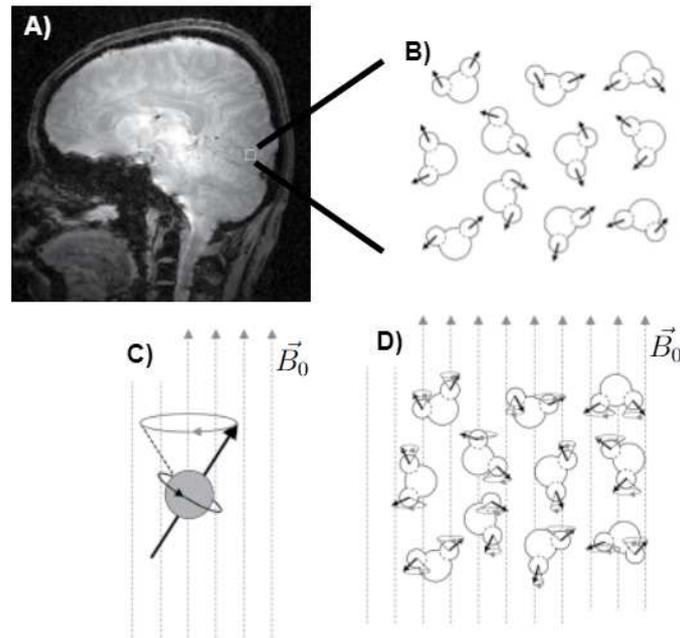


Figura 2.1 – Um pequeno volume de $1 \mu\text{l}$ do cérebro humano contém cerca de 6×10^{19} prótons, localizados principalmente em moléculas de água (A); Moléculas de água com os momentos magnéticos dos núcleos de hidrogênio apontando em direções aleatórias na ausência de um campo magnético externo (B); Visão detalhada de um único núcleo de hidrogênio precessionando paralelamente a um campo magnético externo \vec{B}_0 (C); Moléculas de água com os momentos magnéticos dos núcleos de hidrogênio precessionando de forma paralela ou antiparalela ao campo magnético externo (D). Fonte: Adaptado de (GRAAF, 2019).

Além disso, esses núcleos realizam um movimento de precessão em torno do campo \vec{B}_0 . Esse movimento é semelhante ao de um pião, onde o eixo de rotação descreve um círculo ao redor da direção do campo magnético aplicado (Fig. 2.1). A precessão é uma consequência do torque que o campo magnético \vec{B}_0 exerce sobre os momentos magnéticos dos núcleos, forçando-os a girar em torno da direção do campo. Esse movimento de precessão ocorre com uma frequência específica, conhecida como frequência de Larmor, definida como (BLOCH, 1946; PURCELL, 1953):

$$\omega_0 = \gamma B_0 \quad (2.1)$$

onde ω_0 é a frequência de Larmor e γ é a razão giromagnética do núcleo.

No entanto, apesar da resultante não nula de \vec{M}_0 , o sinal só pode ser medido quando um segundo campo magnético, \vec{B}_1 , é utilizado para mudar a orientação de \vec{M}_0 para o plano transversal (BLOCH, 1946; PURCELL, 1953), comumente referido como xy . Quando a amplitude e a duração do campo \vec{B}_1 são ajustadas para resultar em uma

rotação de 90° da magnetização \vec{M}_0 do eixo longitudinal para o plano transversal, os spins sofrem uma excitação, e o campo \vec{B}_1 é denominado pulso de excitação (GRAAF, 2019).

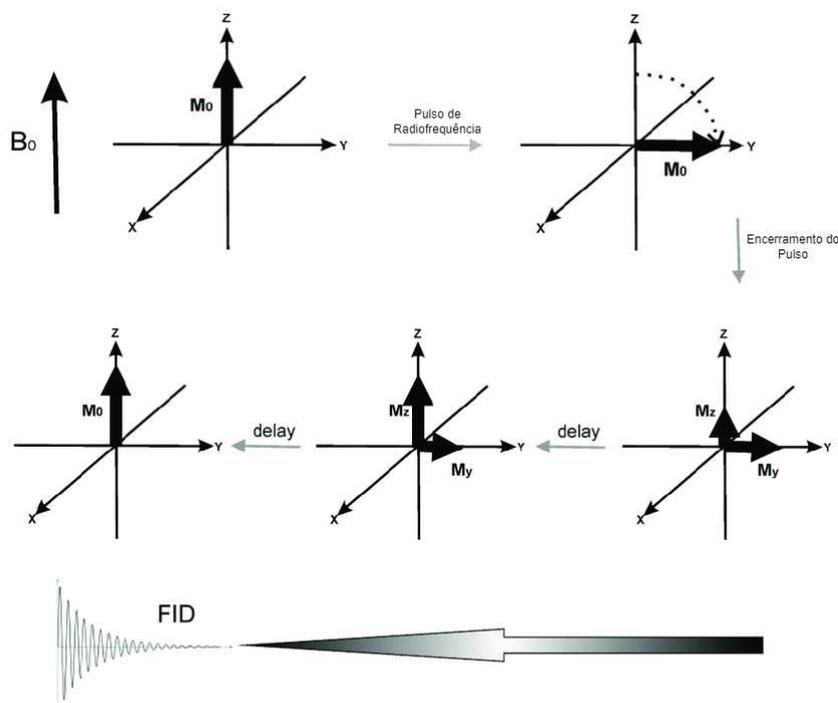


Figura 2.2 – Esquemático ilustrando a obtenção do sinal FID. A parte superior mostra o vetor de magnetização \vec{M}_0 , alinhado com o campo magnético principal \vec{B}_0 , sendo excitado por um pulso de radiofrequência que desloca \vec{M}_0 para a componente transversal no plano xy . A parte inferior ilustra o sinal FID captado após o término do pulso, com a componente transversal \vec{M}_y decaindo ao longo do tempo até atingir o estado inicial de equilíbrio antes da aplicação do pulso, resultando em uma componente total de magnetização apenas longitudinal no eixo z \vec{M}_0 . Fonte: Adaptado de (DULIŃSKA-LITEWKA *et al.*, 2019).

Para que essa excitação ocorra, é fundamental que o fenômeno da ressonância seja alcançado. Isso significa que a frequência de oscilação do campo \vec{B}_1 precisa ser igual à frequência de Larmor do núcleo atômico em questão (BLOCH, 1946; PURCELL, 1953). Quando essa condição de ressonância é satisfeita, a energia de radiofrequência de \vec{B}_1 é eficientemente transferida para os spins nucleares, permitindo que \vec{M}_0 seja rotacionado para o plano xy (Fig .2.2).

Após a excitação, o campo \vec{B}_1 é removido, deixando a magnetização no plano transversal para ser detectada por meio de indução eletromagnética em uma bobina receptora próxima, de acordo com a lei de Faraday. O sinal captado é o Decaimento de Indução Livre (FID, do inglês *Free Induction Decay*), que é o sinal fundamental de RM (GRAAF, 2019). O FID é um sinal complexo regido por uma função senoidal amortecida, cuja intensidade corresponde à tensão elétrica na bobina. O amortecimento do sinal ocorre

devido ao processo de relaxamento. Esse processo é o mecanismo pelo qual a magnetização retorna gradualmente ao estado inicial de equilíbrio antes da aplicação do pulso, resultando em uma componente total de magnetização apenas longitudinal no eixo z (Fig. 2.2) (GRAAF, 2019).

Em resumo, o sinal FID é o sinal fundamental da RM, gerado com apenas três elementos. Primeiro, um campo magnético externo \vec{B}_0 forte é necessário para criar uma magnetização resultante \vec{M}_0 não nula. Segundo, um campo magnético variável no tempo \vec{B}_1 rotaciona a magnetização \vec{M}_0 para o plano transversal. Por fim, a magnetização transversal precessante induz o sinal FID em uma bobina receptora. Embora existam diferentes sequências de RM com variados parâmetros de aquisição, o sinal FID continua sendo o principal elemento a partir do qual todas essas técnicas são derivadas.

2.2 Sinal de MRS

A técnica de MRS pode envolver a aquisição de um único sinal de apenas um voxel (Fig. 2.3), ou de múltiplos sinais provenientes de diversos voxels (MRSI). Na aquisição de voxel único, um volume específico de interesse é selecionado dentro do tecido utilizando uma combinação de pulsos de radiofrequência e gradientes de campo magnético, que permitem a seleção precisa do voxel a ser estimulado. Este estudo de mestrado se limitou apenas às aquisições provenientes de um único voxel.

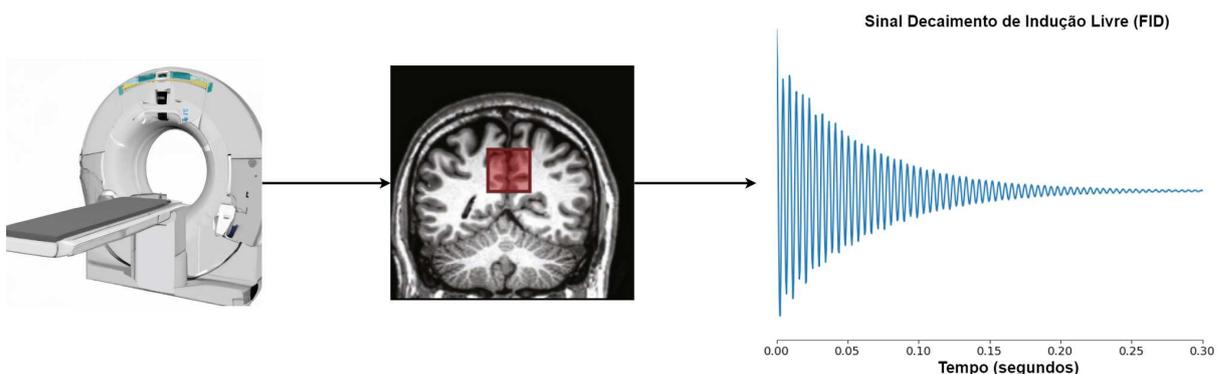


Figura 2.3 – Exemplo de uma aquisição de único voxel com o voxel posicionado no lobo parietal medial do cérebro humano (o diagrama apresenta apenas a vista coronal para fins de simplificação). Todos os dados *in-vivo* utilizados neste trabalho de mestrado, provenientes do repositório Big GABA (MIKKELSEN *et al.*, 2017), foram adquiridos de volumes posicionados nessa região cerebral. Fonte: Adaptado de (MIKKELSEN *et al.*, 2017).

Após a obtenção do sinal FID em uma aquisição de MRS, a Transformada de Fourier é realizada para converter o sinal do domínio do tempo para o domínio da frequência (Fig. 2.4) (ERNST; ANDERSON, 1966). Nesse domínio espectral, é possível

analisar, identificar e distinguir as informações metabólicas presentes no voxel, tanto por inspeção visual quanto pelos métodos tradicionais.

Se todos os núcleos de 1H presentes em uma mistura de moléculas ressoassem na mesma frequência, o espectro resultante apresentaria apenas um único pico. No entanto, devido aos variados ambientes químicos em torno de cada núcleo, eles podem absorver energia em frequências distintas da de Larmor. A circulação dos elétrons ao redor do núcleo gera um campo magnético local que se opõe ao campo aplicado (\vec{B}_0), um efeito conhecido como blindagem. Essa blindagem, resultante do meio eletrônico envolvente, está relacionada à densidade eletrônica do núcleo em seu ambiente químico, fazendo com que a frequência do núcleo de 1H dentro de uma molécula dependa da estrutura molecular (PROCTOR; YU, 1950). Quanto maior a densidade eletrônica ao redor de um núcleo, maior a blindagem. Como resultado, diferentes núcleos de 1H , tanto dentro da mesma molécula quanto em moléculas diferentes, podem apresentar diversas frequências de ressonância, observadas como picos distintos no espectro. Esse fenômeno que provoca desvios na frequência de ressonância de Larmor é conhecido como deslocamento químico (GRAAF, 2019). Embora possa ser medido em Hertz, é convencionalmente normalizado em partes por milhão (ppm):

$$\delta = \frac{\nu - \nu_{\text{ref}}}{\nu_{\text{ref}}} \times 10^6 \quad (2.2)$$

onde δ é referida como o deslocamento químico (em ppm) e ν e ν_{ref} são as frequências de Larmor (em Hz) do composto sob investigação e de um composto de referência. É importante notar que, na visualização gráfica do espectro em MRS, o eixo de ppm é convencionalmente invertido, começando de ordem decrescente (Fig. 2.4). O tetrametilsilano corresponde a um composto químico de referência amplamente empregado (GRAAF, 2019).

Essa normalização serve para tornar o eixo de frequência independente da intensidade magnética dos scanners de RM, facilitando análises e comparações entre diferentes centros de aquisição. Assim, para espectros *in-vivo*, os metabólitos preservam a mesma assinatura espectral. Isso significa que o N-acetilaspártato (NAA) apresentará picos em todos os espectros perto de 2 ppm, colina (Cho) em 3,2 ppm, glutamato (Glu), GABA e glutamina (Gln) formarão um conjunto composto de picos entre 2,2 ppm e 2,6 ppm, com componentes adicionais em 3,6 ppm, creatina (Cr) apresentará picos em 3 ppm e 3,9 ppm, e mioniositol (Myo) em 3,56 ppm (CECIL, 2013) (Fig. 2.4).

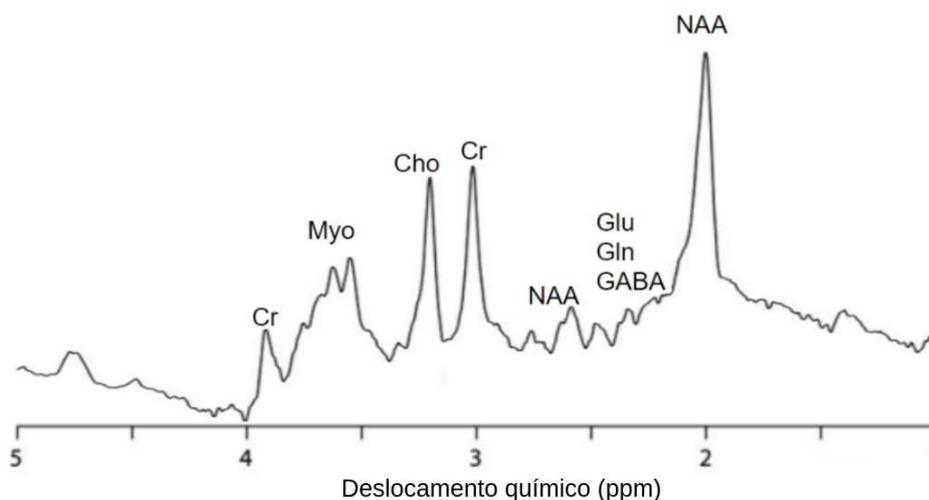


Figura 2.4 – Espectro de MRS do cérebro humano mostrando picos correspondentes a N-acetil aspartato (NAA), creatina (Cr), colina (Cho), mioinositol (Myo), glutamato (Glu), glutamina (Gln) e GABA. Fonte: Adaptado de (PUTS; EDDEN, 2012).

O espectro apresentado na Figura 2.4 destaca o desafio da sobreposição espectral em MRS, especialmente ao estudar o GABA. Os sinais de diferentes metabólitos podem se sobrepor, e os sinais dos metabólitos mais abundantes frequentemente dominam os sinais dos menos abundantes, como o GABA (HARRIS *et al.*, 2017).

2.3 Representação do sinal MRS em Espectrograma

Enquanto a maioria dos estudos na literatura opta por representar o sinal de MRS no domínio convencional da frequência utilizando a tradicional Transformada de Fourier, este estudo utiliza uma representação tempo-frequência conhecida como espectrograma. O espectrograma é uma representação visual amplamente utilizada da Transformada de Fourier de Tempo Curto (STFT, do inglês *Short-time Fourier Transform*) (HLAWATSCH; BOUDREAUX-BARTELS, 1992), uma técnica eficaz para analisar sinais no domínio tempo-frequência. A STFT segmenta o sinal no domínio do tempo e calcula a Transformada de Fourier para cada segmento temporal (KEHTARNAVAZ, 2008) (Fig. 2.5). Essa abordagem permite a identificação de informações de frequência específicas em intervalos de tempo determinados, sendo particularmente útil quando as componentes de frequência de um sinal variam ao longo do tempo, como é o caso do FID. Diferentemente da Transformada de Fourier tradicional, que fornece informações de frequência média sobre toda a duração do sinal, a STFT oferece informações das frequências ao longo do tempo.

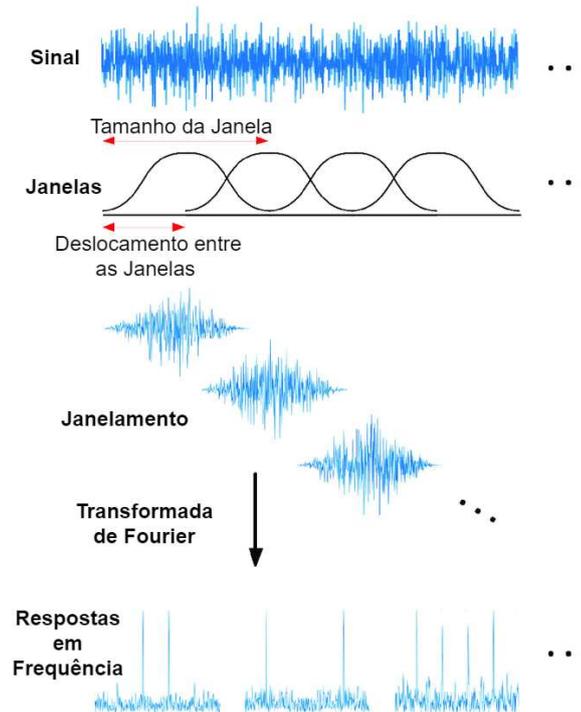


Figura 2.5 – Esquemático da STFT. No topo, tem-se o sinal no domínio do tempo. Esse sinal é dividido em vários segmentos a partir de janelas sobrepostas, cada uma representada por uma forma de janela, como a janela de Hann. Cada segmento é, então, transformado no domínio da frequência usando a Transformada de Fourier, resultando em representações espectrais de cada janela. Essas respostas em frequência são organizados ao longo do tempo, formando uma matriz tempo-frequência que fornece informações sobre como o conteúdo em frequência do sinal varia ao longo do tempo Fonte: Adaptado de (JEON *et al.*, 2020).

Os espectrogramas são visualizações gráficas bidimensionais obtidas a partir da STFT. Nesses espectrogramas, o eixo horizontal representa os instantes de tempo, o eixo vertical representa as frequências e cada valor indica a intensidade de cada componente de frequência em cada instante de tempo (Fig. 2.6). Como a STFT gera valores complexos, os espectrogramas podem ser construídos a partir de projeções dessas componentes complexas, como a parte real, a parte imaginária, a magnitude ou a fase dos valores. Portanto, é intuitivo utilizar modelos de DL explorados na área de visão computacional em espectrogramas para executar diversos tipos de tarefas. No domínio do áudio, essa abordagem é amplamente empregada (MARIOTTI *et al.*, 2018; LE *et al.*, 2019), e áreas clínicas que lidam com sinais biológicos também têm demonstrado sua eficácia em tarefas de classificação (DIKER *et al.*, 2019; ASLAN; AKIN, 2020).

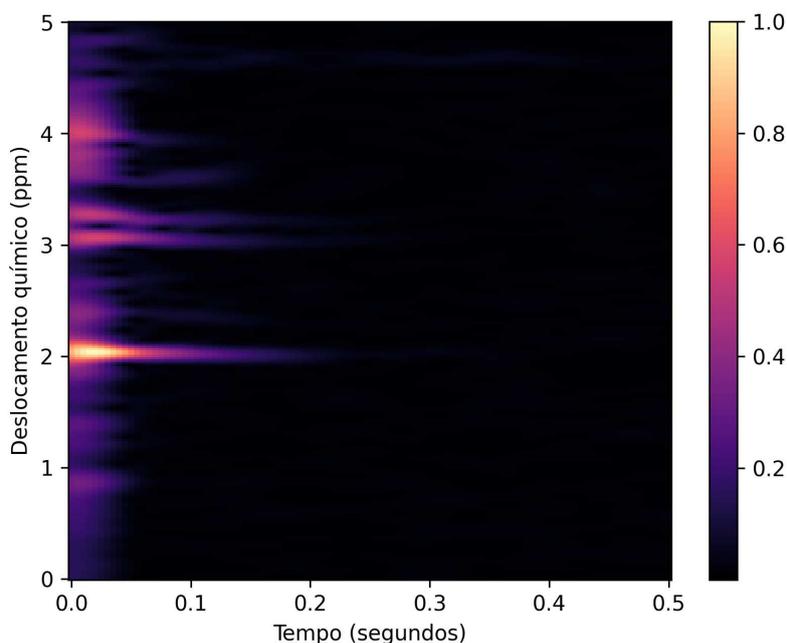


Figura 2.6 – Espectrograma obtido a partir de um sinal de MRS (FID) simulado. O espectrograma foi plotado utilizando a projeção da magnitude de seus valores. O eixo horizontal corresponde ao tempo e o eixo vertical ao deslocamento químico. As intensidades dos valores são representadas por pseudocores, onde cores mais claras indicam maior intensidade e cores mais escuras indicam menor intensidade.

2.4 GABA MEGA-PRESS

Selecionar a sequência de pulsos apropriada constitui a primeira etapa de qualquer aquisição de MRS. Essas sequências determinam a implementação precisa e o tempo dos pulsos de radiofrequência e gradientes magnéticos ao longo do procedimento, gerando dados que elucidam o foco pretendido da investigação. Entre as várias sequências de MRS (BOTTOMLEY *et al.*, 1987; FRAHM *et al.*, 1987; ORDIDGE *et al.*, 1986; BROWN *et al.*, 1982; STAR-LACK *et al.*, 1998; MESCHER *et al.*, 1998), os métodos de edição *J-difference editing* (STAR-LACK *et al.*, 1998; MESCHER *et al.*, 1998; CHOI *et al.*, 2021) surgiram como ferramentas indispensáveis para abordar o desafio da sobreposição espectral, onde sinais mais fortes de diferentes metabólitos podem dificultar a diferenciação e quantificação de outros (HARRIS *et al.*, 2017).

Uma das sequências de MRS mais amplamente utilizadas para medir GABA é GABA MEGA-PRESS, que corresponde a uma implementação de edição *J-difference editing*. O GABA, devido à sua menor concentração e padrão de pico complicado, possui um sinal espectral que é difícil de separar de forma confiável de metabólitos mais abundantes, como a creatina (MULLINS *et al.*, 2014). A técnica de GABA MEGA-PRESS

envolve dois sub-experimentos, geralmente adquiridos de maneira intercalada. Um sub-experimento aplica pulsos de edição em uma frequência de 1,9 ppm para refocar seletivamente a evolução do acoplamento do sinal de GABA em 3 ppm (Edit-ON), enquanto o outro permite a evolução livre do sistema de spins ao longo do tempo de eco (Edit-OFF). Dependendo do pulso de edição, um sinal é classificado como Edit-ON ou Edit-OFF. A subtração do espectro Edit-OFF do espectro Edit-ON revela um espectro editado para GABA, comumente referido como espectro de diferença, removendo os sinais mais fortes sobrepostos dos compostos contendo creatina (GRAAF, 2019; PEEK *et al.*, 2023) (Fig. 2.7). O sinal editado resultante contém sinais co-editados de homocarnosina e macromoléculas (ROTHMAN *et al.*, 1997), sendo comumente referido como GABA+. Nesta dissertação, utiliza-se a notação GABA para referir-se ao sinal do metabólito, mesmo que os dados *in-vivo* utilizados não tenham realizado a supressão de macromoléculas, ciente das possíveis contribuições das macromoléculas ao sinal adquirido.

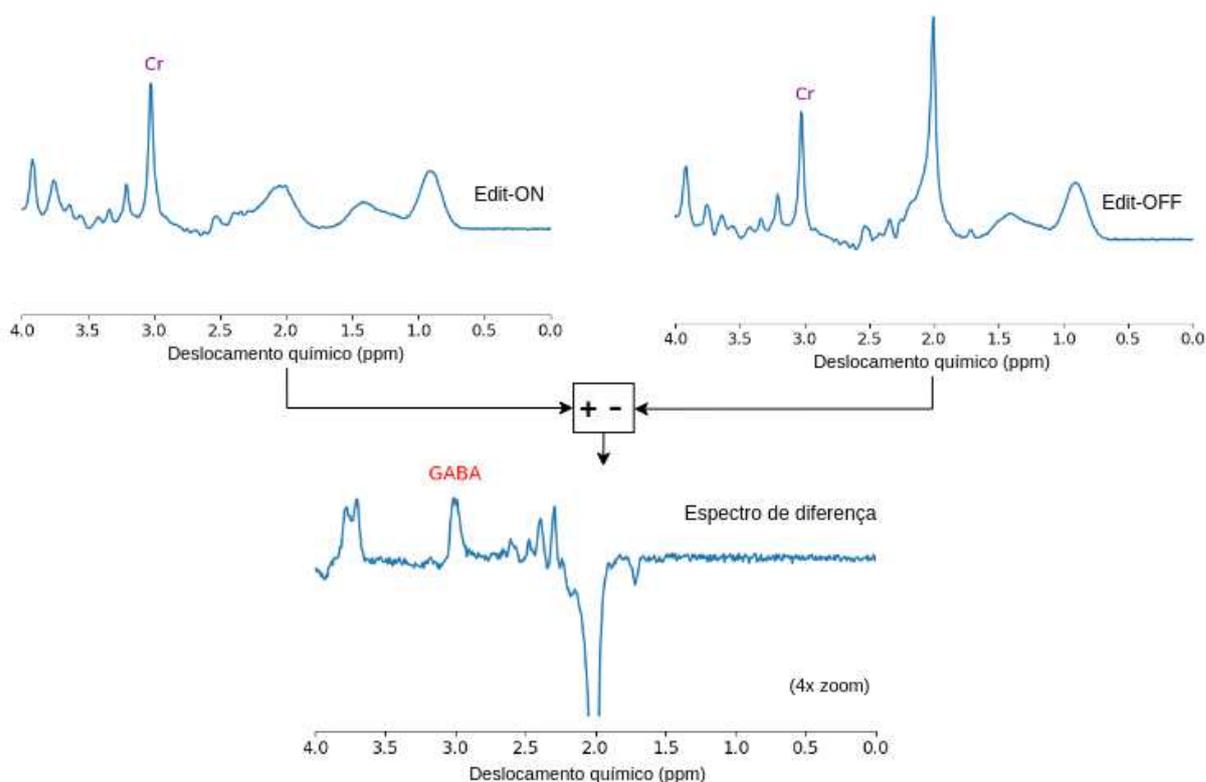


Figura 2.7 – Pulso de edição aplicado a 1,9 ppm modula a forma do sinal de GABA a 3 ppm. Ao subtrair o espectro adquirido sem esse pulso (rotulado como Edit-OFF) do espectro adquirido com o pulso de edição (rotulado como Edit-ON), remove-se significativamente os sinais sobrepostos de creatina, revelando o sinal de GABA no sinal resultante do espectro de diferença.

Embora o método MEGA-PRESS melhore significativamente a separabilidade do sinal de GABA, ele enfrenta desafios técnicos, como baixa SNR e artefatos de subtração. A técnica exige alta estabilidade tanto do sistema quanto do sujeito. Para uma aquisição

bem-sucedida, é consensual manter a estabilidade entre as aquisições.

Artefatos de subtração, resultantes do desalinhamento de frequência e fase dos sinais, são um dos principais problemas em sequências de MRS editadas (EVANS *et al.*, 2013; HARRIS *et al.*, 2014). Esse desalinhamento geralmente é causado por deslocamentos do campo \vec{B}_0 devido ao movimento da cabeça do participante, pois o corpo humano possui propriedades magnéticas que podem influenciar localmente o campo magnético. Além disso, o aquecimento e resfriamento dos elementos do ímã, especialmente após ciclos intensos de gradiente, também podem contribuir para o desalinhamento (ROWLAND *et al.*, 2017; JUCHEM *et al.*, 2021). Na prática, é desafiador determinar a fonte exata dos artefatos em dados *in-vivo*. Varreduras de menor duração podem ajudar a mitigar esses problemas.

No método de aquisição convencional, as etapas essenciais no pré-processamento dos dados brutos de MRS para obter o espectro de diferença que aborda esses principais problemas incluem a técnica *signal averaging* e a correção de frequência e fase (PEEK *et al.*, 2023). Esses procedimentos são fundamentais em qualquer *pipeline* convencional e são cruciais para garantir resultados precisos e confiáveis na detecção e quantificação de GABA utilizando GABA MEGA-PRESS (GRAAF, 2019) (Fig. 2.8). Neste estudo, o software Gannet (EDDEN *et al.*, 2014) foi escolhido como referência para processar os dados brutos, gerar o espectro de diferença e realizar a quantificação, devido à sua robustez e acessibilidade para os pesquisadores. O Gannet, especializado em MRS editada, é um software gratuito e de código aberto, com funções amplamente automatizadas, beneficiando-se de atualizações contínuas e de discussões ativas dentro da comunidade MRS.

Os métodos de reconstrução têm como objetivo recuperar informações perdidas durante a subamostragem, que é utilizada para acelerar a técnica. Em vez de usar a quantidade total de transientes, uma quantidade menor de dados é introduzida no método de reconstrução, esperando-se que a qualidade e a informação essencial dos dados sejam preservadas. Por exemplo, ao utilizar um fator de aceleração de quatro no pipeline convencional, ou seja, utilizando apenas 25% dos dados, o sinal resultante será de má qualidade. No entanto, neste trabalho de mestrado, foi investigada uma abordagem de reconstrução utilizando um modelo de aprendizado profundo, que visa gerar espectros de diferença com métricas de qualidade e quantificação comparáveis às da aquisição convencional utilizando o total de transientes (Fig. 2.9).

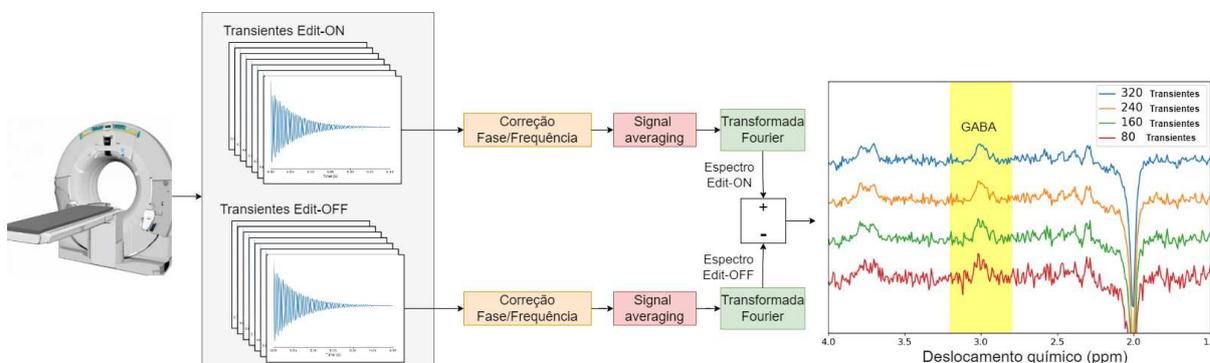


Figura 2.8 – O *pipeline* convencional do GABA MEGA-PRESS envolve a aquisição de múltiplos transientes FID pareados (Edit-ON e Edit-OFF) para obter um sinal com uma SNR adequada. Esse pré-processamento envolve a correção de fase e frequência, *signal averaging* dos transientes de acordo com a quantidade coletada e a computação da Transformada de Fourier para a subtração espectral. Quanto mais transientes, melhor a SNR.

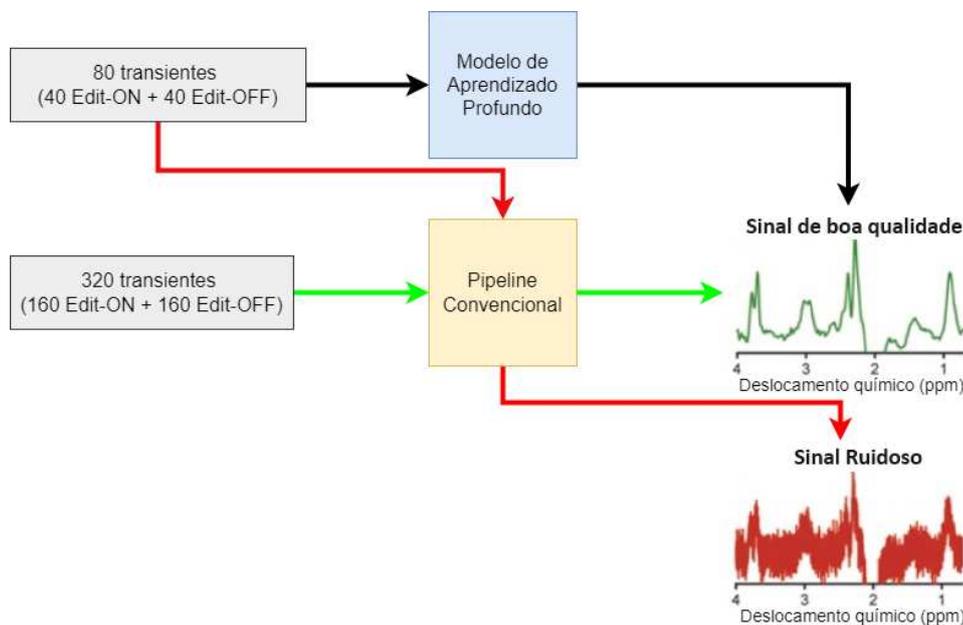


Figura 2.9 – O pipeline convencional da aquisição GABA MEGA-PRESS, que utiliza 320 transientes, gera sinais de boa qualidade (fluxo em verde) em comparação com a aquisição que utiliza apenas 80 transientes (fluxo em vermelho). O objetivo deste trabalho de mestrado foi investigar um modelo de aprendizado profundo capaz de gerar um sinal utilizando apenas 80 transientes, sendo comparável ao adquirido pelo pipeline convencional com o total de 320 transientes (fluxo em preto).

Depois de obter um espectro de diferença de qualidade satisfatória, o próximo

passo é quantificar o espectro para determinar a concentração de GABA. As concentrações em MRS são relatadas como uma razão entre o metabólito primário e outro, sendo crucial especificar qual metabólito é usado como referência no denominador. As concentrações de metabólitos podem ser expressas em unidades padrão, como molar ou molal, utilizando métodos de conversão que geralmente dependem de várias suposições, como um conteúdo tecidual assumido para o componente de referência (LIN *et al.*, 2021).

Nesta dissertação, o software Gannet foi utilizado para a quantificação, que reporta as concentrações em “unidades institucionais”, uma abordagem comum que permite comparações normalizadas. Essa abordagem envolve calcular a razão entre o sinal ajustado do metabólito e o ajuste da ressonância da água não suprimida (LIN *et al.*, 2021). No contexto do Gannet, os resultados quantitativos são apresentados como a razão da integral do GABA em relação à integral da água.

O Gannet utiliza um modelo gaussiano para estimar parâmetros nos espectros de diferença, que são utilizados para o cálculo da quantificação. Este modelo possui um conjunto específico de parâmetros, otimizados através de ajuste não linear de mínimos quadrados (EDDEN *et al.*, 2014). Para facilitar a interpretação, o Gannet gera um erro de ajuste normalizado com base na amplitude do sinal. Um erro de ajuste de 5% indica que o erro residual é 5% da amplitude máxima do sinal. Esse erro de ajuste serve como uma métrica de qualidade, onde sinais ruidosos, com artefatos e alargamento de linha de base, tendem a apresentar um maior erro de ajuste.

Com base nesses conceitos teóricos de MRS, segue-se o desenvolvimento e a avaliação dos modelos de reconstrução em aquisições GABA MEGA-PRESS, que serão abordados nos capítulos seguintes. No próximo capítulo, será detalhada a primeira versão do modelo de reconstrução Spectro-ViT, desenvolvida no contexto do *Edited-MRS Reconstruction Challenge*. Serão apresentados os detalhes dos conjuntos de dados utilizados, a metodologia de desenvolvimento e os resultados obtidos, com ênfase nas métricas de qualidade dos sinais reconstruídos.

3 Spectro-ViT: solução vencedora do *Edited-MRS Reconstruction Challenge*

Com o avanço das soluções de aprendizado de máquina para MRS e motivados pelos bons resultados do estudo de reconstrução em aquisições GABA MEGA-PRESS (BERTO *et al.*, 2023), os autores do trabalho decidiram organizar o desafio *Edited-MRS Reconstruction Challenge* (BERTO *et al.*, 2024), apresentado no Simpósio Internacional de Imagem Biomédica em Cartagena, Colômbia (ISBI 2023).

O desafio proposto buscava soluções baseadas em DL para a reconstrução de espectros de diferença GABA MEGA-PRESS usando 80 transientes, equivalentes a um quarto dos transientes adquiridos em uma aquisição convencional, resultando em aquisições quatro vezes mais rápidas. Este fator de aceleração foi inspirado pelos resultados de desafios de reconstrução em RM (MUCKLEY *et al.*, 2021; BEAUFERRIS *et al.*, 2022) e no estudo de Berto *et al.* (2023).

Os participantes tiveram pouco mais de três meses para desenvolver suas soluções, modelos DL cuja saída consistia na parte real do espectro de diferença GABA MEGA-PRESS utilizando 80 transientes como entrada. Os organizadores do desafio disponibilizaram um modelo DL de referência, dados simulados, dados *in-vivo* e códigos de programação para aumentar os dados através da adição de ruído sintético. O desafio foi dividido em três trilhas independentes:

- Trilha 1: foram disponibilizados 5000 dados simulados para o treinamento do modelo, e a avaliação pelos organizadores do desafio foi realizada em 500 amostras de teste.
- Trilha 2: dados *in-vivo* coletados com parâmetros de aquisição quase idênticos (incluindo fornecedor, número de pontos espectrais e técnica de supressão de água), com exceção da versão do software e do número de canais da bobina de cabeça. Os participantes do desafio receberam um conjunto de treinamento com 12 aquisições. Os modelos foram avaliados usando um conjunto de teste de 24 aquisições.
- Trilha 3: dados *in-vivo* coletados com diferentes parâmetros de aquisição (por exemplo, fornecedor, número de pontos espectrais e técnica de supressão de água). Os participantes receberam um conjunto com 30 aquisições para o treino do modelo. Os modelos foram avaliados usando um conjunto de teste de 48 aquisições.

Nas trilhas com dados *in-vivo* (Trilhas 2 e 3), os espectros reconstruídos com 80 transientes de cada modelo foram comparados com os espectros convencionais gerados a partir de 320 transientes pelo software Gannet (EDDEN *et al.*, 2014). O Gannet foi utilizado para gerar o espectro de referência, sendo uma escolha robusta e acessível para pesquisadores.

A equipe criada no contexto desta pesquisa de mestrado, chamada Deep Spectral Divers, obteve o primeiro lugar com o modelo denominado de Spectro-ViT. A equipe que conquistou o segundo lugar, chamada Spectralligence, foi formada por pesquisadores da Universidade Tecnológica de Eindhoven, Holanda. A equipe que ficou em terceiro lugar, chamada Dolphins, foi formada por pesquisadores da Universidade Imperial College London, Reino Unido.

3.1 Conjunto de Dados

Os dados do desafio incluem espectros de voxel único simulados e *in-vivo* para GABA MEGA-PRESS divididos em três trilhas (Tab. 3.1). Os participantes não tiveram acesso aos espectros de referência do conjunto de teste, apenas aos 80 FIDs brutos para realizarem as reconstruções de seus modelos.

Tabela 3.1 – Conjuntos de dados do desafio, incluindo as fontes dos dados com suas respectivas quantidades de pontos espectrais, bem como o número de amostras disponibilizadas para treino e teste.

Trilha	Tipo	Fonte	Nº pontos	Amostras para Treino	Amostras para Teste
1	Simulado	FID-A	2048	5000	1000
2	<i>in-vivo</i>	Big GABA (General Electric)	2048	12 com 320 transientes	24 com 80 transientes
3	<i>in-vivo</i>	Big GABA (General Electric)	4096	6 com 320 transientes	6 com 80 transientes
3	<i>in-vivo</i>	Big GABA (Phillips)	2048	12 com 320 transientes	24 com 80 transientes
3	<i>in-vivo</i>	Big GABA (Siemens)	4096	12 com 320 transientes	18 com 80 transientes

3.1.1 Dados Simulados (Trilha 1)

Os dados simulados usados no desafio podem ser divididos em duas categorias: transientes ideais e transientes ruidosos. Os dados ideais consistem em FIDs sem ruído, gerados utilizando o FID-A, uma ferramenta de simulação e processamento para MRS em Matlab (SIMPSON *et al.*, 2017). Os FIDs foram gerados com a ressonância de 22 metabólitos (alanina, ascorbato, aspartato, beta-hidroxiubutirato, creatina, GABA, glutamina, glutamato, glutatona, glicina, glicerofosfocolina, glicose, mio-inositol, lactato, N-acetilaspártato, N-acetilaspártato-glutamato, fosfocolina, fosfocreatina, fosfoetanolamina,

scyllo-inositol, serina, taurina), cinco macromoléculas (MM 09, 12, 14, 17, 20) e um lipídio (Lip20). Cada concentração de metabólito foi amostrada independentemente de uma distribuição normal centrada nos valores médios relatados para cérebros de adultos humanos saudáveis na literatura (NEAR *et al.*, 2013) e com desvio padrão de $\pm 10\%$. Os parâmetros de aquisição foram: intensidade do campo magnético de 3 T, sequência GABA MEGA-PRESS, pulso Edit-ON em 1,9 ppm por 14 milissegundos (ms), pulso Edit-OFF em 7,46 ppm por 14 ms, pulsos de edição intercalados entre ON e OFF a cada Tempo de Repetição (TR), TR/TE = 2 s/68 ms, largura de banda espectral de 2 Kilo-Hertz (kHz) e 2048 pontos.

Para gerar os transientes ruidosos, ruído foi adicionado aos FIDs ideais usando um código Python fornecido aos participantes para imitar os condições de ruído encontrados nos dados *in-vivo*. O ruído adicionado consistia em uma combinação aleatória de ruído de amplitude, deslocamentos de frequência e deslocamentos de fase (Fig. 3.1), de acordo com:

$$\vec{S}_{\text{ruidoso}} = \left(\vec{S}_{\text{ideal}} + \vec{N} \right) * e^{i\left(\vec{t}*F*2\pi + P*\frac{\pi}{180}\right)} \quad (3.1)$$

onde \vec{S}_{ideal} é o FID de referência, N é o ruído de amplitude, t é o vetor com os instantes de tempo da aquisição, F é o ruído de frequência e P é o ruído de fase.

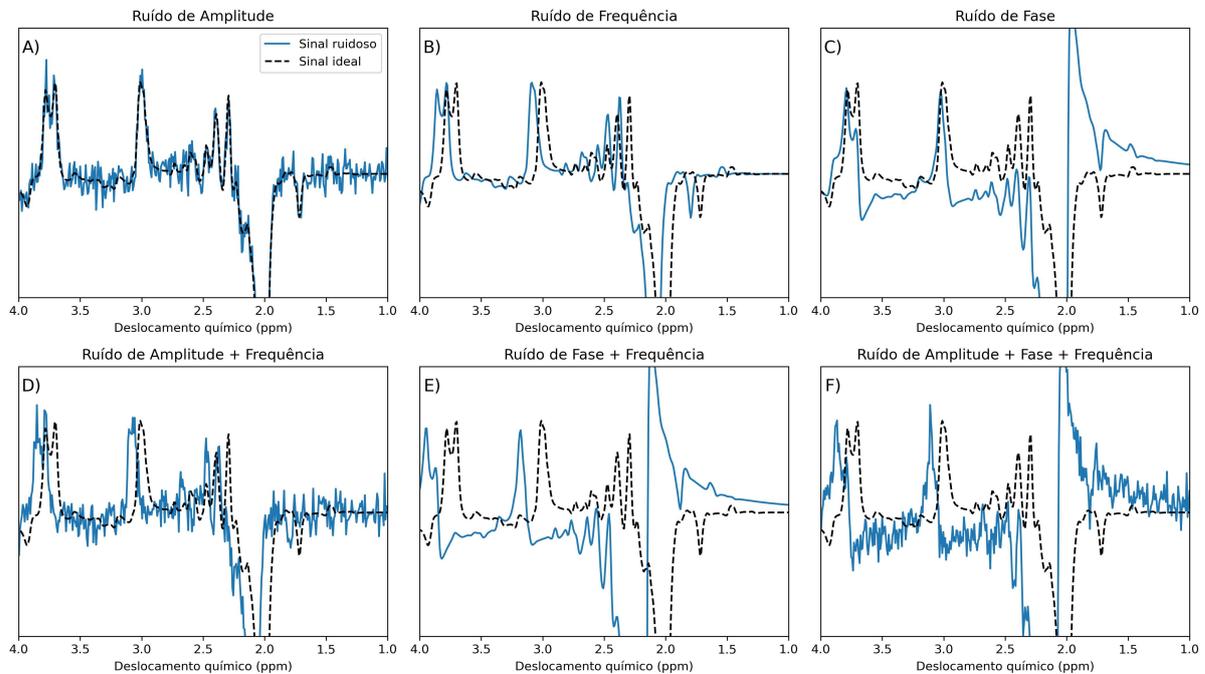


Figura 3.1 – Dados da Trilha 1: variações de espectros ruidosos (em azul) criados a partir dessa simulação ideal (em preto): Apenas ruído de amplitude (A); Apenas ruído de frequência (B); Apenas ruído de fase (C); Combinação de ruídos de amplitude e frequência (D); Combinação de ruídos de fase e frequência (E); Combinação de ruídos de amplitude, fase e frequência (F).

Foram fornecidos 5000 pares ideais de FIDs simulados para GABA MEGA-PRESS (ou seja, 5000 Edit-ON e 5000 Edit-OFF) para o treinamento dos modelos. Os participantes deveriam adicionar ruído a esses dados ideais para gerar os transientes e treinar os modelos. Em contraste, o conjunto de teste consistiu em 1000 dados simulados de aquisições GABA MEGA-PRESS com 80 transientes ruidosos (40 Edit-ON e 40 Edit-OFF), onde o ruído foi aplicado aos espectros simulados de referência pelos organizadores do desafio. Os espectros de diferença ideais constituíram a referência para as reconstruções.

3.1.2 Dados *in-vivo* (Trilha 2 e Trilha 3)

Os dados *in-vivo* utilizados no desafio são provenientes do repositório Big GABA (MIKKELSEN *et al.*, 2017; MIKKELSEN *et al.*, 2019), obtidos e compartilhados de acordo com os padrões éticos e a aprovação dos comitês de ética institucional de cada centro de aquisição. Esses dados consistem em espectros de voxel único de dimensão $30 \times 30 \times 30$ mm no lobo parietal medial de voluntários adultos saudáveis (divididos igualmente entre mulheres e homens), com idades entre 18 e 35 anos e sem nenhuma doença neurológica ou psiquiátrica conhecida. Os dados amostrados foram considerados de qualidade média a alta pelos organizadores do desafio.

Os parâmetros associados aos dados da Trilha 2 foram os seguintes: intensidade do campo magnético de 3 Tesla, fabricante de scanner General Electric (GE), método de supressão da água CHES (HAASE *et al.*, 1985), com pulsos de edição de 15 ms a 1,9 ppm (Edit-ON) e 7,46 ppm (Edit-OFF) intercalados entre ON e OFF em cada TR, $TR/TE = 2$ s/68 ms, largura de banda espectral de 2 kHz, 2048 pontos, ciclo de fase de 8 passos e 320 transientes. O conjunto disponibilizado para treinamento consistiu em 12 aquisições de 320 transientes fornecidos como FIDs brutos, cada aquisição com sua própria reconstrução de referência associada pelo Gannet usando os 320 transientes completos. O conjunto de teste consistiu em 24 aquisições, sendo que os primeiros 80 transientes adquiridos na aquisição foram fornecidos aos participantes.

Para a Trilha 3, os parâmetros de aquisição foram: intensidade do campo magnético de 3 Tesla, fabricantes de scanner GE, Siemens e Philips, método de supressão da água CHES, MOIST (HEER *et al.*, 2016) ou WET (OGG *et al.*, 1994), com pulsos de edição de 15 ms (GE e Philips) a 1,9 ppm (Edit-ON) e 7,46 ppm (Edit-OFF) intercalados entre ON e OFF em cada TR, $TR/TE = 2$ s/68 ms, largura de banda espectral de 2 kHz, 4 kHz ou 5 kHz, 2048 ou 4096 pontos, com 320 transientes. O conjunto disponibilizado para treinamento consistiu em 30 aquisições de 320 transientes fornecidos como FIDs brutos, cada aquisição com sua própria reconstrução de referência associada pelo Gannet usando os 320 transientes completos. O conjunto de teste consistiu em 48 aquisições, sendo que os primeiros 80 transientes foram fornecidos aos participantes.

3.2 Métricas de Avaliação

Como o foco deste desafio foi o metabólito GABA, todas as métricas de avaliação empregadas foram especificamente projetadas para avaliar a qualidade do espectro de diferença reconstruído para GABA. Algumas das métricas também consideraram o pico combinado de Glutamato e Glutamina (Glx), pois sua qualidade pode ser usada como um indicativo de confiabilidade e qualidade geral do espectro.

Quatro métricas foram usadas para avaliar as reconstruções, escolhidas como uma combinação de métricas tradicionais de DL e MRS e uma nova métrica proposta. A escolha das métricas teve como objetivo avaliar diferentes aspectos das reconstruções. Na Trilha 1, os espectros de referência utilizados corresponderam aos espectros de diferença gerados pelos FIDs ideais (sem ruído). Já nas Trilhas 2 e 3, os espectros de referência corresponderam aos espectros de diferença gerados com 320 transientes pelo Gannet.

Erro Quadrático Médio (MSE, do inglês *Mean Squared Error*): calculado entre o espectro de diferença reconstruído pelo modelo e o espectro de referência correspondente, ao longo da faixa espectral de 2,5 ppm a 4 ppm, abrangendo os picos de GABA e Glx (Fig. 3.2):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\text{referência}_{2,5-4 \text{ ppm}, i} - \text{reconstrução}_{2,5-4 \text{ ppm}, i})^2 \quad (3.2)$$

onde o índice i se refere ao i -ésimo ponto de dado dentro do intervalo de ppm especificado.

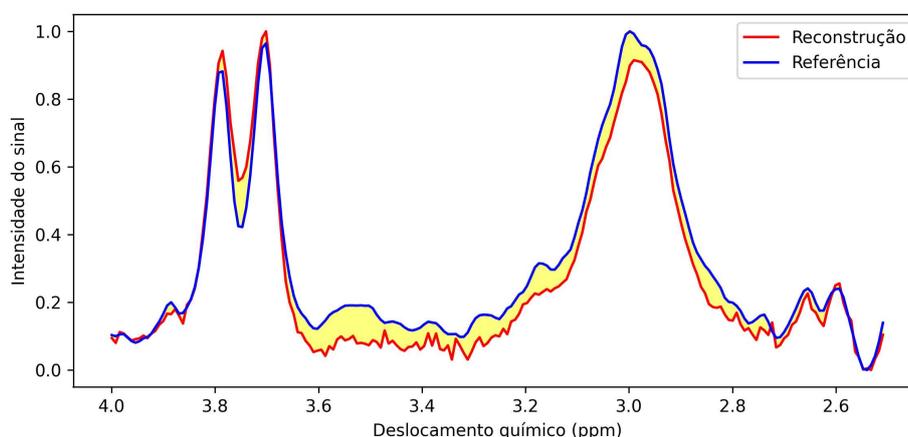


Figura 3.2 – Intervalo no espectro de diferença utilizado para o cálculo do MSE. O sinal em vermelho representa uma reconstrução, enquanto o sinal em azul representa o sinal de referência correspondente, ambos gerados para fins ilustrativos. A região destacada em amarelo corresponde à região de diferença entre os sinais.

SNR: calculado dividindo-se o valor máximo do pico de GABA localizado na faixa de 2,8 ppm a 3,2 ppm no espectro pela desvio padrão do ruído do espectro, onde a faixa de 9,8 ppm a 10,8 ppm é considerada ruído puro:

$$\text{SNR} = \frac{\text{MÁXIMO}[\text{espectro}_{2.8-3.2\text{ppm}}]}{2 \cdot \sigma_{9.8-10\text{ppm}}} \quad (3.3)$$

Largura à meia altura (FWHM, do inglês *Full Width at Half Maximum*): representa a largura do pico espectral na metade de sua intensidade máxima. Nesse estudo, FWHM foi calculado sobre o pico de GABA (Fig. 3.3).

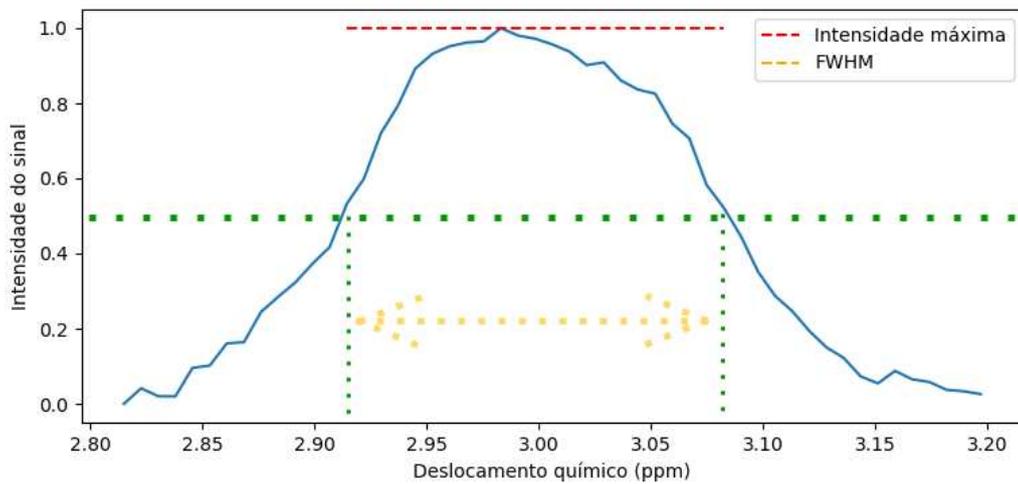


Figura 3.3 – Exemplo ilustrativo para a métrica FWHM, indicando a largura do pico de GABA na metade de sua intensidade máxima (linha pontilhada em amarelo).

Shape Score: calculada computando o coeficiente de correlação de Pearson, ponderando-o com 60% para GABA e 40% para Glx. Os sinais são normalizados utilizando uma abordagem de mínimo-máximo para cada pico. A faixa de ppm para normalizar o pico de GABA é de 2,8 a 3,2 ppm, enquanto para o pico de Glx é de 3,55 a 3,9 ppm (Fig. 3.4). Essa métrica foi proposta pelos organizadores para substituir a avaliação qualitativa dos picos mais relevantes por uma abordagem quantitativa. De maneira semelhante ao MSE, o Shape Score foca na forma dos picos de GABA e Glx, comparando a referência normalizada com a reconstrução normalizada.

$$\text{Shape Score} = 0.6 \times r(x, y)_{2.8-3.2\text{ppm}} + 0.4 \times r(x, y)_{3.55-3.9\text{ppm}} \quad (3.4)$$

O cálculo é feito de acordo com a Equação 3.4, onde r é o coeficiente de correlação de Pearson para a região espectral subscrita, x é a reconstrução do modelo e y é o espectro

de referência. A região de 2,8–3,2 ppm refere-se ao pico de GABA e a região de 3,55–3,9 ppm refere-se ao pico de Glx.

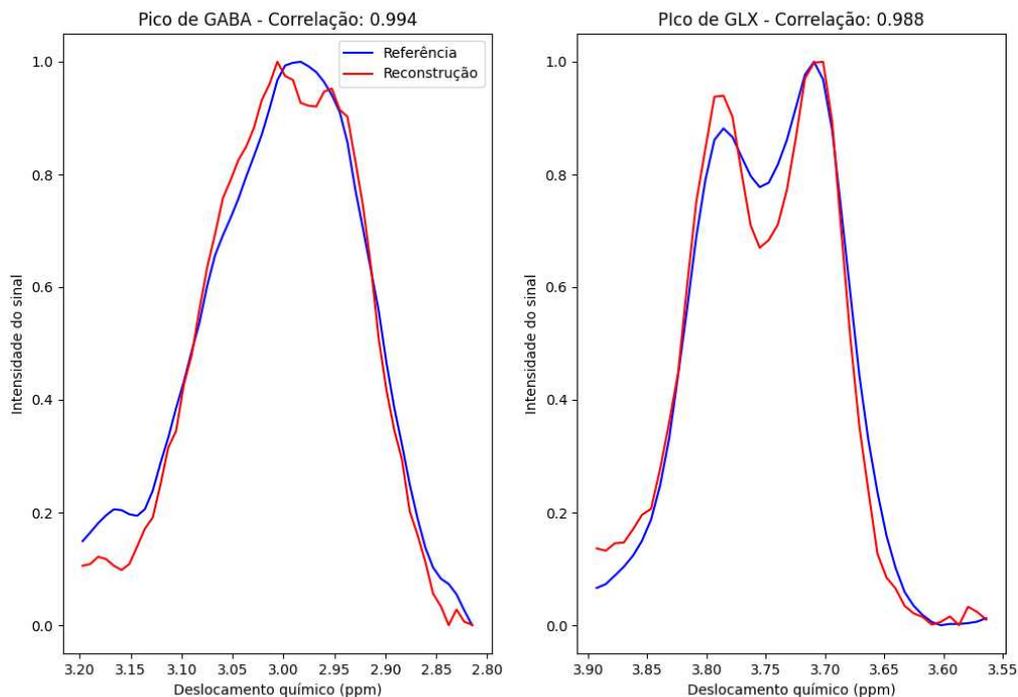


Figura 3.4 – Picos de GABA (à esquerda) e Glx (à direita) usados no cálculo do Shape Score, com seus respectivos valores de correlação calculados entre a reconstrução e a referência destacados no título de cada gráfico. O sinal em vermelho representa a reconstrução, enquanto o sinal em azul representa o sinal de referência correspondente.

Erro de Fit: calculado para GABA/Água utilizando a função de *fitting* do Gannet com configurações padrões. Como essa métrica é calculada com base em um espectro individual, dois resultados suplementares foram adicionados para comparação: referência — que é o espectro gerado pelo Gannet com os 320 transientes completos; e Gannet Quarter — que é o espectro gerado pelo Gannet com apenas 80 transientes. Essa métrica também foi introduzida pelos organizadores do desafio para a interpretação dos resultados, mas não foi empregada para fazer a ordenação dos modelos submetidos pelas equipes.

3.3 Modelo Proposto: Spectro-ViT

3.3.1 Arquitetura do Modelo

A arquitetura proposta foi baseada em *Transformers*, por sua habilidade de modelar dependências de longo alcance por meio de mecanismos de atenção, se destacando na extração de características de dados sequenciais. Eles gerenciam eficientemente dependências entre elementos da sequência de entrada e possibilitam um processamento paralelo eficaz (ISLAM *et al.*, 2023). A capacidade de processar todas as partes dos dados simultaneamente e de ponderar a importância de cada parte com base no contexto da sequência inteira os torna uma ferramenta promissora para trabalhar com espectrogramas (GONG *et al.*, 2021). Além disso, a eficácia da combinação de abordagens de DL em domínio tempo-frequência foi demonstrada no campo de MRS, incluindo quantificação e remoção de artefatos (KYATHANAHALLY *et al.*, 2018a; SHAMAEI *et al.*, 2021; RIZZO *et al.*, 2023).

O Spectro-ViT inicia com o modelo Vision Transformer (ViT), que recebe como entrada (3 canais) a componente real do espectrograma. A saída do ViT é uma representação latente (*embedding*) unidimensional de tamanho 768×1 , que passa por uma sequência de camadas totalmente conectadas com a função de ativação ReLU (do inglês, *Rectified Linear Unit* - ReLU): começando com uma camada de neurônios de tamanho 512×1 , passando para uma de tamanho 1024×1 e, finalmente, alcançando uma de tamanho 2048×1 . Esta última camada resulta em um espectro de diferença GABA MEGA-PRESS contendo 2048 pontos espectrais. Para reconstruções de sinais com 4096 pontos, uma camada adicional de tamanho 4096×1 é anexada à camada existente de tamanho 2048×1 (Fig. 3.5).

3.3.2 Preparação dos Dados

Trilha 1:

Os transientes simulados utilizados como entrada para o modelo foram criados adicionando ruído de amplitude e deslocamentos de fase e frequência aos dados simulados ideais. A partir de cada uma das 5000 amostras simuladas, foram gerados 80 transientes (ou seja, 40 Edit-ON e 40 Edit-OFF). Para isso, foi utilizado um código em Python desenvolvido e compartilhado pelos organizadores do desafio. Com o par ideal de sub-sinais (Edit-ON e Edit-OFF), foi computado o espectro de diferença, que está livre de ruído e artefatos, e foi utilizado como referência na Trilha 1 para o treinamento da reconstrução do modelo (Fig. 3.6).

Trilhas 2 e 3:

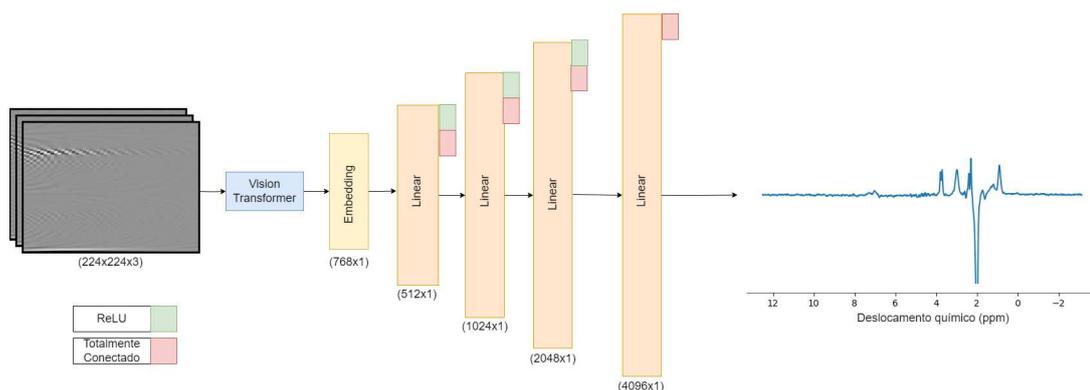


Figura 3.5 – O modelo ViT recebe o espectrograma como entrada em 3 canais e gera uma representação latente unidimensional de tamanho 768×1 , que passa por camadas totalmente conectadas com ativação ReLU (512×1 , 1024×1 , e 2048×1). Esta última camada resulta no espectro de diferença GABA MEGA-PRESS com 2048 pontos. Para a reconstrução do sinal com 4096 pontos, é necessária uma camada de neurônios adicional de tamanho 4096×1 . Fonte: (DIAS *et al.*, 2023)

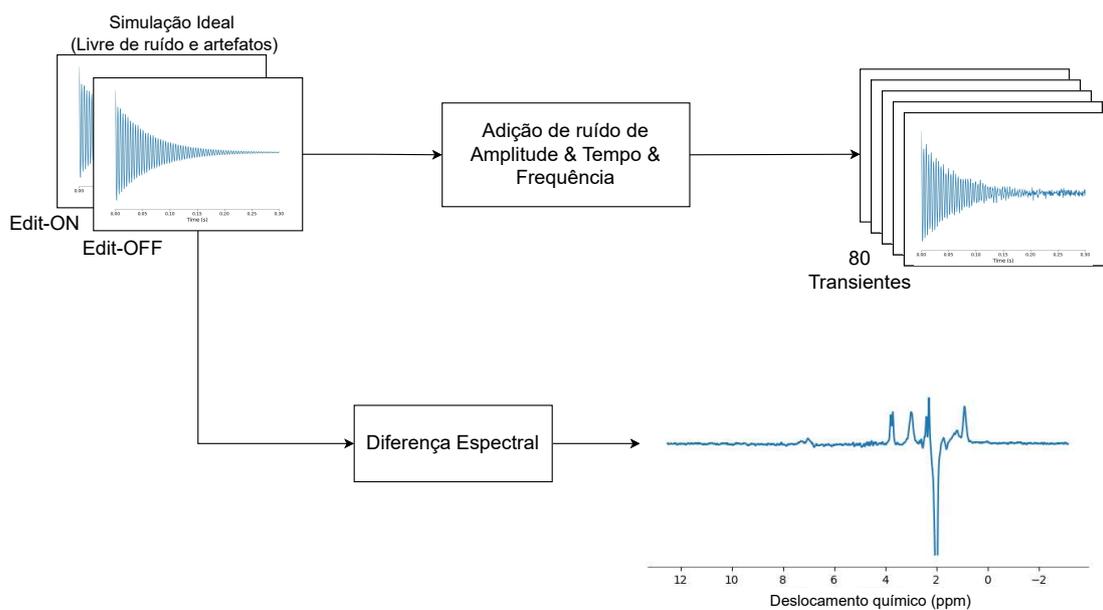


Figura 3.6 – A partir de uma simulação ideal de uma aquisição GABA MEGA-PRESS que gera um par de sinais FID (Edit-ON e Edit-OFF), foram criados 80 transientes (40 por sub-sinal) adicionando ruído de amplitude e deslocamentos de fase e frequência para servir como entrada do modelo. O sinal alvo para reconstrução é o espectro gerado pela diferença espectral do sinal sem ruído.

Como cada conjunto de dados *in-vivo* consistiu em uma aquisição com 320 transientes, e o modelo utiliza apenas 80 transientes, foi empregado um método para gerar novas amostras a partir dos dados *in-vivo* sem simplesmente descartar 240 ($320-80$) transientes. O método utiliza uma janela deslizante que se move ao longo do eixo dos transientes com um passo unitário e um tamanho de 80, criando grupos de 80 transientes

por sub-sinal (ON e OFF). No final, esse método gerou 120 amostras por aquisição (Fig. 3.7).

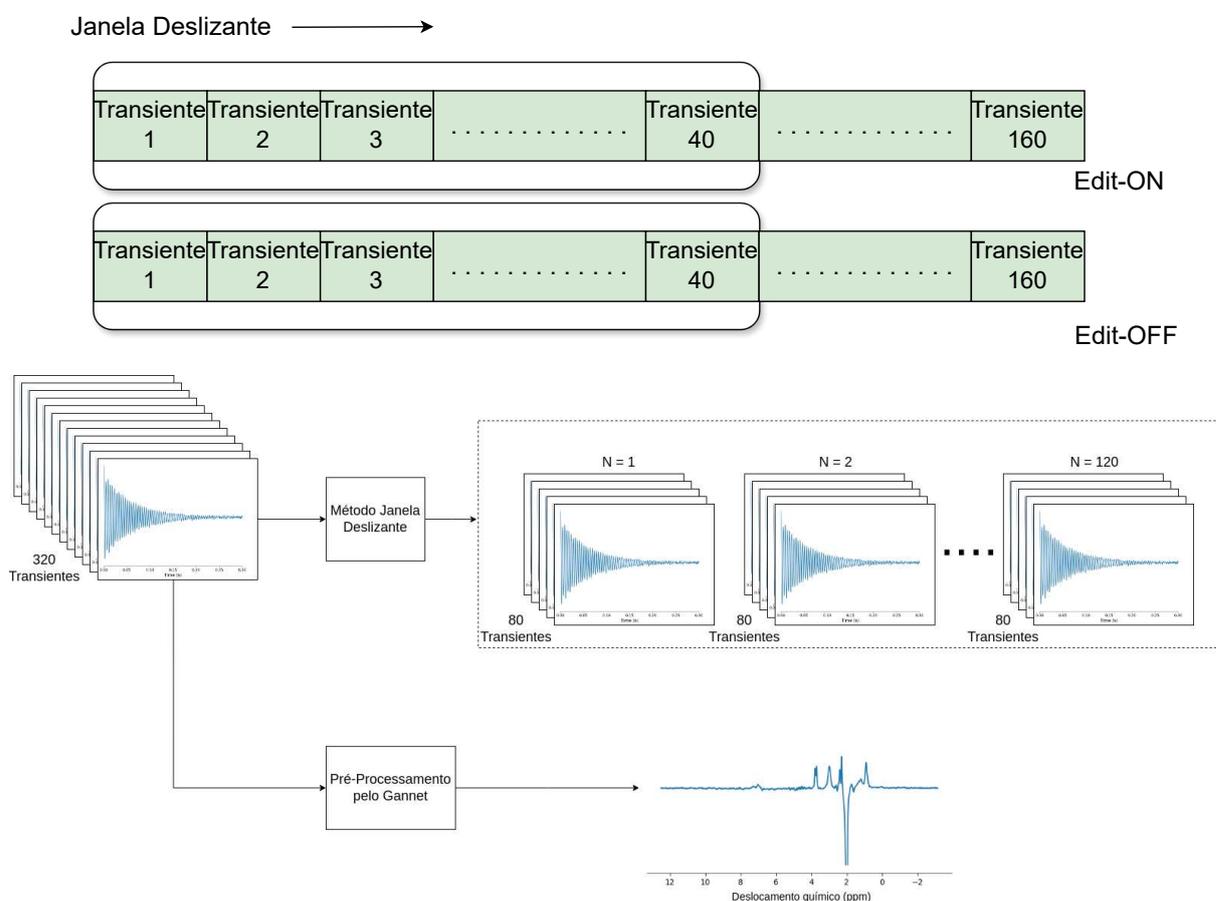


Figura 3.7 – O método da janela deslizante é um processo de amostragem no qual uma janela se move ao longo do eixo dos transientes com um passo unitário, formando grupos de 40 transientes por sub-sinal. Para 320 transientes (160 por sub-sinal), o algoritmo realiza 120 iterações, resultando na criação de 120 conjuntos de 80 transientes.

3.3.3 Geração dos Espectrogramas

O processo de geração do espectrograma a partir de 80 transientes envolveu os seguintes passos: subtração de 40 pares de FIDs, especificamente subtraindo 40 sinais Edit-OFF de 40 sinais Edit-ON; cálculo da média dos 40 sinais diferenciais resultantes; aplicação da STFT selecionando os coeficientes da STFT localizados no eixo positivo do deslocamento químico; normalização dos coeficientes da STFT dividindo pelo valor máximo da magnitude; seleção apenas da componente real; e aplicação de preenchimento de zeros (ou *zero-padding* do inglês) para ajustar o tamanho do espectrograma de 204×206 para 224×224 pixels, garantindo compatibilidade com a arquitetura do ViT utilizado (Fig. 3.8). Cada espectrograma resultante foi replicado ao longo de três canais que servem de entrada para o ViT.

Para o algoritmo da STFT, utilizou-se uma janela Hann com um tamanho de 256, com deslocamento entre as janelas de 10 para sinais de 2048 pontos e 64 para sinais de 4096 pontos.

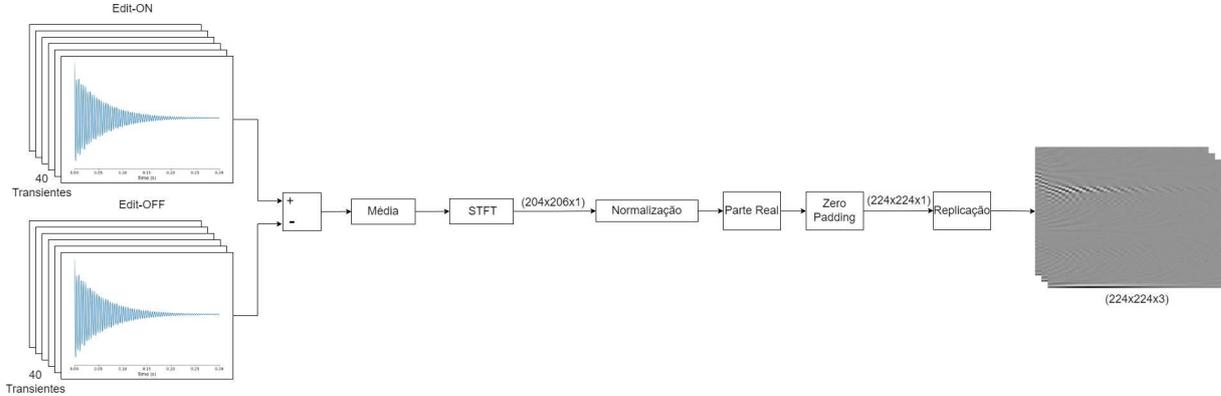


Figura 3.8 – O processo de geração do espectrograma envolveu: subtração de 40 pares de FIDs (Edit-ON - Edit-OFF); cálculo da média dos 40 sinais diferenciais; aplicação da STFT selecionando os coeficientes da STFT localizados no eixo positivo do deslocamento químico; normalização dos coeficientes da STFT pela magnitude máxima; seleção da componente real; e preenchimento de zeros (*zero-padding*) para ajustar o espectrograma para 224×224 pixels.

3.3.4 Treinamento

Cada amostra consiste em 40 transientes por sub-sinal (Edit-ON e Edit-OFF), usados para gerar o espectrograma de entrada do modelo. Na Trilha 1, foram geradas 5000 amostras: 4000 para treinamento e 1000 para validação. Na Trilha 2, cada aquisição gerou 120 amostras, totalizando 1440 amostras, das quais 1200 foram usadas para treinamento e 240 para validação. A mesma metodologia foi aplicada à Trilha 3 com 2048 pontos. Na Trilha 3 com 4096 pontos, cada aquisição gerou 120 amostras, totalizando 2160 amostras, das quais 1800 foram usadas para treinamento e 360 para validação (Tab. 3.2).

Tabela 3.2 – Descrição do conjunto de dados usado para treinamento e validação nas diferentes trilhas do desafio.

Conjunto	Trilha 1 (amostra)	Trilha 2 (amostra/aquisição)	Trilha 3 (2048) 2048 (amostra/aquisição)	Trilha 3 (4096) 4096 (amostra/aquisição)
Treino	4000	1200 / 10	1200 / 10	1800 / 15
Validação	1000	240 / 2	240 / 2	360 / 3

Quatro modelos distintos foram treinados de forma supervisionada para cada conjunto de dados (Trilha 1, Trilha 2, Trilha 3 com 2048 pontos, e Trilha 3 com 4096 pontos). Considerando a crescente complexidade dos conjuntos de dados (Trilha 1: simulado, Trilha 2: *in-vivo* homogêneo, Trilha 3: *in-vivo* heterogêneo), foi utilizada a técnica

de transferência de aprendizado. A ideia foi usar os pesos aprendidos de um conjunto de dados como inicialização para o próximo modelo. Especificamente, o modelo da Trilha 2 foi inicializado com os pesos da Trilha 1, o modelo da Trilha 3 com 2048 pontos com os pesos da Trilha 2, e o modelo da Trilha 3 com 4096 pontos com os pesos da Trilha 3 com 2048 pontos (Fig. 3.9). No treinamento da Trilha 1, os pesos do ViT foram carregados a partir do ViT pré-treinado na famosa base de dados ImageNet (DENG *et al.*, 2009).

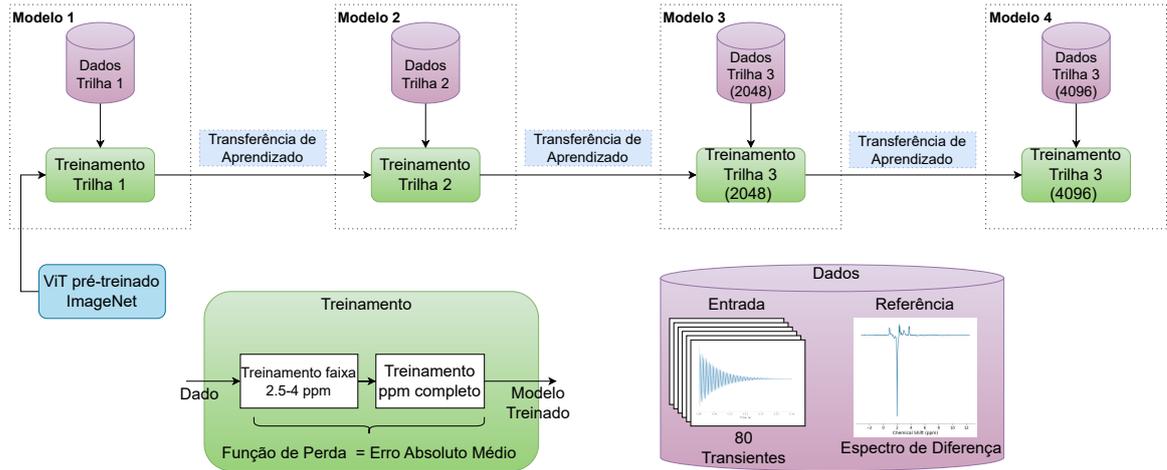


Figura 3.9 – Para cada conjunto de dados, um modelo foi treinado em duas etapas: primeiro, para minimizar o MAE na faixa de ppm de 2,5 a 4; depois, para minimizar o MAE em toda a faixa espectral. Os pesos aprendidos de uma Trilha foram usados para inicializar outra Trilha (transferência de aprendizado) e, em seguida, sofreram um ajuste fino (*fine tuning*).

O treinamento envolveu duas etapas: primeiro, focou-se na região espectral de maior relevância (2,5 a 4 ppm) e, posteriormente, em toda a região espectral utilizando a função de perda correspondente ao erro absoluto médio (MAE, do inglês *Mean Absolute Error*) (Fig. 3.9). Para cada uma das duas etapas de treinamento, o modelo foi treinado por 20 épocas com um tamanho de lote (ou *batch size* do inglês) de 100 e uma taxa de aprendizagem (LR, do inglês *learning rate*) constante de 1×10^{-4} , utilizando o algoritmo de Estimativa de Momento Adaptativo (ADAM, do inglês *Adaptive Moment Estimation*) (KINGMA; BA, 2014). O conjunto de dados foi embaralhado uma vez no início de cada época. Adicionalmente, para as Trilhas 2 e 3, foi aplicado um método de aumento de dados (*data augmentation*) individualmente a cada amostra no momento em que era carregada no modelo. Esse procedimento introduziu ruído de amplitude e desvios de frequência e fase nos transientes. O treinamento foi realizado usando uma GPU NVIDIA GeForce RTX 3060, com a versão do driver 525.147.05 e a versão CUDA 12.0.

3.4 Resultados

Os modelos foram avaliados utilizando as métricas MSE, SNR, FWHM e Shape Score, além de uma análise qualitativa por meio de visualizações gráficas do espectro de diferença. Apenas os organizadores do desafio tinham acesso aos dados de referência do conjunto de teste. Por esse motivo, a maioria dos gráfico não pôde ser gerada por nós e precisou ser trazida do artigo que publicou os resultados do desafio (BERTO *et al.*, 2024). O erro de ajuste também foi considerado para a interpretação dos resultados, embora não tenha sido utilizado para fins de ordenação dos modelos submetidos.

3.4.1 Trilha 1: Dados Simulados

O modelo Spectro-ViT obteve a segunda melhor reconstrução dentre os times participando do desafio. Seus valores foram comparáveis aos do modelo primeiro colocado, Spectralligence, destacando-se na métrica FWHM, onde o Spectro-ViT foi ligeiramente superior (Tab. 3.3). O Spectro-ViT apresentou uma SNR alta, mas dentro de uma faixa mais realista, com uma reconstrução geral mais fiel, demonstrando boa reprodução dos picos na região de 2,5 - 4 ppm, bem como entre 1 e 2,5 ppm (Fig. 3.10).

A equipe Spectralligence obteve os melhores resultados na Trilha 1 em 3 das 4 métricas (Tab. 3.3), sendo que suas reconstruções conseguiram reproduzir bem os picos de GABA e Glx na região de maior interesse (2,5 - 4 ppm) (Fig. 3.10). No entanto, seu modelo apresentou resultados de SNR irrealistas, com um valor médio de 2361,5. A análise qualitativa complementar dos espectros reconstruídos apontou que o modelo aprendeu um sobreajuste para suprimir o ruído predominantemente na região de cálculo de ruído (9,8-10 ppm), resultando em uma SNR artificialmente elevada. Além disso, houve regiões do espectro puramente ruidosas que não foram bem reconstruídas, como, por exemplo, na faixa de 1 até 2,5 ppm (Fig. 3.10).

O modelo da equipe Dolphins apresentou o pior desempenho entre todos, com resultados consideravelmente inferiores, inclusive piores do que o modelo base U-Net. O modelo apresentou um valor médio de SNR de apenas 2,0, um baixo valor médio de Shape Score e valores elevados de MSE e FWHM em comparação com os demais modelos (Tab. 3.3). Esses resultados indicam uma reconstrução de baixa qualidade, com maior presença de ruído e picos menos definidos no espectro (Fig. 3.10).

3.4.2 Trilha 2: Dados *in-vivo* com mesmos parâmetros de aquisição

A maioria dos modelos apresentou uma queda no desempenho ao passar da Trilha 1 para a Trilha 2, o que já era esperado, uma vez que a Trilha 1 era composta de dados simulados, enquanto a Trilha 2 apresentava dados *in-vivo*. O Spectro-ViT foi

Tabela 3.3 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score nos resultados do conjunto de teste para a Trilha 1, comparando as reconstruções obtidas pelos modelos participantes do desafio. Resultados em negrito indicam o melhor desempenho para cada métrica, enquanto o asterisco indica os modelos cujos resultados não foram considerados na ordenação final do desafio. Fonte: (BERTO *et al.*, 2024).

Modelo	MSE	SNR	FWHM	Shape Score
U-Net base*	0,03452 ± 0,00748	91,3 ± 64,2	0,076661 ± 0,006877	0,7000 ± 0,1359
Spectro-ViT	0,00100 ± 0,00078	80,2 ± 21,8	0,076447 ± 0,000242	0,9994 ± 0,0004
Spectralligence	0,00089 ± 0,00077	2361,5 ± 768,2	0,076455 ± 0,000000	0,9996 ± 0,0002
Dolphins	0,08971 ± 0,02920	2,0 ± 0,6	0,343908 ± 0,547723	0,2366 ± 0,2620

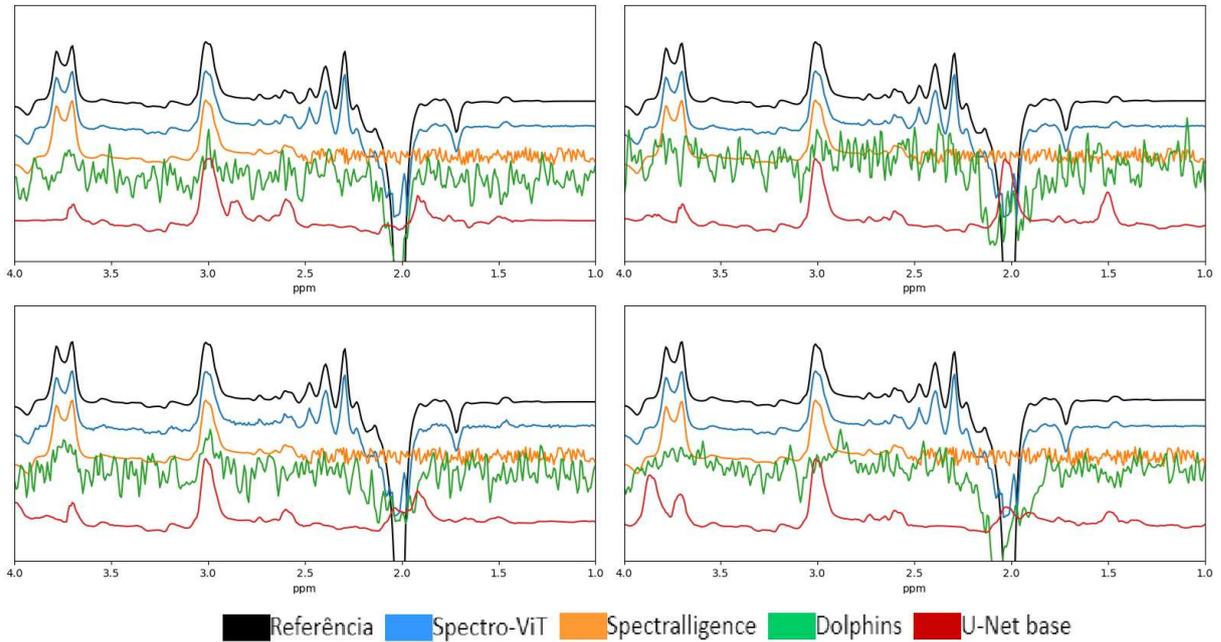


Figura 3.10 – Cada gráfico se refere a uma amostra do conjunto de teste da Trilha 1 e contém cinco espectros: espectro de referência, gerado utilizando 320 transientes com o software Gannet (preto); reconstrução feita pelo Spectro-ViT (azul); reconstrução feita pelo modelo U-Net base (vermelho); reconstrução pelo modelo da equipe Spectralligence (laranja); reconstrução pelo modelo da equipe Dolphins (verde). Os espectros reconstruídos (azul, vermelho, laranja e verde) foram gerados a partir dos primeiros 80 transientes das aquisições do conjunto de teste. O deslocamento vertical entre espectros serve apenas para melhor visualização. Fonte: (BERTO *et al.*, 2024).

o modelo que obteve os melhores resultados da Trilha 2, mesmo com quedas em todas as métricas: a SNR média reduziu de 80,2 para 48,9 (~39%), o MSE médio aumentou de 0,00100 para 0,01529 (~15x), o FWHM médio aumentou de 0,076447 para 0,153865 (~2x), e o Shape Score médio diminuiu de 0,9994 para 0,9809 (~2%) (Tabs. 3.3 e 3.4).

A equipe Spectralligence também apresentou uma queda de desempenho. A SNR média caiu de 2361,5 para 47,1, atingindo valores mais realistas. O MSE médio subiu de 0,00089 para 0,00999 (~11x), o FWHM médio piorou de 0,076455 para 0,163103

($\sim 2x$), e o Shape Score médio diminuiu de 0,9996 para 0,9769 ($\sim 2\%$) (Tabs. 3.3 e 3.4).

Curiosamente, a equipe Dolphins apresentou uma melhora na Trilha 2, embora a SNR média tenha permanecido baixo (6,4). O MSE médio reduziu de 0,08971 para 0,03574 ($\sim 60\%$), o FWHM médio melhorou de 0,343908 para 0,0171704 ($\sim 50\%$), e o Shape Score médio aumentou de 0,2366 para 0,7037 ($\sim 3x$) (Tabs. 3.3 e 3.4).

Tabela 3.4 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score nos resultados do conjunto de teste para a Trilha 2, comparando as reconstruções obtidas pelos modelos participantes do desafio. Resultados em negrito indicam o melhor desempenho para cada métrica, enquanto o asterisco indica os modelos cujos resultados não foram considerados na ordenação final do desafio. Fonte: (BERTO *et al.*, 2024).

Modelo	MSE	SNR	FWHM	Shape Score
U-Net base*	0,03486 \pm 0,03254	45,2 \pm 24,4	0,147175 \pm 0,042553	0,8283 \pm 0,2013
Spectro-ViT	0,01529 \pm 0,02008	48,9 \pm 2,6	0,153865 \pm 0,002583	0,9809 \pm 0,0202
Spectralligence	0,00999 \pm 0,01050	47,1 \pm 23,1	0,163103 \pm 0,006637	0,9769 \pm 0,0233
Dolphins	0,03574 \pm 0,01616	6,4 \pm 1,6	0,171704 \pm 0,044005	0,7037 \pm 0,1312

A análise da distribuição das métricas evidencia que, as métricas em que o Spectro-ViT demonstrou superioridade (SNR, FWHM e Shape Score), apresentam robustez com uma pequena variância (Fig. 3.11). Em relação à análise qualitativa dos espectros reconstruídos, os espectros gerados pelo modelo Spectro-ViT e pelo modelo da equipe Spectralligence demonstraram reproduzir com maior fidelidade a curva de referência na faixa entre 2,5 e 4 ppm. No entanto, o modelo da equipe Spectralligence, assim como ocorreu na Trilha 1, não representa adequadamente os picos esperados entre 1 e 2,5 ppm, deixando essa faixa espectral com predominância de ruído (Fig. 3.12).

3.4.3 Trilha 3: Dados *in-vivo* com diferentes parâmetros de aquisição

Na transição da Trilha 2 para a Trilha 3, todos os modelos tiveram piora no desempenho, o que também já era esperado, já que a Trilha 2 era composta por dados *in-vivo* homogêneos, enquanto a Trilha 3 apresentava dados *in-vivo* heterogêneos (diferentes equipamentos e diferentes protocolos de aquisição). O Spectro-ViT continuou sendo o melhor modelo, com redução do MSE médio de 0,01529 para 0,01271 ($\sim 17\%$). Já a SNR média caiu de 48,9 para 44,2 ($\sim 10\%$), o FWHM médio piorou de 0,153865 para 0,182649 ($\sim 19\%$) e o Shape Score médio caiu de 0,9809 para 0,9486 ($\sim 3\%$) (Tabs. 3.4 e 3.5).

Nas reconstruções da equipe Spectralligence, a SNR média caiu de 47,1 para 33,9 (28%), o MSE médio aumentou de 0,00999 para 0,01697 ($\sim 1,7x$), o FWHM médio subiu de 0,163103 para 0,166503 ($\sim 2\%$), e o Shape Score médio piorou de 0,9769 para 0,9195 ($\sim 6\%$) (Tabs. 3.4 e 3.5).

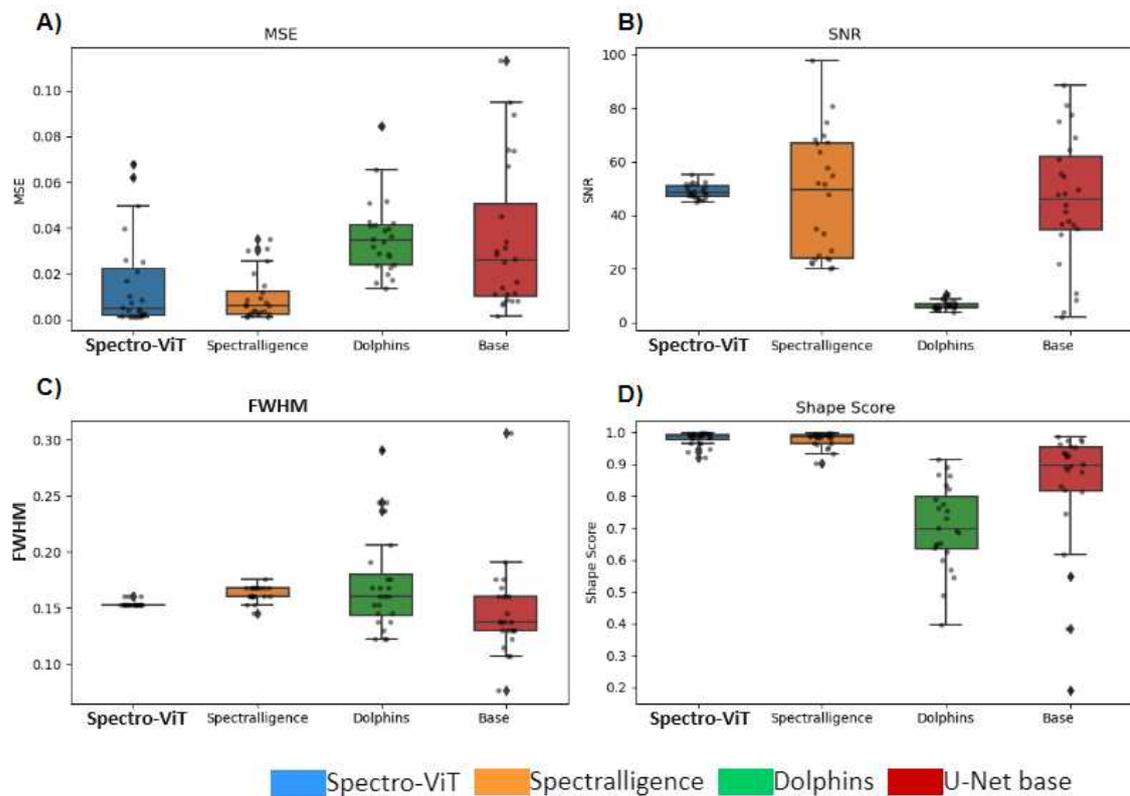


Figura 3.11 – Distribuição de métricas dos resultados de reconstrução na Trilha 2: A) MSE; B) SNR; C) FWHM; D) Shape Score. Resultados obtidos pelos modelos: Spectro-ViT (azul); U-Net (vermelho); Spectralligence (laranja); Dolphins (verde). As reconstruções foram feitas com os primeiros 80 transientes das aquisições do conjunto de teste. Fonte: (BERTO *et al.*, 2024).

O modelo da equipe Dolphins novamente apresentou os piores resultados. Em relação à Trilha 2, o MSE médio aumentou de 0,03574 para 0,07279 ($\sim 2x$), o FWHM médio subiu de 0,171704 para 0,191015 ($\sim 11\%$), e o Shape Score médio caiu de 0,7037 para 0,3700 ($\sim 47\%$) (Tabs. 3.4 e 3.5).

Tabela 3.5 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score nos resultados do conjunto de teste para a Trilha 3, comparando as reconstruções obtidas pelos modelos participantes do desafio. Resultados em negrito indicam o melhor desempenho para cada métrica, enquanto o asterisco indica os modelos cujos resultados não foram considerados na ordenação final do desafio. Fonte: (BERTO *et al.*, 2024).

Modelo	MSE	SNR	FWHM	Shape Score
U-Net base*	0,02420 \pm 0,01711	35,6 \pm 17,5	0,172427 \pm 0,046778	0,8253 \pm 0,1040
Spectro-ViT	0,01271 \pm 0,03592	44,2 \pm 10,6	0,182649 \pm 0,030305	0,9486 \pm 0,2090
Spectralligence	0,01697 \pm 0,02586	33,9 \pm 13,9	0,166503 \pm 0,021886	0,9195 \pm 0,2118
Dolphins	0,07279 \pm 0,05466	5,9 \pm 2,0	0,191105 \pm 0,040382	0,3700 \pm 0,4909

Mais uma vez, a análise da distribuição das métricas obtidas pelos modelos mostram uma variância bem menor para o Spectro-ViT (Fig. 3.13). Na análise qualitativa dos espectros reconstruídos, as reconstruções para as amostras adquiridas em equipamento

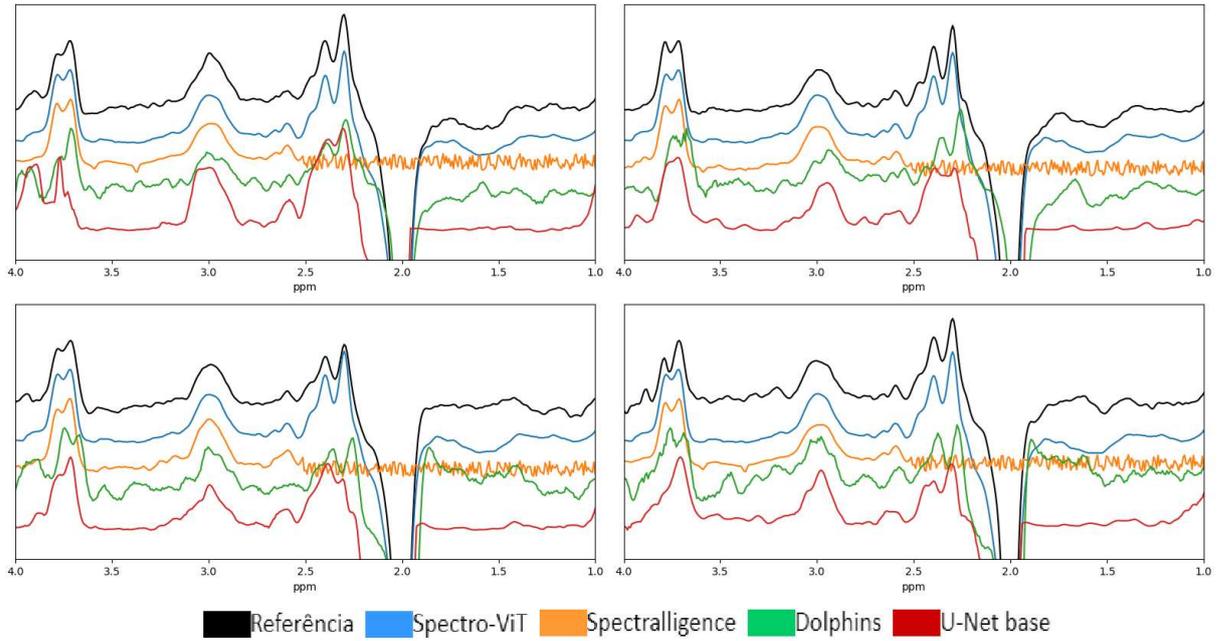


Figura 3.12 – Cada gráfico se refere a uma amostra do conjunto de teste da Trilha 2 e contém cinco espectros: espectro de referência, gerado utilizando 320 transientes com o software Gannet (preto); reconstrução feita pelo Spectro-ViT (azul); reconstrução feita pelo modelo U-Net base (vermelho); reconstrução pelo modelo da equipe Spectralligence (laranja); reconstrução pelo modelo da equipe Dolphins (verde). Os espectros reconstruídos (azul, vermelho, laranja e verde) foram gerados a partir dos primeiros 80 transientes das aquisições do conjunto de teste. O deslocamento vertical entre espectros serve apenas para melhor visualização. Fonte: (BERTO *et al.*, 2024).

GE seguiram melhor a referência do que as reconstruções para amostras dos equipamentos da Philips e Siemens, provavelmente porque a GE possui mais amostras de treinamento para ambas as Trilhas 2 e 3 (Fig. 3.14). Os espectros gerados pelo Spectro-ViT e pela equipe Spectralligence reproduziram com maior fidelidade a curva de referência na faixa entre 2,5 e 4 ppm. No entanto, o modelo da equipe Spectralligence continuou não representando adequadamente os picos esperados entre 1 e 2,5 ppm.

3.4.4 Análise Adicional: Erro de ajuste

A análise de distribuição de erro de ajuste (*Fit error*) revelou que os modelos Spectro-ViT e da equipe Spectralligence apresentaram as melhores distribuições para as Trilhas 2 e 3. Ambos superaram a referência, que utiliza o Gannet processando 320 transientes, em termos de menores valores de erro de ajuste. Para a Trilha 2, o Spectro-ViT se destacou, apresentando menores e mais consistentes erros de ajuste em comparação ao modelo da equipe Spectralligence. Já na Trilha 3, o modelo da equipe Spectralligence teve os melhores resultados, com uma diminuição na variância, enquanto a do Spectro-ViT aumentou (Fig. 3.15).

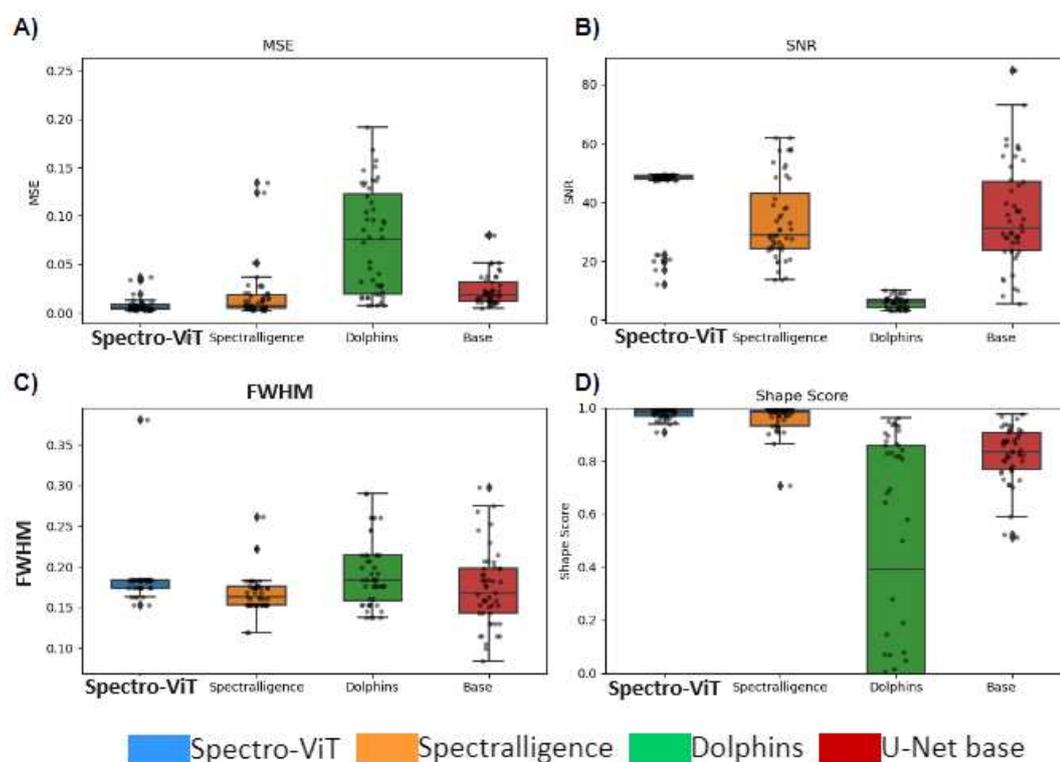


Figura 3.13 – Distribuição de métricas dos resultados de reconstrução na Trilha 3: A) MSE; B) SNR; C) FWHM; D) Shape Score. Resultados obtidos pelos modelos: Spectro-ViT (azul); U-Net (vermelho); Spectralligence (laranja); Dolphins (verde). As reconstruções foram feitas com os primeiros 80 transientes das aquisições do conjunto de teste. Fonte: (BERTO *et al.*, 2024).

O modelo da equipe Dolphins, por outro lado, apresentou resultados mais dispersos com alta variância entre todas as distribuições, especialmente na Trilha 3 (Fig. 3.15). Curiosamente, apesar das métricas ruins para essa Trilha (Tab. 3.5), o modelo da equipe Dolphins obteve os menores valores de erro de ajuste. Isso ressalta a importância de uma análise conjunta com outras métricas, pois um sinal ruidoso e com artefatos pode, embora seja menos provável, ainda apresentar baixos erros de ajuste.

A análise visual dos espectros reconstruídos para algumas amostras da Trilha 2 e Trilha 3 fornece informações complementares (Fig. 3.16). Na aquisição rotulada como “Scan #1” para a Trilha 3 (gráfico inferior esquerdo), o espectro reconstruído pelo modelo da equipe Dolphins (verde) não consegue representar bem o pico de GABA se comparado à referência (roxo), mas mesmo assim obteve um erro de ajuste de 1,7%, enquanto a referência apresentou 4,4%. Além disso, é interessante observar que, em todos os gráficos, os espectros reconstruídos pelo modelo da equipe Spectralligence, embora tenham alcançado baixos valores de erro de ajuste, não aproximam bem a referência fora da faixa de 4 a 2,5 ppm. Isso acontece porque o Gannet realiza a modelagem do ajuste entre 2,79 e 4,1 ppm (Fig. 3.16).

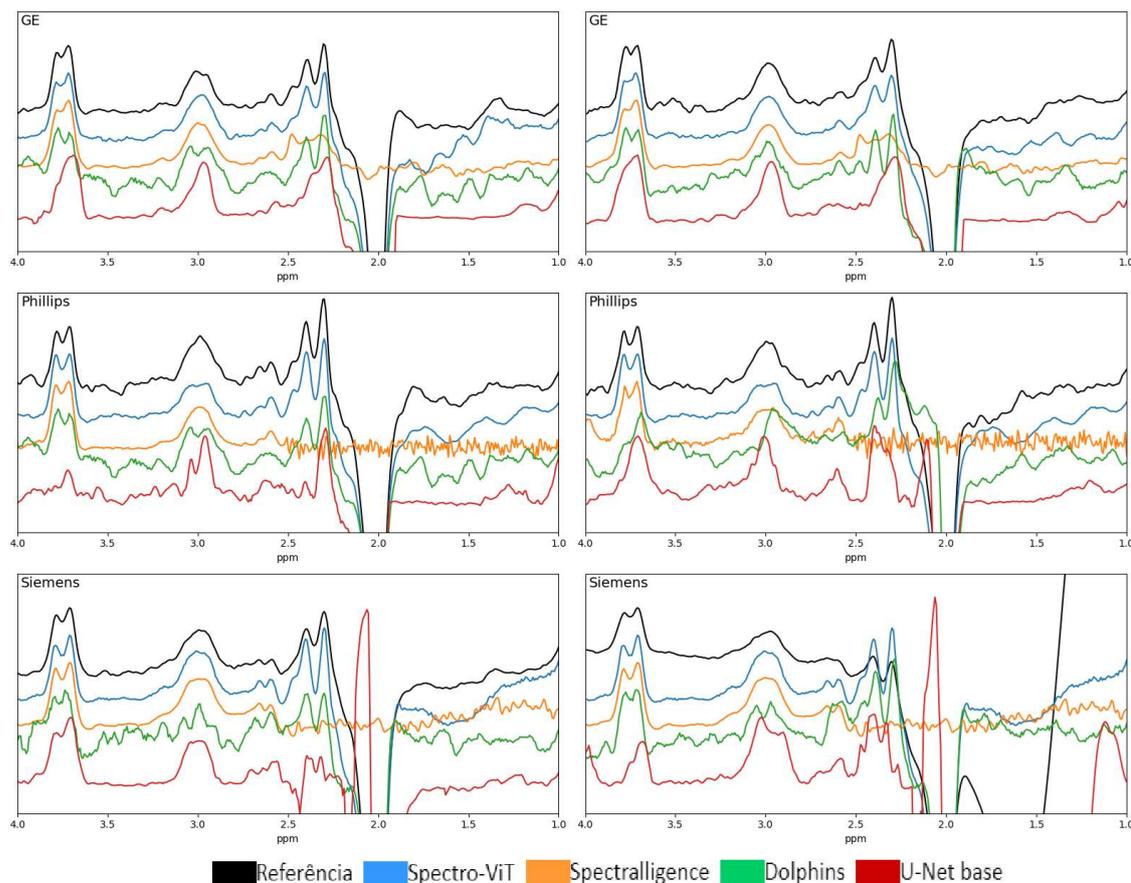


Figura 3.14 – Cada gráfico se refere a uma amostra do conjunto de teste da Trilha 3, tendo sido escolhidas 2 amostras de cada fabricante: GE (fileira superior); Philips (fileira do meio); e Siemens (fileira de baixo). Cada gráfico contém cinco espectros: espectro de referência, gerado utilizando 320 transientes com o software Gannet (preto); reconstrução feita pelo Spectro-ViT (azul); reconstrução feita pelo modelo U-Net base (vermelho); reconstrução pelo modelo da equipe Spectralligence (laranja); reconstrução pelo modelo da equipe Dolphins (verde). Os espectros reconstruídos (azul, vermelho, laranja e verde) foram gerados a partir dos primeiros 80 transientes das aquisições do conjunto de teste. O deslocamento vertical entre espectros serve apenas para melhor visualização. Fonte: (BERTO *et al.*, 2024).

3.5 Discussão

O desafio *Edited-MRS Reconstruction Challenge* propôs investigar e estabelecer um *benchmark* para soluções baseadas em DL na reconstrução de aquisições GABA MEGA-PRESS, utilizando quatro vezes menos dados (80 transientes) em comparação com uma aquisição convencional que utiliza a quantidade total de dados (320 transientes). O desafio foi dividido em três trilhas: a Trilha 1 utilizou dados simulados, a Trilha 2 utilizou dados *in-vivo* com parâmetros de aquisição quase idênticos, e a Trilha 3 utilizou dados *in-vivo* com parâmetros de aquisição diferentes.

Embora não fosse um objetivo explícito do desafio, os participantes puderam

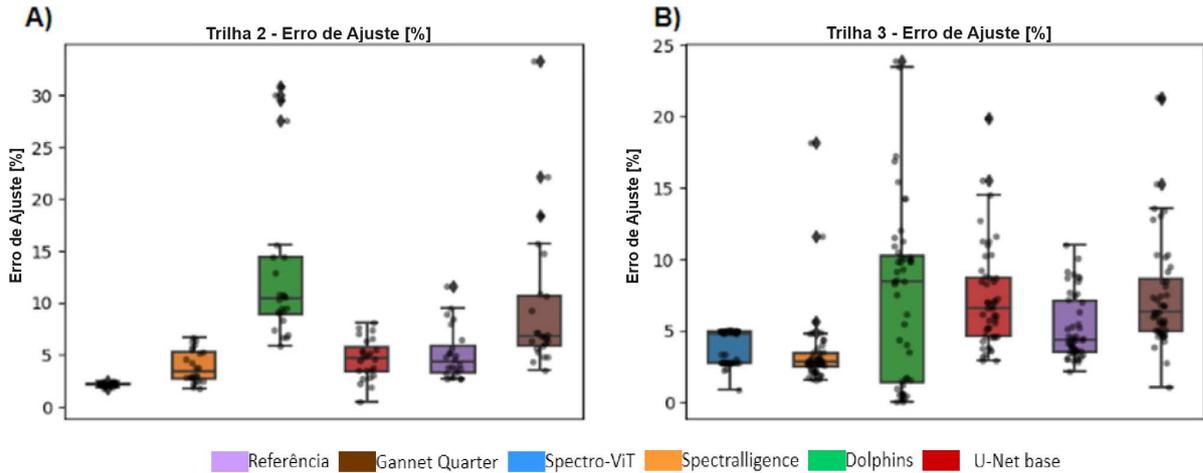


Figura 3.15 – Distribuição dos erros de ajuste de GABA/Água para cada modelo nas Trilha 2 (A) e Trilha 3 (B), incluindo o modelo U-Net base, a referência (Gannet utilizando 320 transientes) e o Gannet Quarter (Gannet utilizando 80 transientes) para a Trilha 2 (A) e a Trilha 3 (B). A distribuição de cada equipe é representada por um gráfico de caixa colorido e cada amostra no conjunto de teste gera um ponto que é marcado por um ponto cinza. A codificação de cores e a ordem são as seguintes: Spectro-ViT (azul); Spectralligence (laranja); Dolphins (verde); U-Net base (vermelho); referência (roxo); e Gannet Quarter (marrom). Fonte: (BERTO *et al.*, 2024).

explorar diferentes representações para os dados espectroscópicos, que são tipicamente unidimensionais. Todas as submissões transformaram essas entradas em representações bidimensionais, o que permitiu aproveitar a vasta literatura de DL sobre modelos bidimensionais. O modelo da equipe Dolphins e o modelo U-Net base utilizam transientes concatenados para formar imagens bidimensionais. O modelo da equipe Spectralligence recebe uma matriz de covariância dos transientes, enquanto o modelo Spectro-ViT utiliza espectrogramas dos transientes como entrada.

Com relação às arquiteturas dos modelos envolvidos no desafio, a maioria teve como base operações de convolução. O modelo base U-Net foi puramente convolucional, os modelos da equipe Spectralligence foram majoritariamente convolucionais, e o modelo da equipe Dolphins foi puramente convolucional. Apenas o modelo Spectro-ViT empregou *Transformers*. A maior parte dos modelos na literatura de DL em MRS empregou CNNs, sendo que nenhum até o momento havia explorado o uso de *Transformers* (SANDE *et al.*, 2023). Dessa forma, a abordagem do Spectro-ViT corresponde à primeira aplicação de *Transformers* em MRS reportada na literatura. O modelo Spectro-ViT foi considerado o vencedor geral do desafio, obtendo o melhor desempenho nas trilhas mais desafiadoras, as Trilhas 2 e 3, que abordaram dados *in-vivo*. O modelo da equipe Spectralligence obteve a melhor performance na Trilha 1, que consistia em dados simulados.

As métricas para o desafio foram selecionadas combinando métricas convencio-

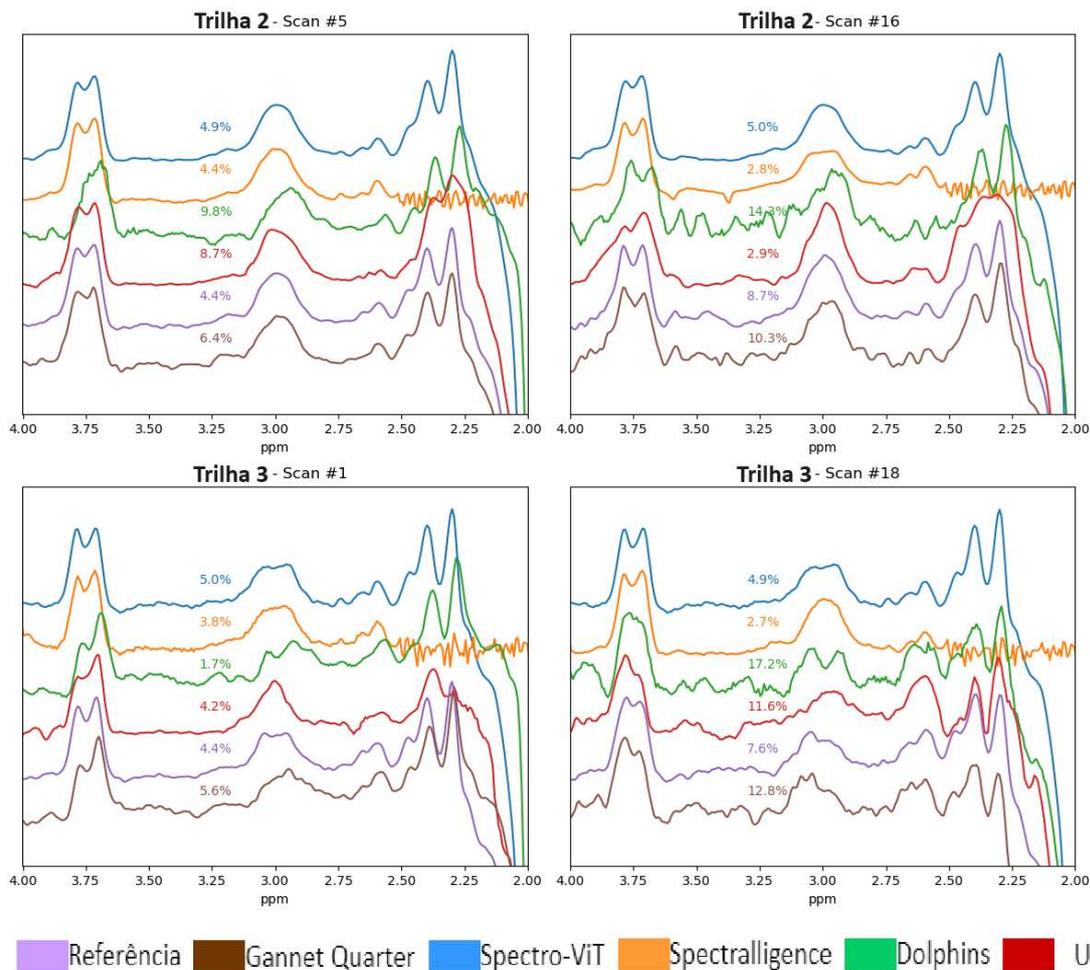


Figura 3.16 – Amostras de espectros da Trilha 2 (fileira superior) e Trilha 3 (fileira inferior) com os respectivos erros de ajuste (*Fit error*). Cada gráfico contém seis espectros: espectro de referência, gerado utilizando 320 transientes com o software Gannet (roxo); Gannet Quarter, gerado utilizando os primeiros 80 transientes com o Gannet (marrom); reconstrução feita pelo modelo Spectro-ViT (azul); reconstrução pelo modelo U-Net base (vermelho); reconstrução pelo modelo da equipe Spectralligence (laranja); reconstrução pelo modelo da equipe Dolphins (verde). Os espectros reconstruídos (azul, vermelho, laranja e verde) foram gerados a partir dos primeiros 80 transientes das aquisições do conjunto de teste. O deslocamento vertical entre espectros serve apenas para melhor visualização. Fonte: (BERTO *et al.*, 2024).

nais utilizadas em MRS (SNR, FWHM, erro de ajuste) e em DL (MSE), além de uma nova métrica proposta pelos organizadores do desafio (Shape Score). Essas métricas incluem tanto aquelas que necessitam de uma referência (MSE e Shape Score) quanto métricas autônomas (SNR, FWHM, erro de ajuste). As escolhas refletem o objetivo de analisar diferentes aspectos dos espectros reconstruídos, mantendo-se alinhadas à literatura existente (LIN *et al.*, 2021).

O MSE foi selecionado como uma métrica mais representativa no campo de

DL em comparação com as métricas tradicionais de qualidade de MRS. O MSE avaliou o erro entre 2,5 e 4 ppm, pois essa faixa contém os picos de GABA e Glx. O Glx foi incluído porque sua qualidade pode ser usada como um indicativo de confiabilidade e qualidade geral do espectro. No entanto, ao avaliar o espectro como um todo para as Trilhas 2 e 3 (Tabs. 3.4 e 3.5), percebeu-se que, embora o modelo da equipe Spectralligence tenha apresentado valores favoráveis de MSE em comparação com os demais valores obtidos no desafio, a reconstrução fora da janela de interesse do modelo não apresentava os picos de metabólitos esperados, divergindo consideravelmente da referência (espectro de diferença gerado com a quantidade completa de 320 transientes) (Figs. 3.10, 3.12, 3.14). Isso pode acontecer dependendo da estratégia de modelagem, como, por exemplo, utilizar uma função de perda que não cubra essas regiões.

Embora modelos que se concentrem apenas em otimizar a qualidade dentro da janela ou região de interesse possam obter melhores resultados para métricas janeladas como o MSE, eles podem estar sacrificando a confiabilidade geral do modelo, pois é incomum para um especialista analisar um espectro que não siga o perfil geral esperado para a aquisição. Além disso, embora o MSE possa fornecer bons *insights* sobre como a reconstrução se relaciona com a aquisição completa, para dados *in-vivo*, o MSE não pode fornecer uma avaliação tão objetiva quanto em um cenário simulado, em que a comparação é feita com uma referência perfeita (*ground truth*). Mesmo a referência *in-vivo*, gerada com a quantidade completa de transientes, apresentará imperfeições nos dados.

A SNR é amplamente utilizado no campo de MRS (LIN *et al.*, 2021; CHOI *et al.*, 2021; NEAR *et al.*, 2021). No entanto, no contexto dessa investigação de reconstrução, ela tem a limitação de considerar apenas uma faixa estreita e específica de ppm onde se espera encontrar ruído. No pré-processamento convencional de MRS, onde o ruído ao longo de todo o espectro é tratado de maneira homogênea, isso não é um problema, pois a faixa de ppm escolhida convencionalmente é uma aproximação adequada do ruído na região dos metabólitos. No entanto, ao considerar modelos de DL (que se comportam como caixa preta), essa suposição não pode ser confirmada, pois um modelo poderia aprender a suprimir o ruído apenas nessa região específica usada para o cálculo da SNR, otimizando a métrica sem melhorar a qualidade geral do sinal. Isso explica o motivo da equipe Spectralligence ter apresentado valores exageradamente altos da SNR para a Trilha 1 e, portanto, irrealis, mesmo tendo intervalos puramente ruidosos no sinal que não preservaram os picos esperados (Figs. 3.10, 3.10, 3.14).

O FWHM é outra métrica amplamente utilizada no campo de MRS (LIN *et al.*, 2021; CHOI *et al.*, 2021; NEAR *et al.*, 2021). Intuitivamente, em MRS, o FWHM é definido como uma métrica a ser minimizada; no entanto, com modelos de DL, é possível obter resultados com larguras de linha menores do que o valor verdadeiro esperado, tornando

difícil interpretar se a reconstrução com um pico mais fino ou mais largo é melhor. Assim como a SNR, essa métrica apresenta limitações ao considerar a natureza de caixa preta dos modelos de DL. Se otimizada para essa métrica, a melhoria significativa pode resultar em valores irrealistas e questionáveis para os especialistas.

A métrica Shape Score requer uma referência como o MSE e, portanto, apresenta limitações em sua aplicação para contextos *in-vivo*, onde não é possível estabelecer um *ground truth* como no caso de dados simulados. Como a métrica foi proposta pela primeira vez para o desafio, são necessárias mais investigações a seu respeito. Ao considerar que a melhor submissão para cada Trilha obteve a melhor classificação para essa métrica (Tabs. 3.3, 3.4 e 3.5), é possível inferir que ela é provavelmente um indicador válido da qualidade de uma reconstrução.

O erro de ajuste foi utilizado como uma métrica adicional, sem influência na ordenação das submissões. Esta métrica foi escolhida por ser amplamente aceita como uma métrica quantitativa para a avaliação da qualidade do espectro e por fornecer uma perspectiva diferente das métricas existentes no desafio (LIN *et al.*, 2021; CHOI *et al.*, 2021; NEAR *et al.*, 2021). Assim como outras métricas, o erro de ajuste também apresenta limitações e seu cálculo envolveu janelamento, uma vez que a faixa de ajuste do Gannet, empregada para o cálculo do erro de ajuste, vai de 2,79 ppm até 4,1 ppm. Dessa forma, a equipe Spectralligence conseguiu obter resultados favoráveis de erro de ajuste mesmo não reconstruindo bem a região fora da faixa de 2,5 ppm até 4 ppm (Fig. 3.16). Também é importante ressaltar que um sinal ruidoso e com artefatos pode, embora seja menos provável, ainda apresentar baixos erros de ajuste, como aconteceu nas reconstruções da equipe Dolphins (Fig. 3.16).

Dadas as métricas do desafio, conclui-se que cada uma tem suas próprias limitações em relação à aplicabilidade no campo oposto (MRS vs. DL). Para as métricas de MSE e Shape Score empregadas em aplicações de DL, quando aplicadas a dados *in-vivo*, a validação dos resultados apresenta limitações devido à falta de um *ground truth*. Para as métricas de SNR e FWHM, tradicionalmente usadas em aplicações *in-vivo*, quando aplicadas a dados reconstruídos por DL, os resultados podem ser inadequados, pois os modelos podem otimizar as métricas de uma forma que não é realista. Mudanças para melhorar a robustez das métricas quando usadas em um campo para o qual não foram originalmente desenvolvidas podem ser necessárias. Por exemplo, com o FWHM, talvez minimizar a diferença absoluta da largura do pico em relação à referência seja mais adequado para aplicações de DL do que minimizar o valor total.

Em geral, os resultados do desafio destacaram que reconstruções de melhor qualidade obtiveram resultados favoráveis em várias métricas, com a solução líder de cada Trilha alcançando os melhores resultados em três das quatro métricas (Tabs. 3.3,

3.4 e 3.5). A métrica proposta, Shape Score, mostrou uma correlação positiva com os melhores resultados da Trilha do desafio, sugerindo sua utilidade na avaliação da qualidade espectral. Além disso, ficou evidente que o uso de métricas deve ser feito com cautela, pois cada métrica individual apresenta suas próprias limitações, a maioria das quais foi superada pela combinação de todas as métricas. Apesar dos bons resultados alcançados e dos esclarecimentos obtidos no desafio, a avaliação dos modelos não abordou a quantificação dos sinais reconstruídos, deixando assim uma questão importante em aberto para ser explorada na literatura.

3.6 Estudos intermediários para a definição do método final

Esta seção aborda os principais experimentos intermediários que levaram à definição do modelo final apresentado na Seção 3.3. Esses experimentos foram conduzidos majoritariamente na Trilha 1, considerada a base para a generalização do modelo para outras trilhas mais complexas e com pouca quantidade de dados. Os principais experimentos focaram na validação do ViT como o melhor codificador (*encoder*) (GOODFELLOW *et al.*, 2016), nos parâmetros da STFT e na normalização do espectrograma.

Para isso, foi definida uma base experimental para os parâmetros da STFT, fixando-os até a realização de experimentos específicos para sua alteração. Essa base foi escolhida conforme referências da literatura. O tamanho da janela foi definido como 128, inspirado no estudo de Kyathanahally *et al.* (2018a), que demonstrou boa performance para dados de MRS. O deslocamento entre as janelas foi estabelecida em 64, considerada um bom valor inicial exploratório para uma janela de 128 (HEINZEL *et al.*, 2002). A normalização inicial utilizada foi a Min-Max, conforme a linha de base na solução dos organizadores do desafio (BERTO *et al.*, 2024).

3.6.1 Modelo codificador (*encoder*)

Foram testados vários modelos codificadores para validar a superioridade do ViT na extração dos atributos dos espectrogramas e na geração da representação latente. Esses modelos foram selecionados com base na literatura de visão computacional, destacando-se por sua alta performance em diversos problemas da área. Entre as arquiteturas exploradas, destacaram-se Wide ResNet (ZAGORUYKO; KOMODAKIS, 2016), TResNet (RIDNIK *et al.*, 2021), EfficientNet Pruned (AFLALO *et al.*, 2020) e RegNet (RADOSAVOVIC *et al.*, 2020), com o ViT apresentando o melhor desempenho (Tab. 3.6). A estrutura das camadas totalmente conectadas, responsáveis pela reconstrução do espectro a partir da representação latente, foi mantida idêntica à utilizada no método final (Fig. 3.5).

Tabela 3.6 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score de resultados do conjunto de validação da Trilha 1 para diferentes modelos codificadores. Resultados em negrito indicam o melhor desempenho para cada métrica.

Modelo	MSE	SNR	FWHM	Shape Score
ViT	0,00070 ± 0,00090	77,9 ± 5,4	0,07775 ± 0,000231	0,9899 ± 0,0001
Wide ResNet	0,00081 ± 0,00094	72,4 ± 1,7	0,07872 ± 0,000569	0,9821 ± 0,0002
TResNet	0,00079 ± 0,00103	75,7 ± 10,4	0,07799 ± 0,000240	0,9741 ± 0,0005
EfficientNet Pruned	0,00099 ± 0,00402	77,1 ± 6,6	0,07988 ± 0,000625	0,9799 ± 0,0003
RegNet	0,00086 ± 0,00097	69,9 ± 4,0	0,07799 ± 0,000294	0,9851 ± 0,0009

Além disso, a arquitetura do ViT utilizada possui pesos pré-treinados no ImageNet (DENG *et al.*, 2009), o que possibilitou um novo experimento. Inspirado pelo estudo de Gong *et al.* (2021), que combinou espectrogramas com um modelo ViT pré-treinado no ImageNet, resultando em desempenho de estado da arte no domínio de áudio, foi realizado um experimento semelhante. Embora os resultados não tenham sido altamente expressivos, superaram todas as métricas, justificando sua utilização na metodologia final (Tab. 3.7).

Tabela 3.7 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score obtidas para o conjunto de validação da Trilha 1 para o ViT pré-treinado e não pré-treinado. Resultados em negrito indicam o melhor desempenho para cada métrica.

ViT pré-treinado	MSE	SNR	FWHM	Shape Score
Sim	0,00065 ± 0,00072	85,2 ± 7,1	0,07655 ± 0,000112	0,9969 ± 0,0002
Não	0,00070 ± 0,00090	77,9 ± 5,4	0,07775 ± 0,000231	0,9899 ± 0,0001

3.6.2 Parâmetros da STFT

Após a seleção do ViT pré-treinado, foram realizados experimentos para determinar o tamanho da janela da STFT, que nos testes anteriores estava fixada em 128. Foram testadas janelas de tamanho 256, 512 e 1024, que respeitam as condições de computação inversa da STFT e representam potências de 2. A STFT se beneficia da eficiência do algoritmo de Transformada Rápida de Fourier (do inglês, *Fast Fourier Transform* - FFT) para quantidades de pontos que são potências de 2, pois envolve a aplicação de FFT em várias janelas de tempo. Os resultados revelaram que a janela de tamanho 256 proporcionou os melhores resultados (Tab.3.8).

Em seguida, foram realizados experimentos para determinar o deslocamento entre as janelas da STFT, que nos testes anteriores estava fixada em 64. Foram testados valores inteiros selecionados no intervalo de 1 a 128, visando garantir uma sobreposição mínima ideal entre as janelas de tamanho 256 (HEINZEL *et al.*, 2002). Devido à alta demanda computacional e ao tempo necessário para explorar todos os valores possíveis,

Tabela 3.8 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score no conjunto de validação da Trilha 1 para diferentes tamanhos de janela. Resultados em negrito indicam o melhor desempenho para cada métrica.

Tamanho	MSE	SNR	FWHM	Shape Score
128	0,00065 ± 0,00072	85,2 ± 7,1	0,07655 ± 0,000112	0,9969 ± 0,0002
256	0,00060 ± 0,00056	89,2 ± 5,4	0,07745 ± 0,000106	0,9972 ± 0,0004
512	0,00095 ± 0,00142	84,2 ± 9,6	0,07846 ± 0,000126	0,9811 ± 0,0003
1024	0,03982 ± 0,00568	70,2 ± 10,3	0,076661 ± 0,005869	0,6995 ± 0,13912

Tabela 3.9 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score no conjunto de validação da Trilha 1 para diferentes deslocamentos entre as janelas da STFT. Estão apresentados apenas os 5 melhores resultados (dentro todos os valores de deslocamento testados). Resultados em negrito indicam o melhor desempenho para cada métrica.

Deslocamento	MSE	SNR	FWHM	Shape Score
10	0,00055 ± 0,00052	95,2 ± 5,4	0,07635 ± 0,000126	0,9988 ± 0,0002
34	0,00062 ± 0,00054	92,2 ± 6,3	0,07785 ± 0,000152	0,9989 ± 0,0002
42	0,00059 ± 0,00073	88,3 ± 2,5	0,07665 ± 0,000138	0,9989 ± 0,0001
64	0,00060 ± 0,00056	89,2 ± 5,4	0,07745 ± 0,000106	0,9972 ± 0,0004
80	0,00057 ± 0,00068	90,1 ± 5,7	0,07655 ± 0,000258	0,9989 ± 0,0004

a exploração foi limitada. Dentre os valores testados, o valor de 10 apresentou o melhor resultado (Tab. 3.9).

3.6.3 Normalização

Foram testados quatro métodos de normalização: Max-Magnitude, Min-Max, Decibel e Z-score. As normalizações Min-Max e Z-score são amplamente utilizadas em ML. A normalização em Decibel foi inspirada por seu uso comum em espectrogramas no domínio de áudio. A normalização Max-Magnitude, nome atribuído neste trabalho, corresponde a dividir o sinal pela sua máxima magnitude. Essa abordagem foi explorada porque a divisão do sinal pela máxima intensidade é uma prática comum em MRS.

Os resultados indicam que a normalização Max-Magnitude obteve o melhor desempenho. No entanto, a Z-score e a Min-Max também apresentaram resultados competitivos. A normalização em Decibel, embora eficiente para a análise de espectrogramas no domínio de áudio, apresentou desempenho notavelmente inferior em relação às outras normalizações (Tab. 3.10).

Tabela 3.10 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score do conjunto de validação para a Trilha 1 para diferentes técnicas de normalização. Resultados em negrito indicam o melhor desempenho para cada métrica.

Normalização	MSE	SNR	FWHM	Shape Score
Max-Magnitude	0,00035 ± 0,00049	107,5 ± 8,3	0,07645 ± 0,000156	0,9996 ± 0,0002
Min-Max	0,00055 ± 0,00052	95,2 ± 5,4	0,07635 ± 0,000126	0,9988 ± 0,0002
Decibel	0,00126 ± 0,0010	70,9 ± 6,9	0,07667 ± 0,005688	0,8253 ± 0,1040
Z-score	0,00049 ± 0,00057	98,5 ± 9,1	0,07632 ± 0,000317	0,9990 ± 0,0001

3.7 Conclusão do Capítulo

O desafio *Edited-MRS Reconstruction Challenge* foi proposto para estabelecer uma referência de qualidade para reconstruções em GABA MEGA-PRESS usando DL com quatro vezes menos dados (80 transientes) do que a quantidade convencionalmente adquirida em uma aquisição completa (320 transientes). O desafio foi composto por três trilhas, sendo que a Trilha 1 utilizou dados simulados, a Trilha 2 utilizou dados *in-vivo* com parâmetros de aquisição quase idênticos, e a Trilha 3 utilizou dados *in-vivo* com parâmetros de aquisição diferentes. Os resultados demonstraram que reconstruções de alta qualidade alcançaram resultados favoráveis na maioria das métricas, com a melhor solução de cada Trilha obtendo os melhores resultados em três das quatro métricas. Como era esperado, os resultados com dados simulados para todas as equipes foram geralmente melhores do que os resultados com dados *in-vivo*, sugerindo que mais trabalho é necessário para generalizar modelos treinados com dados simulados para dados *in-vivo*. A métrica Shape Score proposta pelos organizadores mostrou uma correlação positiva com os resultados de cada Trilha do desafio, indicando ser bem adequada para avaliar a qualidade espectral. Além disso, o uso de métricas para aplicações para as quais não foram originalmente destinadas (MRS *in-vivo* vs DL) deve ser feito com cautela, pois cada métrica individual apresenta suas próprias limitações. Contudo, essas limitações foram, em grande parte, superadas pela combinação de todas as métricas.

Apesar dos bons resultados e dos *insights* obtidos no desafio, este estudo apresentou limitações e deixou questões em aberto. Embora muitas reconstruções tenham alcançado boas métricas de qualidade, a avaliação não incluiu a quantificação metabólica dos sinais reconstruídos para verificar se as concentrações de GABA foram realmente preservadas. Além disso, a avaliação das reconstruções *in-vivo* foi realizada apenas em um subconjunto dos primeiros 80 transientes adquiridos nas aquisições de teste, que contêm um total de 320 transientes. Os transientes são suscetíveis a variações na qualidade espectral ao longo do tempo de aquisição, devido à presença de artefatos de movimento, degradação e perdas de intensidade do sinal causadas por flutuações no campo magnético (ROWLAND *et al.*, 2017; JUCHEM *et al.*, 2021). Portanto, além de não quantificar os

espectros reconstruídos, o estudo não investigou a robustez e consistência dos modelos utilizando diferentes subconjuntos de transientes.

4 Spectro-ViT: versão aprimorada

Após o desenvolvimento do Spectro-ViT submetido ao *Edited-MRS Reconstruction Challenge*, estudos adicionais foram realizados com foco em aprimorar a metodologia de desenvolvimento do modelo. A arquitetura do modelo Spectro-ViT foi preservada, e as alterações se concentraram na expansão do conjunto de dados, na preparação dos dados, no pré-processamento dos espectrogramas e na metodologia de treinamento do modelo. Além disso, o estudo expandiu os experimentos, avaliando não somente a qualidade dos dados reconstruídos, mas também a quantificação de GABA a partir dos espectros reconstruídos, um aspecto relevante ainda não explorado na literatura. Os resultados obtidos pelo modelo Spectro-ViT aprimorado foram comparados com resultados obtidos pelo modelo de reconstrução desenvolvido por Berto *et al.* (2023) e fazem parte de um artigo submetido ao *Magnetic Resonance Imaging*, em fase final de revisão.

4.1 Alterações na metodologia original

4.1.1 Conjunto de dados

Como forma de aprimorar a metodologia, um novo conjunto de dados *in-vivo* foi extraído do repositório Big GABA (MIKKELSEN *et al.*, 2017; MIKKELSEN *et al.*, 2019), o mesmo repositório utilizado pelo desafio. Foram analisadas e amostradas 144 aquisições GABA MEGA-PRESS *in-vivo*, representando 30 aquisições adicionais em relação ao número utilizado no desafio. As aquisições foram obtidas de scanners dos fabricantes GE, Philips e Siemens em doze centros de aquisição: G4, G5, G7, G8, P4, P6, P8, P10, S1, S3, S5 e S8. O número de aquisições por fabricante foi o mesmo, totalizando 48 aquisições por fabricante. Cada local seguiu um protocolo padronizado e usou um voxel de $30 \times 30 \times 30 \text{ mm}^3$ no lobo parietal medial enquanto evitava ventrículos e/ou as superfícies externas do cérebro. Os parâmetros de aquisição foram: TE = 68 ms; TR = 2000 ms; 320 transientes; pulsos de edição ON/OFF em 1,9 e 7,46 ppm; duração do pulso de edição de 15 ms; e número de pontos amostrados igual a 2048 ou 4096. Além disso, os participantes não relataram doenças neurológicas ou psiquiátricas, e a idade média foi de 26,4 com um desvio padrão de $\pm 4,1$ anos.

4.1.2 Preparação dos dados

As aquisições foram divididas em três subconjuntos: 84 para treinamento, 24 para validação e 36 para teste. Cada um desses subconjuntos incluiu um número igual de

aquisições por fabricante de scanner, e a distribuição de idade dos sujeitos foi equilibrada. No entanto, neste estudo, após uma análise da qualidade dos dados de cada subconjunto, uma amostra do centro de aquisição G7 foi removida do conjunto de teste. Avaliações visuais e quantitativas identificaram artefatos no espectro de referência dessa amostra que comprometiam a qualidade do sinal. Dessa forma, o conjunto de teste final apresentou 12 amostras provenientes de scanners da Philips, 12 da Siemens e 11 da GE. O uso dos dados seguiu a metodologia de Berto *et al.* (2023), de forma a permitir uma comparação direta dos resultados.

A metodologia da janela deslizante para formar os subconjuntos de 80 transientes (Fig. 3.7) foi substituída por amostragem aleatória uniforme, selecionando 40 transientes por sub-sinal (Edit-ON e Edit-OFF). O método da janela deslizante gerava subconjuntos com alta correlação entre si. Em contraste, a nova amostragem tem maior potencial para aumentar a robustez do modelo em aquisições com transientes de qualidade espectral variável, como em movimentos da cabeça e variações graduais na precisão do scanner ao longo do tempo (HUI *et al.*, 2021).

Com esse novo método, foram amostrados 200 subconjuntos de 80 transientes a partir de cada aquisição com 320 transientes. Assim, o conjunto de treino consistiu em 16800 amostras (200×84) e o de validação em 4800 amostras (200×24). Além disso, os dados contendo 4096 pontos foram subamostrados para 2048 pontos. Esta abordagem foi adotada para otimizar a eficiência computacional e simplificar o desenvolvimento do modelo, permitindo o uso de uma única arquitetura. Dessa forma, a arquitetura do Spectro-ViT utilizada neste novo estudo foi a com saída de 2048 pontos (Fig. 3.5).

4.1.3 Geração dos Espectrogramas

Em vez de utilizar os 80 transientes para criar um único espectrograma replicado nos 3 canais de entrada da arquitetura, foram gerados 3 espectrogramas distintos a partir de subconjuntos diferentes dos 80 transientes, utilizando o mesmo algoritmo para pré-processar os transientes (Fig. 3.8). Dessa forma, o primeiro canal de entrada do modelo passou a receber o espectrograma derivado dos primeiros 14 transientes Edit-ON e 14 Edit-OFF (28 transientes), o segundo canal recebeu o espectrograma gerado pelos 13 transientes Edit-ON e 13 Edit-OFF (26 transientes) subsequentes, e o terceiro canal recebeu o espectrograma gerado pelos 13 transientes Edit-ON e 13 Edit-OFF (26 transientes) restantes.

Uma comparação do desempenho do modelo treinado utilizando o mesmo espectrograma de 80 transientes replicado para todos os canais, como no desafio original, com o desempenho do modelo treinado utilizando espectrogramas diferentes formados por

subconjuntos de transientes para cada um dos três canais (28 transientes para o primeiro canal, 26 transientes para o segundo canal e 26 transientes para o terceiro canal), mostrou melhorias em todas as métricas (Tab. 4.1).

Tabela 4.1 – Comparação entre o novo método de geração dos espectrogramas, que utiliza 28 transientes para o espectrograma do primeiro canal, 26 para o espectrograma do segundo e 26 para o espectrograma do terceiro, e o método antigo, que usa 80 transientes replicados nos três canais. Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score obtidos no conjunto de validação. Resultados em negrito indicam o melhor desempenho para cada métrica.

Geração Espectrogramas	MSE	SNR	FWHM	Shape Score
Modificado	0,0060 ± 0,0121	31,4 ± 5,6	0,1641 ± 0,0012	0,9831 ± 0,017
Desafio	0,0072 ± 0,0221	29,6 ± 7,4	0,1654 ± 0,0014	0,9810 ± 0,018

4.1.4 Treinamento

O conjunto de dados simulados utilizado na Trilha 1 do desafio não foi mais utilizado, uma vez que a transferência de aprendizado a partir desses dados demonstrou resultados similares aos obtidos sem o uso desses pesos treinados (Tab.4.2).

Tabela 4.2 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score nos resultados do conjunto de validação com e sem pré-treino do modelo usando dados da Trilha 1 do desafio.

Pré-Treino Trilha 1	MSE	SNR	FWHM	Shape Score
Sim	0,0060 ± 0,0121	31,4 ± 5,6	0,1641 ± 0,0012	0,9831 ± 0,017
Não	0,0061 ± 0,0119	32,1 ± 6,1	0,1642 ± 0,0010	0,9832 ± 0,015

Além disso, foi utilizada uma nova função de perda, que consiste em uma soma ponderada de parcelas de MAE focadas em faixas espectrais específicas no espectro. As faixas incluem o pico de GABA ($MAE_{2.8-3.2ppm}$) em 3 ppm, o pico de Glx em 3,75 ppm ($MAE_{3.55-3.95ppm}$), além do erro global considerando a faixa completa de ppm (MAE_{global}):

$$\text{Função de Perda} = \frac{(MAE_{global} + MAE_{3.55-3.95ppm} + 8 \times MAE_{2.8-3.2ppm})}{10} \quad (4.1)$$

O treinamento foi realizado em duas etapas. Na primeira etapa, a única alteração foi na função de perda. Na segunda etapa, além da função de perda, foram introduzidas outras modificações: um treinamento de 10 épocas e a substituição da LR fixa por uma com redução cosseno (período: 10 épocas, LR inicial: 1×10^{-5} , LR mínimo: 1×10^{-6}).

Os resultados utilizando o treinamento modificado foram superiores na maioria das métricas, especialmente MSE e Shape Score (Tab. 4.3).

Tabela 4.3 – Média e desvio padrão das métricas MSE, SNR, FWHM e Shape Score no conjunto de validação para metodologias de treino modificada e do desafio. Resultados em negrito indicam o melhor desempenho para cada métrica.

Metodologia Treino	MSE	SNR	FWHM	Shape Score
Modificada	0,0050 ± 0,0113	36,8 ± 4,4	0,1658 ± 0,0010	0,9845 ± 0,012
Desafio	0,0061 ± 0,0119	32,1 ± 6,1	0,1642 ± 0,0010	0,9832 ± 0,015

4.2 Experimentos e Resultados

Uma série de experimentos foi realizada no conjunto de teste, envolvendo quatro *pipelines* distintos: o modelo Spectro-ViT proposto, Gannet Full (processamento do software Gannet utilizando todos os 320 transientes), Gannet Quarter (processamento do software Gannet utilizando um quarto dos transientes) e o modelo de reconstrução Rdc-UNET2D proposto por Berto *et al.* (2023). Adicionalmente, o espectro obtido pela média dos espectros de referência do conjunto de treinamento foi analisado para avaliar a possibilidade de um viés em direção ao resultado médio do treinamento (DZIADOSZ *et al.*, 2023). Este espectro é referido como Espectro Médio de Referência do Treino (EMRT).

Cada conjunto de experimentos teve um objetivo distinto: avaliação da qualidade dos sinais reconstruídos (Seção 4.2.1); avaliação do impacto da reconstrução na quantificação de metabólitos (Seção 4.2.2); análise da robustez dos modelos de reconstrução frente à variação da qualidade dos transientes (Seção 4.2.3); e avaliação do desempenho da reconstrução para fatores de aceleração maiores que 4 (Seção 4.2.4).

Em alguns dos experimentos os 320 transientes de cada amostra do conjunto de teste foram divididos em múltiplos subconjuntos de 80 transientes, cada um abrangendo um intervalo de transientes organizados conforme a ordem sequencial de sua aquisição, sendo o transiente 1 o primeiro adquirido e o transiente 320 o último. Quatro destes subconjuntos são especificamente referidos nas próximas seções como: Conjunto 1 (transientes do número 1 ao 80), Conjunto 2 (transientes do número 81 ao 160), Conjunto 3 (transientes do número 161 ao 240) e Conjunto 4 (transientes do número 241 ao 320).

4.2.1 Qualidade da Reconstrução

Os espectros de diferença de cada pipeline, bem como o EMRT, foram avaliados usando MSE, SNR, FWHM, Shape Score e erro de ajuste. Inspeções visuais também foram realizadas comparando os três diferentes fabricantes: Siemens, Philips e GE. Os primeiros 80 transientes adquiridos (Conjunto 1) foram usados como entrada para os *pipelines* Spectro-ViT, Gannet Quarter e Rdc-UNET2D.

O Spectro-ViT apresentou o melhor desempenho geral para as métricas dire-

tamente relacionadas com a referência, MSE e Shape Score. Também, quando comparado aos outros *pipelines* (Rdc-UNET2D e Gannet Quarter), o Spectro-ViT obteve a menor média e desvio padrão para as métricas de erro de ajuste e FWHM (Tab. 4.4).

Como esperado, os espectros gerados pelo Gannet Quarter apresentaram as piores métricas de qualidade MRS (SNR, FWHM e erro de ajuste), exibindo maior nível de ruído e distorções na linha de base (Fig. 4.1). Ambos os métodos baseados em DL (Spectro-ViT e Rdc-UNET2D) geraram espectros que, em sua maioria, apresentaram métricas superiores em comparação ao Gannet Quarter. O Spectro-ViT, especificamente, apresentou uma redução de 12,22% no MSE médio, um aumento de 181,48% na SNR médio, uma melhoria de 9,29% no FWHM médio, um aumento de 0,71% no Shape Score médio e uma redução de 59,49% no erro de ajuste médio em comparação com Gannet Quarter. O Rdc-UNET2D também demonstrou melhorias notáveis em comparação com Gannet Quarter: um aumento de 458,33% na SNR médio, uma melhoria de 8,20% no FWHM médio, um aumento de 0,31% no Shape Score médio e uma redução de 44,30% no erro de ajuste médio. Apenas o MSE médio apresentou um aumento de 16,67%, indicando um desempenho inferior nesta métrica.

Os métodos Spectro-ViT e Rdc-UNET2D também foram capazes de gerar espectros com melhores métricas quantitativas de SNR, FWHM e erro de ajuste em comparação com o Gannet Full. O Spectro-ViT teve um aumento de 35,71% na SNR médio, uma melhoria de 1,78% no FWHM médio e uma redução de 36,00% no erro de ajuste médio. O modelo Rdc-UNET2D apresentou um aumento de 169,20% na SNR médio, uma melhoria de 0,59% no FWHM médio e uma redução de 12,00% no erro de ajuste médio.

Tabela 4.4 – Média e desvio padrão das métricas MSE, SNR, FWHM, shape score e erro de ajuste para GABA/Água obtidos no conjunto de teste, comparando o Gannet Full, Gannet Quarter, Rdc-UNET2D, Spectro-ViT e o Espectro Médio de Referência do Treino (EMRT). Os melhores resultados para cada métrica estão destacados em negrito.

Pipeline	MSE	SNR	FWHM (ppm)	Shape Score	Erro de ajuste (%)
Spectro-ViT	0,0079 ± 0,0135	30,4 ± 9,6	0,166 ± 0,012	0,986 ± 0,015	3,2 ± 1,1
Rdc-UNET2D	0,0105 ± 0,0113	60,3 ± 25,7	0,168 ± 0,015	0,982 ± 0,014	4,4 ± 2,4
Gannet Quarter	0,0090 ± 0,0268	10,8 ± 3,4	0,183 ± 0,031	0,979 ± 0,024	7,9 ± 5,3
EMRT	0,0129 ± 0,0153	106,3 ± 0,0	0,161 ± 0,000	0,978 ± 0,021	2,5 ± 0,0
Gannet Full	N/A	22,4 ± 6,5	0,169 ± 0,014	N/A	5,0 ± 2,5

O EMRT atingiu o melhor desempenho geral para SNR, FWHM e erro de ajuste (Tab. 4.4). A média dos espectros-alvo de treinamento levou a uma redução significativa no desvio padrão dentro da região ruidosa, o que, por sua vez, aumentou a SNR. Além disso, a média suavizou a linha de base, contribuindo para uma largura de linha GABA mais estreita. O aumento expressivo na SNR e a redução no FWHM contribuíram para a melhoria no erro de ajuste. No entanto, a melhoria das métricas de qualidade não é

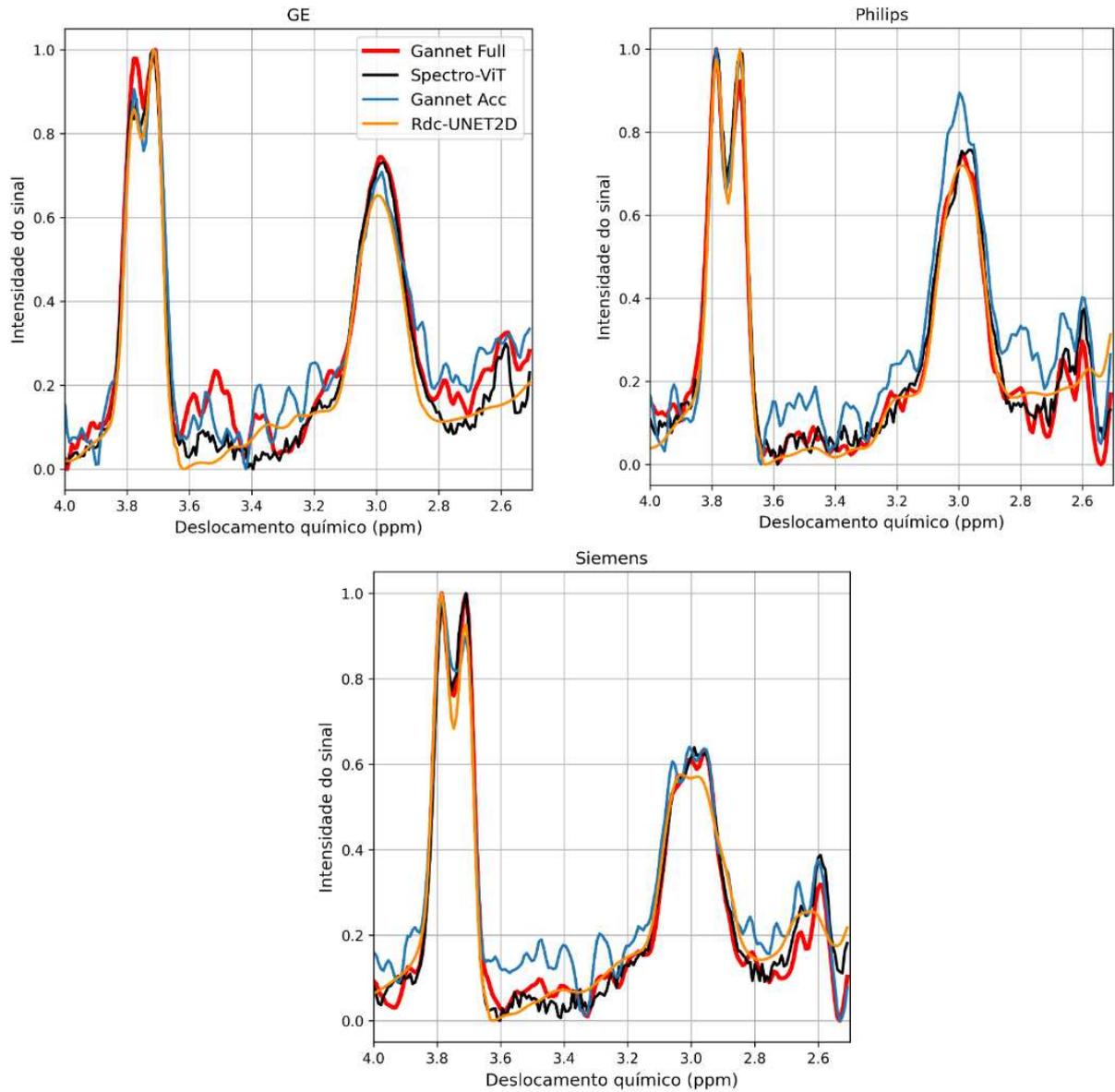


Figura 4.1 – Exemplos dos espectros de diferença GABA MEGA-PRESS do conjunto de teste, apresentados na faixa de 2,5 a 4 ppm. Os espectros foram obtidos a partir de equipamentos Siemens, Philips e GE. Os *pipelines* de DL, Spectro-ViT (preto) e Rdc-UNET2D (amarelo) reconstruíram espectros utilizando o Conjunto 1 (1-80) de transientes. O Gannet Quarter (em azul), que também utiliza o Conjunto 1 de transientes, e o *pipeline* de referência Gannet Full (em vermelho), que usa 320 transientes, também são exibidos para comparação. Fonte: (DIAS *et al.*, 2023)

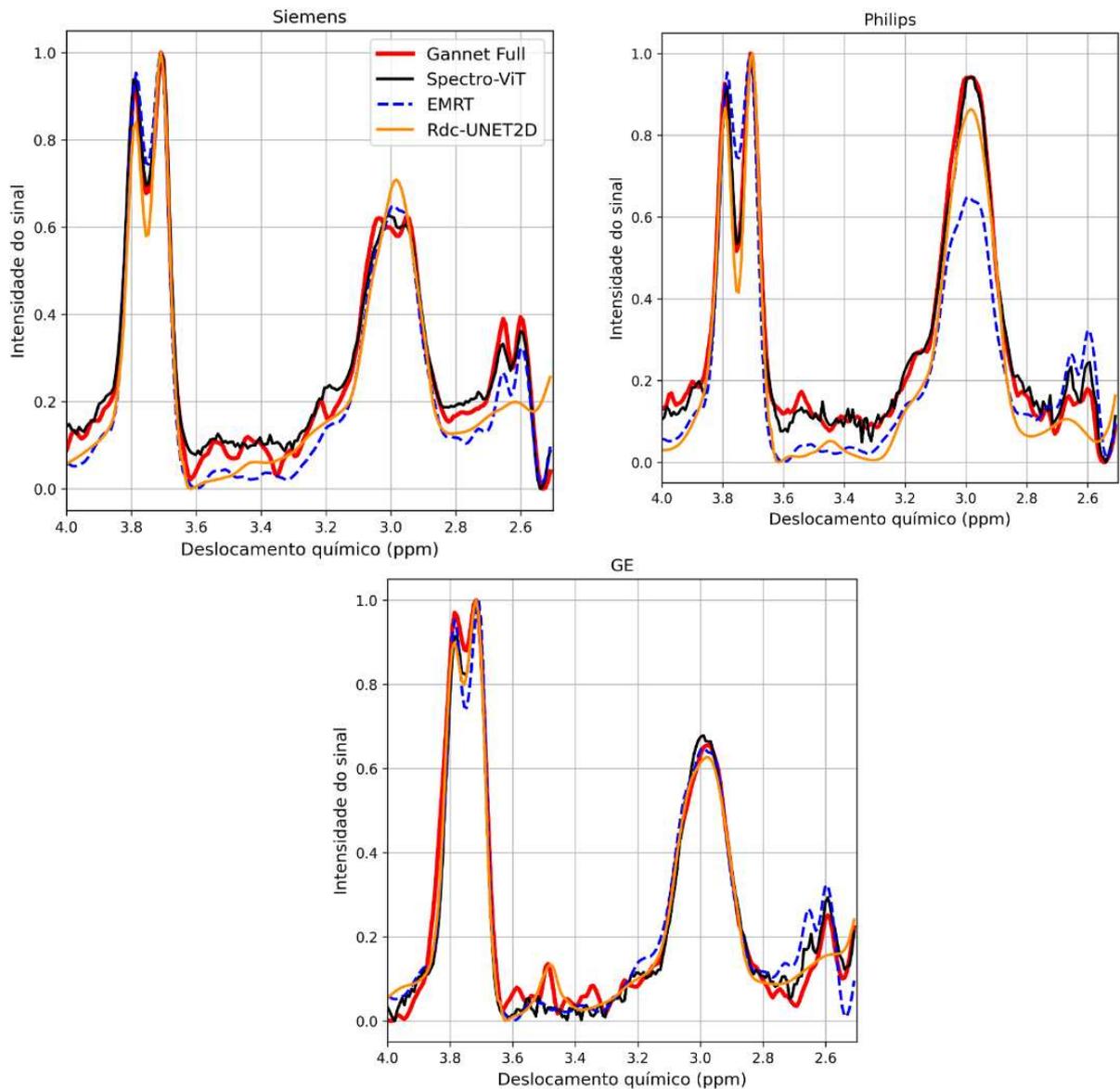


Figura 4.2 – Exemplos dos espectros de diferença GABA MEGA-PRESS do conjunto de teste, mostrados na faixa de 2,5 a 4 ppm. Espectros amostrados de Siemens, Philips e GE. Os *pipelines* de DL Spectro-ViT (preto) e Rdc-UNET2D (amarelo) reconstruíram espectros do Conjunto 1 (1-80) de transientes. O EMRT (em azul) e o *pipeline* de referência Gannet Full (em vermelho), que utiliza 320 transientes, também é exibida. Fonte: (DIAS *et al.*, 2023)

relevante se o espectro não representar bem a referência, como evidenciado pela amostra de teste da Philips (Fig. 4.2) e pelos piores valores de MSE (Tab. 4.4).

4.2.2 Quantificação

A análise para quantificar o GABA utilizou a água como referência (GABA/Água), sendo a relação GABA/Água quantificada em unidades institucionais (i.u) por meio do software Gannet. As medidas foram feitas com base nas saídas dos mesmos *pipelines* utilizadas na avaliação anterior (Spectro-ViT, Rdc-UNET2D, Gannet Quarter, Gannet Full), aplicadas ao mesmo subconjunto de transientes (Conjunto 1). Além disso, o sinal EMRT também foi quantificado.

Como não existem valores conhecidos de concentração real para os dados *in-vivo*, as concentrações obtidas pelos espectros gerados pelo Gannet Full foram utilizadas como referência. Para as avaliações, foram calculados o R^2 (coeficiente de determinação), o erro dessas medições, especificamente, a distribuição do erro absoluto e o erro absoluto percentual médio (MAPE, do inglês *Mean Absolute Percentage Error*) das concentrações.

O teste de hipótese de Wilcoxon (CONOVER, 1999) também foi conduzido utilizando os valores de concentração quantificados a partir da *pipeline* Gannet Full como referência. A hipótese nula testada verificava se havia uma diferença estatisticamente significativa em comparação com a referência, com um nível de significância de $\alpha = 0,05$. Esse teste de significância foi escolhido porque compara dados pareados e não faz suposições sobre a distribuição dos dados.

O Spectro-ViT obteve o melhor desempenho geral na quantificação, alcançando os melhores resultados na quantificação de GABA/Água para o R^2 , o menor valor na métrica MAPE e o menor desvio padrão para o erro absoluto. Além disso, foi o único *pipeline* que não apresentou diferenças estatisticamente significativas na quantificação dos espectros de diferença, quando comparada aos resultados de GABA/Água dos espectros-alvo no teste de Wilcoxon bilateral.

Tabela 4.5 – R^2 , distribuição do erro absoluto (média \pm desvio padrão), MAPE e teste de Wilcoxon bilateral calculados para avaliação de quantificação de GABA/Água para o Spectro-ViT, Rdc-UNET2D, Gannet Quarter, e EMRT. Os resultados foram obtidos a partir do Conjunto 1. As concentrações foram comparadas com as obtidas pela *pipeline* Gannet Full. Resultado com símbolo ‘*’ indica um valor abaixo do nível de significância de 0,05.

Pipeline	R^2	Erro Absoluto (i.u)	MAPE (%)	Wilcoxon (p-valor)
Spectro-ViT	0,68	0,1958 \pm 0,1344	9,74	0,4911
Rdc-UNET2D	0,43	0,2330 \pm 0,2201	11,85	*0,0024
Gannet Quarter	0,65	0,1832 \pm 0,1738	10,04	*0,0174
EMRT	-214,51	3,8913 \pm 4,8740	227,81	*0,0151

O Gannet Quarter apresentou desempenho comparável ao Spectro-ViT, com uma pontuação R^2 de 0,65, o melhor MAE e um valor MAPE próximo. No entanto, o teste de Wilcoxon revelou uma diferença estatisticamente significativa entre as concentrações do Gannet Quarter e do Gannet Full (Tab. 4.5). Testes iniciais bilaterais indicaram essa diferença, e testes adicionais unilaterais sugeriram que o Gannet Quarter tende a apresentar valores mais altos. A inspeção visual das distribuições dos gráficos forneceu informações complementares (Fig. 4.3). A análise do gráfico de dispersão demonstrou que a maioria dos valores de concentração do Gannet Quarter estão acima da linha de referência, e seu boxplot mostrou um desvio para valores mais altos.

O Rdc-UNET2D exibiu a menor pontuação R^2 de 0,43 para GABA/Água e parece subestimar a maioria dos valores, conforme sugerido pelo desvio da sua mediana e pelo posicionamento da maioria dos valores abaixo da linha de referência (Fig. 4.3). Similarmente ao Gannet Quarter, um teste unilateral de Wilcoxon foi realizado para o Rdc-UNET2D, revelando uma diferença estatisticamente significativa em direção à subestimação em comparação com o alvo. O EMRT exibiu resultados altamente discrepantes para GABA/Água em todas as métricas quantitativas. Ele apresentou um MAPE de aproximadamente 228% para ambos os metabólitos e uma diferença estatisticamente significativa em relação ao alvo, conforme o teste bilateral de Wilcoxon (Tab. 4.5).

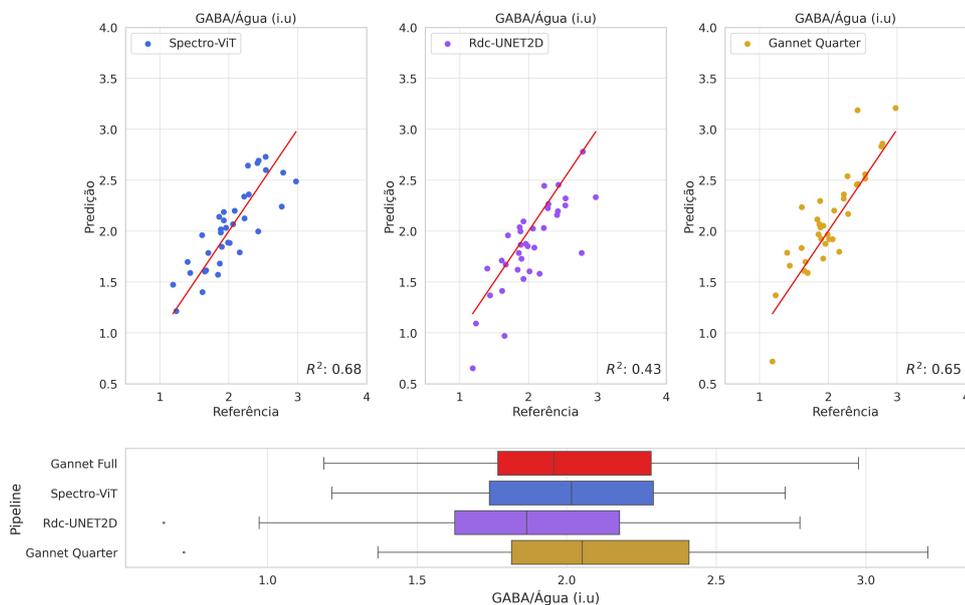


Figura 4.3 – Parte superior: distribuições de concentração para Spectro-ViT (em azul), Rdc-UNET2D (em roxo) e Gannet Quarter (em amarelo). A linha de referência (vermelha) é calculada para os valores do Gannet Full. Parte inferior: distribuição das concentrações para cada *pipeline* também é representada com as mesmas cores. Esses valores foram obtidos a partir de cálculos utilizando o Conjunto 1 de transientes. Fonte: (DIAS *et al.*, 2023)

4.2.3 Análise de Robustez

A robustez da reconstrução foi avaliada analisando a reconstrução realizada com múltiplos subconjuntos de transientes da mesma aquisição. Esse método avaliou o desempenho de cada *pipeline* em diferentes níveis de qualidade na entrada, especialmente em transientes mais afetados por ruído, variações de frequência e deslocamentos de fase.

Esses subconjuntos foram selecionados dividindo os transientes do exame de teste em 25 subconjuntos, cada um composto por 80 transientes. Os subconjuntos foram formados sequencialmente da seguinte maneira: o primeiro subconjunto incluiu os transientes de 1 a 80, o segundo subconjunto de 11 a 90, o terceiro de 21 a 100, e assim por diante, até chegar ao subconjunto final, que incluiu os transientes de 241 a 320. Assim, para cada amostra no conjunto de teste, foram feitas 25 reconstruções usando esses diferentes subconjuntos.

Os espectros reconstruídos foram avaliados quanto à SNR, FWHM, erro de ajuste e MAPE de GABA/Água. Essas métricas foram calculadas e, em seguida, obtidas as médias dentro de cada subconjunto. Subsequentemente, a distribuição dos 25 valores resultantes para cada métrica foi examinada calculando-se a média e o desvio padrão. Além disso, o R^2 foi calculado para cada subconjunto.

Por fim, para o Spectro-ViT, foram utilizados boxplots para facilitar uma análise granular dentro de 4 subconjuntos de transientes não sobrepostos dos 25 subconjuntos: Conjunto 1 (transientes de 1 a 80), Conjunto 2 (transientes de 81 a 160), Conjunto 3 (transientes de 161 a 240) e Conjunto 4 (transientes de 241 a 320).

A análise dos espectros editados dos subconjuntos de transientes revelou que o Spectro-ViT teve o melhor desempenho geral (Tab. 4.6). Em relação às métricas de qualidade (SNR, FWHM e erro de ajuste), o Spectro-ViT manteve desempenho superior. Embora o Spectro-ViT tenha sido superado na SNR pelo Rdc-UNET2D, a SNR do Spectro-ViT ainda representou uma melhoria de aproximadamente 30% em relação à SNR médio dos espectros gerados pelo Gannet Full (Tab. 4.5).

Tabela 4.6 – Distribuição das métricas de SNR, FWHM, erro de ajuste para GABA/Água, MAPE e R^2 nas reconstruções realizadas em 25 subconjuntos de transientes obtidos por um método de amostragem de janela deslizante (passo de 10, 80 transientes cada) no conjunto de teste. O desempenho das *pipelines* Spectro-ViT, Rdc-UNET2D e Gannet Quarter é comparado, com os resultados apresentados como média \pm desvio padrão.

Pipeline	SNR	FWHM (ppm)	Erro de ajuste (%)	MAPE (%)	R^2
Spectro-ViT	28,8 \pm 0,8	0,166 \pm 0,001	3,29 \pm 0,03	9,68 \pm 0,24	0,67 \pm 0,01
Rdc-UNET2D	59,4 \pm 3,2	0,168 \pm 0,002	4,60 \pm 0,34	12,72 \pm 4,14	0,42 \pm 0,16
Gannet Quarter	11,3 \pm 0,7	0,173 \pm 0,005	7,27 \pm 0,74	25,15 \pm 24,58	0,41 \pm 0,33

Com relação à quantificação, o Spectro-ViT também manteve o melhor desempenho para MAPE e R^2 na medição de GABA. O Spectro-ViT se destaca com um valor médio de R^2 de 0,67, acompanhado de um pequeno desvio padrão de 0,01. Este resultado médio é aproximadamente 60% superior ao segundo melhor valor, obtido pelo Rdc-UNET2D (0,42) (Tab. 4.6). O desempenho do Gannet Quarter caiu significativamente na média dos subconjuntos, com o valor médio de MAPE subindo cerca de 15 pontos percentuais em relação aos resultados obtidos do Conjunto 1 (Tab. 4.5). Além disso, o valor médio de R^2 caiu cerca de 36% (de 0,65 para 0,41) (Tab. 4.6).

A análise dos resultados do Spectro-ViT para os conjuntos 1, 2, 3 e 4 dos subconjuntos de transientes revelou que a mediana e os valores máximos de SNR diminuíram nos últimos transientes adquiridos, especificamente no Conjunto 4 (transientes de 241 a 320). No entanto, a maioria dos valores ainda atingiu níveis mais altos em comparação com o Gannet Full. A faixa interquartil da métrica FWHM permaneceu aproximadamente constante do Conjunto 1 ao 4, embora o Conjunto 4 tenha apresentado os piores valores comparados aos outros subconjuntos. Em relação ao erro de ajuste, o Conjunto 1 mostrou a distribuição com os menores valores (melhores), enquanto o Conjunto 4 teve os maiores valores (piores). Apesar disso, todos os quatro subconjuntos apresentaram distribuições melhores do que o Gannet Full (Fig. 4.4).

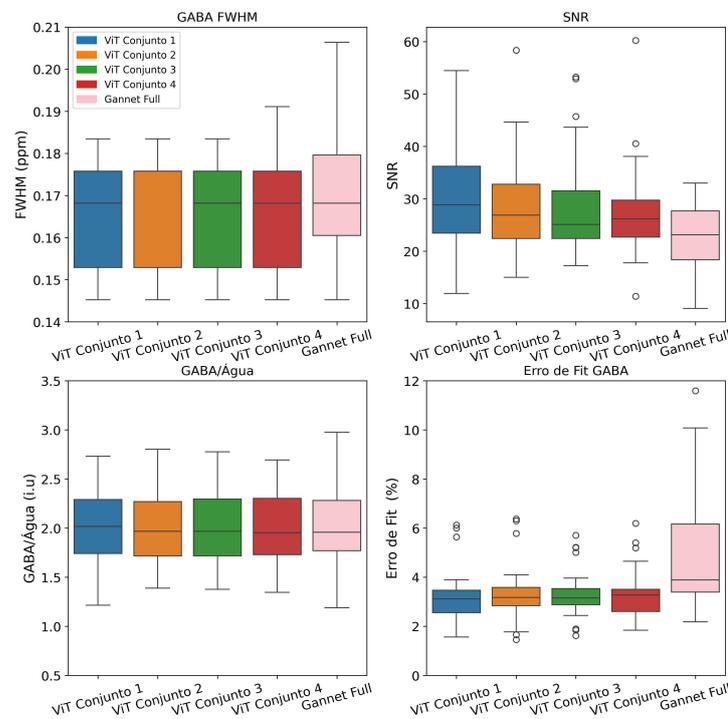


Figura 4.4 – Comparação das métricas de reconstrução SNR, FWHM e os resultados de quantificação para GABA/Água, juntamente com seus respectivos erros de ajuste, para os espectros reconstruídos no conjunto de teste pelo Spectro-ViT: Conjunto 1 (em azul); Conjunto 2 (em laranja); Conjunto 3 (em verde); Conjunto 4 (em vermelho); Gannet Full (em rosa). Fonte: (DIAS *et al.*, 2023)

Os valores de concentração nos subconjuntos do Spectro-ViT foram consistentes entre si e com o Gannet Full para a maioria dos casos, especialmente no intervalo interquartil. No entanto, os valores localizados nas caudas da distribuição-alvo revelaram ser mais desafiadores de serem modelados, com as melhores reconstruções vindo de subconjuntos menos impactados por desvios de frequência (Fig. 4.4). Além disso, testes de hipótese de Wilcoxon bilateral comparando as distribuições de concentração com a referência Gannet Full não revelaram diferenças estatisticamente significativas em todos os 25 subconjuntos (essa análise foi previamente conduzida para o Conjunto 1 na avaliação anterior [Tab. 4.5]).

4.2.4 Redução de Transientes

Para testar como o modelo Spectro-ViT se comporta quando treinado com menos transientes, o modelo Spectro-ViT foi re-treinado usando uma quantidade ainda menor de transientes sob a mesma metodologia de treinamento da configuração original de 80 transientes. Foram conduzidas avaliações usando 70, 60, 50, 40, 30, 20 e 10 transientes. O mesmo número de amostras de treinamento da configuração com 80 transientes foi usado, mas os subconjuntos amostrados da aquisição de 320 transientes foram ajustados de acordo com o número reduzido de transientes, mantendo a estrutura de entrada de três canais e a geração de espectrogramas. O objetivo foi alcançar a distribuição mais equilibrada possível entre os três canais para evitar que qualquer canal dominasse o processo de aprendizado, garantindo que o desempenho do modelo não dependesse excessivamente da qualidade e das características dos dados de um único canal.

Por exemplo, com 70 transientes, 24 transientes (12 Edit-ON + 12 Edit-OFF) foram alocados para o primeiro e segundo canais, e 22 transientes (11 Edit-ON + 11 Edit-OFF) para o terceiro canal. Com 60 transientes, 20 transientes (10 Edit-ON + 10 Edit-OFF) foram usados para cada um dos três canais, enquanto para 50 transientes, 18 transientes (9 Edit-ON + 9 Edit-OFF) foram usados para o primeiro canal, e 16 transientes (8 Edit-ON + 8 Edit-OFF) para o segundo e terceiro canais. À medida que o número de transientes diminuía, essa abordagem de balanceamento continuava.

Cada modelo resultante dessas configurações foi avaliado medindo o R^2 , erro absoluto, MAPE, e o teste de Wilcoxon bilateral para quantificação de GABA/Água na reconstrução de subconjuntos criados ao tomar os primeiros transientes adquiridos das varreduras no conjunto de teste, de acordo com a configuração do modelo. Por exemplo, para o modelo treinado com 50 transientes, a avaliação foi realizada usando os primeiros 50 transientes adquiridos da varredura de 320 transientes. A quantificação obtida usando o Gannet Full foi usada como referência.

Os resultados obtidos a partir dos experimentos realizados usando 70, 60, 50, 40, 30, 20 e 10 transientes revelaram que, embora o teste de Wilcoxon bilateral não tenha revelado nenhuma diferença estatisticamente significativa para nenhum dos experimentos (nível de significância de $\alpha = 0.05$), à medida que o número de transientes diminuía, o desempenho da reconstrução para manter a informação de GABA/Água piorava (Tab. 4.7 e Fig. 4.5). Os melhores valores de R^2 , erro absoluto e MAPE foram alcançados usando um maior número de transientes, com os melhores resultados sendo obtidos com o número de 80 transientes previamente relatado em avaliações anteriores (Tab. 4.7).

Com 70 transientes, o valor de R^2 cai 0,04 unidades para um valor de 0,64. Com 60 transientes, o valor de R^2 teve uma leve queda de 0,01, resultando em um valor final de 0,63. Com 50 transientes, o valor de R^2 caiu para 0,60. Abaixo desse número de transientes, todos os outros resultados obtidos ficaram abaixo de 0,60. Os resultados obtidos para 40 e 30 transientes variaram entre 0,6 e 0,5, especificamente 0,55 para 40 transientes e 0,51 para 30 transientes. Para 20 e 10 transientes, os resultados foram abaixo de 0,5, especificamente 0,47 para 20 transientes e 0,44 para 10 transientes. A maior queda de valor foi de 50 para 40 transientes, que foi de 0,05, e a menor queda de valor foi de 70 para 60 transientes, que foi de 0,01. A queda média por dois experimentos consecutivos de redução de 10 transientes foi de aproximadamente 0.034 (Tab. 4.7 e Fig. 4.5).

Tabela 4.7 – R^2 , distribuição do erro absoluto (média \pm desvio padrão), MAPE e o teste de Wilcoxon bilateral para a quantificação de GABA/Água para diferentes números de transientes com suas respectivas configurações de canal. A configuração de canal é representada como uma tupla (número 1, número 2, número 3), onde o primeiro número corresponde ao número de transientes usados para gerar o espectrograma para o primeiro canal, o segundo número para o segundo espectrograma para o segundo canal, e o terceiro número para o terceiro espectrograma no terceiro canal. Os resultados foram obtidos a partir dos transientes iniciais adquiridos no conjunto de teste. As concentrações foram comparadas com aquelas obtidas usando o *pipeline* Gannet com 320 transientes. Os melhores resultados para cada métrica estão destacados em negrito.

Transientes	Configuração	R^2	Erro Absoluto (u.i)	MAPE (%)	Wilcoxon (p-valor)
80	(28,26,26)	0,68	0,1958 \pm 0,1344	9,74	0,4911
70	(24,24,22)	0,64	0,2013 \pm 0,1555	9,80	0,2449
60	(20,20,20)	0,63	0,2086 \pm 0,1519	10,33	0,0601
50	(18,16,16)	0,60	0,2129 \pm 0,1640	10,48	0,7402
40	(14,14,12)	0,55	0,2286 \pm 0,1668	11,09	0,8907
30	(10,10,10)	0,51	0,2316 \pm 0,1847	11,51	0,0840
20	(8,6,6)	0,47	0,2492 \pm 0,1809	12,72	0,2655
10	(4,4,2)	0,44	0,2628 \pm 0,1788	13,58	0,4036

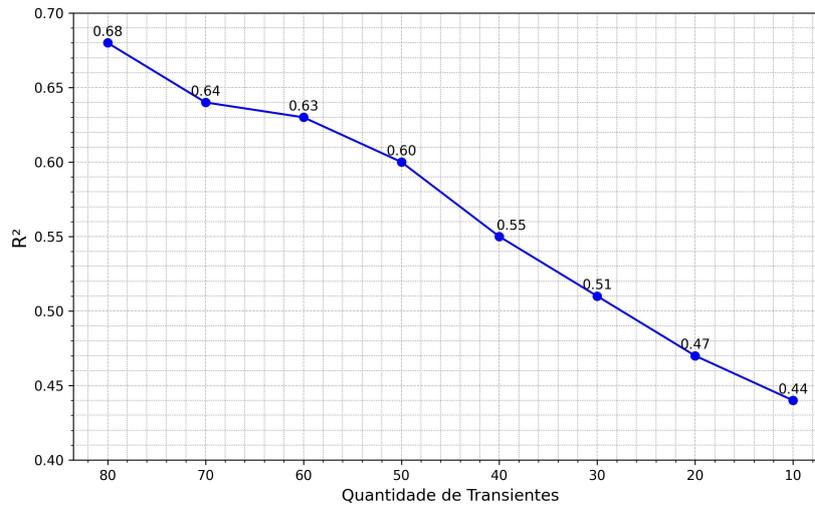


Figura 4.5 – Relação entre o número de transientes utilizados no processo de reconstrução e o valor de R^2 resultante para a quantificação de GABA/Água. Os experimentos foram conduzidos re-treinando o modelo Spectro-ViT usando 70, 60, 50, 40, 30, 20 e 10 transientes sob a mesma metodologia de treinamento da configuração inicial de 80 transientes. Cada modelo resultante dessas configurações foi avaliado medindo o valor de R^2 da quantificação de GABA/Água na reconstrução de subconjuntos criados tomando os primeiros transientes adquiridos das varreduras no conjunto de teste. A quantificação obtida usando Gannet Full foi utilizada como referência. Fonte: (DIAS *et al.*, 2023)

4.3 Discussão

O modelo Spectro-ViT inovou na reconstrução de aquisições GABA MEGA-PRESS com 80 transientes, combinando espectrogramas de sinais MRS com um modelo ViT para extração de características. Testado com dados *in-vivo*, seu desempenho foi comparado a outras abordagens usando métricas quantitativas e valores estimados de concentração de GABA, tomando aquisições típicas de 320 transientes como referência.

Os resultados dos experimentos forneceram evidências concretas da eficácia dessa abordagem em comparação com outras (BERTO *et al.*, 2023; BERTO *et al.*, 2024). A análise da reconstrução realizada nos primeiros 80 transientes (Conjunto 1) das amostras de teste revelou que o Spectro-ViT exibiu o melhor desempenho geral nas métricas de qualidade de reconstrução diretamente relacionadas à referência, como o MSE e o Shape Score, ao mesmo tempo em que manteve um erro de ajuste menor e uma SNR mais alta em comparação com o Gannet Full de referência (Tab. 4.4).

Além de demonstrar o melhor desempenho nas métricas de qualidade espectral para o Conjunto 1 de transientes, o Spectro-ViT alcançou o melhor desempenho para a quantificação. O Spectro-ViT obteve o valor mais alto de R^2 para GABA/Água de 0,68,

juntamente com o menor valor de MAPE (Tab. 4.5). Além disso, o Spectro-ViT foi o único *pipeline* cujas concentrações dos espectros de diferença quantificados para GABA/Água não revelaram diferenças estatisticamente significativas em relação às concentrações de referência (Tab. 4.5).

O Spectro-ViT também se destacou em termos de robustez, mantendo um valor de R^2 consistente de 0,67 para a quantificação de GABA/Água, aproximadamente 60% maior que o segundo melhor valor nos 25 subconjuntos de transientes formados a partir do conjunto de teste. As amostras localizadas nas caudas da distribuição de concentração-alvo foram mais desafiadoras de modelar, com as melhores concentrações provenientes de subconjuntos de transientes menos impactados por desvios de frequência (Fig. 4.4). Testes de hipóteses de Wilcoxon bilaterais foram conduzidos nas concentrações obtidas de cada subconjunto, comparando as distribuições de concentrações dos subconjuntos com a referência gerada por Gannet Full. Os resultados não mostraram diferenças estatisticamente significativas para nenhum subconjunto.

O modelo Rdc-UNET2D exibiu os valores de SNR mais altos em comparação com as *pipelines* Spectro-ViT e Gannet Quarter, alcançando um valor médio aproximadamente 2,65 vezes maior que o valor da SNR obtido pelo Gannet Full de referência. Além disso, o Rdc-UNET2D demonstrou erro de ajuste e FWHM superiores em comparação com o Gannet Full, junto com o segundo melhor shape score, se aproximando do melhor valor alcançado pelo Spectro-ViT (Tab. 4.4). No entanto, o valor médio de R^2 alcançado pelo Rdc-UNET2D foi de 0,42, indicando que as concentrações obtidas por suas reconstruções não conseguiram explicar 58% da variabilidade das concentrações dentro da distribuição-alvo (Tab. 4.6). Além disso, as concentrações obtidas pelo Rdc-UNET2D revelaram diferenças estatisticamente significativas nos transientes inicialmente adquiridos (Tab. 4.5). Isso destaca a importância da investigação inicial conduzida nesta pesquisa, que considerou a quantificação juntamente com as métricas de qualidade de reconstrução.

O *pipeline* Gannet Quarter exibiu as piores métricas de qualidade de SNR, FWHM e erro de ajuste entre todas as *pipelines* (Tab. 4.4). Apesar de seu desempenho de quantificação comparável para os valores de R^2 e MAPE ao Spectro-ViT para os primeiros 80 transientes adquiridos (Tab. 4.5), o Gannet Quarter teve o pior desempenho no experimento de robustez (Tab. 4.6). Seu desempenho caiu consideravelmente na média dos subconjuntos, com o valor médio de MAPE subindo cerca de 15 pontos percentuais (10,04 para 25,15) e o valor médio de R^2 caindo quase 35% em comparação com o resultado do Conjunto 1 (0,65 para 0,41) (Tab. 4.6). Portanto, demonstrou ser o *pipeline* menos robusto e mais vulnerável a possíveis movimentos da cabeça ou flutuações do campo magnético durante a aquisição.

Em relação à análise do espectro derivado pela média dos espectros-alvo de trei-

namento (EMRT), tanto as *pipelines* baseadas em DL Spectro-ViT quanto Rdc-UNET2D demonstraram melhores valores de MSE e shape score, além de desempenho de quantificação em comparação com o EMRT (Tab. 4.4). Embora o EMRT tenha alcançado os melhores valores de SNR, FWHM e erro de ajuste, não teve um bom desempenho em métricas diretamente relacionadas à referência (MSE e shape score). Isso destaca a necessidade de combinar métricas na análise para avaliar a reconstrução (BERTO *et al.*, 2024). Além disso, seus resultados de erro de quantificação foram os piores, com um alto valor de MAPE de aproximadamente 228% (Tab. 4.5).

A análise de sensibilidade do modelo Spectro-ViT ao número de transientes mostrou, como esperado, que o desempenho cai conforme diminuimos o número de transientes disponíveis para o treinamento. Embora o teste bilateral de Wilcoxon não tenha identificado diferenças estatisticamente significativas em nenhum dos experimentos, à medida que o número de transientes na metodologia de treinamento diminuía, o MAPE, a pontuação R^2 e o erro absoluto para GABA/Água pioravam (Tab. 4.7 e Fig. 4.5). Os melhores resultados foram alcançados com 80 transientes, atingindo um valor R^2 de 0,68, conforme apresentado em experimentos anteriores (Tab. 4.5). A queda média na pontuação R^2 , por cada dois experimentos consecutivos, em que o número de transientes era reduzido em 10, foi de aproximadamente 0,034. A faixa com uma quantidade de 50 transientes ou mais demonstrou maior robustez, mantendo valores R^2 acima de 0,6. Notavelmente, 50 transientes correspondem ao uso de 6,4 vezes menos dados em comparação com a aquisição convencional de 320 transientes. Experimentos conduzidos com menos de 30 transientes resultaram em valores R^2 abaixo de 0,5, indicando que essas configurações foram incapazes de explicar pelo menos 50% da variabilidade dentro da distribuição alvo.

4.4 Conclusão do Capítulo

O modelo Spectro-ViT representa uma abordagem inovadora para a reconstrução de aquisições GABA MEGA-PRESS com apenas 80 transientes, combinando a representação de espectrograma dos sinais de MRS com um modelo ViT para a extração de características relevantes. Esse modelo foi testado com dados GABA MEGA-PRESS *in-vivo*, e seu desempenho foi comparado com outras abordagens da literatura, utilizando métricas quantitativas e valores estimados de concentração de GABA. As aquisições típicas de 320 transientes serviram como referência para comparação.

O Spectro-ViT superou os outros métodos da literatura, alcançando performance estado da arte na reconstrução de espectros GABA MEGA-PRESS com apenas 80 transientes, apresentando métricas de qualidade superiores e estimativas de concentração de GABA mais precisas. Ele demonstrou o melhor desempenho, caracterizado pela

menor média e desvio padrão do erro de ajuste, menor MSE, maior Shape Score, além de superar o método convencional em relação à SNR e FWHM. Com relação aos resultados de quantificação, o modelo alcançou um valor médio consistente de R^2 de 0,67 para GABA/Água. Notavelmente, não houve diferenças significativas na distribuição de concentração em comparação com a referência.

Os resultados sobre o desempenho do Spectro-ViT são promissores para o campo emergente de reconstrução de aquisições GABA MEGA-PRESS. No entanto, esta pesquisa apresenta limitações e pontos que ainda precisam ser investigados.

O desenvolvimento do Spectro-ViT foi conduzido à luz dos achados de Dzia-dosz *et al.* (2023) e Berto *et al.* (2024), que destacaram possíveis vieses no desenvolvimento de modelos de DL em MRS, que podem afetar a generalização e resultar em melhorias de qualidade espectral não representativas. Dessa forma, este estudo procurou abordar as limitações relatadas na literatura utilizando dados *in-vivo*, considerando o sinal médio de treinamento na análise, utilizando um conjunto abrangente de métricas para descrever a qualidade e a precisão dos resultados obtidos, empregando software de MRS de código aberto e disponibilizando o código do modelo desenvolvido para promover a reprodutibilidade e a transparência dos resultados. Além disso, investigou-se, pela primeira vez na literatura, a quantificação dos espectros reconstruídos.

Este trabalho também priorizou o uso de um conjunto de dados balanceado proveniente de múltiplos fornecedores e técnicas de aumento de dados. No entanto, o conjunto de dados utilizado teve uma distribuição limitada, considerando apenas participantes de controle, parâmetros de aquisição específicos e a mesma região cerebral com tamanho de voxel fixo, baseando-se na disponibilidade de dados *in-vivo* de acesso aberto. Adicionalmente, amostras cujos valores de concentrações estavam nas caudas da distribuição-alvo revelaram-se mais desafiadoras para o modelo. Portanto, recomenda-se que estudos futuros explorem um conjunto de dados mais diversificado para verificar a consistência dos resultados obtidos pela modelagem atual.

Aumentar a complexidade do modelo por meio de alterações na arquitetura tem o potencial de melhorar os resultados. Em relação à arquitetura do modelo, o uso de três canais foi uma extensão da estrutura do modelo pré-treinado e ajudou a reduzir a complexidade computacional. Foi realizado um experimento em que a média de todos os 80 transientes foi calculada e o mesmo espectrograma foi replicado para os três canais, comparado com o método final em que os 80 transientes foram divididos em grupos de 28, 26 e 26, calculando a média dentro de cada grupo para gerar os espectrogramas dos três canais. Os resultados foram melhores em todas as métricas no segundo experimento. Dessa forma, aumentar a complexidade do modelo por meio do aumento da quantidade de canais processados apresenta um potencial de melhorar a performance atual do modelo.

A agregação dos transientes por média ajuda a reduzir a complexidade computacional, porém resulta em uma perda de detalhes e informações específicas de cada transiente, limitando o modelo na exploração de detalhes mais sutis nos dados. Estudos futuros poderiam envolver o desenvolvimento de um modelo ViT codificador tridimensional mais complexo que processe espectrogramas individuais de cada transiente simultaneamente e faça operações ao longo do eixo dos transientes. Isso poderia aprimorar o processo de aprendizagem, permitindo que o modelo capturasse mais variações e interações entre os transientes. Além disso, a complexidade do modelo pode ser aumentada pela adição de mais camadas totalmente conectadas para reconstruir os *embeddings* codificados pelo ViT. No entanto, é importante ressaltar que essas abordagens exigiriam mais recursos computacionais e demandariam mais investigação e exploração de técnicas adicionais de regularização para lidar com o potencial sobreajuste do modelo.

Por fim, embora o foco deste trabalho tenha sido a utilização de 80 transientes para alcançar um fator de aceleração que torna as aquisições quatro vezes mais rápidas, conforme trabalhos prévios da literatura, novos estudos podem explorar o uso de uma quantidade maior de transientes. O objetivo seria avaliar o *trade-off* entre a performance da reconstrução para uma determinada aceleração e as demandas computacionais.

5 Conclusão

Este trabalho apresenta uma abordagem inovadora que utiliza um modelo ViT treinado com espectrogramas para a reconstrução de espectros de diferença provenientes de aquisições GABA MEGA-PRESS. A metodologia emprega apenas um quarto do número total de transientes convencionais, resultando em um fator de aceleração que tornaria as aquisições quatro vezes mais rápidas. Essa solução foi comparada com outros métodos presentes na literatura, alcançando performance de estado da arte.

A primeira versão do Spectro-ViT, desenvolvida no contexto do desafio *Edited-MRS Reconstruction Challenge*, apresentou resultados promissores, confirmando o potencial da abordagem sendo proposta, que aliava o uso do espectrograma para representar os sinais com um modelo do tipo *Transformer*. O modelo proposto foi testado durante o desafio não só em dados simulados, mas também em dados *in-vivo*, sendo estes provenientes de equipamentos de ressonância de três fabricantes distintos e adquiridos com parâmetros distintos. O Spectro-ViT foi considerado o vencedor geral do desafio, tendo sido premiado durante o *IEEE International Symposium on Biomedical Imaging (ISBI 2023)*, realizado na Colômbia.

A versão aprimorada do Spectro-ViT, com entradas distintas em seus 3 canais, uma função de perda customizada para o treinamento apresentou uma performance ainda melhor que a versão original premiada. O novo modelo Spectro-ViT, treinado e testado em um conjunto ainda maior de dados, exibiu o melhor desempenho nas métricas de qualidade espectral entre todos os métodos comparados, apresentando as métricas de qualidade SNR, FWHM e erro de ajuste comparáveis à aquisição com a quantidade convencional de transientes. Além disso, a quantificação das reconstruções do Spectro-ViT para GABA obtiveram um valor médio de R^2 de 0,67 e um valor médio de MAPE de 9,68%, sem diferenças estatísticas significativas em comparação com a referência.

O Spectro-ViT pode servir como *benchmark* para pesquisas futuras, com seu código compartilhado em um repositório público para facilitar a reutilização, aperfeiçoamento e incentivar investigações similares. A utilização deste repositório e a abordagem inovadora deste trabalho, que combina o modelo *Transformer* com a representação de dados de MRS em espectrogramas, já se provaram fundamentais na pesquisa de Almeida *et al.* (2023) envolvendo quantificação.

Apesar dos resultados promissores, pesquisas futuras devem validar a capacidade de generalização do modelo para uma distribuição de dados não explorada neste estudo. O conjunto de dados utilizado teve uma distribuição limitada, considerando apenas

participantes de controle, parâmetros de aquisição específicos e a mesma região cerebral com tamanho de voxel fixo, baseando-se na disponibilidade de dados *in-vivo* de acesso aberto.

Há também espaço para estudos que visem aumentar a complexidade do modelo Spectro-ViT com alterações na arquitetura, com o objetivo de aprimorar os resultados das reconstruções. Agregar transientes por média em um canal reduz a demanda computacional, mas pode perder detalhes específicos. Um modelo ViT tridimensional que processe simultaneamente espectrogramas individuais e também opere ao longo do eixo dos transientes pode capturar mais variações e interações em todos os eixos, otimizando assim a performance. Adicionar mais camadas para reconstruir a representação latente codificada pelo ViT também pode melhorar o desempenho. Contudo, essas alterações demandam mais recursos computacionais e investigações de técnicas adicionais de regularização para evitar o potencial sobreajuste do modelo.

5.1 Premiações

- Primeiro lugar no desafio *Edited-MRS Reconstruction Challenge*, apresentado na conferência *IEEE International Symposium on Biomedical Imaging*, realizada em Cartagena, Colômbia, em 18 de abril de 2023.

5.2 Publicações resultantes

Artigos em periódicos

DIAS, G.; BERTO, R. P.; OLIVEIRA, M.; UEDA, L.; DERTKIGIL, S.; COSTA, P. D.; SHAMAEI, A.; SOUZA, R.; HARRIS, A.; RITTNER, L. Spectro-ViT: A Vision Transformer Model for GABA-edited MRS Reconstruction Using Spectrograms. *Magnetic Resonance Imaging* (submetido).

Artigo que descreve o modelo Spectro-ViT, comparando seus resultados com outros modelos de reconstrução da literatura. Submetido ao periódico *Magnetic Resonance Imaging*, atualmente em processo de revisão.

BERTO, R. P.; BUGLER, H.; **DIAS, G.**; OLIVEIRA, M.; UEDA, L.; DERTKIGIL, S.; COSTA, P. D.; RITTNER, L.; MERKOFER, J. P.; SANDE, D. M. van de et al. Results of the 2023 ISBI challenge to reduce GABA-edited MRS acquisition time. *Magnetic Resonance Materials in Physics, Biology and Medicine*, Springer, p. 1–15, 2024

Artigo que descreve o desafio de reconstrução *Edited-MRS Reconstruction Challenge*, incluindo os resultados do modelo SpectroViT.

Trabalhos em anais de eventos

ALMEIDA, M. V. D. J.; DIAS, G.; SILVA, M. O. D.; RITTNER, L. Magnetic resonance spectroscopy fitting through spectrogram and vision transformers. In: IEEE. 2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)., 2023. p. 1–5

Trabalho que apresenta um modelo de fitting em MRS utilizando ViT como codificador e representações em espectrogramas do sinal. Usando um conjunto de dados com 13128 sinais sintéticos de MRS (PRESS, eco curto), o modelo prevê 24 parâmetros, incluindo amplitudes de 20 metabólitos e macromoléculas, amortecimento global, frequência e deslocamento de fase. Nosso modelo superou a abordagem recente de Shamaei *et al.* (2023), especialmente para metabólitos com sobreposição espectral.

DIAS, G.; OLIVEIRA, M.; ALMEIDA, M.; DERTKIGIL, S.; RITTNER, L. Improving Magnetic Resonance Spectroscopy Fitting with Vision Transformers and Spectrograms through Enhanced Localized Processing. In: BRAINN. 10th BRAINN Congress Journal of Epilepsy and Clinical Neurophysiology, 2024. v. 29, n. 1

Este estudo aprimora o modelo de fitting para MRS de Almeida *et al.* (2023), melhorando a eficiência computacional e adicionando filtros de tempo e frequência no espectrograma que otimizam a modelagem do ViT. O espectrograma é filtrado temporalmente para reter os primeiros 60% da duração do sinal e em frequência para incluir a região de 0 a 6 ppm, focando nas principais informações de metabólitos. O novo modelo apresentou melhores resultados de R^2 e MAPE comparado aos de Almeida *et al.* (2023) e Shamaei *et al.* (2023).

5.3 Repositórios de código público

Repositório GitHub *Spectro-ViT*: permite a reprodução da pesquisa desenvolvida no trabalho *Spectro-ViT: A Vision Transformer Model for GABA-edited MEGA-PRESS Reconstruction Using Spectrograms* (DIAS *et al.*, 2023) <<https://github.com/MICLab-Unicamp/Spectro-ViT>> .

Repositório GitHub *MICLab-2023-ISBI-MRS-Challenge* : permite a reprodução da solução vencedora do desafio *Edited-MRS Reconstruction Challenge* (BERTO *et al.*, 2024) <<https://github.com/MICLab-Unicamp/MICLab-2023-ISBI-MRS-Challenge>> .

Repositório GitHub *MR SpectroFitDL*: permite a reprodução da pesquisa desenvolvida no trabalho *Improving Magnetic Resonance Spectroscopy Fitting with Vision Transformers and Spectrograms through Enhanced Localized Processing* (DIAS *et al.*, 2024) <<https://github.com/gsantosmdias/MRSpectroFitDL>>.

Repositório GitHub *Spectral_fitting_SIPAIM*: permite a reprodução da pesquisa desenvolvida no trabalho *Magnetic Resonance Spectroscopy Fitting through Spectrogram and Vision Transformers* (ALMEIDA *et al.*, 2023) (<https://github.com/MICLab-Unicamp/Spectral_fitting_SIPAIM>).

5.4 Bolsas de Estudo

- *Emerging Leaders in the Americas Program* (Governo do Canadá):
 - **Instituição:** Universidade de Calgary, Calgary, Alberta, Canadá.
 - **Supervisores:** Dr. Roberto Souza (Laboratório de Imagem Avançada e Inteligência Artificial) e Dra. Ashley Harris (Harris Imaging Lab).
 - **Tipo:** Bolsa de Intercâmbio
 - **Principais Atividades:**
 - * Aperfeiçoamento de habilidades de programação e compreensão da arquitetura de Transformes.
 - * Expansão do conhecimento em MRS.
 - * Apresentação da pesquisa de mestrado na série de Seminários Avançados de Imagem Médica da Universidade de Calgary.
 - * Participação no curso de Programação de Pulsos de Ressonância Magnética da GE HealthCare.
 - * Redação de artigo para submissão em revista científica.
- *Scholarship Program - Google DeepMind*:
 - **Agradecimento:** O presente trabalho foi realizado com apoio do processo nº 2020/09838-0, da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).
 - **Benefícios:** Financiamento integral para o mestrado, orientação com um engenheiro de pesquisa da empresa Google DeepMind e acesso exclusivo a sessões e seminários da empresa, focados em aprimorar habilidades essenciais em IA e promover o sucesso acadêmico.

Referências

- AFLALO, Y.; NOY, A.; LIN, M.; FRIEDMAN, I.; ZELNIK, L. Knapsack pruning with inner distillation. *arXiv preprint arXiv:2002.08258*, 2020. Citado na página 59.
- ALMEIDA, M. V. D. J.; DIAS, G. S. M.; SILVA, M. O. D.; RITTNER, L. Magnetic resonance spectroscopy fitting through spectrogram and vision transformers. In: *IEEE. 2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)*. [S.l.], 2023. p. 1–5. Citado 5 vezes nas páginas 17, 23, 82, 84 e 85.
- ASLAN, Z.; AKIN, M. Automatic detection of schizophrenia by applying deep learning over spectrogram images of eeg signals. *Traitement du Signal*, v. 37, p. 235–244, 2020. Citado na página 30.
- BEAUFERRIS, Y.; TEUWEN, J.; KARKALOUSOS, D.; MORIAKOV, N.; CAAN, M.; YIASEMIS, G. *et al.* Multi-coil MRI reconstruction challenge—assessing brain MRI reconstruction models and their generalizability to varying coil configurations. *Front Neurosci*, Frontiers, v. 16, p. 919186, 2022. Citado 3 vezes nas páginas 16, 20 e 36.
- BELL, T.; STOKOE, M.; HARRIS, A. D. Macromolecule suppressed GABA levels show no relationship with age in a pediatric sample. *Sci Rep*, Nature Publishing Group UK London, v. 11, p. 722, 2021. Citado na página 14.
- BERTO, R.; BUGLER, H.; SOUZA, R.; HARRIS, A. An Investigation of Different Deep Learning Pipelines for GABA-Edited MRS Reconstruction. In: *SPRINGER. International Workshop on Machine Learning in Medical Imaging*. [S.l.], 2023. p. 72–82. Citado 9 vezes nas páginas 16, 20, 21, 22, 36, 64, 65, 67 e 77.
- BERTO, R. P.; BUGLER, H.; DIAS, G.; OLIVEIRA, M.; UEDA, L.; DERTKIGIL, S.; COSTA, P. D.; RITTNER, L.; MERKOFER, J. P.; SANDE, D. M. van de *et al.* Results of the 2023 ISBI challenge to reduce GABA-edited MRS acquisition time. *Magnetic Resonance Materials in Physics, Biology and Medicine*, Springer, p. 1–15, 2024. Citado 16 vezes nas páginas 16, 36, 48, 49, 50, 51, 52, 53, 54, 55, 56, 59, 77, 79, 80 e 84.
- BLOCH, F. Nuclear induction. *Physical review*, APS, v. 70, n. 7-8, p. 460, 1946. Citado 3 vezes nas páginas 24, 25 e 26.
- BOTTOMLEY, P. A. *et al.* Spatial localization in nmr spectroscopy in vivo. *Ann NY Acad Sci*, v. 508, p. 333–348, 1987. Citado 2 vezes nas páginas 14 e 31.
- BREUER, F. A.; BLAIMER, M.; MUELLER, M. F.; SEIBERLICH, N.; HEIDEMANN, R. M.; GRISWOLD, M. A.; JAKOB, P. M. Controlled aliasing in volumetric parallel imaging (2d caipirinha). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Wiley Online Library, v. 55, n. 3, p. 549–556, 2006. Citado na página 20.
- BROWN, T. R.; KINCAID, B.; UGURBIL, K. Nmr chemical shift imaging in three dimensions. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 79, p. 3523–3526, 1982. Citado 3 vezes nas páginas 14, 19 e 31.

- CECIL, K. M. Proton magnetic resonance spectroscopy: technique for the neuroradiologist. *Neuroimaging Clinics*, Elsevier, v. 23, p. 381–392, 2013. Citado na página 28.
- CHOI, I.-Y.; ANDRONESI, O. C.; BARKER, P.; BOGNER, W.; EDDEN, R. A.; KAISER, L. G.; LEE, P.; MARJAŃSKA, M.; TERPSTRA, M.; GRAAF, R. A. de. Spectral editing in 1h magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR in Biomedicine*, Wiley Online Library, v. 34, n. 5, p. e4411, 2021. Citado 5 vezes nas páginas 14, 15, 31, 57 e 58.
- CONOVER, W. J. *Practical nonparametric statistics*. 3th. ed. [S.l.]: John Wiley & Sons, 1999. Citado na página 71.
- DEMAYO, M. M.; HARRIS, A. D.; SONG, Y. J. C.; POKORSKI, I.; THAPA, R.; PATEL, S. *et al.* Age-related parietal GABA alterations in children with autism spectrum disorder. *Autism Res*, Wiley Online Library, v. 14, p. 859–872, 2021. Citado na página 14.
- DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. Imagenet: A large-scale hierarchical image database. In: IEEE. *IEEE Comput Soc Conf Comput Vis Pattern Recognit*. [S.l.], 2009. p. 248–255. Citado 2 vezes nas páginas 47 e 60.
- DHAMALA, E.; ABDELKEFI, I.; NGUYEN, M.; HENNESSY, T. J.; NADEAU, H.; NEAR, J. Validation of in vivo MRS measures of metabolite concentrations in the human brain. *NMR in Biomedicine*, Wiley Online Library, v. 32, n. 3, p. e4058, 2019. Citado na página 14.
- DIAS, G.; BERTO, R. P.; OLIVEIRA, M.; UEDA, L.; DERTKIGIL, S.; COSTA, P. D.; SHAMAEI, A.; SOUZA, R.; HARRIS, A.; RITTNER, L. Spectro-ViT: A Vision Transformer Model for GABA-edited MRS Reconstruction Using Spectrograms. *arXiv preprint arXiv:2311.15386*, 2023. Citado 7 vezes nas páginas 44, 69, 70, 72, 74, 77 e 84.
- DIAS, G.; OLIVEIRA, M.; ALMEIDA, M.; DERTKIGIL, S.; RITTNER, L. Improving Magnetic Resonance Spectroscopy Fitting with Vision Transformers and Spectrograms through Enhanced Localized Processing. In: BRAINN. *10th BRAINN Congress Journal of Epilepsy and Clinical Neurophysiology*. [S.l.], 2024. v. 29, n. 1. Citado 2 vezes nas páginas 23 e 85.
- DIKAIOS, N. Deep learning magnetic resonance spectroscopy fingerprints of brain tumours using quantum mechanically synthesised data. *NMR in Biomedicine*, Wiley Online Library, v. 34, n. 4, p. e4479, 2021. Citado na página 17.
- DIKER, A.; CÖMERT, Z.; AVCı, E.; TOĞAÇAR, M.; ERGEN, B. A novel application based on spectrogram and convolutional neural network for ecg classification. In: *2019 1st International Informatics and Software Engineering Conference (UBMYK)*. [S.l.: s.n.], 2019. p. 1–6. Citado na página 30.
- DOSOVITSKIY, A.; BEYER, L.; KOLESNIKOV, A.; WEISSENBORN, D.; ZHAI, X.; UNTERTHINER, T. o. *An image is worth 16x16 words: Transformers for image recognition at scale*. 2020. Citado na página 21.
- DULIŃSKA-LITEWKA, J.; ŁAZARCZYK, A.; HAŁUBIEC, P.; SZAFRAŃSKI, O.; KARNAS, K.; KAREWICZ, A. Superparamagnetic iron oxide nanoparticles—current

and prospective medical applications. *Materials*, MDPI, v. 12, n. 4, p. 617, 2019. Citado na página 26.

DZIADOSZ, M.; RIZZO, R.; KYATHANAHALLY, S. P.; KREIS, R. Denoising single MR spectra by deep learning: Miracle or mirage? *Magn Reson Med*, Wiley Online Library, v. 90, p. 1749–1761, 2023. Citado 2 vezes nas páginas 67 e 80.

EDDEN, R. A.; PUTS, N. A.; HARRIS, A. D.; BARKER, P. B.; EVANS, C. J. Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *J Magn Reson Imaging*, Wiley Online Library, v. 40, p. 1445–1452, 2014. Citado 5 vezes nas páginas 21, 22, 33, 35 e 37.

ERNST, R. R.; ANDERSON, W. A. Application of fourier transform spectroscopy to magnetic resonance. *Review of Scientific Instruments*, American Institute of Physics, v. 37, n. 1, p. 93–102, 1966. Citado na página 27.

EVANS, C. J.; PUTS, N. A.; ROBSON, S. E.; BOY, F.; MCGONIGLE, D. J.; SUMNER, P.; SINGH, K. D.; EDDEN, R. A. Subtraction artifacts and frequency (mis-) alignment in j-difference gaba editing. *Journal of Magnetic Resonance Imaging*, Wiley Online Library, v. 38, n. 4, p. 970–975, 2013. Citado 2 vezes nas páginas 15 e 33.

FRAHM, J.; MERBOLDT, K.-D.; HÄNICKE, W. Localized proton spectroscopy using stimulated echoes. *Journal of Magnetic Resonance (1969)*, Elsevier, v. 72, p. 502–508, 1987. Citado 2 vezes nas páginas 14 e 31.

FUCHS, A.; BOESIGER, P.; SCHULTE, R. F.; HENNING, A. Profit revisited. *Magnetic Resonance in Medicine*, Wiley Online Library, v. 71, n. 2, p. 458–468, 2014. Citado na página 19.

GHIT, A.; ASSAL, D.; AL-SHAMI, A. S.; HUSSEIN, D. E. E. Gabaa receptors: structure, function, pharmacology, and related disorders. *Journal of Genetic Engineering and Biotechnology*, Springer, v. 19, n. 1, p. 123, 2021. Citado na página 14.

GONG, Y.; CHUNG, Y.-A.; GLASS, J. Ast: Audio spectrogram transformer. *arXiv preprint arXiv:2104.01778*, 2021. Citado 2 vezes nas páginas 43 e 60.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016. Citado 2 vezes nas páginas 16 e 59.

GRAAF, R. A. D. *In vivo NMR spectroscopy: principles and techniques*. 3th. ed. [S.l.]: John Wiley & Sons, 2019. Citado 8 vezes nas páginas 15, 24, 25, 26, 27, 28, 32 e 33.

GRISWOLD, M. A.; JAKOB, P. M.; HEIDEMANN, R. M.; NITTKA, M.; JELLUS, V.; WANG, J.; KIEFER, B.; HAASE, A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Wiley Online Library, v. 47, p. 1202–1210, 2002. Citado na página 20.

HAASE, A.; FRAHM, J.; HANICKE, W.; MATTHAEI, D. 1h nmr chemical shift selective (chess) imaging. *Physics in Medicine & Biology*, IOP Publishing, v. 30, p. 341, 1985. Citado na página 39.

- HARRIS, A. D.; GILBERT, D. L.; HORN, P. S.; CROCETTI, D.; CECIL, K. M.; EDDEN, R. A. *et al.* Relationship between GABA levels and task-dependent cortical excitability in children with attention-deficit/hyperactivity disorder. *Clin Neurophysiol*, Elsevier, v. 132, p. 1163–1172, 2021. Citado na página 14.
- HARRIS, A. D.; GLAUBITZ, B.; NEAR, J.; EVANS, C. J.; PUTS, N. A.; SCHMIDT-WILCKE, T.; TEGENTHOFF, M.; BARKER, P. B.; EDDEN, R. A. Impact of frequency drift on gamma-aminobutyric acid-edited MR spectroscopy. *Magn Reson Med*, Wiley Online Library, v. 72, p. 941–948, 2014. Citado 2 vezes nas páginas 15 e 33.
- HARRIS, A. D.; SALEH, M. G.; EDDEN, R. A. Edited 1h magnetic resonance spectroscopy in vivo: Methods and metabolites. *Magn Reson Med*, Wiley Online Library, v. 77, p. 1377–1389, 2017. Citado 2 vezes nas páginas 29 e 31.
- HEER, P. de; BIZINO, M. B.; LAMB, H. J.; WEBB, A. G. Parameter optimization for reproducible cardiac 1H-MR spectroscopy at 3 Tesla. *Journal of Magnetic Resonance Imaging*, Wiley Online Library, v. 44, n. 5, p. 1151–1158, 2016. Citado na página 39.
- HEINZEL, G.; RÜDIGER, A.; SCHILLING, R. Spectrum and spectral density estimation by the discrete fourier transform (dft), including a comprehensive list of window functions and some new at-top windows. 2002. Citado 2 vezes nas páginas 59 e 60.
- HERNÁNDEZ-VILLEGAS, Y.; ORTEGA-MARTORELL, S.; ARÚS, C.; VELLIDO, A.; JULIÀ-SAPÉ, M. Extraction of artefactual MRS patterns from a large database using non-negative matrix factorization. *NMR in Biomedicine*, Wiley Online Library, v. 35, n. 4, p. e4193, 2022. Citado na página 17.
- HLAWATSCH, F.; BOUDREAUX-BARTELS, G. F. Linear and quadratic time-frequency signal representations. *IEEE signal processing magazine*, IEEE, v. 9, p. 21–67, 1992. Citado na página 29.
- HUANG, G.; LIU, Z.; MAATEN, L. V. D.; WEINBERGER, K. Q. Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 4700–4708. Citado na página 18.
- HUI, S. C.; MIKKELSEN, M.; ZÖLLNER, H. J.; AHLUWALIA, V.; ALCAUTER, S.; BALTUSIS, L.; BARANY, D. A.; BARLOW, L. R.; BECKER, R.; BERMAN, J. I. *et al.* Frequency drift in mr spectroscopy at 3t. *NeuroImage*, Elsevier, v. 241, p. 118430, 2021. Citado na página 65.
- HYBERTS, S. G.; MILBRADT, A. G.; WAGNER, A. B.; ARTHANARI, H.; WAGNER, G. Application of iterative soft thresholding for fast reconstruction of nmr data non-uniformly sampled with multidimensional poisson gap scheduling. *Journal of biomolecular NMR*, Springer, v. 52, p. 315–327, 2012. Citado na página 18.
- IQBAL, Z.; NGUYEN, D.; HANGEL, G.; MOTYKA, S.; BOGNER, W.; JIANG, S. Super-resolution 1h magnetic resonance spectroscopic imaging utilizing deep learning. *Frontiers in oncology*, Frontiers Media SA, v. 9, p. 1010, 2019. Citado na página 19.
- IQBAL, Z.; NGUYEN, D.; THOMAS, M. A.; JIANG, S. Deep learning can accelerate and quantify simulated localized correlated spectroscopy. *Scientific reports*, Springer, v. 11, p. 1–13, 2021. Citado na página 19.

- ISLAM, S.; ELMEEKKI, H.; ELSEBAI, A.; BENTAHAR, J.; DRAWEL, N.; RJOUB, G. *et al.* A comprehensive survey on applications of transformers for deep learning tasks. *Expert Syst Appl*, Elsevier, v. 241, p. 122666, 2023. Citado 3 vezes nas páginas 17, 22 e 43.
- JANG, J.; LEE, H. H.; PARK, J.-A.; KIM, H. Unsupervised anomaly detection using generative adversarial networks in 1h-MRS of the brain. *Journal of Magnetic Resonance*, Elsevier, v. 325, p. 106936, 2021. Citado na página 17.
- JEON, H.; JUNG, Y.; LEE, S.; JUNG, Y. Area-efficient short-time fourier transform processor for time–frequency analysis of non-stationary signals. *Applied Sciences*, MDPI, v. 10, n. 20, p. 7208, 2020. Citado na página 30.
- JUCHEM, C.; CUDALBU, C.; GRAAF, R. A. de; GRUETTER, R.; HENNING, A.; HETHERINGTON, H. P.; BOER, V. O. B0 shimming for in vivo magnetic resonance spectroscopy: experts' consensus recommendations. *NMR in Biomedicine*, Wiley Online Library, v. 34, n. 5, p. e4350, 2021. Citado 2 vezes nas páginas 33 e 62.
- KAZIMIERCZUK, K.; OREKHOV, V. Y. Accelerated nmr spectroscopy by using compressed sensing. *Angewandte Chemie International Edition*, Wiley Online Library, v. 50, p. 5556–5559, 2011. Citado na página 18.
- KEHTARNAVAZ, N. Chapter 7 - frequency domain processing. In: KEHTARNAVAZ, N. (Ed.). *Digital Signal Processing System Design (Second Edition)*. Second edition. Burlington: Academic Press, 2008. p. 175–196. ISBN 978-0-12-374490-6. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780123744906000076>>. Citado na página 29.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. Citado na página 47.
- KREIS, R.; BOER, V.; CHOI, I.-Y.; CUDALBU, C.; GRAAF, R. A. de; GASPAROVIC, C. *et al.* Terminology and concepts for the characterization of in vivo MR spectroscopy methods and MR spectra: background and experts' consensus recommendations. *NMR Biomed*, Wiley Online Library, v. 34, p. e4347, 2021. Citado 2 vezes nas páginas 15 e 20.
- KYATHANAHALLY, S. P.; DÖRING, A.; KREIS, R. Deep learning approaches for detection and removal of ghosting artifacts in MR spectroscopy. *Magn Reson Med*, Wiley Online Library, v. 80, p. 851–863, 2018. Citado 2 vezes nas páginas 43 e 59.
- KYATHANAHALLY, S. P.; MOCIOIU, V.; BARROS, N. Pedrosa de; SLOTBOOM, J.; WRIGHT, A. J.; JULIÀ-SAPÉ, M.; ARÚS, C.; KREIS, R. Quality of clinical brain tumor mr spectra judged by humans and machine learning tools. *Magnetic resonance in medicine*, Wiley Online Library, v. 79, n. 5, p. 2500–2510, 2018. Citado na página 17.
- LE, L.; KABIR, A. N. M.; JI, C.; BASODI, S.; PAN, Y. Using transfer learning, svm, and ensemble classification to classify baby cries based on their spectrogram images. In: *2019 IEEE 16th International Conference on Mobile Ad Hoc and Sensor Systems Workshops (MASSW)*. [S.l.: s.n.], 2019. p. 106–110. Citado na página 30.
- LEE, H.; LEE, H. H.; KIM, H. Reconstruction of spectra from truncated free induction decays by deep learning in proton magnetic resonance spectroscopy. *Magnetic resonance in medicine*, Wiley Online Library, v. 84, n. 2, p. 559–568, 2020. Citado 2 vezes nas páginas 17 e 18.

- LI, J.; CHEN, J.; TANG, Y.; WANG, C.; LANDMAN, B. A.; ZHOU, S. K. Transforming medical imaging with Transformers? A comparative review of key properties, current progresses, and future perspectives. *Med Image Anal*, v. 85, p. 102762, 2023. ISSN 1361-8415. Citado na página 17.
- LIN, A.; ANDRONESI, O.; BOGNER, W.; CHOI, I.-Y.; COELLO, E.; CUDALBU *et al.* Minimum reporting Standards for in vivo Magnetic Resonance Spectroscopy (MRSinMRS): Experts' consensus recommendations. *NMR Biomed*, Wiley Online Library, v. 34, p. e4484, 2021. Citado 4 vezes nas páginas 35, 56, 57 e 58.
- LUO, J.; ZENG, Q.; WU, K.; LIN, Y. Fast reconstruction of non-uniform sampling multidimensional nmr spectroscopy via a deep neural network. *Journal of Magnetic Resonance*, Elsevier, v. 317, p. 106772, 2020. Citado na página 18.
- LUSTIG, M.; DONOHO, D.; PAULY, J. M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Wiley Online Library, v. 58, n. 6, p. 1182–1195, 2007. Citado na página 19.
- MARIOTTI, O.; CORD, M.; SCHWANDER, O. Exploring deep vision models for acoustic scene classification. In: *Workshop on Detection and Classification of Acoustic Scenes and Events*. [S.l.: s.n.], 2018. Citado na página 30.
- MAUDSLEY, A. A.; ANDRONESI, O. C.; BARKER, P. B.; BIZZI, A.; BOGNER, W.; HENNING, A.; NELSON, S. J.; POSSE, S.; SHUNGU, D. C.; SOHER, B. J. Advanced magnetic resonance spectroscopic neuroimaging: Experts' consensus recommendations. *NMR Biomed*, Wiley Online Library, v. 34, p. e4309, 2021. Citado na página 19.
- MCCORMICK, D. A. GABA as an inhibitory neurotransmitter in human cerebral cortex. *J Neurophysiol*, American Physiological Society Bethesda, MD, v. 62, p. 1018–1027, 1989. Citado na página 14.
- MESCHER, M.; MERKLE, H.; KIRSCH, J.; GARWOOD, M.; GRUETTER, R. Simultaneous in vivo spectral editing and water suppression. *NMR Biomed*, Wiley Online Library, v. 11, p. 266–272, 1998. Citado 2 vezes nas páginas 14 e 31.
- MIKKELSEN, M.; BARKER, P. B.; BHATTACHARYYA, P. K.; BRIX, M. K.; BUUR, P. F.; CECIL, K. M. *et al.* Big GABA: Edited MR spectroscopy at 24 research sites. *Neuroimage*, Elsevier, v. 159, p. 32–45, 2017. Citado 3 vezes nas páginas 27, 39 e 64.
- MIKKELSEN, M.; LOO, R. S.; PUTS, N. A.; EDDEN, R. A.; HARRIS, A. D. Designing GABA-edited magnetic resonance spectroscopy studies: considerations of scan duration, signal-to-noise ratio and sample size. *J Neurosci Meth*, Elsevier, v. 303, p. 86–94, 2018. Citado na página 15.
- MIKKELSEN, M.; RIMBAULT, D. L.; BARKER, P. B.; BHATTACHARYYA, P. K.; BRIX, M. K.; BUUR, P. F. *et al.* Big GABA ii: Water-referenced edited MR spectroscopy at 25 research sites. *Neuroimage*, Elsevier, v. 191, p. 537–548, 2019. Citado 2 vezes nas páginas 39 e 64.
- MUCKLEY, M. J.; RIEMENSCHNEIDER, B.; RADMANESH, A.; KIM, S.; JEONG, G.; KO, J. *et al.* Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. *IEEE T Med Imaging*, IEEE, v. 40, p. 2306–2317, 2021. Citado 3 vezes nas páginas 16, 20 e 36.

- MULLINS, P. G.; MCGONIGLE, D. J.; O’GORMAN, R. L.; PUTS, N. A.; VIDYASAGAR, R.; EVANS, C. J.; EDDEN, R. A. *et al.* Current practice in the use of mega-press spectroscopy for the detection of gaba. *Neuroimage*, Elsevier, v. 86, p. 43–52, 2014. Citado na página 31.
- NASSIRPOUR, S.; CHANG, P.; HENNING, A. MultiNet PyGRAPPA: Multiple neural networks for reconstructing variable density GRAPPA (a 1H FID MRSI study). *Neuroimage*, Elsevier, v. 183, p. 336–345, 2018. Citado na página 20.
- NEAR, J.; ANDERSSON, J.; MARON, E.; MEKLE, R.; GRUETTER, R.; COWEN, P.; JEZZARD, P. Unedited in vivo detection and quantification of γ -aminobutyric acid in the occipital cortex using short-te MRS at 3 t. *NMR Biomed*, Wiley Online Library, v. 26, p. 1353–1362, 2013. Citado na página 38.
- NEAR, J.; HARRIS, A. D.; JUCHEM, C.; KREIS, R.; MARJAŃSKA, M.; ÖZ, G. *et al.* Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: experts’ consensus recommendations. *NMR Biomed*, Wiley Online Library, v. 34, p. e4257, 2021. Citado 2 vezes nas páginas 57 e 58.
- OGG, R. J.; KINGSLEY, R.; TAYLOR, J. S. Wet, a t1-and b1-insensitive water-suppression method for in vivo localized 1h nmr spectroscopy. *Journal of Magnetic Resonance, Series B*, Elsevier, v. 104, p. 1–10, 1994. Citado na página 39.
- ORDIDGE, R. J.; CONNELLY, A.; LOHMAN, J. A. Image-selected in vivo spectroscopy (isis). a new technique for spatially selective nmr spectroscopy. *Journal of Magnetic Resonance (1969)*, Elsevier, v. 66, p. 283–294, 1986. Citado 2 vezes nas páginas 14 e 31.
- PAZ, L. D. Ochoa-de la; GULIAS-CAÑIZO, R.; RUÍZ-LEYJA, E. D.; SÁNCHEZ-CASTILLO, H.; PARODÍ, J. *et al.* The role of GABA neurotransmitter in the human central nervous system, physiology, and pathophysiology. *Revista mexicana de neurociencia*, Academia Mexicana de Neurología AC, v. 22, n. 2, p. 67–76, 2021. Citado na página 14.
- PEEK, A. L.; REBBECK, T. J.; LEAVER, A. M.; FOSTER, S. L.; REFSHAUGE, K. M.; PUTS, N. A. *et al.* A comprehensive guide to MEGA-PRESS for GABA measurement. *Anal Biochem*, Elsevier, v. 669, p. 115113, 2023. Citado 3 vezes nas páginas 15, 32 e 33.
- PIGOTT, T.; MCPEAK, A.; CHASTELAIN, A. de; DEMAYO, M. M.; RASIC, N.; RAYNER, L. *et al.* Changes in Brain GABA and Glutamate and Improvements in Physical Functioning Following Intensive Pain Rehabilitation in Youth With Chronic Pain. *J Pain Res*, Elsevier, v. 24, p. 1288–1297, 2023. Citado na página 14.
- PROCTOR, W.; YU, F. The dependence of a nuclear magnetic resonance frequency upon chemical compound. *Physical Review*, APS, v. 77, n. 5, p. 717, 1950. Citado na página 28.
- PRUESSMANN, K. P.; WEIGER, M.; SCHEIDEGGER, M. B.; BOESIGER, P. Sense: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Wiley Online Library, v. 42, n. 5, p. 952–962, 1999. Citado na página 20.

- PURCELL, E. M. Research in nuclear magnetism. *Science*, American Association for the Advancement of Science, v. 118, n. 3068, p. 431–436, 1953. Citado 3 vezes nas páginas 24, 25 e 26.
- PUTS, N. A.; EDDEN, R. A. In vivo magnetic resonance spectroscopy of gaba: a methodological review. *Progress in nuclear magnetic resonance spectroscopy*, Elsevier, v. 60, p. 29–41, 2012. Citado na página 29.
- RADOSAVOVIC, I.; KOSARAJU, R. P.; GIRSHICK, R.; HE, K.; DOLLÁR, P. Designing network design spaces. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 10428–10436. Citado na página 59.
- RIDNIK, T.; LAWEN, H.; NOY, A.; BARUCH, E. B.; SHARIR, G.; FRIEDMAN, I. Tresnet: High performance gpu-dedicated architecture. In: *proceedings of the IEEE/CVF winter conference on applications of computer vision*. [S.l.: s.n.], 2021. p. 1400–1409. Citado na página 59.
- RITTER, C.; BUCHMANN, A.; MÜLLER, S. T.; VOLLEBERG, M.; HAYNES, M.; GHISLENI, C. *et al.* Evaluation of Prefrontal γ -Aminobutyric Acid and Glutamate Levels in Individuals With Major Depressive Disorder Using Proton Magnetic Resonance Spectroscopy. *JAMA Psychiat*, American Medical Association, v. 79, p. 1209–1216, 2022. Citado na página 14.
- RIZZO, R.; DZIADOSZ, M.; KYATHANAHALLY, S. P.; SHAMAEI, A.; KREIS, R. 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. *Magn Reson Med*, Wiley Online Library, v. 89, p. 1707–1727, 2023. Citado na página 43.
- ROTHMAN, D. L.; BEHAR, K. L.; PRICHARD, J. W.; PETROFF, O. A. Homocarnosine and the measurement of neuronal pH in patients with epilepsy. *Magn Reson Med*, Wiley Online Library, v. 38, p. 924–929, 1997. Citado na página 32.
- ROWLAND, B. C.; LIAO, H.; ADAN, F.; MARIANO, L.; IRVINE, J.; LIN, A. P. Correcting for frequency drift in clinical brain mr spectroscopy. *Journal of Neuroimaging*, Wiley Online Library, v. 27, n. 1, p. 23–28, 2017. Citado 2 vezes nas páginas 33 e 62.
- SANDE, D. M. van de; MERKOFER, J. P.; AMIRRAJAB, S.; VETA, M.; SLOUN, R. J. van; VERSLUIS, M. J. *et al.* A review of machine learning applications for the proton MR spectroscopy workflow. *Magn Reson Med*, Wiley Online Library, 2023. Citado 6 vezes nas páginas 16, 17, 18, 20, 22 e 55.
- SCHMID, N.; BRUDERER, S.; PARUZZO, F.; FISCHETTI, G.; TOSCANO, G.; GRAF, D.; FEY, M.; HENRICI, A.; ZIEBART, V.; HEITMANN, B. *et al.* Deconvolution of 1d nmr spectra: A deep learning-based approach. *Journal of Magnetic Resonance*, Elsevier, v. 347, p. 107357, 2023. Citado na página 17.
- SCHMIEDER, P.; STERN, A. S.; WAGNER, G.; HOCH, J. C. Application of nonlinear sampling schemes to cosy-type spectra. *Journal of biomolecular NMR*, Springer, v. 3, p. 569–576, 1993. Citado na página 18.

SHAMAEI, A.; STARCUKOVÁ, J.; JR, Z. S. A Wavelet Scattering Convolutional Network for Magnetic Resonance Spectroscopy Signal Quantitation. In: SCIENCE AND TECHNOLOGY PUBLICATIONS. *BIOSIGNALS*. [S.l.], 2021. p. 268–275. Citado na página 43.

SHAMAEI, A.; STARCUKOVA, J.; JR, Z. S. Physics-informed deep learning approach to quantification of human brain metabolites from magnetic resonance spectroscopy data. *Computers in Biology and Medicine*, Elsevier, v. 158, p. 106837, 2023. Citado 2 vezes nas páginas 17 e 84.

SHYU, C.; CHAVEZ, S.; BOILEAU, I.; FOLL, B. L. Quantifying GABA in addiction: a review of proton magnetic resonance spectroscopy studies. *Brain Sciences*, MDPI, v. 12, n. 7, p. 918, 2022. Citado na página 14.

SIMPSON, R.; DEVENYI, G. A.; JEZZARD, P.; HENNESSY, T. J.; NEAR, J. Advanced processing and simulation of MRS data using the fid appliance (fid-a)—an open source, matlab-based toolkit. *Magn Reson Med*, Wiley Online Library, v. 77, p. 23–33, 2017. Citado na página 37.

STAR-LACK, J.; SPIELMAN, D.; ADALSTEINSSON, E.; KURHANEWICZ, J.; TERRIS, D. J.; VIGNERON, D. B. In vivo lactate editing with simultaneous detection of choline, creatine, naa, and lipid singlets at 1.5 t using press excitation with applications to the study of brain and head and neck tumors. *Journal of magnetic resonance*, Elsevier, v. 133, p. 243–254, 1998. Citado 2 vezes nas páginas 14 e 31.

THOMAS, M. A.; YUE, K.; BINESH, N.; DAVANZO, P.; KUMAR, A.; SIEGEL, B.; FRYE, M.; CURRAN, J.; LUFKIN, R.; MARTIN, P. *et al.* Localized two-dimensional shift correlated MR spectroscopy of human brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Wiley Online Library, v. 46, p. 58–67, 2001. Citado na página 18.

YING, J.; DELAGLIO, F.; TORCHIA, D. A.; BAX, A. Sparse multidimensional iterative lineshape-enhanced (smile) reconstruction of both non-uniformly sampled and conventional nmr data. *Journal of biomolecular NMR*, Springer, v. 68, p. 101–118, 2017. Citado na página 18.

ZAGORUYKO, S.; KOMODAKIS, N. Wide residual networks. *arXiv preprint arXiv:1605.07146*, 2016. Citado na página 59.

ZARINABAD, N.; ABERNETHY, L. J.; AVULA, S.; DAVIES, N. P.; GUTIERREZ, D. R.; JASPAN, T.; MACPHERSON, L.; MITRA, D.; ROSE, H. E.; WILSON, M. *et al.* Application of pattern recognition techniques for classification of pediatric brain tumors by in vivo 3t 1h-mr spectroscopy—a multi-center study. *Magnetic Resonance in Medicine*, Wiley Online Library, v. 79, n. 4, p. 2359–2366, 2018. Citado na página 17.

ZHAO, D.; GRIST, J. T.; ROSE, H. E.; DAVIES, N. P.; WILSON, M.; MACPHERSON, L.; ABERNETHY, L. J.; AVULA, S.; PIZER, B.; GUTIERREZ, D. R. *et al.* Metabolite selection for machine learning in childhood brain tumour classification. *NMR in Biomedicine*, Wiley Online Library, v. 35, n. 6, p. e4673, 2022. Citado na página 17.