

ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
MESTRADO EM CIÊNCIA DA COMPUTAÇÃO

MARILIA ROSA SILVEIRA

**UMA ABORDAGEM MULTITAREFA PARA AVALIAÇÃO
CÉREBRO-COMPORTAMENTO A PARTIR DE
RESSONÂNCIA MAGNÉTICA FUNCIONAL**

Porto Alegre
2022

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**UMA ABORDAGEM
MULTITAREFA PARA
AVALIAÇÃO
CÉREBRO-COMPORTAMENTO
A PARTIR DE RESSONÂNCIA
MAGNÉTICA FUNCIONAL**

MARILIA ROSA SILVEIRA

Dissertação apresentada como requisito parcial à obtenção do grau de Mestra em Ciência da Computação na Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Prof. Afonso Sales
Co-Orientador: Prof. Dr. Prof. Felipe Meneguzzi

**Porto Alegre
2022**

Ficha Catalográfica

S587a Silveira, Marília Rosa

Uma abordagem multitarefa para avaliação cérebro-comportamento a partir de ressonância magnética funcional / Marília Rosa Silveira. – 2022.

75.

Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Computação, PUCRS.

Orientador: Prof. Dr. Afonso Henrique Correa de Sales.

Coorientador: Prof. Dr. Felipe Rech Meneguzzi.

1. Aprendizado Multitarefa. 2. Ressonância Magnética Funcional. 3. Redes Cerebrais. 4. Abordagem Preditiva. I. Sales, Afonso Henrique Correa de. II. Meneguzzi, Felipe Rech. III. , . IV. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).

Bibliotecária responsável: Loiva Duarte Novak CRB-10/2079

MARILIA ROSA SILVEIRA

**UMA ABORDAGEM MULTITAREFA PARA
AVALIAÇÃO CÉREBRO-COMPORTAMENTO A
PARTIR DE RESSONÂNCIA MAGNÉTICA
FUNCIONAL**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestra em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação, Escola Politécnica da Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovado(a) em 20 de Abril de 2022.

BANCA EXAMINADORA:

Dr. Alexandre Franco (NKI)

Prof. Dr. Rodrigo C. Barros (PPGCC/PUCRS)

Prof. Dr. Prof. Felipe Meneguzzi (PPGCC/PUCRS- Co-Orientador)

Prof. Dr. Prof. Afonso Sales (PPGCC/PUCRS - Orientador)

DEDICATÓRIA

Dedico este trabalho a todas as vítimas da COVID-19, em especial a meu estimado padrinho Wilson.

AGRADECIMENTOS

Àqueles que me mostraram o conhecimento como fonte eterna de riqueza e pertencimento: meus pais, Marcos Antônio e Maria Jussara. Dedico este trabalho a eles e a Deus por permanecer ao lado deles quando não havia o que eu pudesse fazer por eles em suas lutas contra o câncer. Agradeço às minhas irmãs, Julia e Cecília pela compreensão de minha ausência e pelo suporte que me deram para acreditar neste trabalho. Ao meu namorado, Eduardo Andrade, pela escuta e paciência em todos os momentos.

A meu orientador Dr Afonso Sales pela dedicação a este trabalho desafiador, buscando formas de torná-lo possível e por me suportar durante todo o processo. Obrigada também ao meu coorientador Dr Felipe Meneguzzi pelas dúvidas sanadas durante este trajeto.

Aos colegas de academia Nathalia Bianchini, Matheus Marcon e professor Augusto Buchweitz por me introduzirem aos conhecimentos de neuroimagem. Ao meu primeiro orientador e amigo, professor Ulisses Corrêa, me apoiando nos passos da profissão e dos estudos.

A equipe DBServer na qual trabalho, principalmente ao time de ciência de dados, por possibilitar minha dupla jornada, pela compreensão e apoio neste período.

Por fim, agradeço a esta Universidade pela qualidade e excelência do ensino.

UMA ABORDAGEM MULTITAREFA PARA AVALIAÇÃO CÉREBRO-COMPORTAMENTO A PARTIR DE RESSONÂNCIA MAGNÉTICA FUNCIONAL

RESUMO

O exame de ressonância magnética funcional é utilizado para verificar o comportamento das redes cerebrais em pesquisas de desordens cerebrais e distúrbios cognitivos. Para compor essa avaliação, dados comportamentais e testes clínicos padronizados podem ser associados. Algoritmos lineares são utilizados na literatura para encontrar correlação entre os dados das imagens e uma variável comportamental alvo, valor final de um teste cognitivo, por exemplo. Quando há mais de uma variável a ser predita, pode ser utilizado algoritmos com aprendizado multitarefa. Considerando a premissa de que quando tarefas são relacionadas podem ser preditas em simultâneo, o presente trabalho visa desenvolver uma abordagem multitarefa utilizando redes baseadas em grafos para responder a seguinte questão de pesquisa: é possível prever múltiplas variáveis comportamentais com aprendizado multitarefa e verificar regiões cerebrais em comum considerando como entrada dados de rs-fMRI? Nossa abordagem demonstrou performance competitiva em relação com a literatura de abordagens *single*. Além disso, validamos que regiões do cérebro tiveram suas importâncias alteradas entre os modelos, sendo 53 regiões com mudança significativa na população pareada e escores submetidos ao teste de Wilcoxon na importância atribuída pelos modelos *single* e *multi*.

Palavras-Chave: aprendizado multitarefa, ressonância magnética funcional, redes cerebrais, abordagem preditiva.

A MULTI-TASK LEARNING APPROACH TO BRAIN-BEHAVIOR ANALYSIS USING FUNCTIONAL MAGNETIC RESONANCE IMAGING

ABSTRACT

Functional magnetic resonance imaging is used to verify the behavior of brain networks in cognitive disorders. To compose this assessment, behavioral data and standardized clinical tests can be associated. Linear algorithms are used in the literature to find a correlation between image data and a target behavioral variable, the final value of a cognitive test, for example. When there is more than one variable to be predicted, algorithms with multitasking learning can be used. Considering the premise that when tasks are related they can be predicted simultaneously, the present work aims to develop a multitasking approach using graph convolutional networks to answer the following research question: is it possible to predict multiple behavioral variables with multitasking learning and verify brain regions in common considering how to input data from rs-fMRI? Our approach demonstrated competitive performance in relation to the literature on *single* approaches. In addition, we validated which brain regions had their importance altered between the models, with 53 brain regions with a significant change in the matched population and scores submitted to the Wilcoxon test in the importance attributed by the *single* and *multi* models.

Keywords: Multitasking, fMRI, Neural Networks, Predictive Modeling.

LISTA DE FIGURAS

2.1	Ilustração de arquitetura utilizando o compartilhamento <i>hard</i> de parâmetros no aprendizado multitarefa em redes neurais profundas [113].	18
2.2	Ilustração de arquitetura utilizando compartilhamento <i>soft</i> de parâmetros no aprendizado multitarefa em redes neurais profundas [113].	19
2.3	Arquitetura clássica Lenet-5 para reconhecimento de dígitos [57].	19
2.4	Representação gráfica da operação de convolução (adaptada de [94])	20
2.5	Representação de grafo e matriz de adjacência por valores binários e correlação de pearson entre os nodos.	22
2.6	Etapas do protocolo CPM original [84]	24
4.1	Processo de identificação, triagem e elegibilidade dos estudos primários para revisão sistemática.	32
5.1	Macro etapas da metodologia de pesquisa.	37
5.2	Transformação das séries temporais em grafos ponderados para entrada nas redes baseadas em grafos.	41
5.3	Etapas da metodologia aplicada aos experimentos.	41
5.4	Representação das arquiteturas e suas diferenças para obtenção dos escores por ROI.	44
5.5	Exemplo gráfico do processo adotado para validar a diferença estatística dos escores por região entre arquiteturas MTL e <i>single</i>	46
6.1	Perda média por época entre os <i>folds</i> de validação cruzada.	49
6.2	Matriz de scores dos sujeitos de teste e escores atribuído às ROIs no modelo BrainGNN-MTL último <i>fold</i>	52
6.3	Distribuição dos escores atribuídos às ROIs pelos modelos <i>single</i> (laranja) e <i>multi</i> (azul) agrupadas pela significância no teste de Wilcoxon.	53
6.4	Superfície cortical esquerda ressaltando as redes que não apresentaram mudança significativa nos escores atribuídos pelos modelos. As regiões em cinza equivalem aquelas que não apresentaram mudança significativa no teste de Wilcoxon ($p > 0.05$). As destacadas em verde apresentaram mudanças significativas ($p < 0.05$).	54
B.1	Visualização da mediana dos escores das regiões que rejeitaram a hipótese nula pelo teste de Wilcoxon.	75

LISTA DE TABELAS

2.1	Exemplos de funções de ativação não lineares.	20
4.1	String de busca definida a partir da estratégia PICO para pesquisas não-clínicas.	30
4.2	Critérios de inclusão e exclusão.	31
4.3	Resumo dos estudos	33
5.1	Descrição do conjunto de dados	42
5.2	Configurações de hiperparâmetros utilizadas nos experimentos.	42
6.1	Arquitetura das camadas compartilhadas entre todas as tarefas no modelo BrainGNN- MTL.	47
6.2	Arquitetura das camadas compartilhadas entre todas as tarefas no modelo GCN-MTL.	48
6.3	Arquitetura das camadas específicas correspondente às saídas das tarefas	48
6.4	Métricas média e desvio padrão por modelo e suas variações.	50
6.5	Ranking dos 6 principais componentes com maior escore atribuído durante predição do modelo BrainGNN-MTL (colocação no ranking, escore acumulado).	56
6.6	Exemplos de componentes com diferença estatística significativa na população de teste entre os escores BrainGNN-MTL e <i>single</i> (colocação no ranking do modelo, escore acumulado).	57
A.1	Dados resultantes dos testes de Wilcoxon dos componentes que apresentaram mudanças superiores significativas ao modelo <i>single</i>	74

LISTA DE SIGLAS

AAL – *Automated anatomical labelling*
ABIDE – *Autism Brain Imaging Data Exchange*
ACM – *Association for Computing Machinery*
ADAS-COG – *Alzheimer's Disease Assessment Scale–Cognitive Subscale*
ADNI – *Alzheimer's Disease Neuroimaging Initiative*
AI – *Artificial Intelligence*
ASD – *Autism spectrum disorder*
BOLD – *Blood Oxygen Level Dependent*
CAM – *Class activation mapping*
CC – *Coeficiente de Correlação*
CIFTI – *Connectivity Informatics Technology Initiative*
CPM – *Connectome-based Predictive Modeling*
DMN – *Default Mode Network*
HCP – *Human Brain Project*
ICA – *Independent Component Analysis*
IEEE – *Institute of Electrical and Electronic Engineers*
fMRI – *Functional Magnetic Resonance Imaging*
GCN – *Graph Convolutional Neural Network*
GM – *Gray Matter*
ML – *Machine Learning*
MMSE – *Mini-Mental State Examination*
MRI – *Magnetic Resonance Imaging*
MTL – *Multi-task learning*
PMAT – *Penn Matrix Reasoning Test*
RAVLT – *Rey auditory-verbal learning test*
RMSE – *Root Mean Squared Error*
RS-fMRI – *Resting-state Functional Magnetic Resonance Imaging*
ROIS – *Regions of Interest*
SVM/R – *Support Vector Machine/Regression*
TRAILS – *Trails Making Test*
VBM – *Voxel-based morphometry*
WM – *White Matter*

SUMÁRIO

1	INTRODUÇÃO	13
1.1	CONTRIBUIÇÕES	14
1.2	OBJETIVOS	14
1.3	ORGANIZAÇÃO DO TEXTO	15
2	FUNDAMENTAÇÃO DE APRENDIZADO DE MÁQUINA	16
2.1	APRENDIZADO DE MÁQUINA	16
2.2	APRENDIZADO MULTITAREFA	17
2.2.1	COMPARTILHAMENTO <i>HARD</i> DE PARÂMETROS	18
2.2.2	COMPARTILHAMENTO <i>SOFT</i> DE PARÂMETROS	18
2.3	REDES NEURAIS CONVOLUCIONAIS	18
2.3.1	CAMADA DE CONVOLUÇÃO	19
2.3.2	CAMADA DE <i>POOLING</i>	20
2.3.3	CAMADA TOTALMENTE CONECTADA	21
2.4	REDES NEURAIS CONVOLUCIONAIS BASEADAS EM GRAFOS	21
2.5	MÉTODOS DE TREINAMENTO E MÉTRICAS	23
2.6	MODELOS TRADICIONAIS NA NEUROIMAGEM	23
2.6.1	CONNECTOME-BASED PREDICTIVE MODELING (CPM)	24
3	FUNDAMENTAÇÃO DE NEUROIMAGENS	26
3.1	IMAGEM POR RESSONÂNCIA MAGNÉTICA	26
3.1.1	RESSONÂNCIA MAGNÉTICA FUNCIONAL	26
3.2	REGIÕES CEREBRAIS DE INTERESSE	27
4	REVISÃO DA LITERATURA	29
4.1	QUESTÕES DE PESQUISA	29
4.2	ESTRATÉGIA DE BUSCA	30
4.3	CRITÉRIOS DE INCLUSÃO E EXCLUSÃO	30
4.4	EXTRAÇÃO DE DADOS	31
4.5	RESULTADOS PRINCIPAIS	31
4.6	CONCLUSÃO	36
5	METODOLOGIA	37

5.1	TRANSFORMAÇÃO DOS DADOS	38
5.1.1	CONJUNTO DE DADOS HCP	38
5.1.2	FORMULAÇÃO DAS TAREFAS	39
5.1.3	REPRESENTAÇÃO DE SÉRIE TEMPORAL ATRAVÉS DE GRAFO	40
5.2	MODELAGEM E TREINAMENTO	41
5.2.1	ARQUITETURAS DE BASE	42
5.2.2	ABORDAGEM PROPOSTA	43
5.3	ANÁLISE <i>POST-HOC</i> E VISUALIZAÇÕES	45
6	RESULTADOS	47
6.1	ARQUITETURAS EXPERIMENTAIS	47
6.2	TREINAMENTO E MÉTRICAS NO CONJUNTO DE TESTE	48
6.2.1	IMPORTÂNCIA DAS REGIÕES CEREBRAIS	51
6.2.2	ANÁLISE <i>POST-HOC</i>	52
7	CONCLUSÃO	58
7.1	LIMITAÇÕES	59
7.2	TRABALHOS FUTUROS	59
	REFERÊNCIAS BIBLIOGRÁFICAS	61
	APÊNDICE A – Resultados do teste de Wilcoxon	73
	APÊNDICE B – Superfícies cerebrais com escores (<i>stats map</i>)	75

1. INTRODUÇÃO

A partir de dados de neuroimagem é possível obter informações importantes que podem ajudar a compreender distúrbios cognitivos [96, 41]. A Ressonância magnética ou Imagem por ressonância magnética, do inglês *Magnetic Resonance Imaging* (MRI) é um dos principais métodos de avaliação da cognição humana para entender como as conexões cerebrais funcionam. A MRI pode ser realizada com objetivo de avaliar o cérebro estruturalmente ou funcionalmente, chamada ressonância magnética funcional (fMRI) [89]. A ressonância magnética estrutural visa investigar as diferentes formas e anatomias presentes em um órgão. Por outro lado, a ressonância magnética funcional permite mapear, espacialmente e temporalmente, diferentes redes de atividade cerebral. O exame pode ser realizado pelo método do estado de repouso, do inglês *Resting-state Functional Magnetic Resonance Imaging* (rs-fMRI), ou baseado em tarefas. No estado de repouso, o paciente é orientado a ficar imóvel e não pensar em algo específico, ou seja, livre de tarefas [33]. Embora não seja necessário que o usuário execute uma tarefa, vários estudos indicam a qualidade desse teste para avaliar o estado funcional do cérebro, incluindo distúrbios neuropsiquiátricos [53].

A imagem funcional é uma ferramenta importante para evidenciar alterações relacionadas a um contexto clínico em observação [43]. Para algumas desordens cerebrais, as alterações comportamentais têm sua causa relacionada a diferenças neurobiológicas, como modificações nas conexões cerebrais, que podem ser identificadas neste categoria de neuroimagem [76]. Distinguir e detectar essas mudanças em dados de rs-fMRI implica em utilizar diferentes métodos capazes de extrair padrões a partir das conexões e regiões ativadas, por exemplo. Esse dado pode ser utilizado como entrada de algoritmos computacionais que busquem padrões nos sinais cerebrais e que correlacionem ou não com a mudança comportamental em questão.

Outros dados são frequentemente associados às neuroimagens para compor estudos clínicos, como dados demográficos, histórico de doenças e testes clínicos padronizados [43]. Enquanto alguns algoritmos tentam prever diagnósticos como “saudável” e “não saudável”, estudos buscando prever escores clínicos a partir dos dados cerebrais têm mostrado relevância no estudo de biomarcadores associados a comportamento, como a predição de Quociente de Inteligência (QI) e regiões de ativação distintas entre gêneros, a partir da predição de um modelo. Geralmente, para relacionar neuroimagem e dados comportamentais utilizam-se algoritmos como *Support Vector Machines/Regression* (SVM/R) [67, 84], redes neurais, ou abordagens de regressão penalizada (e.g., *ridge*, *Least Absolute Shrinkage and Selection Operator* - LASSO) [19], entre outros. O que ambos tem em comum é o principal foco em classificação dicotômica ou regressão de um escore cognitivo específico, sob aprendizado de uma única tarefa por vez, denominado aprendizado *single-*

task. Contudo, estudos reforçam [90, 79] o potencial de modelos que possam considerar mais de uma tarefa em simultâneo.

Um método para avaliar múltiplas variáveis em modelos de aprendizado de máquina, do inglês *Machine Learning* (ML), é o aprendizado multitarefa, do inglês *multi-task learning* (MTL). O MTL tem mostrado grande potencial para fornecer *insights* sobre variações individuais em traços comportamentais e cognitivos [113, 100, 93]. O princípio do aprendizado multitarefa baseia-se em otimizar múltiplas tarefas em simultâneo dado que exista uma relação entre elas e compartilhando conhecimento. Associando a necessidade de observar múltiplas variáveis, a utilização de análise de padrões baseados em dados como *Machine Learning* podem apresentar bons resultados nesta categoria de análise [70]. Alguns autores [51, 112, 114, 61, 14] afirmam que a arquitetura de aprendizagem multitarefa, por ser otimizada para realizar todas as tarefas em simultâneo, pode favorecer a extração de características cerebrais relevantes para mais de uma tarefa.

1.1 Contribuições

Como contribuição principal, esta pesquisa desenvolve uma abordagem que se utiliza de múltiplas variáveis comportamentais a partir de imagens de ressonância magnética funcional em repouso (rs-fMRI) de modo a verificar marcadores neurobiológicos que possam estar relacionados. Além disso, este trabalho pode auxiliar em pesquisas futuras que venham a abranger outras arquiteturas multitarefas. Apresentamos também uma revisão da literatura de estudos que aplicaram a modelagem multitarefa em dados de neuroimagem.

1.2 Objetivos

O presente trabalho tem como objetivo principal aplicar uma abordagem de aprendizado multitarefa para predizer múltiplas variáveis dependentes que possa externar relações de marcadores biológicos. Além disso, estabeleceram-se os seguintes objetivos específicos para este trabalho:

- Revisar a literatura de aprendizado multitarefa no cenário de neuroimagens e comportamento;
- Avaliar o desempenho de arquiteturas multitarefa em dados públicos e restritos do *Human Connectome Project*;
- Realizar validação *post-hoc* entre as abordagens *single* e *multi*;

- Destacar marcadores biológicos e suas diferenças quando submetidos a diferentes arquiteturas *single* e *multi*.

1.3 Organização do Texto

Este trabalho apresenta no Capítulo 2 os principais conceitos teóricos para o contexto de aprendizado de máquina envolvido nesta pesquisa. O Capítulo 3 apresenta os conceitos utilizados relacionados à neuroimagem. Logo após, o Capítulo 4 sumariza a revisão da literatura em torno do tema e sua relevância na pesquisa. Além disso, o Capítulo 5 descreve a metodologia aplicada para o objetivo de pesquisa. Por fim, o Capítulo 6 descreve os resultados obtidos, bem como a conclusão e trabalhos futuros são apresentados no Capítulo 7.

2. FUNDAMENTAÇÃO DE APRENDIZADO DE MÁQUINA

2.1 Aprendizado de Máquina

Aprendizado de Máquina, do inglês *Machine learning* (ML) é um subcampo da Inteligência Artificial (IA) que permite simular a inteligência humana a partir do reconhecimento de padrões [66]. Usando funções matemáticas para representar modelos, ML pode ser aplicada em contextos com problemas complexos os quais não é possível representá-los com regras lógicas simples. Além disso, possui um viés multidisciplinar envolvendo desde a probabilidade, computação, passando pela neurociência e psicologia, por exemplo [81].

Os algoritmos de aprendizado de máquina são tipicamente divididos em dois campos com base na abordagem de aprendizado de modelo: *supervisionado* e *não supervisionado* [81]. A aprendizagem supervisionada consiste em usar dados rotulados como entrada de modelos de ML, ou seja, é possível medir erros e acertos tendo uma referência de rótulo correto. Por outro lado, na aprendizagem não supervisionada não há dados rotulados, sendo necessário tentar identificar padrões e tendências, sem ter uma referência de rótulo correto. Esse tipo de dado está mais disponível do que o dado rotulado, ao mesmo tempo que é mais complexo trabalhar com esse tipo de aprendizado. Embora os métodos supervisionados forneçam resultados mais assertivos, em algumas aplicações é difícil ou caro obter dados rotulados de alta qualidade [21].

Um conceito de aprendizado de máquina proposto por Mitchell [66] é definido como: “Diz-se que um programa de computador aprende com a experiência **E** no que diz respeito a alguma classe de tarefas **T** e medida de desempenho **P**, se seu desempenho nas tarefas em **T**, conforme medido por **P**, melhora com a experiência **E**”. Nesse caso, a experiência **E** pode ser o conjunto de dados com histórico ou dados relacionados para melhorar o aprendizado. A tarefa **T** pode ser compreendida não como o processo de aprendizado em si, mas sim como o algoritmo vai processar as informações para poder aprender, dependendo do problema em questão. Conforme Goodfellow et al. [34], se quisermos que um robô seja capaz de andar, então caminhar é a tarefa em análise. Nesse sentido, a forma de aprendizado também pode ser definida quanto a tarefa a ser realizada, podendo ser uma única tarefa por vez, do inglês *single-task learning* (em alguns momentos deste trabalho simplesmente *single*) ou aprendizado com múltiplas tarefas em simultâneo, do inglês *multi-task learning*, em alguns momentos deste trabalho simplesmente *multi*.

2.2 Aprendizado Multitarefa

Aprendizado multitarefas, do inglês *Multitask Learning* (MTL), é uma técnica a qual o objetivo é otimizar mais de uma tarefa em simultâneo [15, 114, 112]. De forma geral, quando as tarefas possuem características em comum é possível utilizar essa técnica associando a uma tarefa principal, tarefas auxiliares. Estas tarefas auxiliares ajudam o modelo a convergir ao passo em que possuem informações úteis compartilhadas [113, 15]. O paradoxo de Stein é considerado como uma teoria que antecedeu ao aprendizado multitarefa [28]. Nesse paradoxo, três ou mais parâmetros quando foram estimados simultaneamente obtiveram melhores resultados que individualmente. Logo, por ser uma teoria estatística de decisão é assemelhada ao processo multitarefa, embora difiram no ponto em que para o paradoxo, as variáveis precisam seguir uma distribuição normal obrigatoriamente. Outra premissa que deve ser observada é que as tarefas precisam estar relacionadas, caso contrário, haverá uma influência negativa entre elas. Para esse caso, aprender as tarefas separadamente apresentarão melhor resultado [15, 113].

A definição formal para o aprendizado multitarefas pode ser tomado dado m tarefas $\{T_i\}_{i=1}^m$ onde as tarefas estão relacionadas, busca-se otimizar o modelo F utilizando o conhecimento adquirido nas m tarefas [114]. Uma tarefa T_i tem sua função de aprendizado definida como na Equação 2.1 [112], onde $W = (W_1, \dots, W_m)$ é o coeficiente de x e $b = (b_1, \dots, b_m)^T$ bias. Assim como no aprendizado *single-task*, o objetivo do modelo é minimizar o erro para se ajustar a essa função. Para isso, uma função de perda \mathcal{L} é utilizada para minimizá-la o máximo possível. Na multitarefa, cada tarefa tem sua função de perda \mathcal{L}_i , e para um resultado unificado elas podem ser associadas de diferentes formas para obter o melhor ajuste do modelo [75]. O método mais comum busca-se minimizar não cada perda individualmente, mas sim o somatório das perdas entre as m tarefas conforme a Equação 2.2, onde θ são os parâmetros do modelo e \mathcal{D}_i o *dataset* de treinamento.

$$f_i(x) = W_i^T x + b_i \quad (2.1)$$

$$\min_{\theta} \sum_{i=1}^T W_i \mathcal{L}_i(\theta, \mathcal{D}_i) \quad (2.2)$$

Ambas as formas de aprendizado podem ser aplicadas em diferentes categorias de algoritmos. Neste contexto, abordaremos Redes Neurais Convolucionais e Redes Neurais Convolucionais baseadas em grafos. A adequação do aprendizado multitarefa poderá se dar alterando, principalmente, as camadas de saída, ou também camadas intermediárias que possam compartilhar parâmetros entre as tarefas, podendo ser compartilhamento dito *hard* ou *soft* de parâmetros [113, 93]. Vale ressaltar que existem outras formas avança-

das de forçar o compartilhamento de parâmetros entre tarefas [114, 63] que não serão alvo deste estudo.

2.2.1 Compartilhamento *hard* de Parâmetros

O modelo de arquitetura de compartilhamento de parâmetros *hard* é o mais comum aplicado em MTL. Como exemplificado na Figura 2.1, camadas são compartilhadas para todas as tarefas seguindo de camadas específicas para cada tarefa. O benefício desta arquitetura é a redução de *overfitting*, e o risco diminui inversamente ao número de tarefa [113]. Essa abordagem é mais utilizada pela sua efetividade no ganho entre as tarefas e por ser mais eficiente computacionalmente [111].

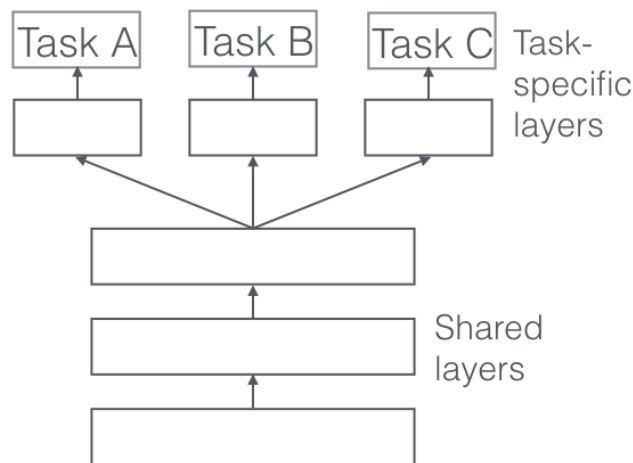


Figura 2.1: Ilustração de arquitetura utilizando o compartilhamento *hard* de parâmetros no aprendizado multitarefa em redes neurais profundas [113].

2.2.2 Compartilhamento *soft* de Parâmetros

No compartilhamento de parâmetros *soft*, cada tarefa possui seus próprios parâmetros reunidos em um modelo para cada tarefa, como exemplifica a Figura 2.2. O relacionamento entre as tarefas é feito pela regularização aplicada entre os modelos para tornar os parâmetros ajustados pela semelhança entre as tarefas aprendidas separadamente.

2.3 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (RNCs) permitem a criação de redes com diversas camadas de funções não-lineares sensíveis a detalhes [34]. As RNCs são bem

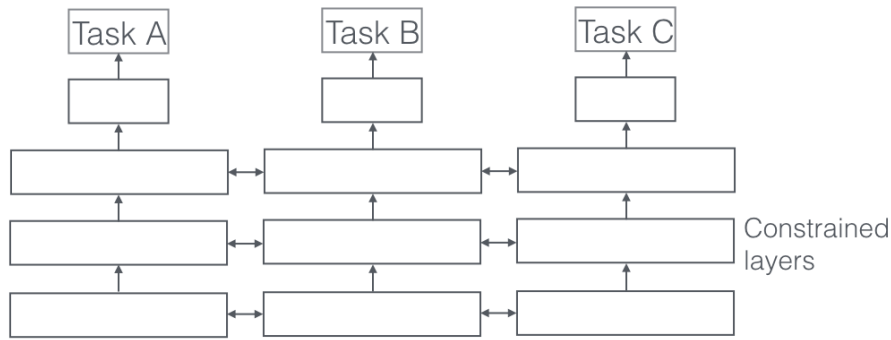


Figura 2.2: Ilustração de arquitetura utilizando compartilhamento *soft* de parâmetros no aprendizado multitarefa em redes neurais profundas [113].

aplicadas a dados que apresentam duas ou mais dimensões como as imagens. Neste tipo de técnica, não é necessário ensinar ao algoritmo quais características na imagem correspondem a determinada classe. Há técnicas que identificam características na imagem de forma a gerar um vetor de atributos [56]. Uma arquitetura clássica de RNC é a LeNet criada por [57] para reconhecimento de dígitos conforme apresentado na Figura 2.3.

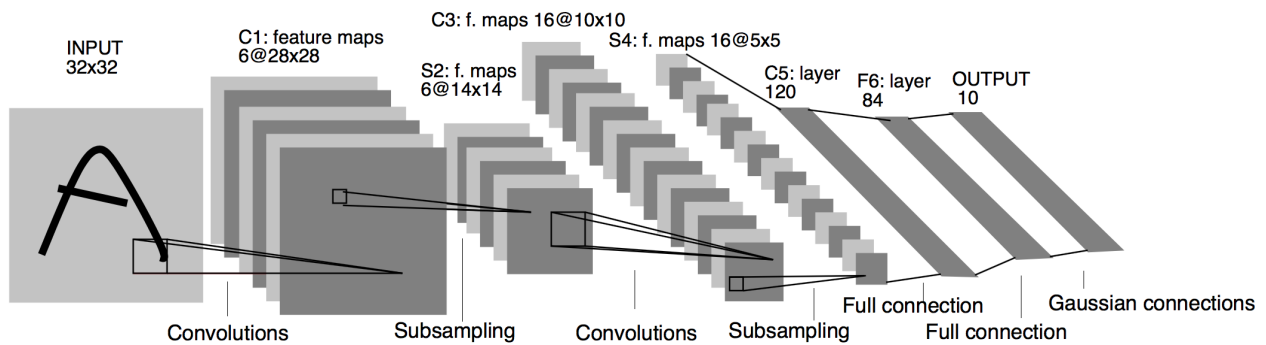


Figura 2.3: Arquitetura clássica LeNet-5 para reconhecimento de dígitos [57].

Uma rede neural convolucional é composta principalmente por 3 tipos de camadas: camada de convolução, camada de *pooling* e camada *fully-connected*. A seguir, serão descritas como se constituem as referidas camadas.

2.3.1 Camada de Convolução

A camada de convolução tem como parâmetros, N filtros de tamanho $m \times n$ que vão ser convoluídos pelo dado de entrada. A convolução pode ser matematicamente descrita como equação $s = x * w$, sendo x a entrada (uma imagem, por exemplo) e w o filtro ou *kernel*. O símbolo $*$ representa a operação de convolução entre *arrays* e não de multiplicação escalar. Por essas características que este modelo é adequado para dados como imagens, as quais são *arrays* multidimensionais cujo os valores representam pixels. O resultado desta operação de convolução produz uma matriz denominada *feature map*

representada na equação por s que possui, por exemplo, as bordas detectadas. Uma representação gráfica da operação é apresentada na Figura 2.4. Em imagens coloridas, w tem a mesma profundidade da imagem de entrada, no entanto, o *feature map* resulta do somatório das convoluções de cada canal da imagem.

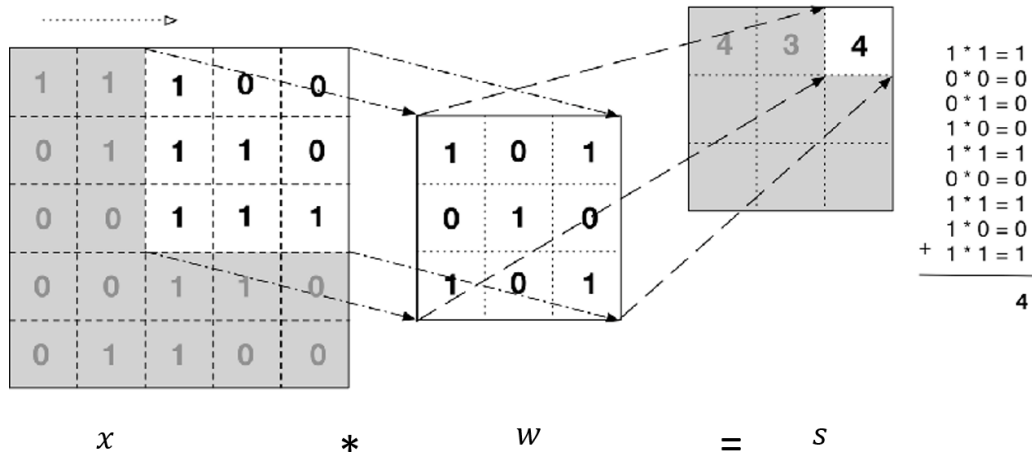


Figura 2.4: Representação gráfica da operação de convolução (adaptada de [94])

Além disso, em cada neurônio é utilizada uma função de ativação, responsável por inserir a não-linearidade às redes convolucionais, mapeando as saídas dos neurônios para um conjunto de valores o qual função está definida. A função de ativação mais utilizada é a *Rectified Linear Unit* (ReLU), que insere a não linearidade aos neurônios retornando 0 para $x \leq 0$ e x para $x > 0$. Também há outras funções conforme descritas na Tabela 2.1 como *Sigmoid* e *Tanh* (Tangente hiperbólica).

Tabela 2.1: Exemplos de funções de ativação não lineares.

Nome	Função	Representação
Sigmoid	$\sigma(x) = \frac{1}{1+e^{-x}}$	
tanh	$\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	
ReLU	$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0. \end{cases}$	

2.3.2 Camada de *Pooling*

A camada de *Pooling* ou *subsampling* tem como objetivo a redução de dimensionalidade do *feature map* gerado pela camada anterior unindo características semelhantes

em uma só [56]. As funções *depooling* mais utilizadas em dados de imagem são *Maxpooling* e *Average pooling* obtendo, respectivamente, o maior valor ou a média de valores uma determinada região. Ao passo em que a camada de *pooling* tem a propriedade de reduzir a escala do dado de entrada, é possível utilizar esta etapa para obter as regiões mais relevantes para o aprendizado do modelo.

2.3.3 Camada Totalmente Conectada

A camada totalmente conectada, do inglês *fully-connected* ou *dense layers* é responsável pela classificação após o processo de extração de características realizado pelas camadas anteriores. Como entrada, a camada deve receber um *array* unidimensional, sendo necessária uma etapa de *flattening* para transformar os dados multidimensionais da camada anterior. Nesta camada também deve-se aplicar uma função de ativação para realizar a classificação dos dados por meio de uma função não linear, geralmente a função *softmax* a qual gera valores entre 0 e 1, podendo ser utilizada para multi-classes gerando valores decimais neste intervalo.

2.4 Redes Neurais Convolucionais baseadas em Grafos

Em 1877, pela primeira vez, o termo grafo foi utilizado em um artigo publicado na Nature por James Joseph Sylvester [3]. Entretanto, a primeira utilização da teoria dos grafos é classicamente associada ao problema das pontes de Koningsberg, em 1736 por Leonhard Euler. Apesar disso, a primeira definição formal só surgiu no século XX e tem sido utilizado em estudos desde então. Um grafo G pode ser definido como um conjunto de vértices ou nodos η e arestas ε , denotado por $G = (\eta, \varepsilon)$. A quantidade de nodos η é o que determina a ordem do grafo, e o número de arestas ε o tamanho do grafo.

Um grafo pode ser representado por uma matriz de adjacência A de tamanho $n \times n$ tal que os elementos da matriz $A_{i,j}$ podem ser valores binários representando ou não a ligação entre os nodos η_i e η_j , mas também podem conter um valor que representa o peso entre η_i e η_j . Esse peso pode ser dado por $W_{i,j}$, tal que W possa representar um valor de correlação de Pearson ou a distância entre dois pontos ou número de nós vizinhos (*K-nearest neighbor* - KNN), por exemplo [50].

Os princípios utilizados em dados pertencentes ao domínio euclidiano que podem ser representados por eixos X e Y, como as imagens, pode ser transposto para grafos, do domínio não euclidiano, fazendo algumas analogias. Uma imagem bidimensional pode ser interpretada como um grafo se pensarmos que cada nó representa um valor do pixel da imagem, enquanto as arestas significam a distância euclidiana entre cada pixel. Assim,

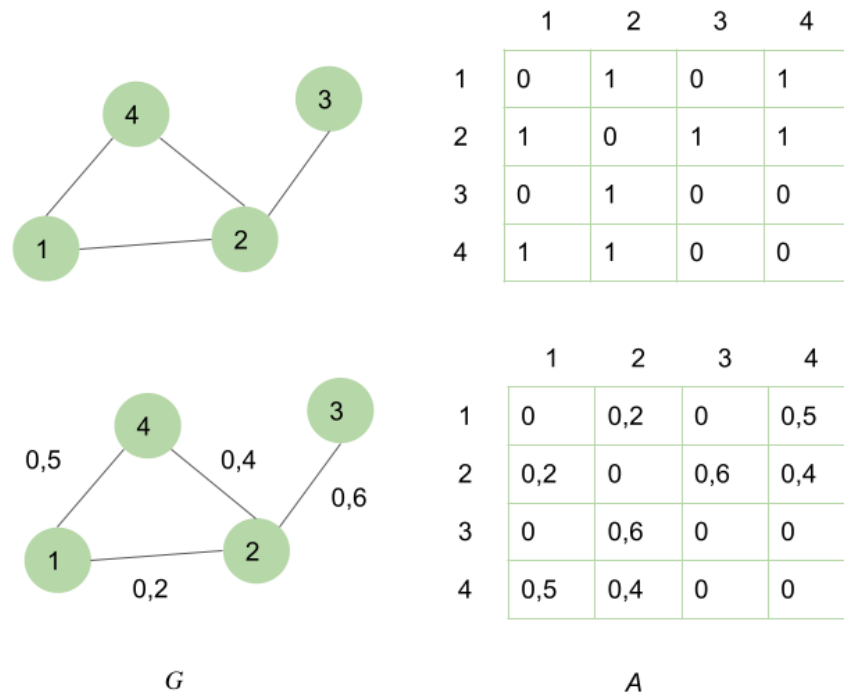


Figura 2.5: Representação de grafo e matriz de adjacência por valores binários e correlação de pearson entre os nodos.

quanto maior o valor da aresta, mais próximos estarão os pixels interconectados. Do mesmo modo, um grafo pode ser representado em uma matriz de adjacência para relacionar os nós, como um grafo que apresenta 6 nós, é representado por uma matriz quadrada de adjacência de dimensão 6. Entretanto, um grafo é uma forma não estruturada tornando a operação de convolução não tão direta como a aplicada em imagens, por exemplo. Dessa forma, as redes neurais convolucionais baseada em grafos, do inglês *Graph Convolutional Neural Networks* (GCNs), seguem duas abordagens convolucionais: Redes Convolucionais de Grafos Espaciais e Redes Convolucionais de Grafos Espectrais [3].

Nas Redes Convolucionais de Grafos Espectrais a propagação do sinal ocorre ao longo dos nós. Os GCNs espectrais recorrem à decomposição espectral da matriz Laplaciana de grafo para implementar este método de propagação de informação. Além disso, a convolução nesse domínio é definida por meio da transformada de Fourier do gráfico. Já no modelo Espacial, os nodos se relacionam a partir da vizinhança. A tarefa de classificação no contexto dos grafos também pode ser observada ao nível de nodos ou ao nível de grafo. No nosso contexto, a previsão ao nível de grafo faz sentido já que a partir de toda a estrutura das conexões se quer inferir uma ou mais classes específicas.

2.5 Métodos de treinamento e métricas

O método de treinamento mais utilizado em estudos de neuroimagem é validação cruzada *k-fold* [72, 96]. O processo se dá pela divisão dos dados em K grupos, treinando e validando o modelo K vezes. Por fim, calcula-se as métricas médias entre os K modelos. As métricas utilizadas dependem do tipo de tarefa que estamos realizando, regressão ou classificação. Para classificação, consideremos acurácia e área sobre a curva. Para regressão, utilizaremos coeficiente de determinação (R^2) conforme a Equação 2.3 e *root mean squared error* (rMSE) seguindo a Equação 2.4, tal que y representa a variável a ser predita enquanto x representam as features que poderão ajudar a prever o valor de y . Assim, dado n o total de observações ou sujeitos, x_i e y_i correspondem a cada i observação até n , enquanto \bar{x} e \bar{y} equivalem às médias para todas as observações.

$$R^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right)\left(\sum_{i=1}^n (y_i - \bar{y})^2\right)}} \quad (2.3)$$

$$rMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - x_i)^2} \quad (2.4)$$

2.6 Modelos tradicionais na Neuroimagem

O termo “modelagem preditiva” (do inglês “*predictive modeling*”) tem sido amplamente utilizado em estudos com abordagens preditivas na área da saúde, e por vezes com objetivo de não apenas prever, como a interpretação literal do método, mas compreender características que levaram a esta predição [78, 68]. Geralmente a aplicação da predição tende a permear uma abordagem descritiva, dependendo da pergunta de pesquisa. Por exemplo, Pereira et al. [73] abordam “distinção de padrões” (do inglês “*pattern discrimination*”), “localização de padrões” (do inglês “*pattern localization*”) e “caracterização de padrões” (do inglês “*pattern characterization*”) como principais utilizações de abordagens preditivas.

Revisões da literatura [72, 96, 67] resumizam algoritmos lineares como *Support Vector Machines/Regression* (SVM/R), os considerando principais abordagens utilizadas em neuroimagem. Uma abordagem que aplica a linearidade diretamente ao uso de conectomas e comportamento têm apresentado relevância nos estudos da área chamado mode-

lagem preditiva baseada em conectoma, do inglês *Connectome-based predictive modeling* (CPM).

2.6.1 Connectome-based predictive modeling (CPM)

Geralmente, para relacionar dados de neuroimagem e comportamentais são usados modelos de correlação ou regressão [71, 10]. Esses métodos têm mostrado boas métricas de previsão, mas tendem a sobreajustar e falhar na generalização. No entanto, uma nova abordagem tem sido utilizada descrita por Shen et al. [84] chamado modelagem preditiva baseada em conectoma (do inglês *Connectome-based predictive modeling* - CPM). Essa abordagem consiste em um protocolo para modelos preditivos que, com base na matriz de correlação obtida pelos dados de rs-fMRI, podem se relacionar com variáveis comportamentais conforme apresentado na Figura 2.6.

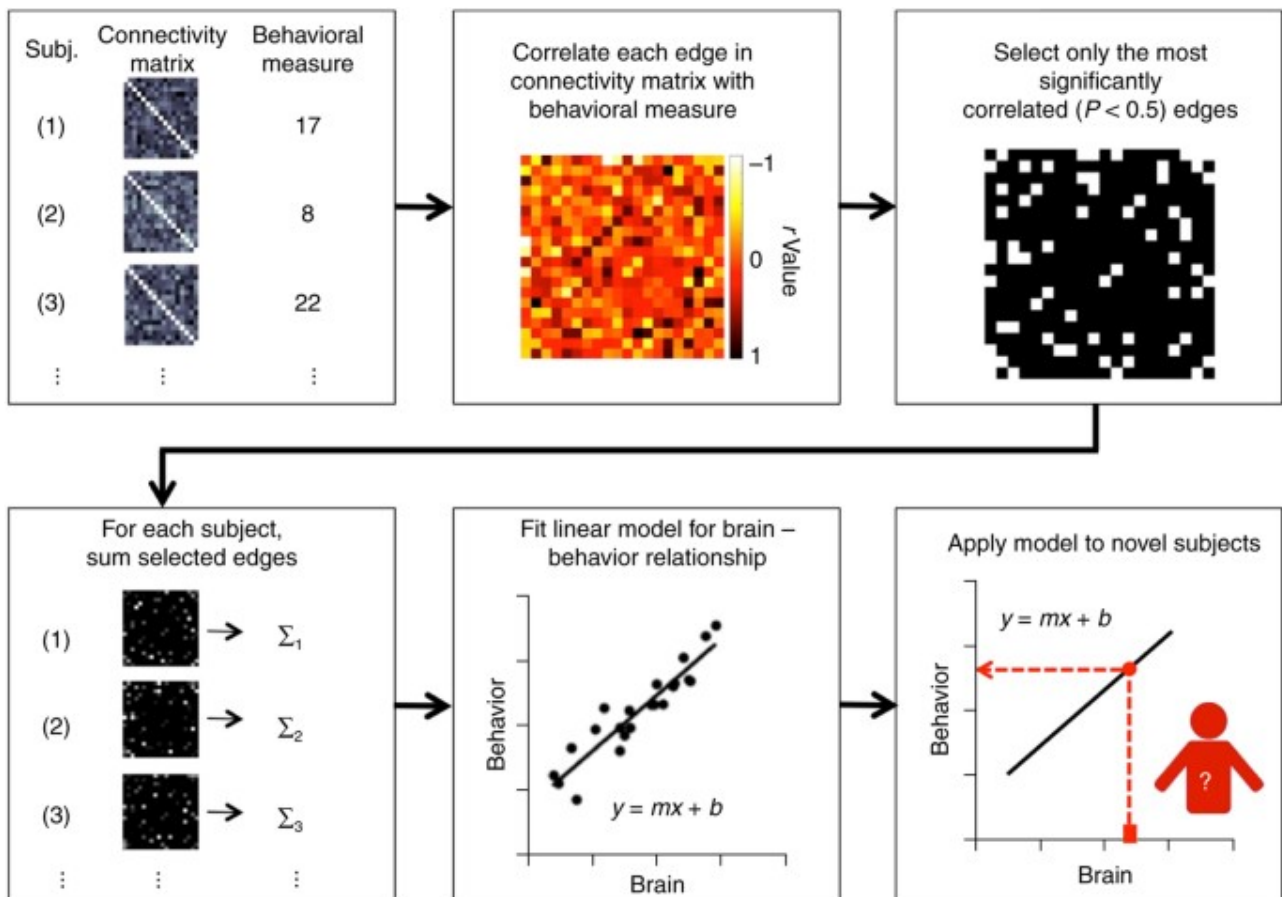


Figura 2.6: Etapas do protocolo CPM original [84]

Este protocolo inclui 4 etapas: (i) seleção de recursos, (ii) resumo de recursos, (iii) construção de modelos e (iv) avaliação da significância da previsão [84]. Em resumo, a relevância de cada conexão entre o valor da variável comportamental de interesse é calculada, mantendo apenas as relações mais significativas a partir de um limite. Depois disso,

os valores das conexões são somados. Por fim, essa soma e as variáveis comportamentais servem de entrada para um modelo de regressão linear usando validação cruzada.

Este método apresenta bons resultados na literatura em análises de personalidade individual [31, 80, 8], sem delimitar regiões de interesse específicas, como pelo método *whole-brain*. Esse aspecto é uma vantagem do modelo, uma vez que não é necessário selecionar regiões de interesse, ele é capaz de lidar com todo o cérebro, aumentando as possibilidades de observar novas áreas relacionadas à variável comportamental [42].

3. FUNDAMENTAÇÃO DE NEUROIMAGENS

O estudo de neuroimagem abrange um campo interdisciplinar que visa compreender o cérebro por meio de técnicas de diagnóstico por imagem como a Ressonância Magnética, um método capaz de produzir dados estruturais e funcionais do cérebro [33]. A seguir, essas duas técnicas são detalhadas.

3.1 Imagem por Ressonância Magnética

A Ressonância Magnética é considerada um exame não invasivo por não utilizar radiação, como a radiografia ou a tomografia computadorizada [89]. O método consiste em aplicar ondas de rádio para ordenar os átomos de hidrogênio que constituem naturalmente o corpo humano. Esses átomos, ao retornarem em sua ordem normal, emitem energia e o scanner gera uma imagem a partir disso. O processo de ressonância magnética consiste em fazer múltiplas fatias do corpo criando matrizes, o que cada posição na matriz é chamada voxel [89].

O foco desse método é diferenciar tecidos e suas modificações, como tecido doente e tecido normal [108]. Logo, visam investigar as diferentes formas e anatomias presentes em um órgão, quando se trata de uma aplicação médica. As diferentes sequências aproveitam os variados tempos de relaxamento dos tecidos para destacar algumas características específicas como a presença ou ausência de líquidos, lesões. Imagens ponderadas em tempo de relaxamento T1, por exemplo, ressaltam tecido adiposo, enquanto as ponderadas em T2 tem maior sinal para regiões com água, permitindo visualizar como edemas, infecções e tumores [16].

Outro tipo de ressonância magnética é a imagem por ressonância magnética funcional (fMRI), usada para analisar aspectos metabólicos dentro das estruturas anatômicas [33]. Usando os mesmos princípios da ressonância magnética, a ressonância magnética também considera pequenas alterações no fluxo sanguíneo.

3.1.1 Ressonância Magnética Funcional

A Ressonância Magnética Funcional (fMRI) é adquirida devido ao mapeamento do sinal dependente do nível de oxigênio no sangue ou sinal BOLD, do inglês *Blood-oxygen-level-dependent*. O sinal BOLD é ativado quando há aumento da concentração de desoxihemoglobina por alguma atividade específica ou modulação espontânea do metabolismo neuronal [33]. Enquanto neuroimagens estruturais podem ser usadas para avaliar aspectos

volumétricos e morfológicos do cérebro, neuroimagens funcionais permitem mapear, espacialmente e temporalmente, diferentes redes de atividade cerebral relacionadas à visão, audição ou apenas ativação cerebral padrão, do inglês *Default Mode Network* (DMN). O exame funcional pode ser realizado pelo método do estado de repouso e baseado em tarefas. No estado de repouso, o paciente é orientado a ficar quieto e não pensar em algo específico, ou seja, tarefas livres. Embora não seja necessário para o usuário para realizar uma tarefa, estudos indicam a qualidade desse teste para avaliar o estado funcional do cérebro, incluindo distúrbios neuropsiquiátricos [53, 32].

As conexões obtidas a partir das interações cerebrais formam uma unidade chamada conectoma. Este sistema de interação pode ser usado para compreender e investigar redes cerebrais. Por correlação, as áreas de interação podem ser mapeadas como um grafo representando as redes neurais funcionais [100].

3.2 Regiões Cerebrais de Interesse

A conectividade funcional pode ser descrita por uma matriz C , tal que cada elemento da matriz $C(i, j)$ corresponde a uma medida entre os nodos i e j . No caso das imagens funcionais, essa medida equivale ao sinal BOLD entre os nodos, descrevendo a conectividade funcional entre eles. A medida mais comum para obter a conectividade entre os nodos é pelo coeficiente de correlação de Pearson. As matrizes de correlação apresentam tanto conexões diretas quanto indiretas entre regiões do cérebro. Para visualizar apenas as interações diretas, autores sugerem que seja utilizado o cálculo da covariância esparsa inversa [99, 87].

A definição de quais nodos i e j estão sendo observados contempla o conceito de região de interesse (ROI, do inglês *Region Of Interest*). As regiões de interesse podem ser obtidas a partir do parcelamento das séries temporais dos voxels a partir de atlas anatómicos a partir da literatura, como o *Automated anatomical labeling* (AAL) [95] ou *Harvard-Oxford probabilistic atlas* [86]. Porém, estes métodos podem não sobrepor corretamente as regiões anatómicas e funcionais [82, 88]. Como alternativa, é possível utilizar análise de componentes independentes (ICA, do inglês *Independent Component Analysis*) [47, 20] em dados de rs-fMRI tanto para nível individual quanto em nível de grupo. Vários estudos utilizaram este parcelamento para medir a conectividade funcional em contextos clínicos como doença de Alzheimer [36], Esquizofrenia [46] e depressão [115].

A definição do parcelamento que será utilizado é uma etapa importante em estudos de análise cerebral e não há um consenso de qual é o melhor parcelamento [1, 82]. Neste estudo, utilizaremos o parcelamento espacial ICA a partir de arquivos de conectividade CIFTI, do inglês *Connectivity Informatics Technology Initiative*, contendo as respectivas localizações das regiões a partir de *brainordinates* [22]. A utilização dessa abordagem

de parcelamento irá auxiliar na visualização das regiões ressaltadas no Capítulo 6 de resultados.

Brainordinate é o nome dado à localização de um nodo ou voxel, composta por *grayordinates*, correspondente aos voxels da região cinzenta do cérebro, e *whiteordinates*, correspondente a região branca do cérebro. Essas coordenadas são armazenadas em arquivos do tipo CIFTI, o qual contem uma lista de *brainordinates* para visualizações de volume ou superfície [22].

4. REVISÃO DA LITERATURA

Estudos aplicando *machine learning* e aprendizado *single-task* já apresentam resultados consolidados, conforme revisões da literatura já publicadas [45, 96, 27, 54]. Entretanto, estudos utilizando aprendizado multitarefa têm sido explorados [103, 104, 65] de diferentes formas e não foi encontrada um estudo secundário que pudesse trazer o panorama de *multi-task learning* em neuroimagens. O mais próximo é o estudo de Sui et al. [90] que destaca na sua revisão que houve artigos com abordagem multitarefa, mas que não seriam relatados no estudo por representar menor quantidade. Dessa forma, uma etapa desta pesquisa consistiu em uma revisão sistemática da literatura, a qual será resumida nesta seção para os aspectos relevantes que contextualizem a presente dissertação.

Revisão sistemática é um estudo secundário com o objetivo de identificar os principais estudos e materiais de um tópico específico, sendo estes denominados estudos primários [52]. A presente revisão constitui seu método a partir de protocolo estabelecido nas seguintes etapas baseado em Kitchenham [52] e Dybå [23]: (1) Questões de pesquisa, (2) Estratégia de busca, (3) Critérios de inclusão e exclusão, (4) Definição e aplicação dos critérios de qualidade, (5) Extração de dados e (6) Análise e síntese dos dados, sendo esta última apresentada na Seção 4.5 (Resultados Principais).

4.1 Questões de Pesquisa

As questões de pesquisa que nortearam a revisão foram definidas considerando os aspectos técnicos referentes aos métodos de aprendizado de máquina aplicados e como foram aplicados no contexto das neuroimagens e distúrbios neurológicos. As questões de pesquisas selecionadas são descritas a seguir:

1. QP1: Como o aprendizado multitarefa pode ser utilizado em neuroimagem e distúrbios neurológicos?
2. QP2: É possível perceber ganhos quando comparado o aprendizado *single-task* e *multi-task*?
3. QP3: Como podem ser interpretados os achados clínicos e regiões cerebrais consideradas em conjunto no aprendizado multitarefa?

4.2 Estratégia de Busca

O planejamento da estratégia de busca consistiu em definir as bases Pubmed, IEEE, ACM e Scopus buscando melhor cobertura dos artigos. Além disso, definiu-se como *string* de busca a partir da estratégia PICO conforme a Tabela 4.1.

Tabela 4.1: String de busca definida a partir da estratégia PICO para pesquisas não-clínicas.

String PICO	População AND Interesse AND Contexto
População	(“functional magnetic resonance imaging” OR “Neuroimaging” OR “connectome” OR “fMRI” OR “MRI” OR “magnetic resonance imaging”)
Interesse	(“multi-task learning” OR “multitask learning” OR “multi-task” OR “multitask”) (“machine learning” OR “graph neural network” OR “graph convolutional neural network” OR “geometric Deep Learning” OR “convolutional neural network”)
Contexto	AND NOT (“EEG” AND “MEG” AND “multitask fMRI”)

No contexto de comportamento cerebral, a palavra multitarefa tem um significado semelhante ao aprendizado multitarefa no contexto computacional, visto que um pode explicar o outro de certa forma. Entretanto, ao realizar busca na literatura sobre artigos com foco em modelos de aprendizado de máquina multitarefa foi possível encontrar também estudos que abordavam o aspecto multitarefa do ser humano, como estudos analisando indivíduos realizando duas tarefas ao mesmo tempo. Assim, foi necessário filtrar os estudos o qual a multitarefa se trata sobre a abordagem do ML e não sobre os dados cerebrais terem sido observados no indivíduo realizando múltiplas tarefas. Para evitar isso, na string utilizada adicionamos ao escopo “NOT” a expressão “multitask fMRI” a qual se refere a dados os quais o usuário realiza multitarefa e não o algoritmo propriamente dito. Optou-se por selecionar apenas artigos completos, dos últimos 5 anos a partir de Outubro/2021.

4.3 Critérios de Inclusão e Exclusão

Após a exclusão dos artigos duplicados, os critérios de inclusão e exclusão apresentados na Tabela 4.2 foram aplicados no resumo de cada artigo selecionado. Em casos

de dúvida, foi realizada a leitura do texto completo do artigo. O critério de exclusão é eliminatório, ou seja, bastava apenas um critério de exclusão associado ser considerado satisfatório para que o artigo não fosse selecionado para a próxima etapa.

Tabela 4.2: Critérios de inclusão e exclusão.

Critérios de Inclusão	Critérios de Exclusão
1. O estudo aplica algoritmos de aprendizado de máquina.	1. Artigo apresenta revisões sistemáticas e mapeamentos sistemáticos.
2. O artigo considera um atlas conhecido e não probabilístico.	2. Artigo não aborda o uso de redes neurais convolucionais.
3. O artigo utiliza mais de uma tarefa ao mesmo tempo, ou seja, aplica aprendizado multitarefa.	3. O artigo não realiza aprendizado multitarefa.
4. O artigo está disponível na língua inglesa e publicado nos últimos 5 anos.	4. Utiliza como entrada outros tipos de imagens senão de ressonância magnética, como EEG, DTI, DWI ou MEG.

Foram excluídos, por exemplo, artigos que utilizavam outros tipos de neuroimagem ou exames como Eletroencefalograma (EEG), Magnetoencefalografia (MEG), imagens funcionais baseadas em difusão (DWI) e imagem de tensor de difusão (DTI). Artigos que tinham por objetivo proposição de novos atlas (parcelamento do cérebro) também foram excluídos.

4.4 Extração de Dados

A extração dos dados que nortearam a discussão e os resultados foi feita utilizando os seguintes itens de cada estudo: doença ou contexto clínico abordado, *datasets*, Atlas utilizado, tipo de neuroimagem, dados comportamentais e scores clínicos. Esta extração será resumida a título informativo na próxima seção.

4.5 Resultados principais

Como resultado da busca, na base IEEE obtiveram-se 28 artigos, 64 no Pubmed, 144 no Scopus e 85 na ACM. A Figura 4.1 apresenta a quantidade de artigos encontrada inicialmente em casa base e as quantidades obtidas a partir dos refinamentos utilizados. Após a seleção dos artigos a partir da string de busca e filtragem por últimos 5 anos, foi necessário remover os duplicados. Essa etapa foi realizada com o auxílio da ferramenta Rayyan.ai. A remoção dos duplicados diminui o total de estudos para 235 estudos submetidos aos critérios de exclusão e inclusão a partir de seus resumos. Ao final, resultou em

47 artigos para serem lidos de forma completa e confirmar se apresentavam informações relevantes para responder às perguntas de pesquisa desta revisão. Por conseguinte, 16 artigos foram selecionados para a etapa final de extração e resumo dos dados. Uma visão geral sobre os artigos selecionados é sumarizado na Tabela 4.3.

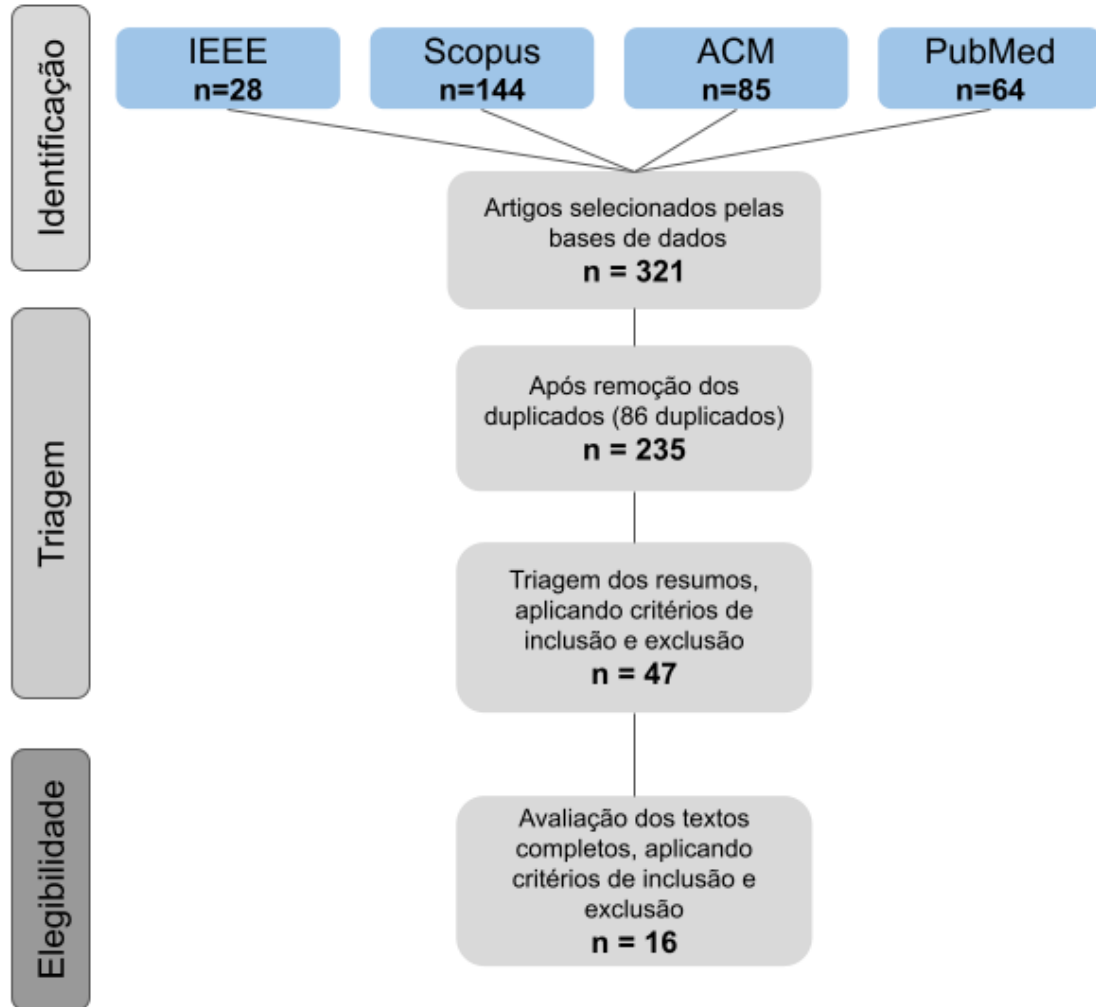


Figura 4.1: Processo de identificação, triagem e elegibilidade dos estudos primários para revisão sistemática.

Tabela 4.3: Resumo dos estudos

Estudo	Doença	Datasets	Alias	Imagem	Dados comportamentais	Scores
[45]	Alzheimer	ADNI	AAL	MRI	Idade, categorias (NC, MCI, AD)	ADAS-Cog
[109]	Schizophrenia	privado (6 hospitais da China)	264 MNI	MRI, fMRI	Idade, gênero	Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID-I/P), The Positive and Negative Syndrome Scale (PANSS)
[14]	Alzheimer	ADNI	-	MRI	-	ADAS, MMSE, RAVLT-TOTAL, RAVLT-T30, RAVLT-RECOG
[62]	Alzheimer	ADNI-1, ADNI-2 [39], MIRIAD, AIBL	ALL	MRI	Idade, gênero, educação	CDRSB, ADAS11, ADAS13, MMSE
[92]	Glioblastoma	privado(120 pacientes)	-	MRI	Idade, gênero, tamanho do tumor e localização do tumor	-
[102]	Autismo	ABIDE	ALL	rs-fMRI,	idade, genero	-
[97]	Alzheimer	Mayo Clinic Study of Aging (MCSA)	Mayo Clinic Adult Lifespan Template (MCALT)	MRI	Idade, gênero, dados cardiovasculares e metabólicos	MMSE
[39]	Autismo	ABIDE-I	-	fMRI	-	Bayley Scales of Infant and Toddler Development III
[103]	Alzheimer	ADNI	93 manualmente anotadas	MRI	-	MMSE e ADAS-Cog
[65]	Autismo	ABIDE-II e ADHD-200	AAL3	MRI	Idade, gênero, classificação da saúde mental	-
[104]	-	HCP	Yeo parcellation	rs-fMRI	Idade, gênero, cognição, uso de substâncias ilícitas, uso de tabaco, dado emocional, saúde mental	-
[61]	Dementia/Alzheimer	ADNI-1 e ADNI-2	ALL	MRI	idade, genero, educação	CDRSB, ADAS-Cog, MMSE
[91]	Alzheimer	ADNI	Desikan-Killiany cortical	MRI	-	ADAS-cog, Memory from Wechsler Memory Scale IV, Neurobattery scores, Neuropsychiatric Inventory
[102]	Alzheimer	ADNI	-	MRI	Idade, gênero, educação, genótipo ApoE	ADAS, MMSE, RAVLT, FLU, TRAILS
[38]	Alzheimer	ADNI	-	MRI	-	ADAS, MMSE
[59]	Autismo	ABIDE	ALL	MRI	-	Autism Measurement Observation Program (ADOS) measurements

O estudo [14] relata um método chamado *feature-aware*, que não somente melhorou a performance do modelo, mas permitiu entender melhor quais eram as relações entre as *features* observadas. O termo *aware* é frequentemente utilizado em estudos de ML para descrever que existe um processo de considerar determinado fator no processo de aprendizado, como na equação de regularização. Neste caso, *feature-aware* foi possível pelo método de regularização que aumentava a dispersão entre todas as *features* e a não-dispersão entre as tarefas, dando maior peso às características que eram importantes para todas as tarefas. Segundo o mesmo autor, o método *multi-task* foi capaz de lidar com *features* irrelevantes ou redundantes para as tarefas ao mesmo tempo.

Gênero e idade frequentemente são alvos de modelos para o estudo do envelhecimento do cérebro. Nos aprendizados multitarefa, essas características também podem ser vistas cada uma como uma tarefa. No estudo [102], foi aplicado um método chamado Sparse-MVTC que seleciona uma tarefa central e as demais são ditas auxiliares. Os dados foram divididos em grupos conforme as tarefas de gênero e idade. O artigo explora a não linearidade com um modelo *multi-kernel* multitarefa baseado na *mixed sparsity-inducing norm* a qual supõem conseguir perceber relações entre os escores cognitivos e as imagens. Os autores apostam também no compartilhamento de features entre os sujeitos que possam auxiliar os kernels na classificação. O modelo tem como objetivo classificar 10 escores cognitivos como ADAS, MMSE, RAVLT, FLU and TRAILS.

No estudo [97], as relações entre as tarefas foram associadas por meio de um grafo de dependência ilustrando a associação entre os *inputs* do modelo e os *outputs* possíveis. Como *input*, foram incluídos dados cardiológicos, idade, tempo de estudo, e expressão do gene APOE-E4 considerado fator de risco para desenvolvimento de Alzheimer.

O artigo [39] propõem um modelo multitarefa para prever aspectos do neurodesenvolvimento em crianças prematuras. Como entrada, dados de conectoma e clínicos foram associados com expectativa de otimizar o modelo para prever escores de Bayley-III sobre a cognição, linguagem e aspecto motor. Foram avaliadas as métricas para as predições individuais (*single-task*) e multi-tarefa. As predições em multitarefa apresentaram melhores performances comparadas as mesmas tarefas preditas separadamente.

A abordagem descrita por [103] assume que o modelo irá aprender as relações intrínsecas aos dados longitudinais automaticamente. Modelo denominado MTERL modelo de aprendizado exclusivo multitarefa proverá associações entre marcadores fenotípicos e dados cognitivos dos indivíduos. Para comparar o desempenho, outras 4 abordagens foram avaliadas sendo elas: *Lasso*, *Multi-task feature learning*, *Multi-task Exclusive Lasso* e *Multi-task Relationship Learning*. Vale ressaltar que o modelo tem como foco a análise longitudinal, avaliando os sujeitos ao longo do tempo, neste caso dados de MRI data e dois tipos de escores clínicos (*i.e.*, MMSE and ADAS-Cog) em 4 momentos (*i.e.*, baseline, 6-month, 12-month and 24-month).

Existem conexões compartilhadas entre os indivíduos que são comuns entre os sujeitos do estudo, mas também existem aquelas que são indivíduos-específico que varia bastante dentro dos sujeitos. Por isso, [104] propõem o modelo MTL-sCASO *multi-task learning-based sparse convex alternating structure optimization* que a partir da conectividade funcional prediz 5 variáveis comportamentais como cognição, uso de substância ilícita, uso de tabaco, personalidade e traços emocionais. A relação entre os sujeitos foi calculada por meio de correlação de Pearson a cada par de indivíduos. A média desses valores significava a similaridade entre sujeitos.

Apesar da abordagem supervisionada ser amplamente utilizada pela facilidade em convergir modelos com dados rotulados, o artigo de [61] propõe uma abordagem *weakly-supervised* com atenção para regressão de escores MMSE, CDRSB, and ADAS-Cog a partir da imagem de MRI. O modelo é dividido em módulos, o qual a primeira etapa consiste em classificar demência e obter mapas de ativação e obter as regiões cerebrais relevantes ao diagnóstico. Em seguida, esse mapa de ativação compõe a entrada para a etapa seguinte que prevê os escores clínicos. O estudo analisa os resultados do modelo multitarefa com um modelo idêntico predizendo apenas um escore por vez. O resultado mostra o modelo multitarefa com melhor performance quanto ao *single-task*.

A maioria dos dados que compõem um diagnóstico de desordens cerebrais possui uma distribuição não linear e isso é bem avaliado pelo estudo [91]. O estudo propõe um modelo chamado “*Subspace Network*” não linear multitarefa e ressalta que utilizar multitarefa nesses tipos de dados precisaram considerar o espaço limitado e específico das variáveis de regressão envolvidas. Alguns escores que são alvos de modelos preditivos possuem uma escala definida sendo quantitativos geralmente ordinais como o *Mini-Mental State Examination* (MMSE) que varia de 0 a 30 ou *Clinical Dementia Rating Scale* (CDR) e *Alzheimer’s Disease Assessment Scale-Cog* (ADAS- Cog) que possuem outras escalas. Para os autores, isso não é considerado quando um modelo de regressão linear é aplicado. Para isso, a arquitetura também aplica um tipo específico de regressão chamado regressão censurada, a qual assume que existe uma relação não linear entre as *features*.

Outro aspecto que esta revisão busca detalhar é a forma como os achados clínicos em termos neurobiológicos foram relacionados pelos modelos multitarefa. A forma mais simples de descrever a relação entre as tarefas e as regiões cerebrais é ponderar os pesos ou importância das regiões para as predições. O estudo [59] seleciona apenas aquelas que foram correspondentes em todas as tarefas, a partir das 100 principais de cada modalidade de imagem (estrutural e funcional) para cada centro de pesquisa, e verificou-se aquelas que se repetiam em todos. Resultou em 24 características de imagens funcionais e 18 das estruturais.

Poucos estudos diferenciaram explicitamente as análises pelo nível de observação quanto ao sujeito individualmente ou quanto ao grupo, ou população. O nível de grupo é associado pela variável comportamental que os une, como grupo caso e grupo controle.

A análise ao nível de grupo relaciona características comuns a todos ou maioria dos indivíduos proporcionando conhecimento sobre as conexões cerebrais daquela população. O método de ativação de classes, do inglês *Class Activation Mapping* (CAM), foi usado no estudo [61] com intenção de avaliar o grau de atenção do modelo para cada sujeito quanto a classificação quanto ao grau de demência. Para isso, o estudo avalia o método CAM versus o seu modelo com módulo de atenção MWAN. Para isso, 6 sujeitos foram selecionados aleatoriamente para terem seus mapas de ativação de classes visualizados. Os autores concluem que seu próprio modelo destacou regiões cerebrais biologicamente mais relacionadas com a literatura e suas associações já conhecidas para a demência, como regiões subcorticais onde se encontram o hipocampo e a amígdala responsáveis pela memória e tomada de decisões.

4.6 Conclusão

A abordagem multitarefa é considerada pelos estudos, capaz de avaliar múltiplas tarefas e extrair relações entre elas. A aplicação desse método de aprendizado em dados neurológicos foi alvo do estudo sistemático para compreender os benefícios dessa aplicação.

A partir da revisão apresentada, nota-se ainda oportunidades no campo do aprendizado multitarefa para explorar novas arquiteturas, principalmente para realçar as regiões cerebrais relacionadas ao diagnóstico clínico em questão. Embora, pelos resultados observados, é notório o benefício do MTL em relação às métricas de abordagens *single*.

A partir das tarefas utilizadas nos estudos, foi possível selecionar dentro do nosso escopo quais seriam melhores associadas e passíveis de comparação com a literatura. Logo, essas observações corroboram para a motivação desta pesquisa, bem como também para a metodologia aplicada. Os métodos de regularização para este categoria de aprendizado deve ser apropriado ao objetivo da pesquisa, considerando as tarefas em questão.

Além disso, a forma de manter a coerência dos pesos entre as tarefas, nos orienta a delimitar, com base na literatura, quais são as melhores abordagens de generalizar modelos MTL. Os artigos contribuíram também nas comparações entre modelos *single* e *multi*, demonstrando oportunidades e ganhos em cada um.

Por fim, o aprendizado obtido a partir da literatura norteou as definições metodológicas deste estudo, bem como direcionou para as oportunidades em aberto no campo de estudo de modelos MTL e cérebro-comportamento.

5. METODOLOGIA

O presente trabalho visa um estudo de abordagem mista, ou seja, qualitativo e quantitativo, o qual avalia semelhanças e diferenças entre indivíduos do estudo através de variáveis, relações estatísticas e padronizadas entre os grupos. A metodologia segue a visão de mundo pós-positivista [37] por um estudo de caso exploratório e experimento controlado.

Os estudos de caso exploratórios são usados como investigações iniciais de fenômenos podendo derivar novas hipóteses. Experimentos controlados possibilitam determinar como as variáveis estão relacionadas e se existe uma relação de causa-efeito entre elas. Uma condição para conduzir um experimento controlado é uma hipótese a qual guia todas as etapas do projeto experimental [24]. Logo, este trabalho tem a seguinte pergunta de pesquisa: é possível prever múltiplas variáveis comportamentais com aprendizado multitarefa e verificar regiões cerebrais em comum considerando como entrada dados de rs-fMRI?

De maneira geral, a metodologia desta pesquisa pode ser vista em três macro etapas, representadas na Figura 5.1, e cada etapa será detalhada nas seções a seguir. É detalhado como foram formuladas as tarefas e as arquiteturas experimentadas. Além disso, descreve-se o conjunto de dados, bem como método de treinamento e amostragem dos dados. Por fim, a implementação e as impressões dos experimentos serão discutidas em seguida.

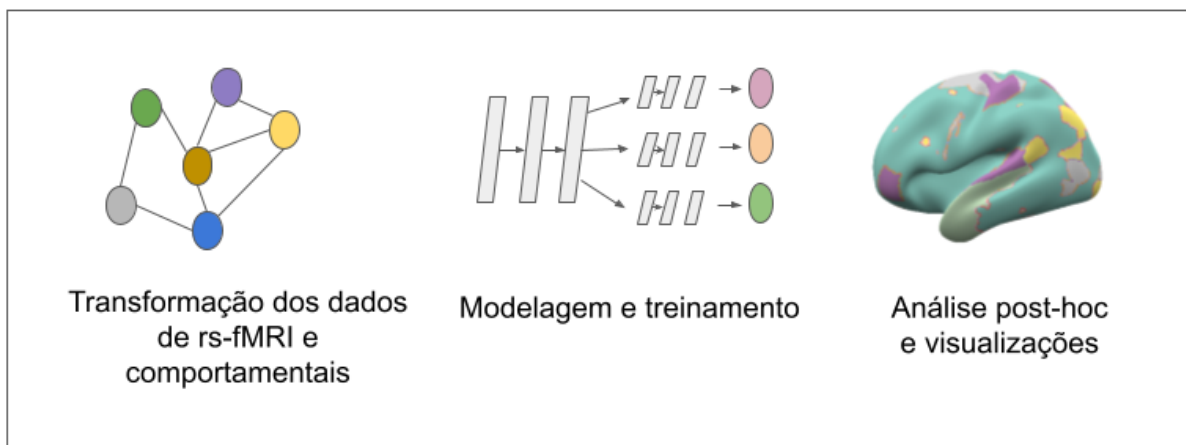


Figura 5.1: Macro etapas da metodologia de pesquisa.

5.1 Transformação dos dados

5.1.1 Conjunto de dados HCP

O *dataset* utilizado corresponde ao *Human Connectome Project* (HCP)[25] “PTN” (*Parcellation, Timeseries, Netmats*)¹ o qual contempla dados processados e reconstruídos pela última versão conforme processo descrito por Smith et al. [88], assim como utilizado em outros estudos de referência [84, 12]. Esta base de dados é composta por dados de neuroimagem de 812 jovens adultos de 22 a 35 anos, além de testes comportamentais e genéticos. Os dados obtidos a partir de rs-fMRI devem ser lidos como séries temporais com diversas dimensões, sendo uma série temporal para cada região do cérebro. Essas regiões podem ser divididas a partir de um atlas que, no contexto de neuroimagens corresponde a um mapa de regiões do cérebro denominadas regiões de interesse, do inglês *Regions of Interest* (ROIs), em análises de dados. A quantidade e o método para separar essas regiões podem ser conduzidas de diferentes maneiras dependendo do objetivo da análise [18].

Além dos atlas já existentes para parcelamento do cérebro em regiões de interesse, também é possível parcelar em componentes utilizando análise de componentes independentes (ICA). A versão HCP PTN utiliza ICA ao nível de grupo, e será utilizada neste experimento para definir as ROIs, pois essa versão fornece dados mais homogêneos e ressaltam diferenças individuais do que aqueles definidos por atlas genéricos [13]. A versão disponível foi parcelada usando a ferramenta MELODIC da FSL [9] e o ICA espacial foi aplicado em várias dimensionalidades diferentes (15, 25, 50, 100, 200 e 300), e para este estudo utilizamos 200 componentes, de forma que pudéssemos comparar com outros estudos que usam esse número de componentes. Para cada dimensionalidade têm-se os arquivos das matrizes de conectividade (*netmats*) em dois formatos: arquivo *netmats1*, contendo a correlação completa entre todas as séries temporais; e arquivo *netmats2* contendo a correlação parcial, capturando apenas as interações diretas entre os nodos removendo associações de nós e cofatores que possam ser correlacionados conforme demonstrou Smith (2012) [85]. Foi aplicada uma transformação gaussiana para converter os valores de escores de correlação de Pearson em escores padronizados (z-scores).

Alguns dados comportamentais não estão acessíveis abertamente, sendo necessário submeter o termo de acesso aos dados restritos. Nosso estudo obteve esta permissão e seguiu conforme os termos de uso de dados restritos do *WU-Minn HCP Consortium*².

¹<http://www.humanconnectome.org>

²<https://www.humanconnectome.org/study/hcp-young-adult/document/wu-minn-hcp-consortium-restricted-data-use-terms>

5.1.2 Formulação das tarefas

Escores cognitivos são frequentemente associados junto a dados de neuroimagem para ressaltar conexões cerebrais que tenham relações com o valor obtido pelo questionário padronizado [55, 32, 39]. Tendo como premissa do aprendizado multitarefa a preexistência de relação entre as tarefas para poderem ser associadas em um mesmo aprendizado [93, 26, 113], observamos nos estudos da literatura quais variáveis são possivelmente associadas. Assim, o primeiro passo da metodologia abrange selecionar as tarefas que, associadas, possam revelar conexões cerebrais ou biomarcadores relevantes entre elas. Primeiramente, com base na literatura e conhecimento prévio de fatores clínicos e comportamentais que possam ser relacionados direciona uma gama menor de tarefas.

A escolha das tarefas que serão aprendidas em simultâneo é crucial para ser percebível o benefício desta abordagem. A forma direta de escolha é dada a partir da busca exaustiva de possibilidades de tarefas, testando os seus efeitos no aprendizado a partir das métricas do modelo. Porém, o número de possibilidades aumenta exponencialmente conforme o domínio de tarefas disponíveis, além de ser um processo custoso e demorado. Deve-se observar ao longo do treinamento as métricas do modelo, verificando se houve benefício na associação das tarefas ou percebendo um antagonismo entre elas. A partir do *dataset Human Connectome Project* (HCP) foram selecionadas as principais tarefas a serem avaliadas neste experimento. A seguir, descrevem-se as tarefas que serão avaliadas nos experimentos, bem como a justificativa da sua inclusão.

MMSE Score

Mini-Exame do Estado Mental, ou em inglês *Mini-Mental State Examination* (MMSE) é um instrumento de rastreio do funcionamento cognitivo geral. Ao final do teste, obtém-se um score variando de 0 a 30, e algumas faixas de valores podem ser tendência para diagnósticos clínicos, como a faixa de 23/24 para o diagnóstico de demência. Este teste é utilizado em inúmeros estudos como variável dependente a partir das conexões cerebrais. Além disso, valor do escore pode ser covariável para predição de distúrbios cognitivos, como Alzheimer [55, 4].

Gênero

A classificação de gênero a partir de neuroimagem é amplamente estudada e por este fato, incluiu-se tal tarefa no estudo. Trata-se de uma tarefa de classificação binária, tendo os valores possíveis “feminino” ou “masculino”. Esta tarefa esta frequentemente associada às características morfológicas do cérebro que diferem entre os sexos, embora estudos também avaliem diferenças conectivas relacionadas [65, 5, 106].

Penn Matrix Reasoning Test (PMAT)

A variável *PMAT24_A_CR* corresponde ao escore do teste PMAT que avalia a inteligência fluída baseada em uma versão curta das matrizes progressivas de Raven (24 itens). Estudos anteriores [?, 69] já comprovaram ser possível prever inteligência fluída a partir das matrizes de conectividade dos sujeitos na base HCP. Assim, essa tarefa foi incluída como referência nas análises.

Histórico familiar

Diferentes variáveis sobre histórico familiar estão disponíveis nos dados comportamentais de acesso restrito. Entretanto, alguns apresentam baixa variabilidade e são mais difíceis de obter uma proporção balanceada entre as classes disponíveis. Por exemplo, questionamento direto sobre histórico de algum distúrbio mental específico na família, como ansiedade, depressão, esquizofrenia. Uma variável que resume de forma genérica de há ou não algum histórico de desordem mental é mais generalizável do que as específicas, e pode representar suficiente sentido ao modelo quanto a relação histórica familiar no escopo de desordens psiquiátricas e neurológicas. Separadas por relação materna ou paterna, as variáveis selecionadas deste núcleo foram *FamHist_Moth_None* e *FamHist_Fath_None* com valores booleanos, sendo *True* (Verdadeiro) caso não apresente nenhum histórico ou *False* (Falso) caso apresente algum dos distúrbios mencionados na pesquisa: Parkinson, Alzheimer, síndrome de Tourette, problema com álcool ou drogas, ansiedade, depressão, transtorno bipolar, esquizofrenia ou psicose.

5.1.3 Representação de série temporal através de grafo

As redes cerebrais comumente são observadas a partir de grafos os quais os nós correspondem a ROIs, ou componentes se extraídos pelo ICA, e as arestas sendo a conectividade funcional entre os nós [7]. Assim, para cada sujeito é necessário extrair as regiões de interesse, assim como cada região interagiu com as demais a partir de correlação em pares entre as séries temporais de cada ROI. A Figura 5.2 exemplifica o processo de transformação dos componentes em grafos. Para realizar esta transformação, utilizamos os dados de correlação parcial *netmats2* que representam os pesos entre os nodos dos grafos.

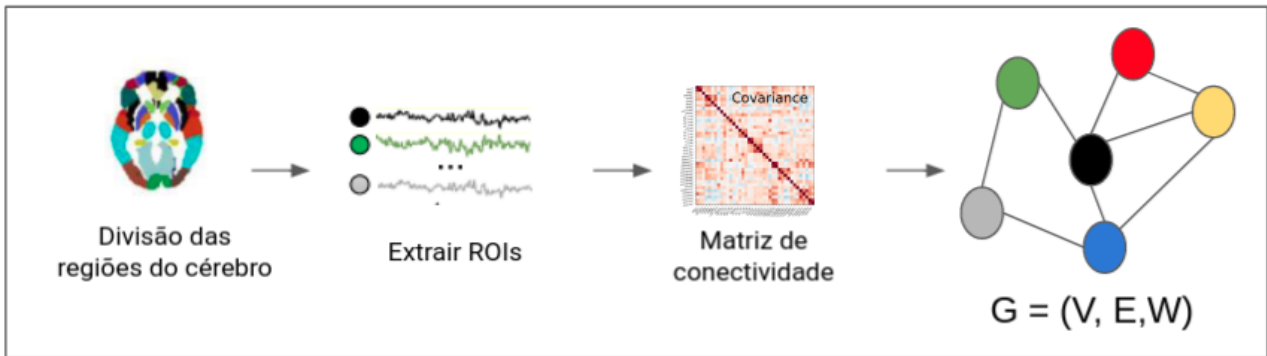


Figura 5.2: Transformação das séries temporais em grafos ponderados para entrada nas redes baseadas em grafos.

5.2 Modelagem e treinamento

De modo a resumir os passos metodológicos aplicados aos experimentos, a Figura 5.3 ilustra as principais etapas deste estudo. Em seguida, detalharemos as definições sobre método de treinamento, arquiteturas de base para fins comparativos bem como as justificativas dessas etapas.

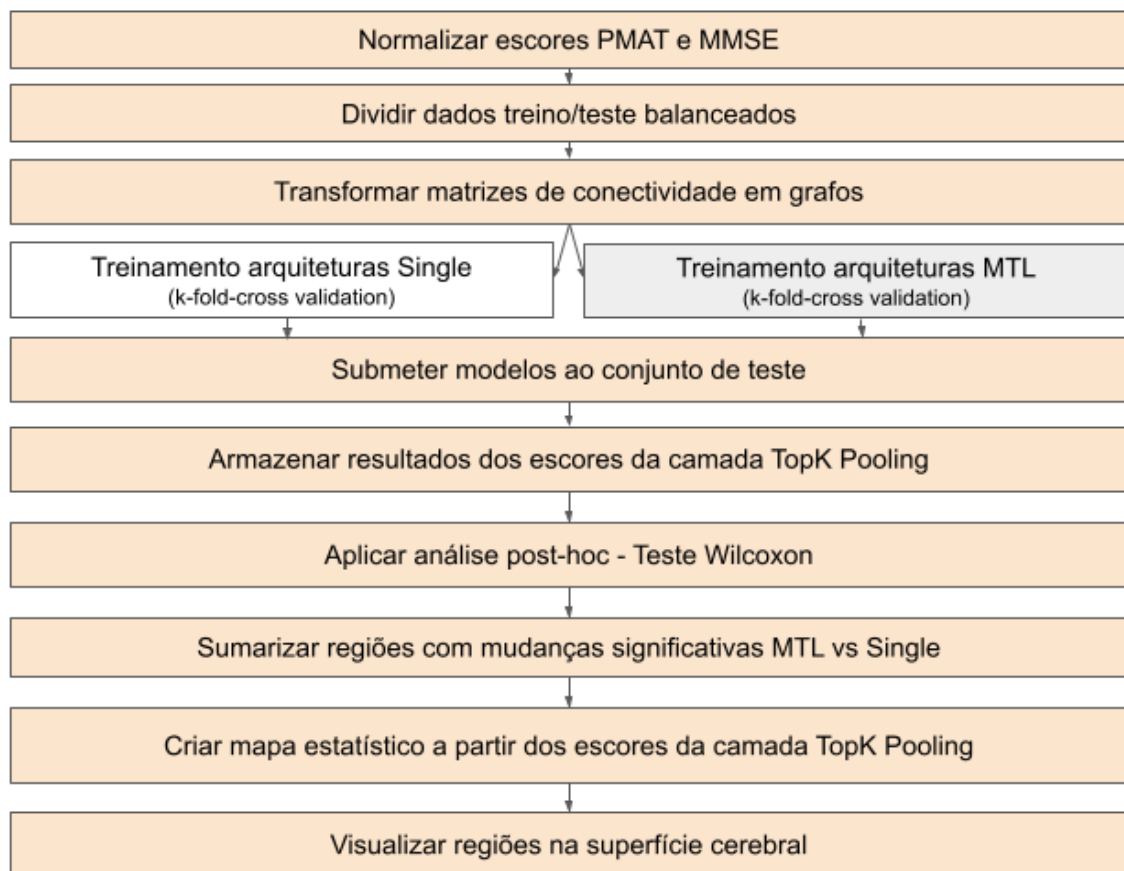


Figura 5.3: Etapas da metodologia aplicada aos experimentos.

Para validar o modelo, a maioria dos artigos no contexto de neuroimagem usa validação cruzada *5-fold*, conforme utilizado em outras metodologias semelhantes [71, 101, 29]. Entre os 812 sujeitos, 2 foram removidos por possuir dados comportamentais faltantes. Um percentual de 20% da amostra é reservado para testes e não será incluído no treinamento. O conjunto de dados foi balanceado entre gênero e proporção de histórico familiar de transtornos mentais materno e paterno. A distribuição das variáveis entre os dois conjuntos está descrita na Tabela 5.1. As variáveis Fam_Hist são booleanas, representadas na tabela pelas letras T (True) e F (False).

Tabela 5.1: Descrição do conjunto de dados

Dataset	N	Fem.	Masc.	Idade	MMSE Score	PMAT24_A_CR	FamHist_Moth	FamHist_Fath
Treino	648	334	314	28.3 ± 3.76	29.0 ± 1.04	17.0 ± 4.77	548±84.8%(T), 98±15.2%(F)	532±82.3%(T), 114±17.7%(F)
Teste	162	75	87	28.7 ± 3.59	29.0 ± 1.06	17.3 ± 4.52	148±91.9%(T), 13±8.1%(F)	130±80.7%(T), 31±19.3%(F)

Quanto às especificações técnicas dos modelos treinados, foram testados dois otimizadores, Adam, SGD e RMSProp, com a mesma taxa de aprendizado inicial de 0.001. Essa taxa foi reduzida em um fator 0.1 a cada 3 épocas sem melhora na perda no valor da função de custo do conjunto de validação. O procedimento de parada antecipada, do inglês *early stopping*, foi utilizado para interromper o treinamento ao não perceber melhora em até N épocas no valor da perda no conjunto de validação. As configurações associadas às variações foram agrupadas de modo a definir um conjunto de hiperparâmetros por vez e estão descritas na Tabela 5.2, e serão testados em todas as arquiteturas de base, bem como na abordagem proposta.

Tabela 5.2: Configurações de hiperparâmetros utilizadas nos experimentos.

Hiperparâmetro	Config 1	Config 2	Config 3
<i>Learning rate</i>	0.01	0.001	0.01
<i>Weight decay</i>	0.0005	0.005	0.0005
Otimizador	Adam	Adam	SGD
Redução da taxa de aprendizado	2 épocas	3 épocas	2 épocas
Parada Antecipada	5 épocas	7 épocas	5 épocas

5.2.1 Arquiteturas de base

A arquitetura BrainGNN [60] descreve uma GCN incluindo uma *ROI-aware graph convolutional layer* (Ra-GNN) com objetivo de auxiliar na interpretação salientando quais ROIs, representadas pelos nodos dos grafos, foram mais utilizadas para a predição. O objetivo dessa abordagem é não realizar uma interpretação com base em outras ferramentas após o treinamento do modelo [6], mas sim conseguir que o próprio modelo seja capaz de ressaltar as regiões cerebrais que possam conter biomarcadores importantes para um

distúrbio, por exemplo. Para isso, camada Ra-GNN modifica seu *kernel* de convolução conforme a ROI ao invés de manter o mesmo *kernel* para todos os nodos. Além disso, na etapa de *pooling*, chamada *R-pool*, a distância entre os nodos após a convolução Ra-GNN é atribuída conforme a importância do nodo. Em seguida, mantêm-se apenas os nodos com maiores distâncias. Essa abordagem pode ser promissora para um modelo multitarefa, visto que as ROIs podem ter comportamentos diferentes tanto em relação a uma tarefa, quanto para múltiplas.

Uma versão parcial da BrainGNN [60] está disponível no github³. Por meio disso, foi possível compreender a arquitetura para realizar o experimento em formato multitarefa. Entretanto, a versão disponível não apresenta nenhum recurso de visualização que possa ser de fato interpretável. A partir das descrições no artigo, entende-se que a relevância das regiões cerebrais para as classificações vêm da função de *pooling* aplicada chamada *TopKPooling* aplicada as ROIs, denominada *ROI-topK Pooling layer*.

Além disso, utilizamos uma arquitetura simples GCN adicionando camadas de *pooling* com a função *TopKPooling* de modo a contrapor com o modelo BrainGNN. Utilizamos as classes pré-definidas da biblioteca *Pytorch Geometric 2.0* para arquiteturas baseadas em grafos [30], Nilearn [2], Nibabel [11], HPC-utils⁴, entre outras.

5.2.2 Abordagem proposta

A Figura 5.4 apresenta uma representação resumida de como as arquiteturas foram implementadas nesta pesquisa. A Figura 5.4(a) destaca o método MTL com suas camadas compartilhadas, do inglês *Shared layers*, no quadrado destacado, as quais serão substituídas pelas camadas conforme a arquitetura base em análise, como descritas nas Tabelas 6.1 e 6.2. Enquanto isso, cada conjunto de camadas específicas referem-se às camadas da Tabela 6.3. Em paralelo, as mesmas arquiteturas de base foram treinadas separadamente para cada tarefa, como ilustra a Figura 5.4(b).

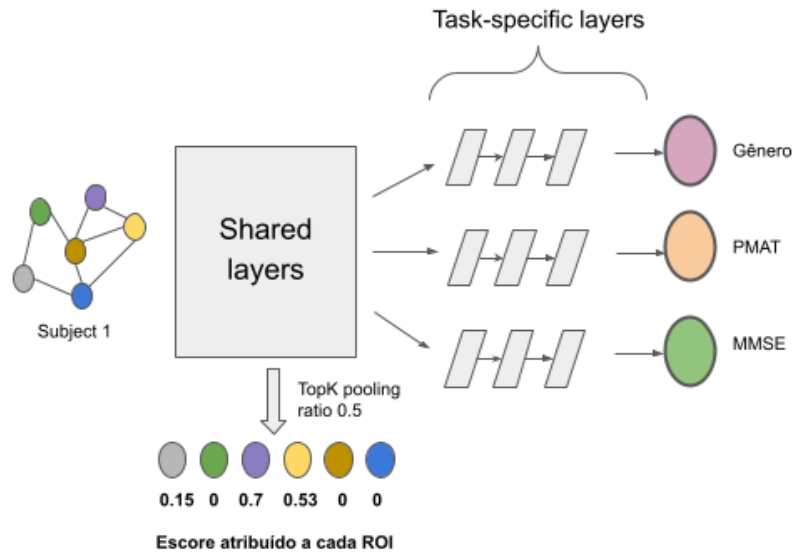
Avaliar o custo nas arquiteturas MTL, aplicamos o método de treinamento conhecido como *task-based regularization*, o qual pretende minimizar a soma das funções de perda de cada tarefa [26, 75]. Formalmente pode-se definir como:

$$L_{MTL}(X) = \sum_{i=0}^N L_{task}(\tau_i) \quad (5.1)$$

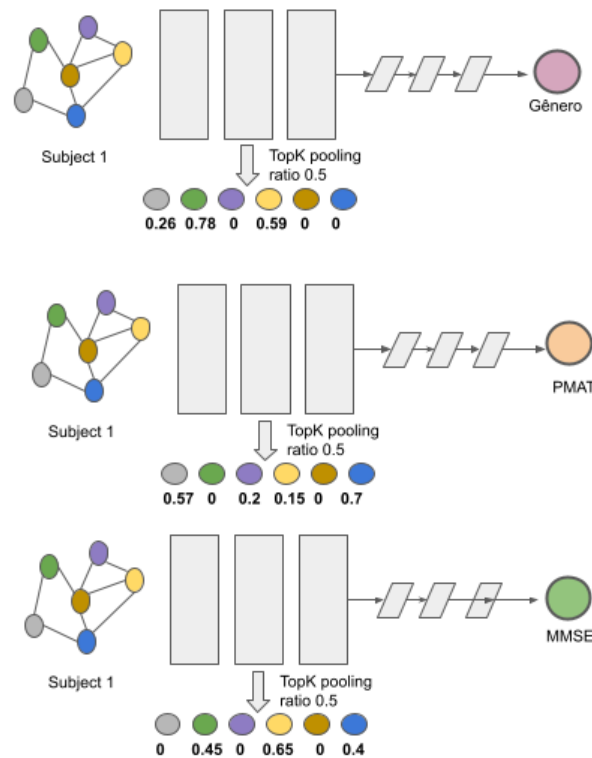
Com objetivo de obter quais nodos do grafo foram mais relevantes para a predição, adaptamos a camada *ROI-topK* utilizada na operação de *pooling*. Esta camada foi utilizada

³https://github.com/xxlyya/BrainGNN_Pytorch

⁴<https://rmlmj.github.io/hcp-utils>



(a) Esquema representando arquiteturas MTL e processo de obtenção dos escores por ROI.



(b) Esquema representando arquiteturas *single* para cada tarefa separadamente e a obtenção dos escores por ROI.

Figura 5.4: Representação das arquiteturas e suas diferenças para obtenção dos escores por ROI.

no estudo para reduzir a dimensionalidade do grafo entre as camadas de convolução, mantendo apenas os nodos mais relevantes ao longo do aprendizado. A definição da relevância de um nodo é armazenada como um vetor H de atributos dos nodos podendo ser definido

conforme a Equação 5.2, em que l se refere à camada do modelo e N o número de nodos (ROIs) do grafo na camada l . Duas camadas de *ROI-topK pooling* foram usadas, utilizando um fator de 0.5 para diminuição da dimensionalidade. Logo, se a entrada possui 200 nodos, o primeiro *pooling* resulta em 100 nodos relevantes, e na segunda 50 nodos relevantes.

$$H(l - 1) = [h_1^{(l-1)}, \dots, h_{N^{(l-1)}}^{(l-1)}] \quad (5.2)$$

A premissa dos escores para cada ROI baseia-se em que os mais relevantes se aproximam de 1 e os menos significativos mais próximos de 0. Para isso, aplica-se a função *sigmoid* ao vetor de escores dos nodos. Além disso, para manter a dispersão dos escores entre os nodos, o cálculo da perda considera também uma perda associada aos escores. Para que a relevância dos nodos seja distribuída corretamente por mais de uma tarefa, a função de perda utilizada é definida na Equação 5.3, onde $topK_{s1}$ corresponde a perda calculada sobre os escores:

$$L_{final} = \lambda_1 L_{MTL} + \lambda_2 L_{topK_{s1}} + \lambda_3 L_{topK_{s2}} \quad (5.3)$$

5.3 Análise *Post-hoc* e visualizações

Para interpretar a relação das tarefas com os dados de entrada, neste caso, as conexões cerebrais, é necessário observar ao nível de população. Os escores atribuídos as regiões cerebrais nos levam a relevância da respectiva região para a predição da variável. Os escores são atribuídos a cada região a cada sujeito, logo, tem-se 162 sujeitos de teste e 200 regiões cerebrais, para cada modelo. Nesse caso, as amostras não são independentes, dado que são os mesmos indivíduos testados em vários momentos. Os modelos podem ser interpretados como medidas repetidas de cada sujeito, sugerindo uma análise pareada. Porém, a variabilidade dos escores entre os algoritmos não tem garantia de atender o Teorema do Limite Central e possível de realizar um teste T pareado. Logo, o teste de Wilcoxon foi considerado mais adequado.

A hipótese nula (h_0) deste estudo pode ser definida tal que os escores das regiões são iguais nas abordagens *multi* e *single*, enquanto a hipótese alternativa (h_a) supõe que os escores para as regiões cerebrais são diferentes nas abordagens *multi* e *single*. A Figura 5.5 demonstra um exemplo hipotético de como serão comparadas cada região na população dos sujeitos de teste. Após realizar o teste estatístico para cada, agrupamos àquelas que rejeitam e aceitam h_0 .

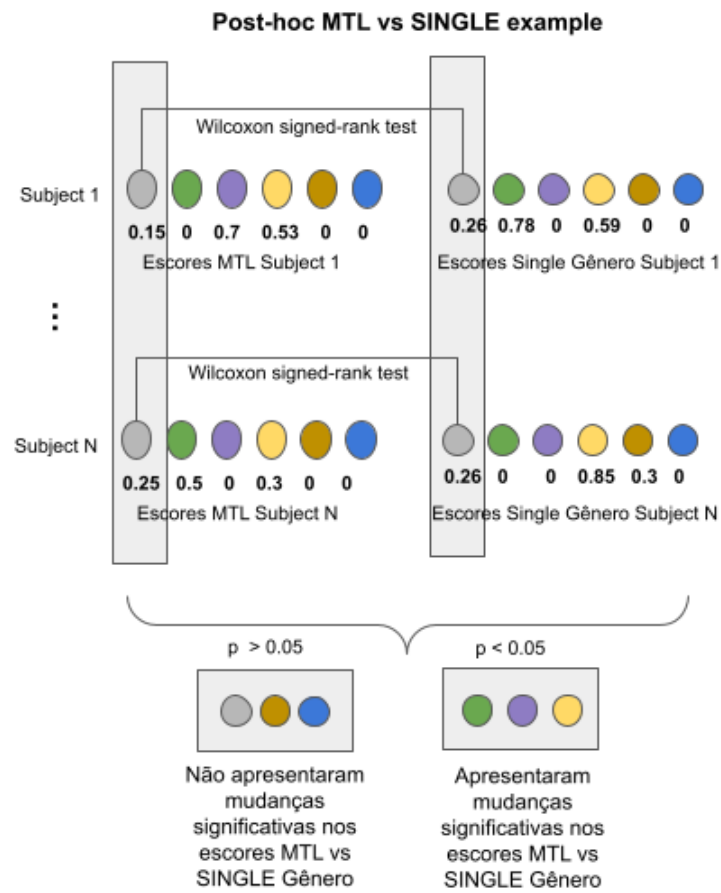


Figura 5.5: Exemplo gráfico do processo adotado para validar a diferença estatística dos escores por região entre arquiteturas MTL e *single*.

6. RESULTADOS

Dadas as tarefas Inteligência Fluída (PMAT), Escore Cognitivo (MMSE) e Gênero, neste capítulo será descrito os resultados dos experimentos na base de dados HCP, detalhando o processo de treinamento, validação e teste. Os modelos foram implementados na linguagem *Python* utilizando a biblioteca *PyTorch* e pacotes específicos como *PyTorch Geometric 2.0* para arquiteturas baseadas em grafos [30], *Nilearn*, *Nibabel*, *HPC-utils*, entre outras. A implementação completa está disponível no repositório GitHub¹.

6.1 Arquiteturas experimentais

Foi aplicada um modelo de base BrainGNN *single* para classificação de gênero e um para regressão da inteligência fluída PMAT e escore cognitivo MMSE, separadamente. Em seguida, as mesmas tarefas foram preditas no modelo multitarefa adaptando a arquitetura BrainGNN [60], que chamaremos de BrainGNN-MTL. A arquitetura BrainGNN-MTL resultou na configuração das camadas compartilhadas listadas na Tabela 6.1. Também aplicamos uma arquitetura GCN comum conforme as camadas compartilhadas descritas na Tabela 6.2. Para ambas, as camadas específicas de cada tarefa são definidas na Tabela 6.3. Tratando-se de 3 tarefas, as camadas específicas foram anexadas ao fim das camadas compartilhadas utilizando a classe *ModuleList* do *Pytorch*, resultando em 3 módulos, um para cada tarefa.

Tabela 6.1: Arquitetura das camadas compartilhadas entre todas as tarefas no modelo BrainGNN- MTL.

Layer	Dimensão
Linear	(200, 500)
ReLU	-
Linear	(50, 12800)
GC	(200, 64)
TopKPooling	64, ratio=0.5
Linear	(200, 500)
ReLU	-
Linear	(50, 4096)
GC	(64, 64)
TopKPooling	64, ratio=0.5
Linear	(256, 64)
BatchNorm1d	64, eps=1e-05
Linear	(64, 32)
BatchNorm1d	32, eps=1e-05

¹<https://github.com/mariliarosa4/multitask-learning-fmri>

Tabela 6.2: Arquitetura das camadas compartilhadas entre todas as tarefas no modelo GCN-MTL.

Layer	Dimensão
GC	(200, 100)
TopKPooling	100, ratio=0.5
GC	(100, 100)
TopKPooling	100, ratio=0.5
Linear	(200, 64)
BatchNorm1d	64, eps=1e-05
Linear	(64, 64)
BatchNorm1d	64, eps=1e-05

Para as tarefas de regressão, PMAT e MMSE a função de perda escolhida foi *Mean Squared Error* (MSE) e para tarefa de gênero a função de perda *binary crossentropy*. Os valores das variáveis de regressão foram normalizados entre 0 e 1, para evitar que a diferença entre as perdas fosse desequilibrar o aprendizado enquanto a perda da classificação binária é menor. Essas definições servem para ambos os modelos. A associação das funções de perda no modelo BrainGNN-MTL resultou na seguinte Equação 6.1, tal que W_{genero} , W_{PMAT} e W_{MMSE} igual a 1. Recapitulando a Equação 5.3 (página 45), os valores $\lambda_1, \lambda_2, \lambda_3$ foram inicialmente seguindo os valores sugeridos na *baseline* 1, 0.1 e 0.1 respectivamente.

Tabela 6.3: Arquitetura das camadas específicas correspondente às saídas das tarefas

Layer	Dimensão
Linear	(32, 12)
Dropout	p=0.2
ReLU	-
Linear	(12, 1)

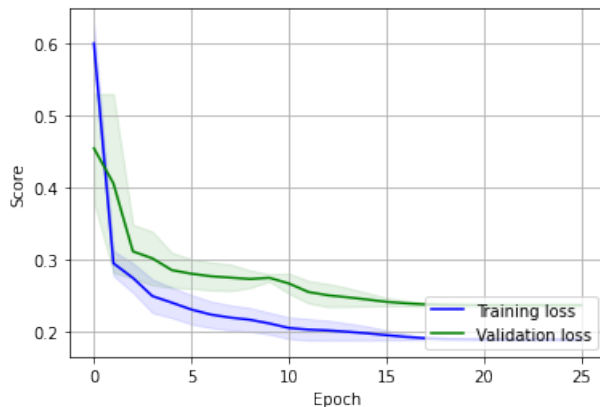
$$\begin{aligned}
L_{MTL} = & W_{genero} * binary_crossentropy(y_{genero}, \check{y}_{genero}) \\
& + W_{PMAT} * MSE(y_{PMAT}, \check{y}_{PMAT}) \\
& + W_{MMSE} * MSE(y_{MMSE}, \check{y}_{MMSE})
\end{aligned} \tag{6.1}$$

6.2 Treinamento e métricas no conjunto de teste

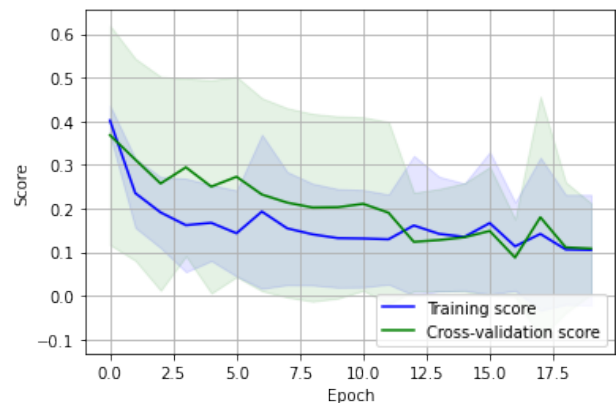
Durante o treinamento dos modelos, foram registrados os valores de perda para o conjunto de treinamento e validação de cada *fold*. A Figura 6.1 apresenta os gráficos de perda para as abordagens *BrainGNN Single* para as tarefas MMSE (Figura 6.1(a)), PMAT (Figura 6.1(b)) e Gênero (Figura 6.1(c)). Para representar as métricas de todos os *folders*, as

linhas do gráfico representam a média das perdas e as áreas sombreadas o desvio padrão. A escala do eixo x representa as épocas de treinamento, enquanto o eixo y corresponde a função de custo da qual desejamos minimizar. Ao passo que a foi aplicado o método de parada antecipada, o qual interrompe o treinamento ao não perceber melhoria na perda, cada modelo teve seu limite de épocas. Além disso, a Figura 6.1(d) mostra o histórico de perda para o modelo BrainGNN-MTL, em que se observa mais épocas de treinamento comparado aos demais. Destacamos que os valores de perda do modelo MTL (eixo y) são superiores dado que a função definida neste estudo considera o somatório da perda para todas as tarefas treinadas em simultâneo.

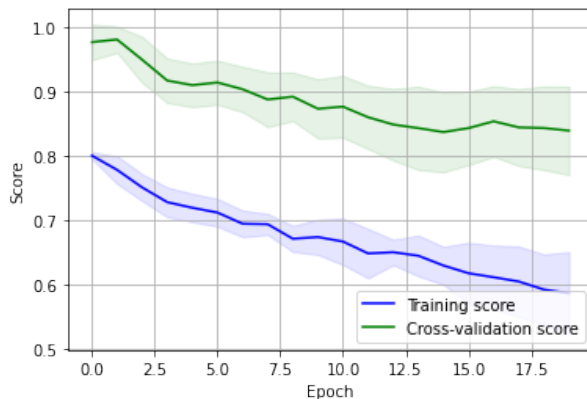
Após o treinamento e ajustes dos parâmetros para os modelos *single* e *multi*, os 162 sujeitos reservados para teste foram submetidos à avaliação dos escores atribuídos às ROIs ao melhor modelo de cada *fold*. Para cada métrica e modelo, apresentamos os resultados da média e desvio padrão entre os 5 *folds* executados na Tabela 6.4.



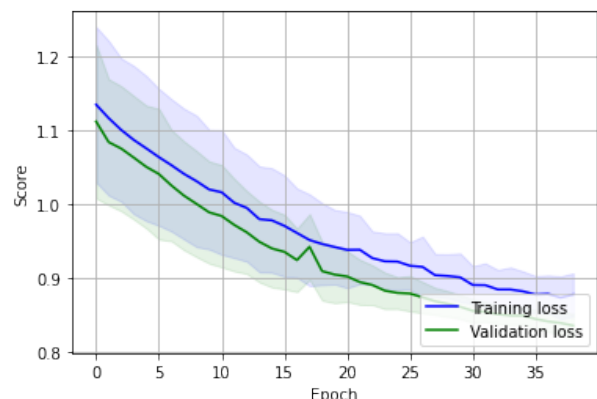
(a) MMSE BrainGNN Single



(b) PMAT BrainGNN Single



(c) Gênero BrainGNN Single



(d) BrainGNN-MTL

Figura 6.1: Perda média por época entre os *folds* de validação cruzada.

O experimento com arquitetura MTL que apresentou melhores resultados foi o *BrainGNN-MTL* com a configuração 3. Enquanto para as abordagens *single*, o melhor modelo foi *BrainGNN* com configuração 1. Não houve uma configuração em que todas as métricas fossem melhores em simultâneo, então pesou-se aquelas que 2 delas apresen-

tassem ganhos em contraponto às configurações com uma determinada métrica otimizada, como a GCN-MTL configuração 1 com menor rMSE para a tarefa MMSE, porém menores valores de AUC para Gênero e R^2 negativo para MMSE.

Tabela 6.4: Métricas média e desvio padrão por modelo e suas variações.

Arquitetura	Variações	PMAT	Gênero	MMSE
CPM Single	Regressão Linear	0.165 ± 0.011 (rMSE)	-	0.026 ± 0.002 (rMSE)
		0.032 ± 0.057 (R^2)		0.019 ± 0.017 (R^2)
BrainGNN single	Config 1	0.198 ± 0.022 (rMSE)	0.869 ± 0.050 (AUC)	0.138 ± 0.01 (rMSE)
		0.07 ± 0.003 (R^2)		0.087 ± 0.002 (R^2)
	Config 2	0.230 ± 0.19 (rMSE)	0.56 ± 0.052 (AUC)	0.111 ± 0.092 (rMSE)
		0.177 ± 0.042 (R^2)		0.00 ± 0.00 (R^2)
	Config 3	0.298 ± 0.009 (rMSE)	0.881 ± 0.161 (AUC)	0.286 ± 0.001 (rMSE)
		0.177 ± 0.069 (R^2)		0.087 ± 0.011 (R^2)
BrainGNN MTL	Config 1	0.440 ± 0.070 (rMSE)	0.58 ± 0.087 (AUC)	0.360 ± 0.006 (rMSE)
		-10.177 ± 0.700 (R^2)		-0.708 ± 0.0440 (R^2)
	Config 2	0.377 ± 0.008 (rMSE)	0.64 ± 0.05 (AUC)	0.289 ± 0.080 (rMSE)
		-0.177 ± 0.045 (R^2)		0.14 ± 0.070 (R^2)
	Config 3	0.259 ± 0.023 (rMSE)	0.901 ± 0.023 (AUC)	0.144 ± 0.017 (rMSE)
		0.177 ± 0.020 (R^2)		0.015 ± 0.004 (R^2)
GCN MTL	Config 1	0.258 ± 0.184 (rMSE)	0.881 ± 0.161 (AUC)	0.132 ± 0.105 (rMSE)
		-2.400 ± 4.923 (R^2)		-10.364 ± 8.522 (R^2)
	Config 2	0.186 ± 0.021 (rMSE)	0.662 ± 0.200 (AUC)	0.072 ± 0.034 (rMSE)
		0.187 ± 0.274 (R^2)		0.064 ± 0.022 (R^2)
	Config 3	0.337 ± 0.02 (rMSE)	0.68 ± 0.03 (AUC)	0.243 ± 0.07 (rMSE)
		-0.047 ± 0.005 (R^2)		0.015 ± 0.0018 (R^2)

A métrica de R^2 ou coeficiente de determinação avalia se a variância da variável dependente pode ser explicada pelas variáveis preditoras [17]. Os valores dessa métrica podem variar de 0 a 1, quando o melhor valor é 1, e também em alguns casos pode atingir valores negativos. Na Tabela 6.4, podemos ver que todos os modelos ficaram próximo de zero, além de alguns, apresentarem R^2 negativos. Isso significa que, dada uma reta horizontal que intercepta a média da variável dependente se ajusta melhor do que os modelos treinados [17]. Apesar disso, essa métrica isoladamente não captura o poder preditivo do modelo [48, 17], e sim se a variância de Y pode ser explicada pelo modelo. Por isso, a métrica rMSE foi considerada, tal qual avalia a raiz quadrada da média dos erros entre a variável alvo e o valor predito, elevando essa diferença ao quadrado para penalizar os valores muito distantes dos reais. Logo, quanto mais próximo de zero, menos o modelo está

errando. Os modelos destacados em negrito apresentaram melhor valor de rMSE, embora o R^2 não.

O artigo de He et al. [40] também encontrou valores de R^2 negativos, apontando que estaria correspondente com a literatura. No artigo os autores citam que já é conhecido sobre valores negativos de R^2 em dados de HCP. Os resultados do projeto HCP MegaTrawl apresentou $R^2 = -0.17$, sugerindo que as variáveis não são fáceis de prever usando dados de resting-state. Para realizar as comparações finais, os autores usaram a correlação de Pearson para determinar qual foi o melhor modelo para inteligência fluída (PMAT) ($R^2 = 0.05$ e R (Pearson's correlation) = 0.257) e idade. Além disso, os estudos de Rahim et al. [74] e Dubois et al. (2018) [22] também reportaram valores de R^2 negativos.

Observando as métricas por tarefa, quanto a variável de Inteligência Fluída (PMAT) foram condizentes com a literatura que utilizaram dados de rs-fMRI da base HCP, como Greene et al. [35] ($R^2 = 0.196$) e Dubois et al. [22] ($R^2 = 0.044$). Para a variável de Escore cognitivo (MMSE), a qual é geralmente utilizada para o diagnóstico de Alzheimer, as métricas obtidas seguem concorrentes com a literatura [74] ($R^2 = 0.22$). Ainda que não utilizando a mesma base de dados, o estudo associa a melhora na predição dos escores MMSE ao modelo associado com a regularização da idade. Apesar de perceber uma diferença importante entre as métricas do CPM *Single* para os modelos BrainGNN *single*, conseguimos alcançar métricas similares ao CPM no modelo BrainGNN-MTL. Logo, há um ganho considerando a possibilidade de obter as mesmas métricas em um único modelo.

6.2.1 Importância das Regiões Cerebrais

As regiões cerebrais de interesse foram definidas como nodos dos grafos $\eta = \{\eta_1, \dots, \eta_N\}$, onde N é o total de ROIs em análise, neste caso 200 componentes. Após a primeira camada de *pooling*, o tamanho das regiões relevantes reduz pelo fator de 0.5 retornando a metade do tamanho. Logo, tendo um *input* de tamanho 200x200, nesse ponto se tem 100 principais nodos, enquanto as demais regiões assume-se escore zero.

Avaliamos cada uma das tarefas em modo *single* e extraímos as regiões mais relevantes de cada tarefa. Cada sujeito recebe um vetor de escores de tamanho P referente ao número de ROIs após a camada de *pooling*. Ao final da avaliação de todo o conjunto de testes, obtêm-se uma matriz $N \times P$, onde N é o total de sujeitos de teste, neste caso 162x200. Visualmente esta matriz está representada na Figura 6.2(a), como exemplo do modelo BrainGNN-MTL em um único fold. Salienta-se que, como o método de validação *k-fold cross-validation* foi adotado, tal que cada *fold* gerou um melhor modelo. Para sumarizar os resultados de todos os *folders*, utilizamos o método de peso absoluto acumulado [107], assim como outros estudos o fizeram [83, 44, 103], para contabilizar os escores atribuídos

as regiões para os mesmos sujeitos nos *5-folds*. A equação do escore final $c^k = \sum_{i=1}^N |w_i^k|$, tal que de cada região k acumula o somatório dos pesos w^k atribuídos em cada N fold. A Figura 6.2(b) exemplifica o resultado desta operação no conjunto de teste para todos os *folds* do modelo BrainGNN-MTL. Vale ressaltar que, por estar somando todos os folds, a Figura 6.2(b) apresenta maior escala de escores comparado à escala da Figura 6.2(a) da esquerda correspondente a um único *fold*.

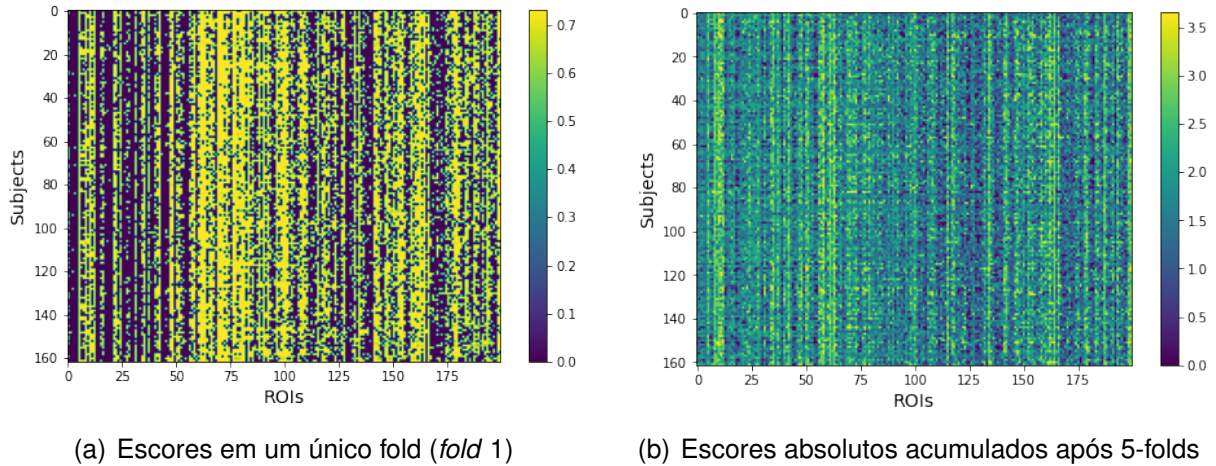


Figura 6.2: Matriz de scores dos sujeitos de teste e escores atribuído às ROIs no modelo BrainGNN-MTL último *fold*.

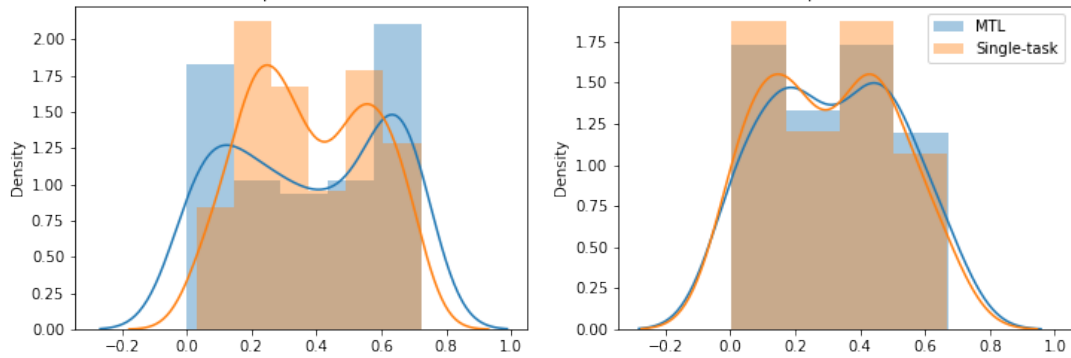
Apesar de ser perceptível pela Figura 6.2 a existência de um padrão de regiões que não foi considerada para nenhum sujeito, enquanto outras para todos os sujeitos, essa informação não revela a diferença ao nível de significância entre os escores atribuídos entre os modelos.

6.2.2 Análise *Post-hoc*

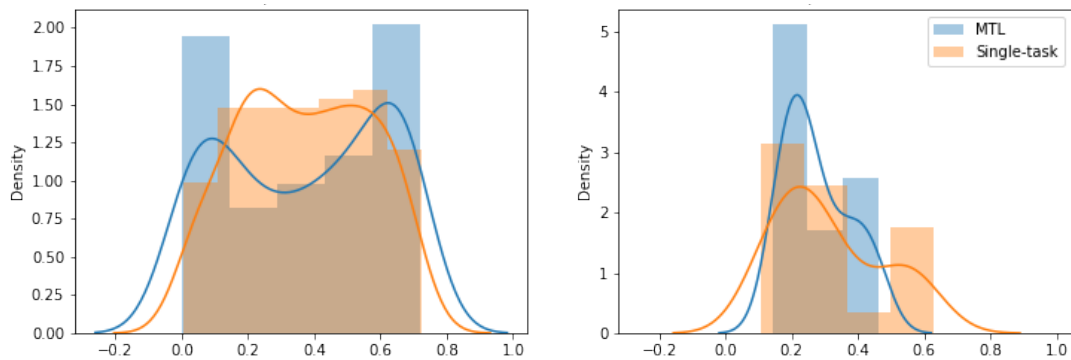
O teste paramétrico Wilcoxon *signed rank* [77] foi utilizado para comparar os escores atribuídos às regiões pelo método *single* e o método *multi-task*, BrainGNN-MTL em suas melhores configurações. O teste foi realizado região a região para cada sujeito, verificando quais regiões apresentam mudança de escores significativa entre os dois algoritmos para o mesmo sujeito.

A distribuição de *p-value* para os algoritmos *single* e BrainGNN-MTL é apresentada na Figura 6.3, subdivididas para cada tarefa sendo Figura 6.3(a) para Inteligência Fluida, Figura 6.3(b) MMSE e Figura 6.3(c) Gênero. As figuras a direita representam *p-values* que aceitam H_0 para o teste de Wilcoxon ($p > 0.05$), onde é possível notar maior proximidade entre as densidades dos valores de p para ambos os modelos. Em contrapartida, à es-

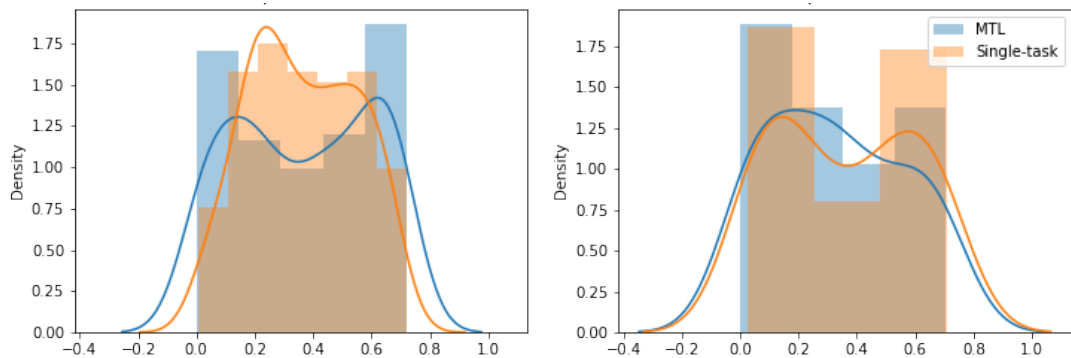
querda das figuras tem-se a distribuição de p -values os quais rejeitam h_0 para o teste de Wilcoxon ($p < 0.05$) e demonstram mudanças significativas nos escores atribuídos às ROIs.



(a) Inteligência fluida



(b) MMSE



(c) Gênero

Figura 6.3: Distribuição dos escores atribuídos às ROIs pelos modelos *single* (laranja) e *multi* (azul) agrupadas pela significância no teste de Wilcoxon.

Ao nível de ROIs ressaltadas para cada tarefa, o teste de Wilcoxon resultou em 95, 78 e 80 regiões para Gênero, PMAT e MMSE, respectivamente, sem mudança significativa ($p > 0.05$) entre os sujeitos nas duas abordagens. Logo, essas regiões receberam escores semelhantes para os mesmos sujeitos e regiões tanto no modelo *single* quanto no *MTL*. Não houve regiões comuns a todas as tarefas que não tiveram mudança significativa entre as abordagens *single* e *multi*, ou seja, todas as regiões que se mantiveram iguais em ambos

os modelos são exclusivamente relativas a cada tarefa. Esta evidência pode ser relevante quando a abordagem MTL preservou as regiões específicas das tarefas e que não foram suficientemente relevantes para mostrar alteração nos escores.

Para diferenciar visualmente as regiões conforme sua significância estatística, atribuímos valores simbólicos aos voxels conforme a região respectiva e o significado. Aquelas que não apresentaram mudança significativa (aceitam h_0) no teste de Wilcoxon ($p > 0.05$) receberam valor escalar de 0.5 (cor cinza), enquanto as em verde apresentaram mudanças significativas (rejeitam h_0) no teste ($p < 0.05$) com valor escalar -0.5. A Figura 6.4 ressalta essas mudanças destacando as conexões que não apres na superfície cortical esquerda.

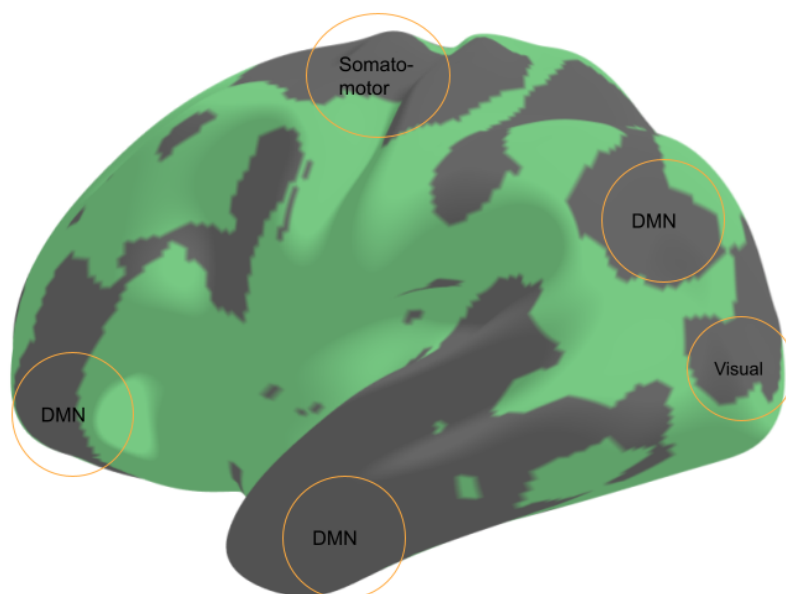


Figura 6.4: Superfície cortical esquerda ressaltando as redes que não apresentaram mudança significativa nos escores atribuídos pelos modelos. As regiões em cinza equivalem aquelas que não apresentaram mudança significativa no teste de Wilcoxon ($p > 0.05$). As destacadas em verde apresentaram mudanças significativas ($p < 0.05$).

Como utilizamos o parcelamento das regiões por componentes, para reagrupar os escores atribuídos por rede funcional, as quais são mais significativas semanticamente, agrupamos os escores por correlação dos componentes ao atlas Yeo17 [110] utilizando a ferramenta *FreeSurfer*, realizamos a correlação entre o atlas e os componentes a fim de obter o respectivo rótulo mais próximo àquele componente, com *threshold* de 0.2.

A Tabela 6.5 apresenta apenas os 6 primeiros componentes com maiores escores absolutos acumulados no modelo BrainGNN-MTL (na ordem decrescente), bem como as respectivas posições e escores em cada modelo *single*. É possível notar um equilíbrio na colocação de alguns componentes quando visto na abordagem MTL quando esta foi importante em 2 tarefas, como o componente 58, tal que individualmente ocupava a 13ª posição de relevância no modelo *BrainGNN Single* PMAT e MMSE, mas sobe para 3ª posição na

abordagem MTL. A região fronto-parietal é relacionada a inteligência fluída na literatura, à cognição superior como aplicação de regras, capacidade de síntese e geralmente apresentam declividade na função com o aumento da idade [105, 109]. Nessa pesquisa, essa região cerebral foi relacionada também ao gênero e escore cognitivo. As regiões demarcadas com o símbolo ▲ apresentaram maior posição do ranking *single* comparado ao MTL. Nota-se que, algumas regiões apesar de mudarem de posição, apresentaram escores acumulados similares entre MTL e *single*. Porém, a ordenação dentro de cada modelo foi respeitada em relação ao valor total acumulado.

Avaliando apenas àquelas regiões que apresentaram significância no teste de Wilcoxon, foram encontradas 53 componentes neste caso. A Tabela 6.6 mostra 8 regiões de exemplo, assim como sua posição no ranking de relevância em cada modelo e seu escore absoluto acumulado. Nesse ponto, podemos avaliar quais delas ganharam mais importância quando avaliadas pelo modelo *multi-task* ou *single*. Nota-se que alguns componentes tiveram diminuição no escore acumulado e desceram no *ranking*, como o componente 134, decaindo na importância pelo modelo MTL em comparação ao modelo *single* para MMSE. Porém, é um componente correspondente a ruído de sinal, e pode ter ganhado essa importância por estar presente em todas as imagens. No Apêndice B, apresenta-se visualmente os escores acumulados nos 53 componentes ressaltados no teste Wilcoxon os quais apresentaram mudança significativa entre o método *BrainGNN Single* e *BrainGNN-MTL* (Figura B.1).

O estudo de Mendes et al. [65], observado na revisão da literatura, relatou um modelo multitarefa a partir de uma rede convolucional 3D com imagens de MRI estrutural. Os autores utilizaram duas tarefas de classificação (gênero, diagnóstico de Autismo) e uma de regressão (idade), enquanto nossa proporção é maior de regressão, se tratando de duas tarefas de regressão e uma tarefa de classificação em conjunto. No estudo, o autor relata a dificuldade em balancear as perdas dado que a perda de regressão foi maior que as outras. No nosso caso, a percepção se inverte, sendo a perda de classificação (gênero) mais difícil de decrementar do que as de regressão (PMAT, MMSE). No processo de regularização que aplicamos, se houver características que se associam a apenas uma tarefa, ela tende a ser desconsiderada por não compartilhar essa relevância com as demais.

Tabela 6.5: Ranking dos 6 principais componentes com maior escore atribuído durante predição do modelo BrainGNN-MTL (colocação no ranking, escore acumulado).















	IC	BrainGNN-MTL	BrainGNN <i>Single</i> PMAT	BrainGNN <i>Single</i> Gênero	BrainGNN <i>Single</i> MMSE	Região Yeo17	Atlas
	63	1º (2.924)	109º (1.46)	192º (0.72)	15º (2.19)	Somatomotor	
	42	2º (2.923)	61º (2.02)	104º (1.46)	3º (2.73)	Atenção ventral	
	8	3º (2.922)	30º (2.16)	50º (1.99)	2º (2.74) ▲	Atenção ventral	
	58	4º (2.89)	13º (2.18)	145º (1.39)	13º (2.19)	Frontoparietal	
	61	5º (2.87)	10º (2.19)	193º (0.72)	53º (2.05)	Temporal Parietal	
	10	6º (2.76)	53º (2.06)	108º (1.46)	11º (2.19)	Somatomotor	

Tabela 6.6: Exemplos de componentes com diferença estatística significativa na população de teste entre os escores BrainGNN-MTL e *single* (colocação no ranking do modelo, escore acumulado).

	IC	BrainGNN-MTL	BrainGNN <i>Single</i> PMAT	BrainGNN <i>Single</i> Gênero	BrainGNN <i>Single</i> MMSE	Região Atlas Yeo17
	132	72 ^o (2.089)	168 ^o (1.45)	64 ^o (1.83) ▲	73 ^o (1.93)	Ruído
	5	17 ^o (2.75)	101 ^o (1.69)	50 ^o (2.00)	4 ^o (2.702) ▲	Control
	134	10 ^o (2.92)	54 ^o (2.06)	45 ^o (2.19)	5 ^o (2.694)	Ruído
	8	3 ^o (2.922)	30 ^o (2.16)	50 ^o (1.99)	3 ^o (2.74) ▲	Ventral Attention
	9	39 ^o (2.19)	126 ^o (1.46)	164 ^o (1.235)	140 ^o (1.45)	Frontoparietal
	10	6 ^o (2.76) ▲	53 ^o (2.06)	108 ^o (1.46)	11 ^o (2.19)	Somatomotor
	16	71 ^o (2.09)	145 ^o (1.46)	142 ^o (1.42)	145 ^o (1.45)	Ventral Attention
	28	35 ^o (2.19)	117 ^o (1.46)	197 ^o (0.649)	100 ^o (1.462)	Ventral Attention

7. CONCLUSÃO

A compreensão de marcadores biológicos para distúrbios cerebrais a partir de algoritmos de detecção de padrões são objetivos de pesquisas recentes. Modelos preditivos vêm sendo explorados principalmente por aprendizado *single-task*. A modelagem *single-task* prediz apenas uma variável comportamental por vez e este método corresponde a maioria dos estudos de análise cérebro-comportamento. Enquanto isso, a abordagem multitarefa compõe o método de alguns poucos estudos dessa área. Entretanto, esses estudos, a exemplo dos apresentados na revisão da literatura deste trabalho, mostraram o poder de relacionar múltiplas tarefas e extrair relações entre elas com modelos MTL.

Nossa pesquisa aplicou o aprendizado multitarefa em arquiteturas de redes neurais baseadas em grafos a partir de dados neurológicos. Exploramos a correlação entre as tarefas por meio de uma método de regularização apropriado para transferir os pesos entre as camadas reavaliando a função de perda a qual considerava todas as tarefas. Em contraponto com os métodos *single* testados, alcançamos métricas equivalentes nas abordagens MTL, bem como similares à literatura.

A partir dos resultados alcançados, conseguimos demonstrar quais regiões foram consideradas para as predições a partir da redução de dimensionalidade nas camadas de *pooling* com a função de perda TopKLoss. Entre os 200 componentes observados, nem todos são relevantes para todas as tarefas. Por isso, o aprendizado multitarefa permite, a partir da função de regularização, identificar um subconjunto de componentes compartilhados entre as tarefas. Além disso, os testes estatísticos aplicados ressaltaram 53 componentes com escores de importância ao modelo significativamente diferentes entre as abordagens multitarefa e *single*. Dentre as regiões que não apresentaram mudança no escore *single* e *multi*, não houve alguma em comum a todas as tarefas. Logo, todas as regiões que mantiveram seus escores em ambos os modelos, não são importantes/associadas entre as 3 tarefas, na visão do algoritmo MTL. Este fato destaca-se não apenas na diferença no valor bruto atribuído às regiões, mas também foi perceptível ao ordenar as regiões por importância, visto que os componentes sofreram alterações nas posições do *ranking* multitarefa quando ambas eram importantes ou menos importantes nos modelos *single*.

Ao passo que a análise por componentes é mais adequada para perceber nuances em uma população, transpor os componentes às regiões cerebrais já conhecidas foi importante para comparar com a literatura clínica. Logo, a correlação com o atlas Yeo17 [98] permitiu observar que redes associadas à cognição humana (por exemplo, a região fronto-parietal) a qual teve aumento na relevância quando prevendo duas tarefas relacionadas à capacidade cognitiva.

7.1 Limitações

Entende-se que há outras tarefas e arquiteturas que possam ser exploradas, bem como métodos avançados de regularização para abordagens multitarefa. Nesse sentido, como limitação do presente estudo podemos mencionar a pequena amplitude de configurações testadas nos experimentos, assim como a utilização de apenas um método de seleção de características pela abordagem de *pooling TopK-pooling*.

Vale ressaltar como limitação a existência de novos artigos que poderiam ser incluídos na revisão da literatura após a data de busca. Neste caso, caberia uma atualização do resultado das buscas pela *string* definida, além da verificação do impacto nos estudos após todas as etapas do processo de seleção de artigos.

A observação de marcadores biológicos no dataset selecionado também é limitada por não ser direcionada a um grupo de diagnóstico e controle, os quais poderiam ser mais perceptíveis de ressaltar regiões e associar a literatura. Porém, o tamanho da base HCP garantiu amplitude da pesquisa, desviando a possibilidade de os resultados serem enviesados pelo tamanho do dataset, por exemplo.

Quanto ao tamanho da amostra, um estudo recente de Marek et al. [64] tratando especificamente da reprodutibilidade de análises cerebrais demonstrou que a garantia de generalização das análises é melhor quando o conjunto apresenta amplitude na casa de milhares. Por isso, este trabalho tem também como limitação o tamanho do conjunto avaliado, mas também a população utilizada. Dado que todos os indivíduos do conjunto são saudáveis, é possível notar valores mais centralizados sobre as medidas e diminuindo a variabilidade de padrões "não saudáveis".

7.2 Trabalhos futuros

Como trabalhos a serem desenvolvidos complementares a este estudo, recomenda-se testar arquiteturas de redes neurais convolucionais bidimensionais similares às aplicadas em imagens, tais como as mencionadas na revisão da literatura, BrainNetCNN [49], por exemplo. A partir disso, verificar se os ganhos do aprendizado MTL se equivalem neste tipo de arquitetura, tal qual superem ou não as redes de grafo aqui testadas. Ao ampliar o escopo de arquiteturas convolucionais bidimensionais, outras formas de obter a relevância das ROIs se mostram interessantes a serem avaliadas, tal como mapas de ativação de classes, *SmoothGrad*, por exemplo, referenciado em alguns artigos da revisão da literatura [65, 6].

Apesar dos resultados desta pesquisa terem mostrado potencial na sua aplicação, é interessante considerar que outras características comportamentais podem ser adiciona-

das aos nodos dos grafos para equilibrar as relações de covariáveis. A partir disso, verificar a diferença nas métricas antes e depois desta melhoria. Também pode-se testar outras formas de obter as importâncias das regiões de interesse em redes baseadas em grafos como o método de pooling de autoatenção, do inglês *Self-Attention Graph Pooling* [58].

Como oportunidade de trabalhos futuros a partir deste, também pode-se explorar outras tarefas como os escores do teste “Cinco Grandes Fatores da Personalidade” (*Big Five*) comumente medido em estudos cognitivos e presente na literatura [22, 29], uso de substância ilícita, uso de tabaco, entre outras. Os dados da base *Human Brain Project* (HCP) não estão direcionados a um grupo de diagnóstico clínico alvo, contudo seria interessante avaliar também a multitarefa em outras bases da literatura específica como dados de Autismo (*Autism Brain Imaging Data Exchange - ABIDE*) ou Alzheimer (*Alzheimer’s Disease Neuroimaging Initiative - ADNI*).

Neste estudo avaliamos a abordagem de compartilhamento de parâmetros mais utilizada em artigos multitarefa denominada *hard*. Porém, recomenda-se testar também o compartilhamento do tipo *soft*, o qual pode vir a otimizar os resultados aqui apresentados.

A partir da revisão da literatura, alguns artigos sugerem verificar diferenças de significância dos modelos quando modificado o método de parcelamento do cérebro. Neste caso, utilizamos a dimensão de 200 componentes, mas pode-se avaliar a mudança do *ranking* das regiões cerebrais e performance dos modelos quando diminuído para 100 componentes ou aumentado para 300.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Abraham, A.; Milham, M. P.; Martino, A. D.; Craddock, R. C.; Samaras, D.; Thirion, B.; Varoquaux, G. “Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example”, *NeuroImage*, vol. 147, Fev 2017, pp. 736–745.
- [2] Abraham, A.; Pedregosa, F.; Eickenberg, M.; Gervais, P.; Mueller, A.; Kossaifi, J.; Gramfort, A.; Thirion, B.; Varoquaux, G. “Machine learning for neuroimaging with scikit-learn”, *Frontiers in Neuroinformatics*, vol. 8, Fev 2014, pp. 14.
- [3] Ahmedt-Aristizabal, D.; Armin, M. A.; Denman, S.; Fookes, C.; Petersson, L. “Graph-based deep learning for medical diagnosis and analysis: Past, present and future”, *Sensors*, vol. 21–14, Jul 2021, pp. 41.
- [4] Arevalo-Rodriguez, I.; Smailagic, N.; i Figuls, M. R.; Ciapponi, A.; Sanchez-Perez, E.; Giannakou, A.; Pedraza, O. L.; Cosp, X. B.; Cullum, S. “Mini-mental state examination (MMSE) for the detection of alzheimer's disease and other dementias in people with mild cognitive impairment (MCI)”, *Cochrane Database of Systematic Reviews*, vol. 3, Mar 2015, pp. 76.
- [5] Arslan, S.; Ktena, S. I.; Glocker, B.; Rueckert, D. “Graph saliency maps through spectral convolutional networks: Application to sex classification with brain connectivity”. In: *Proceedings of Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities*, Stoyanov, D.; Taylor, Z.; Ferrante, E.; Dalca, A. V.; Martel, A.; Maier-Hein, L.; Parisot, S.; Sotiras, A.; Papiez, B.; Sabuncu, M. R.; Shen, L. (Editores), 2018, pp. 3–13.
- [6] Baldassarre, F.; Azizpour, H. “Explainability techniques for graph convolutional networks”, *arXiv Preprint*, vol. 1905.13686, Jun 2019, pp. 21.
- [7] Bassett, D. S.; Sporns, O. “Network neuroscience”, *Nature Neuroscience*, vol. 20–3, Fev 2017, pp. 353–364.
- [8] Beaty, R. E.; Kenett, Y. N.; Christensen, A. P.; Rosenberg, M. D.; Benedek, M.; Chen, Q.; Fink, A.; Qiu, J.; Kwapil, T. R.; Kane, M. J.; Silvia, P. J. “Robust prediction of individual creative ability from brain functional connectivity”. In: *Proceedings of the National Academy of Sciences*, 2018, pp. 1087–1092.
- [9] Beckmann, C.; Smith, S. “Probabilistic independent component analysis for functional magnetic resonance imaging”, *IEEE Transactions on Medical Imaging*, vol. 23–2, Fev 2004, pp. 137–152.

- [10] Bielschky, N. Z.; Uithol, S.; van Mourik, T.; Anderson, P.; Glennon, J. C.; Buitelaar, J. K. “Disentangling causal webs in the brain using functional magnetic resonance imaging: A review of current approaches”, *Network Neuroscience*, vol. 3–2, Jan 2019, pp. 237–273.
- [11] Brett, M.; Markiewicz, C. J.; Hanke, M.; Côté, M.-A.; Cipollini, B.; McCarthy, P.; Cheng, C. P.; Halchenko, Y. O.; Cottaar, M.; Ghosh, S.; Larson, E.; Wassermann, D.; Gerhard, S.; Lee, G. R.; Kastman, E.; Rokem, A.; Madison, C.; Morency, F. C.; Moloney, B.; Burns, C.; Millman, J.; Gramfort, A.; Leppäkangas, J.; Markello, R.; van den Bosch, J. J.; Vincent, R. D.; Subramaniam, K.; Raamana, P. R.; Nichols, B. N.; Baker, E. M.; Goncalves, M.; Hayashi, S.; Pinsard, B.; Haselgrove, C.; Hymers, M.; Koudoro, S.; Oosterhof, N. N.; Amirbekian, B.; Nimmo-Smith, I.; Nguyen, L.; Reddigari, S.; St-Jean, S.; Garyfallidis, E.; Varoquaux, G.; Kaczmarzyk, J.; Legarreta, J. H.; Hahn, K. S.; Hinds, O. P.; Fauber, B.; Panfilov, E.; Poline, J.-B.; Stutters, J.; Jordan, K.; Cieslak, M.; Moreno, M. E.; Haenel, V.; Schwartz, Y.; Thirion, B.; Papadopoulos Orfanos, D.; Pérez-García, F.; Solovey, I.; Gonzalez, I.; Palasubramaniam, J.; Lecher, J.; Leinweber, K.; Raktivan, K.; Fischer, P.; Gervais, P.; Gadde, S.; Ballinger, T.; Roos, T.; Reddam, V. R.; freec84. “nipy/nibabel: 2.5.0”. Capturado em: <https://doi.org/10.5281/zenodo.3360650>, Ago 2019.
- [12] Cai, H.; Chen, J.; Liu, S.; Zhu, J.; Yu, Y. “Brain functional connectome-based prediction of individual decision impulsivity”, *Cortex*, vol. 125, Abr 2020, pp. 288–298.
- [13] Calhoun, V.; Adali, T.; Pearlson, G.; Pekar, J. “A method for making group inferences from functional MRI data using independent component analysis”, *Human Brain Mapping*, vol. 14–3, Nov 2001, pp. 140–151.
- [14] Cao, P.; Tang, S.; Huang, M.; Yang, J.; Zhao, D.; Trabelsi, A.; Zaiane, O. “Feature-aware multi-task feature learning for predicting cognitive outcomes in alzheimer’s disease”. In: *Proceedings of 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019, pp. 1–5.
- [15] Caruana, R. “Multitask learning”, *Machine Learning*, vol. 28–1, 1997, pp. 41–75.
- [16] Chen, Y.; Almarzouqi, S. J.; Morgan, M. L.; Lee, A. G. “T1-weighted image”. In: *Encyclopedia of Ophthalmology*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, vol. 1, pp. 1747–1750.
- [17] Chicco, D.; Warrens, M. J.; Jurman, G. “The coefficient of determination r-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation”, *PeerJ Computer Science*, vol. 7, Jul 2021, pp. 623.

- [18] Craddock, R. C.; James, G.; Holtzheimer, P. E.; Hu, X. P.; Mayberg, H. S. “A whole brain fMRI atlas generated via spatially constrained spectral clustering”, *Human Brain Mapping*, vol. 33–8, Jul 2011, pp. 1914–1928.
- [19] Cui, Z.; Gong, G. “The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features”, *NeuroImage*, vol. 178, Set 2018, pp. 622–637.
- [20] Dosenbach, N. U. F.; Fair, D. A.; Miezin, F. M.; Cohen, A. L.; Wenger, K. K.; Dosenbach, R. A. T.; Fox, M. D.; Snyder, A. Z.; Vincent, J. L.; Raichle, M. E.; Schlaggar, B. L.; Petersen, S. E. “Distinct brain networks for adaptive and stable task control in humans”, *Proceedings of the National Academy of Sciences*, vol. 104–26, Jun 2007, pp. 11073–11078.
- [21] Dougherty, J.; Kohavi, R.; Sahami, M. “Supervised and unsupervised discretization of continuous features”. In: *Proceedings of Twelfth International Conference on Machine Learning 1995*, Prieditis, A.; Russell, S. (Editores), 1995, pp. 194–202.
- [22] Dubois, J.; Galdi, P.; Han, Y.; Paul, L. K.; Adolphs, R. “Resting-state functional brain connectivity best predicts the personality dimension of openness to experience”, *Personality Neuroscience*, vol. 1, Jul 2018, pp. 6.
- [23] Dybå, T.; Dingsøy, T. “Empirical studies of agile software development: A systematic review”, *Information and Software Technology*, vol. 50–9, Fev 2008, pp. 833 – 859.
- [24] Easterbrook, S.; Singer, J.; Storey, M.-A.; Damian, D. “Selecting Empirical Methods for Software Engineering Research”. London: Springer London, 2008, pp. 285–311.
- [25] Essen, D. V.; Ugurbil, K.; Auerbach, E.; Barch, D.; Behrens, T.; Bucholz, R.; Chang, A.; Chen, L.; Corbetta, M.; Curtiss, S.; Penna, S. D.; Feinberg, D.; Glasser, M.; Harel, N.; Heath, A.; Larson-Prior, L.; Marcus, D.; Michalareas, G.; Moeller, S.; Oostenveld, R.; Petersen, S.; Prior, F.; Schlaggar, B.; Smith, S.; Snyder, A.; Xu, J.; Yacoub, E. “The human connectome project: A data acquisition perspective”, *NeuroImage*, vol. 62–4, Out 2012, pp. 2222–2231.
- [26] Evgeniou, T.; Pontil, M. “Regularized multi-task learning”. In: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 109–117.
- [27] Farahani, F. V.; Karwowski, W.; Lighthall, N. R. “Application of graph theory for identifying connectivity patterns in human brain networks: A systematic review”, *Frontiers in Neuroscience*, vol. 13, Jun 2019, pp. 585.

- [28] Feldman, S.; Gupta, M. R.; Frigiyik, B. A. “Revisiting stein’s paradox: Multi-task averaging”, *Journal of Machine Learning Research*, vol. 15–106, Nov 2014, pp. 3621–3662.
- [29] Feng, C.; Wang, L.; Li, T.; Xu, P. “Connectome-based individualized prediction of loneliness”, *Social Cognitive and Affective Neuroscience*, vol. 14–4, Mar 2019, pp. 353–365.
- [30] Fey, M.; Lenssen, J. E. “Fast graph representation learning with pytorch geometric”, *arXiv Preprint*, vol. 1903.02428, Mar 2019, pp. 9.
- [31] Finn, E. S.; Shen, X.; Scheinost, D.; Rosenberg, M. D.; Huang, J.; Chun, M. M.; Papademetris, X.; Constable, R. T. “Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity”, *Nature Neuroscience*, vol. 18–11, Out 2015, pp. 1664–1671.
- [32] Folstein, M. F.; Robins, L. N.; Helzer, J. E. “The Mini-Mental State Examination”, *Archives of General Psychiatry*, vol. 40–7, Jul 1983, pp. 1.
- [33] Glover, G. H. “Overview of functional magnetic resonance imaging”, *Neurosurgery Clinics of North America*, vol. 22–2, Abr 2011, pp. 133–139.
- [34] Goodfellow, I.; Bengio, Y.; Courville, A. “Deep Learning”. MIT Press, 2016, pp. 800.
- [35] Greene, A. S.; Gao, S.; Scheinost, D.; Constable, R. T. “Task-induced brain state manipulation improves prediction of individual traits”, *Nature Communications*, vol. 9–1, Jul 2018, pp. 13.
- [36] Greicius, M. D.; Srivastava, G.; Reiss, A. L.; Menon, V. “Default-mode network activity distinguishes alzheimer’s disease from healthy aging: Evidence from functional MRI”, *Proceedings of the National Academy of Sciences*, vol. 101–13, Mar 2004, pp. 4637–4642.
- [37] Guba, E. G.; Lincoln, Y. S. “Competing paradigms in qualitative research”, *Handbook of qualitative research*, vol. 2, Jun 1994, pp. 163–194.
- [38] Hao, S.; Chen, T.; Wang, Y.; Guo, Y.; Wang, M. “Adaptive multi-task dual-structured learning with its application on alzheimer’s disease study”, *ACM Trans. Internet Technol.*, vol. 21–2, Mai 2021, pp. 16.
- [39] He, L.; Li, H.; Wang, J.; Chen, M.; Gozdas, E.; Dillman, J. R.; Parikh, N. A. “A multi-task, multi-stage deep transfer learning model for early prediction of neurodevelopment in very preterm infants”, *Scientific Reports*, vol. 10–1, Set 2020, pp. 13.

- [40] He, T.; Kong, R.; Holmes, A. J.; Nguyen, M.; Sabuncu, M. R.; Eickhoff, S. B.; Bzdok, D.; Feng, J.; Yeo, B. T. “Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics”, *NeuroImage*, vol. 206, Fev 2020, pp. 116276.
- [41] Hoeft, F.; Ueno, T.; Reiss, A. L.; Meyler, A.; Whitfield-Gabrieli, S.; Glover, G. H.; Keller, T. A.; Kobayashi, N.; Mazaika, P.; Jo, B.; Just, M. A.; Gabrieli, J. D. E. “Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures.”, *Behavioral Neuroscience*, vol. 121–3, Jun 2007, pp. 602–613.
- [42] Hsu, W.-T.; Rosenberg, M.; Scheinost, D.; Constable, R.; Chun, M. “Resting-state functional connectivity predicts neuroticism and extraversion in novel individuals”, *Social cognitive and affective neuroscience*, vol. 13, Fev 2018, pp. 224–232.
- [43] Hulme, C.; Snowling, M. J. “Reading disorders and dyslexia”, *Current Opinion in Pediatrics*, vol. 28–6, Dez 2016, pp. 731–735.
- [44] Ichikawa, N.; Lisi, G.; Yahata, N.; Okada, G.; Takamura, M.; Hashimoto, R.-i.; Yamada, T.; Yamada, M.; Suhara, T.; Moriguchi, S.; Mimura, M.; Yoshihara, Y.; Takahashi, H.; Kasai, K.; Kato, N.; Yamawaki, S.; Seymour, B.; Kawato, M.; Morimoto, J.; Okamoto, Y. “Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants”, *Scientific Reports*, vol. 10–1, Fev 2020, pp. 3542.
- [45] Imani, V.; Prakash, M.; Zare, M.; Tohka, J. “Comparison of single and multitask learning for predicting cognitive decline based on mri data”, *IEEE Access*, vol. 9, Set 2021, pp. 154275–154291.
- [46] Jafri, M. J.; Pearlson, G. D.; Stevens, M.; Calhoun, V. D. “A method for functional network connectivity among spatially independent resting-state components in schizophrenia”, *NeuroImage*, vol. 39–4, Fev 2008, pp. 1666–1681.
- [47] Joel, S. E.; Caffo, B. S.; van Zijl, P. C. M.; Pekar, J. J. “On the relationship between seed-based and ICA-based measures of functional connectivity”, *Magnetic Resonance in Medicine*, vol. 66–3, Mar 2011, pp. 644–657.
- [48] Kasuya, E. “On the use of r and r squared in correlation and regression”, *Ecological Research*, vol. 34–1, Dez 2019, pp. 235–236.
- [49] Kawahara, J.; Brown, C. J.; Miller, S. P.; Booth, B. G.; Chau, V.; Grunau, R. E.; Zwicker, J. G.; Hamarneh, G. “BrainNetCNN: Convolutional neural networks for brain networks towards predicting neurodevelopment”, *NeuroImage*, vol. 146, Fev 2017, pp. 1038–1049.

- [50] Kazi, A.; Farghadani, S.; Navab, N. "la-gcn: Interpretable attention based graph convolutional network for disease prediction", *arXiv Preprint*, vol. 2103.15587, Mar 2021, pp. 1–10.
- [51] Khan, M. R.; Blumenstock, J. E. "Multi-gcn: Graph convolutional networks for multi-view networks, with applications to global poverty", Jan 2019, pp. 8.
- [52] Kitchenham, B. A.; Budgen, D.; Brereton, P. "Evidence-Based Software Engineering and Systematic Reviews". Chapman & Hall/CRC, 2015, pp. 433.
- [53] Ko, C.-H.; Hsieh, T.-J.; Wang, P.-W.; Lin, W.-C.; Yen, C.-F.; Chen, C.-S.; Yen, J.-Y. "Altered gray matter density and disrupted functional connectivity of the amygdala in adults with internet gaming disorder", *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 57, Mar 2015, pp. 185–192.
- [54] Krishnan, A.; Williams, L. J.; McIntosh, A. R.; Abdi, H. "Partial least squares (PLS) methods for neuroimaging: A tutorial and review", *NeuroImage*, vol. 56–2, Mai 2011, pp. 455–475.
- [55] Lama, R. K.; Kwon, G.-R. "Diagnosis of alzheimer's disease using brain network", *Frontiers in Neuroscience*, vol. 15, Fev 2021, pp. 605115.
- [56] LeCun, Y.; Bengio, Y.; Hinton, G. "Deep learning", *Nature*, vol. 521–7553, Mai 2015, pp. 436–444.
- [57] Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE*, 1998, pp. 2278–2324.
- [58] Lee, J.; Lee, I.; Kang, J. "Self-attention graph pooling", *arXiv Preprint*, vol. 1904.08082, Abr 2019, pp. 10.
- [59] Li, S.; Huang, H. "Predicting severity of autism spectrum disorder based on multi-center multi-modality data". In: *Proceedings of the 2019 2nd International Conference on Digital Medicine and Image Processing*, 2019, pp. 12–16.
- [60] Li, X.; Zhou, Y.; Dvornek, N.; Zhang, M.; Gao, S.; Zhuang, J.; Scheinost, D.; Staib, L. H.; Ventola, P.; Duncan, J. S. "Braingnn: Interpretable brain graph neural network for fmri analysis", *Medical Image Analysis*, vol. 74, Mar 2021, pp. 102233.
- [61] Lian, C.; Liu, M.; Wang, L.; Shen, D. "Multi-task weakly-supervised attention network for dementia status estimation with structural mri", *IEEE Transactions on Neural Networks and Learning Systems*, Ago 2021, pp. 1–13.
- [62] Liu, M.; Zhang, J.; Adeli, E.; Shen, D. "Joint classification and regression via deep multi-task multi-channel learning for alzheimer's disease diagnosis", *IEEE Transactions on Biomedical Engineering*, vol. 66–5, Mai 2019, pp. 1195–1206.

- [63] Long, M.; Wang, J. “Learning multiple tasks with deep relationship networks”, *arXiv Preprint*, vol. abs/1506.02117, Nov 2015, pp. 10.
- [64] Marek, S.; Tervo-Clemmens, B.; Calabro, F. J.; Montez, D. F.; Kay, B. P.; Hatoum, A. S.; Donohue, M. R.; Foran, W.; Miller, R. L.; Hendrickson, T. J.; Malone, S. M.; Kandala, S.; Feczko, E.; Miranda-Dominguez, O.; Graham, A. M.; Earl, E. A.; Perrone, A. J.; Cordova, M.; Doyle, O.; Moore, L. A.; Conan, G. M.; Uriarte, J.; Snider, K.; Lynch, B. J.; Wilgenbusch, J. C.; Pengo, T.; Tam, A.; Chen, J.; Newbold, D. J.; Zheng, A.; Seider, N. A.; Van, A. N.; Metoki, A.; Chauvin, R. J.; Laumann, T. O.; Greene, D. J.; Petersen, S. E.; Garavan, H.; Thompson, W. K.; Nichols, T. E.; Yeo, B. T. T.; Barch, D. M.; Luna, B.; Fair, D. A.; Dosenbach, N. U. F. “Reproducible brain-wide association studies require thousands of individuals”, *Nature*, vol. 603–7902, Mar 2022, pp. 654–660.
- [65] Mendes, S. L.; Pinaya, W. H. L.; Pan, P.; Sato, J. R. “Estimating gender and age from brain structural MRI of children and adolescents: A 3d convolutional neural network multitask learning model”, *Computational Intelligence and Neuroscience*, vol. 2021, Mai 2021, pp. 1–12.
- [66] Mitchell, T. M. “Machine learning”. New York: McGraw-Hill, 1997, pp. 447.
- [67] Mohanty, R.; Sinha, A.; Remsik, A.; Dodd, K.; Young, B.; Jacobson, T.; McMillan, M.; Thoma, J.; Advani, H.; Nair, V.; Kang, T.; Caldera, K.; Edwards, D.; Williams, J.; Prabhakaran, V. “Early findings on functional connectivity correlates of behavioral outcomes of brain-computer interface stroke rehabilitation using machine learning”, *Frontiers in Neuroscience*, vol. 12, Set 2018, pp. 19.
- [68] Möttus, R.; Wood, D.; Condon, D. M.; Back, M. D.; Baumert, A.; Costantini, G.; Epskamp, S.; Greiff, S.; Johnson, W.; Lukaszewski, A.; Murray, A.; Revelle, W.; Wright, A. G.; Yarkoni, T.; Ziegler, M.; Zimmermann, J. “Descriptive, predictive and explanatory personality research: Different goals, different approaches, but a shared need to move beyond the big few traits”, *European Journal of Personality*, vol. 34–6, Dez 2020, pp. 1175–1201.
- [69] Noble, S.; Spann, M. N.; Tokoglu, F.; Shen, X.; Constable, R. T.; Scheinost, D. “Influences on the Test–Retest Reliability of Functional Connectivity MRI and its Relationship with Behavioral Utility”, *Cerebral Cortex*, vol. 27–11, Set 2017, pp. 5415–5429.
- [70] O’Sullivan, S.; Jeanquartier, F.; Jean-Quartier, C.; Holzinger, A.; Shiebler, D.; Moon, P.; Angione, C. “Developments in AI and machine learning for neuroimaging”. In: *Artificial Intelligence and Machine Learning for Digital Pathology*, Springer International Publishing, 2020, pp. 307–320.

- [71] Pagnozzi, A. M.; Conti, E.; Calderoni, S.; Fripp, J.; Rose, S. E. “A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective”, *International Journal of Developmental Neuroscience*, vol. 71–1, Ago 2018, pp. 68–82.
- [72] Pellegrini, E.; Ballerini, L.; del C. Valdes Hernandez, M.; Chappell, F. M.; González-Castro, V.; Anblagan, D.; Danso, S.; Muñoz-Maniega, S.; Job, D.; Pernet, C.; Mair, G.; MacGillivray, T. J.; Trucco, E.; Wardlaw, J. M. “Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review”, *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 10–1, Jan 2018, pp. 519–535.
- [73] Pereira, F.; Mitchell, T.; Botvinick, M. “Machine learning classifiers and fMRI: A tutorial overview”, *NeuroImage*, vol. 45–1, Mar 2009, pp. S199–S209.
- [74] Rahim, M.; Thirion, B.; Bzdok, D.; Buvat, I.; Varoquaux, G. “Joint prediction of multiple scores captures better individual traits from brain images”, *NeuroImage*, vol. 158, Set 2017, pp. 145–154.
- [75] Ranjan, R.; Sankaranarayanan, S.; Castillo, C. D.; Chellappa, R. “An all-in-one convolutional neural network for face analysis”. In: Proceedings of 2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017), 2017, pp. 17–24.
- [76] Rapuano, K. M.; Rosenberg, M. D.; Maza, M. T.; Dennis, N. J.; Dorji, M.; Greene, A. S.; Horien, C.; Scheinost, D.; Constable, R. T.; Casey, B. “Behavioral and brain signatures of substance use vulnerability in childhood”, *Developmental Cognitive Neuroscience*, vol. 46, Dez 2020, pp. 100878.
- [77] Rey, D.; Neuhäuser, M. “Wilcoxon-Signed-Rank Test”. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 1658–1659.
- [78] Rosenberg, M.; Finn, E.; Scheinost, D.; Constable, R.; Chun, M. “Characterizing attention with predictive network models”, *Trends in Cognitive Sciences*, vol. 21–4, Abr 2017, pp. 290–302.
- [79] Rosenberg, M. D.; Casey, B. J.; Holmes, A. J. “Prediction complements explanation in understanding the developing brain”, *Nature Communications*, vol. 9–1, Fev 2018.
- [80] Rosenberg, S. A.; Restifo, N. P. “Adoptive cell transfer as personalized immunotherapy for human cancer”, *Science*, vol. 348–6230, Abr 2015, pp. 62–68.
- [81] Russell, S.; Norvig, P. “Artificial Intelligence: A Modern Approach”. Upper Saddle River: Prentice Hall, 2010, pp. 1132.

- [82] Sala-Llonch, R.; Smith, S. M.; Woolrich, M.; Duff, E. P. “Spatial parcellations, spectral filtering, and connectivity measures in fmri: Optimizing for discrimination”, *Human Brain Mapping*, vol. 40–2, Fev 2019, pp. 407–419.
- [83] Shao, L.; Fu, C.; You, Y.; Fu, D. “Classification of ASD based on fMRI data with deep learning”, *Cognitive Neurodynamics*, vol. 15–6, Mai 2021, pp. 961–974.
- [84] Shen, X.; Finn, E. S.; Scheinost, D.; Rosenberg, M. D.; Chun, M. M.; Papademetris, X.; Constable, R. T. “Using connectome-based predictive modeling to predict individual behavior from brain connectivity”, *Nature Protocols*, vol. 12–3, Fev 2017, pp. 506–518.
- [85] Smith, S. M. “The future of FMRI connectivity”, *NeuroImage*, vol. 62–2, Ago 2012, pp. 1257–1266.
- [86] Smith, S. M.; Jenkinson, M.; Woolrich, M. W.; Beckmann, C. F.; Behrens, T. E.; Johansen-Berg, H.; Bannister, P. R.; Luca, M. D.; Drobnjak, I.; Flitney, D. E.; Niazy, R. K.; Saunders, J.; Vickers, J.; Zhang, Y.; Stefano, N. D.; Brady, J. M.; Matthews, P. M. “Advances in functional and structural MR image analysis and implementation as FSL”, *NeuroImage*, vol. 23, Jan 2004, pp. 208–S21.
- [87] Smith, S. M.; Miller, K. L.; Salimi-Khorshidi, G.; Webster, M.; Beckmann, C. F.; Nichols, T. E.; Ramsey, J. D.; Woolrich, M. W. “Network modelling methods for fmri”, *NeuroImage*, vol. 54–2, Jan 2011, pp. 875–891.
- [88] Smith, S. M.; Vidaurre, D.; Beckmann, C. F.; Glasser, M. F.; Jenkinson, M.; Miller, K. L.; Nichols, T. E.; Robinson, E. C.; Salimi-Khorshidi, G.; Woolrich, M. W.; Barch, D. M.; Uğurbil, K.; Essen, D. C. V. “Functional connectomics from resting-state fMRI”, *Trends in Cognitive Sciences*, vol. 17–12, Dez 2013, pp. 666–682.
- [89] Sprawls, P. “Magnetic Resonance Imaging: Principles, Methods, and Techniques”. Medical Physics Publishing Corporation, 2000, pp. 246–255.
- [90] Sui, J.; Jiang, R.; Bustillo, J.; Calhoun, V. “Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: Methods and promises”, *Biological Psychiatry*, vol. 88–11, Dez 2020, pp. 818–828.
- [91] Sun, M.; Baytas, I. M.; Zhan, L.; Wang, Z.; Zhou, J. “Subspace network: Deep multi-task censored regression for modeling neurodegenerative diseases”. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 2259–2268.
- [92] Tang, Z.; Xu, Y.; Jin, L.; Aibaidula, A.; Lu, J.; Jiao, Z.; Wu, J.; Zhang, H.; Shen, D. “Deep learning of imaging phenotype and genotype for predicting overall survival time

- of glioblastoma patients”, *IEEE Transactions on Medical Imaging*, vol. 39–6, Jan 2020, pp. 2100–2109.
- [93] Thung, K.-H.; Wee, C.-Y. “A brief review on multi-task learning”, *Multimedia Tools Appl.*, vol. 77–22, Nov 2018, pp. 29705–29725.
- [94] Tyshchenko, Y. “Depression and anxiety detection from blog posts data”, Dissertação de Mestrado, University of Tartu, 2018.
- [95] Tzourio-Mazoyer, N.; Landeau, B.; Papathanassiou, D.; Crivello, F.; Etard, O.; Delcroix, N.; Mazoyer, B.; Joliot, M. “Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain”, *NeuroImage*, vol. 15–1, jan. 2002, pp. 273–289.
- [96] Vandermosten, M.; Hoeft, F.; Norton, E. S. “Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: a review and quantitative meta-analysis”, *Current Opinion in Behavioral Sciences*, vol. 10, Ago 2016, pp. 155–161.
- [97] Varatharajah, Y.; Saboo, K.; Iyer, R.; Przybelski, S.; Schwarz, C.; Petersen, R.; Jack, C.; Vemuri, P. “A joint model for predicting structural and functional brain health in elderly individuals”. In: *Proceedings of 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019, pp. 1657–1664.
- [98] Varoquaux, G.; Gramfort, A.; Pedregosa, F.; Michel, V.; Thirion, B. “Multi-subject dictionary learning to segment an atlas of brain spontaneous activity”. In: *Proceedings of Information Processing in Medical Imaging*, 2011, pp. 562–573.
- [99] Varoquaux, G.; Gramfort, A.; Poline, J. B.; Thirion, B. “Brain covariance selection: better individual functional connectivity models using population prior”. In: *Proceedings of Advances in Neural Information Processing Systems*, 2010, pp. 1–11.
- [100] Varoquaux, G.; Thirion, B. “How machine learning is shaping cognitive neuroimaging”, *GigaScience*, vol. 3, Dez 2014, pp. 28.
- [101] Wan, J.; Zhang, Z.; Yan, J.; Li, T.; Rao, B. D.; Fang, S.; Kim, S.; Risacher, S. L.; Saykin, A. J.; Shen, L. “Sparse bayesian multi-task learning for predicting cognitive outcomes from neuroimaging measures in alzheimer’s disease”. In: *Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 940–947.
- [102] Wang, J.; Wang, Q.; Zhang, H.; Chen, J.; Wang, S.; Shen, D. “Sparse multiview task-centralized ensemble learning for asd diagnosis based on age- and sex-related functional connectivity patterns”, *IEEE Transactions on Cybernetics*, vol. 49–8, Ago 2019, pp. 3141–3154.

- [103] Wang, M.; Zhang, D.; Shen, D.; Liu, M. “Multi-task exclusive relationship learning for alzheimer’s disease progression prediction with longitudinal data”, *Medical Image Analysis*, vol. 53, Abr 2019, pp. 111–122.
- [104] Wang, X.; Li, Q.; Zhao, Y.; He, Y.; Ma, B.; Fu, Z.; Li, S. “Decomposition of individual-specific and individual-shared components from resting-state functional connectivity using a multi-task machine learning method”, *NeuroImage*, vol. 238, Set 2021, pp. 118252.
- [105] Wu, E. X.; Liaw, G. J.; Goh, R. Z.; Chia, T. T.; Chee, A. M.; Obana, T.; Rosenberg, M. D.; Yeo, B. T.; Asplund, C. L. “Overlapping attentional networks yield divergent behavioral predictions across tasks: Neuromarkers for diffuse and focused attention?”, *NeuroImage*, vol. 209, Abr 2020, pp. 116535.
- [106] Xu, M.; Liang, X.; Ou, J.; Li, H.; Jia Luo, Y.; Tan, L. H. “Sex differences in functional brain networks for language”, *Cerebral Cortex*, vol. 30–3, Set 2019, pp. 1528–1537.
- [107] Yahata, N.; Morimoto, J.; Hashimoto, R.; Lisi, G.; Shibata, K.; Kawakubo, Y.; Kuwabara, H.; Kuroda, M.; Yamada, T.; Megumi, F.; Imamizu, H.; Sr, J. E. N.; Takahashi, H.; Okamoto, Y.; Kasai, K.; Kato, N.; Sasaki, Y.; Watanabe, T.; Kawato, M. “A small number of abnormal brain connections predicts adult autism spectrum disorder”, *Nature Communications*, vol. 7–1, Abr 2016, pp. 11254.
- [108] Yan, J.; Huang, H.; Risacher, S. L.; Kim, S.; Inlow, M.; Moore, J. H.; Saykin, A. J.; Shen, L. “Network-guided sparse learning for predicting cognitive outcomes from mri measures”. In: *Multimodal Brain Image Analysis*, 2013, pp. 202–210.
- [109] Yang, Y.; Cui, Y.; Xu, K.; Liu, B.; Song, M.; Chen, J.; Wang, H.; Chen, Y.; Guo, H.; Li, P.; Lu, L.; Lv, L.; Wan, P.; Wang, H.; Yan, H.; Yan, J.; Zhang, H.; Zhang, D.; Jiang, T. “Distributed functional connectivity impairment in schizophrenia: A multi-site study”. In: *Proceedings of 2nd IET International Conference on Biomedical Image and Signal Processing*, 2017, pp. 1–6.
- [110] Yeo, B. T. T.; Krienen, F. M.; Sepulcre, J.; Sabuncu, M. R.; Lashkari, D.; Hollinshead, M.; Roffman, J. L.; Smoller, J. W.; Zöllei, L.; Polimeni, J. R.; Fischl, B.; Liu, H.; Buckner, R. L. “The organization of the human cerebral cortex estimated by intrinsic functional connectivity”, *Journal of Neurophysiology*, vol. 106–3, Set 2011, pp. 1125–1165.
- [111] Zhang, L.; Yang, Q.; Liu, X.; Guan, H. “Rethinking hard-parameter sharing in multi-task learning”, *arXiv Preprint*, vol. 2107.11359, Jul 2021, pp. 6.
- [112] Zhang, Y.; Wei, Y.; Yang, Q. “Learning to multitask”. In: *Proceedings of Advances in Neural Information Processing Systems*, 2018, pp. 745.

- [113] Zhang, Y.; Yang, Q. “An overview of multi-task learning”, *National Science Review*, vol. 5–1, Set 2017, pp. 30–43.
- [114] Zhang, Y.; Yang, Q. “A survey on multi-task learning”, *arXiv Preprint*, vol. 1707.08114, Mar 2017, pp. 1–3.
- [115] Zuo, X.-N.; Kelly, C.; Adelstein, J. S.; Klein, D. F.; Castellanos, F. X.; Milham, M. P. “Reliable intrinsic connectivity networks: Test–retest evaluation using ICA and dual regression approach”, *NeuroImage*, vol. 49–3, Fev 2010, pp. 2163–2177.

APÊNDICE A – RESULTADOS DO TESTE DE WILCOXON

IC	Single PMAT - MTL		Single MMSE Score - MTL		Single Gender - MTL	
	W	p	W	p	W	p
5	12564.0	1.0189e-23	7461.5	2.0171e-2	10978.0	2.6421e-14
8	12791.0	2.0599e-25	9819.0	3.0828e-11	11601.0	1.0024e-19
9	10469.0	7.9590e-13	12227.0	2.5390e-21	12476.0	4.4388e-23
10	12201.0	4.8446e-24	11729.5	1.3020e-20	12523.5	2.0088e-23
11	12249.5	2.1305e-23	12111.0	1.5780e-20	13054.0	1.8746e-27
12	11857.0	1.8188e-22	10719.5	6.8381e-13	12065.5	5.0096e-23
16	9173.0	8.5281e-6	8963.0	1.8733e-5	8435.0	6.1595e-4
28	11452.0	2.5004e-16	11224.0	5.3696e-15	13089.0	9.9377e-28
31	8688.0	6.4119e-5	7753.0	1.2647e-2	8232.5	1.9286e-3
34	9503.0	2.4016e-7	11053.5	4.8452e-14	11613.5	2.6120e-17
35	12224.0	3.0853e-22	11698.0	7.7838e-18	13007.0	4.4745e-27
37	11619.0	9.5055e-21	12733.0	5.6918e-25	13186.0	1.6835e-28
40	11052.0	4.9382e-14	9396.0	6.0676e-7	12311.0	7.3131e-23
42	12280.0	1.0891e-21	9284.0	3.5866e-9	11987.0	1.3771e-20
48	10657.0	2.5834e-19	13013.0	4.0061e-27	13203.0	1.2251e-28
50	11769.0	4.0482e-19	9697.0	1.1298e-7	12733.0	5.0172e-26
51	11081.0	6.9408e-15	9107.0	1.3952e-5	12758.0	3.6920e-25
56	10665.0	5.4019e-12	11127.5	1.8834e-14	10685.0	1.0397e-12
57	10474.0	2.1580e-16	11479.0	2.8925e-17	11312.0	3.0278e-16
58	12588.5	6.4331e-25	11843.5	9.2425e-19	13139.0	4.0238e-28
61	13150.0	3.2833e-28	12847.0	7.7787e-26	13195.0	1.4242e-28
63	13193.0	1.4805e-28	12909.0	2.5937e-26	13195.0	1.3963e-28
70	10447.0	6.3464e-11	8842.0	2.1368e-5	12460.0	5.7912e-23
73	8896.0	6.2253e-5	11053.0	4.8761e-14	12896.0	3.2692e-26
77	9238.0	5.1924e-6	11101.5	5.2917e-15	13054.0	1.9090e-27
79	11039.0	5.8205e-14	9785.0	5.0825e-8	11445.0	4.7123e-17
98	8342.0	1.0546e-3	8730.0	1.8580e-4	12249.0	1.7893e-21
100	11725.5	1.8044e-22	8998.0	6.5631e-6	8892.5	6.6451e-6
108	8949.5	4.3087e-5	11142.0	1.5623e-14	12879.0	3.5923e-27
109	8824.0	1.0093e-4	10378.0	1.3472e-10	9245.5	2.1186e-6
115	11657.0	5.8261e-22	12314.5	6.8988e-23	12948.0	1.2933e-26
117	10788.0	1.2702e-12	9831.0	3.3196e-8	10472.0	4.8161e-11
118	11342.0	1.1179e-15	11979.0	1.2055e-19	9623.0	2.1765e-7

Tabela A.1 continuação da página anterior

IC	Single PMAT - MTL		Single MMSE Score - MTL		Single Gender - MTL	
	W	p	W	p	W	p
132	10505.0	3.3371e-11	8462.0	9.3132e-4	8233.0	3.1826e-3
134	12599.0	5.6496e-24	10244.0	5.5988e-10	11617.0	2.4853e-17
141	10040.0	4.4573e-9	9829.0	3.3820e-8	9161.0	9.3347e-6
142	12458.0	5.9861e-23	8834.0	9.4457e-5	7612.0	4.5528e-2
146	8805.0	1.1440e-4	11394.0	5.5307e-16	12896.0	2.6330e-27
147	11828.0	1.1630e-18	8360.0	1.6372e-3	11590.0	3.6449e-17
152	11365.0	8.1965e-16	9501.0	6.2100e-7	11965.0	1.4914e-19
153	9433.0	1.0947e-6	7979.0	1.0623e-2	11555.0	5.9699e-17
154	7627.0	4.3178e-2	8396.0	1.3458e-3	12567.0	9.3593e-25
158	10859.0	5.4045e-13	8658.0	2.9186e-4	11500.0	1.2875e-16
160	9984.0	7.7230e-9	8720.0	1.9799e-4	9173.0	8.5282e-6
162	12398.0	1.6074e-22	12520.0	2.1125e-24	12956.0	8.7283e-28
164	12953.0	1.1758e-26	8786.5	3.1986e-5	8967.0	1.5523e-6
166	12153.0	8.1816e-21	12407.0	1.3869e-22	12599.0	4.6624e-26
175	10217.0	7.4161e-10	8534.0	6.1526e-4	8206.0	3.6461e-3
179	10985.0	1.1465e-13	8585.0	4.5496e-4	8023.0	5.6064e-3
183	11661.0	2.0426e-18	7698.0	3.3351e-2	9718.0	9.3557e-8
184	8096.0	6.2229e-3	10181.0	1.0755e-9	12482.0	4.0211e-23
187	12410.0	1.3202e-22	9112.0	6.0977e-6	9590.0	2.9017e-7
193	11673.0	1.1160e-17	11642.0	1.7402e-17	12481.0	4.1318e-24
199	12948.0	1.2935e-26	12385.0	1.9883e-22	12461.0	5.6962e-23

Tabela A.1: Dados resultantes dos testes de Wilcoxon dos componentes que apresentaram mudanças superiores significativas ao modelo *single*.

APÊNDICE B – SUPERFÍCIES CEREBRAIS COM ESCORES (*STATS MAP*)

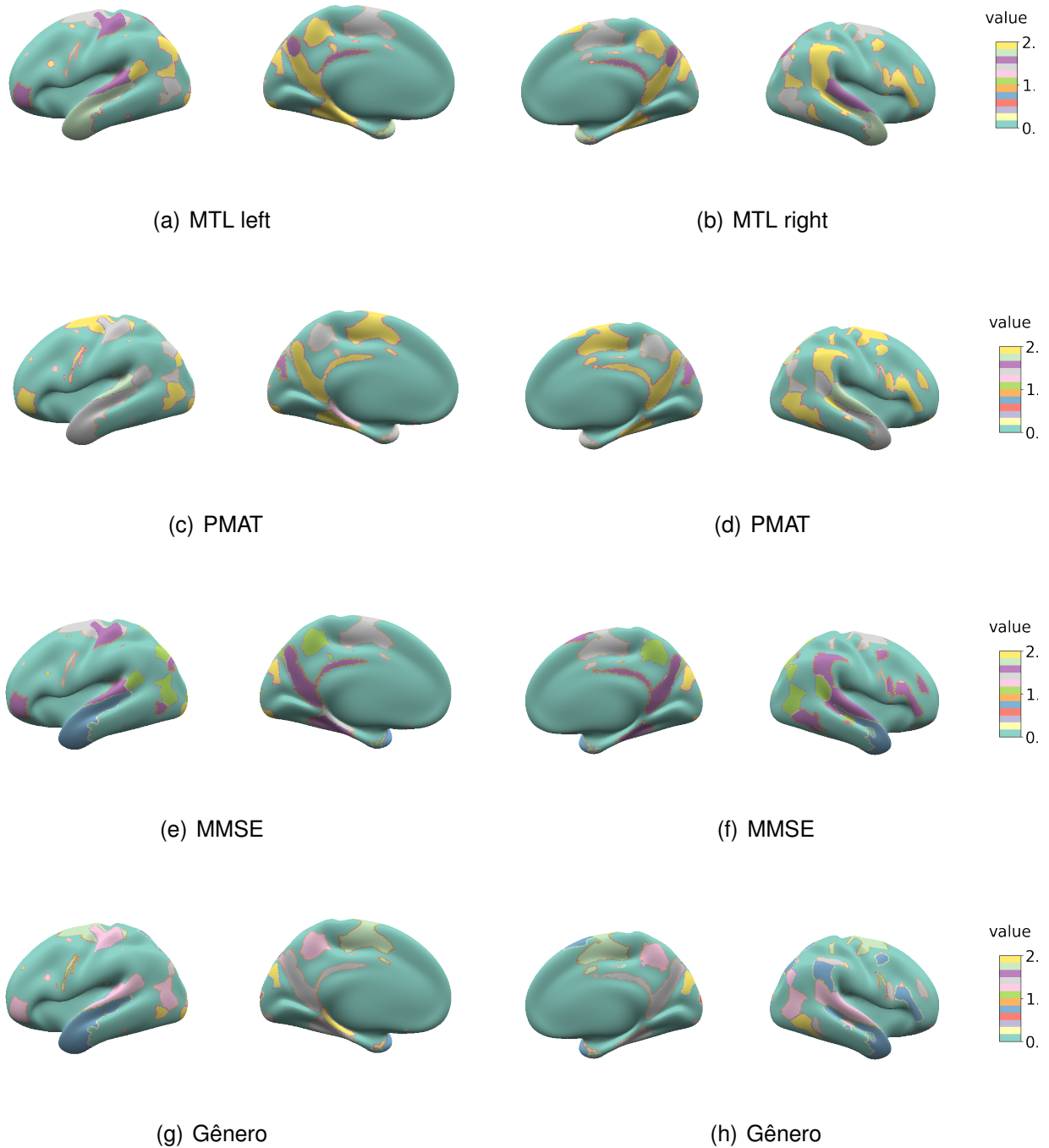


Figura B.1: Visualização da mediana dos escores das regiões que rejeitaram a hipótese nula pelo teste de Wilcoxon.



Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria de Graduação
Av. Ipiranga, 6681 - Prédio 1 - 3º. andar
Porto Alegre - RS - Brasil
Fone: (51) 3320-3500 - Fax: (51) 3339-1564
E-mail: prograd@pucrs.br
Site: www.pucrs.br