

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

GUILHERME RODRIGUES PEDROLLO

**INVESTIGAÇÃO DE MARCADORES
BIOLÓGICOS DE TRANSTORNO DO
DÉFICIT DE ATENÇÃO COM
HIPERATIVIDADE EM DADOS DE
ELETROENCEFALOGRAFIA E DE
RESSONÂNCIA MAGNÉTICA
FUNCIONAL COM INTELIGÊNCIA
COMPUTACIONAL**

Porto Alegre
2024

GUILHERME RODRIGUES PEDROLLO

**INVESTIGAÇÃO DE MARCADORES
BIOLÓGICOS DE TRANSTORNO DO
DÉFICIT DE ATENÇÃO COM
HIPERATIVIDADE EM DADOS DE
ELETROENCEFALOGRAFIA E DE
RESSONÂNCIA MAGNÉTICA
FUNCIONAL COM INTELIGÊNCIA
COMPUTACIONAL**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Eletrônica

ORIENTADOR: Prof. Dr. Alexandre Balbinot

CO-ORIENTADOR: Prof. Dr. Alexandre Rosa Franco

Porto Alegre
2024

GUILHERME RODRIGUES PEDROLLO

**INVESTIGAÇÃO DE MARCADORES
BIOLÓGICOS DE TRANSTORNO DO
DÉFICIT DE ATENÇÃO COM
HIPERATIVIDADE EM DADOS DE
ELETROENCEFALOGRAFIA E DE
RESSONÂNCIA MAGNÉTICA
FUNCIONAL COM INTELIGÊNCIA
COMPUTACIONAL**

Esta tese foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dr. Alexandre Balbinot, UFRGS

Doutor pela UFRGS – Porto Alegre, Brasil

Banca Examinadora:

Prof. Dr. Fernando Santos Osório, ICMC – USP

Doutor pela Institut National Polytechnique de Grenoble – Grenoble, França

Prof. Dr. Francisco José Fraga da Silva, PPGNCG – UFABC

Doutor pela Instituto Tecnológico de Aeronáutica – São José dos Campos, Brasil

Prof. Dr. Tiago Oliveira Weber, PPGEE – UFRGS

Doutor pela Universidade de São Paulo – São Paulo, Brasil

Coordenador do PPGEE: _____

Prof. Dr. Jeferson Vieira Flores

Porto Alegre, Março de 2024.

DEDICATÓRIA

Dedico este trabalho aos meus pais, em especial pela dedicação e apoio em todos os momentos difíceis.

AGRADECIMENTOS

Aos meus pais, Olavo e Márcia, pelo apoio, companhia, incentivo e conselhos que foram essenciais para a minha formação.

À UFRGS e ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de realização de trabalhos em minha área de pesquisa.

Ao orientador Professor Alexandre Balbinot e ao co-orientador Alexandre Rosa Franco por me apresentarem à esta área de pesquisa, pela oportunidade, passagem de conhecimento, pelos conselhos e recomendações e pela atenção ao meu trabalho.

Ao Professor Tiago Weber pela oportunidade de realizar o estágio docência junto à sua disciplina.

Aos colegas do Laboratório de Instrumentação Eletroeletrônica e Inteligência Artificial (IEE-IA) da UFRGS, em especial, Maurício Tosin, Ana Karen Apolo, Vinicius Cene pela convivência, incentivo, troca de conhecimentos e recomendação.

Aos familiares e amigos pela compreensão e pelo apoio emocional.

À CAPES pela provisão da bolsa de doutorado.

RESUMO

O transtorno de déficit de atenção e hiperatividade (TDAH) aflige cerca de 5 a 7% da população infantil no mundo. O tratamento adequado do TDAH, exige que este seja identificado corretamente, assim como o seu subtipo. Para tanto, investigaram-se sinais de Eletroencefalografia (EEG) e imagens de ressonância magnética funcional (fMRI) de 856 e de 809 indivíduos de cada modalidade de dados, respectivamente. Esta Tese apresenta um novo método para identificar TDAH e seus subtipos através de meta-classificadores que combinam informações de possíveis biomarcadores obtidos por classificadores primários. As classificações dos classificadores primários foram utilizadas como características enviadas à entrada dos meta-classificadores. Experimentou-se o uso de quatro algoritmos como meta-classificadores: *XGBoost*, *AdaBoost*, *Ridge* e Voto Majoritário. As otimizações dos pesos foram utilizadas para a identificação dos biomarcadores. Em EEG, utilizaram-se, como classificadores primários, florestas aleatórias (RF) para processar características extraídas em cada faixa de frequência. O desempenho da RFs foi comparado à máquinas de vetor de suporte (SVMs). Gerou-se matrizes de correlação a partir das imagens de fMRI, as quais foram enviadas às Redes Neurais Convolucionais (CNN) que atuaram como classificadores primários. O cérebro foi dividido em oito áreas que foram avaliadas separadamente como possíveis biomarcadores de fMRI. A Votação Majoritária obteve $97,6 \pm 2,1\%$ de *accuracy* na identificação de TDAH a partir de fMRI. Os meta-classificadores apontaram as redes de conectividade límbica, frontoparietal, visual e *default* como os principais biomarcadores de TDAH. Para a distinção de subtipos, o meta-classificador *XGBoost* que atua a partir dos sinais de EEG se destacou, obtendo uma *accuracy* de $97,8 \pm 0,3\%$. O estudo apontou a importância de observar as bandas Delta e Teta para o diagnóstico de subtipos. A metodologia apresentada neste trabalho provou-se eficaz para a detecção de TDAH e na identificação dos subtipos deste transtorno. Além disto, os meta-classificadores apontaram biomarcadores que podem ser relevantes para estudos futuros e o desenvolvimento de novos tratamentos de TDAH.

Palavras-chave: Meta-classificadores, Redes Neurais Convolucionais, Biomarcadores, *Random Forest*, Diagnóstico Automatizado.

ABSTRACT

An estimate of 5 to 7 percent of children worldwide suffer from attention deficit hyperactivity disorder (ADHD). Accurate identification of ADHD and its subtype are essential for effective therapy. In order to achieve this, indicators in functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) signals were examined from 856 and 809 subjects, respectively. In this thesis, a novel approach to identifying ADHD and its subtypes is presented. The ensembles (AdaBoost, Ridge, XGBoost, and Majority Voting) receives the predictions of the base-classifiers (each working with a distinct potential biomarker) as input features and combine them to enhance the accuracy of the system. The weight optimizations of the ensembles pointed out the biomarkers. In EEG, random forests (RF) were used as base-classifiers to process features extracted in each frequency band. The performance of RFs was compared to support vector machines (SVMs). Convolutional Neural Networks (CNN) (fMRI base-classifiers) received correlation matrix maps extracted from the fMRI. Eight regions of the brain maps were identified and assessed independently as possible fMRI biomarkers. With fMRI data, the Majority Vote identified ADHD with $97.6\pm 2.1\%$ accuracy. The ensembles identified the limbic, frontoparietal, visual, and default connectivity networks as biomarkers for ADHD. With an accuracy of $97.8\pm 0.3\%$, the XGBoost meta-classifier that works with EEG signals proved to be effective in differentiating subtypes. The study emphasized the importance of monitoring the delta and theta bands in order to identify ADHD subgroups. The methodology presented in this thesis has been shown to be successful in recognizing ADHD and identifying the subtypes of this disorder. The ensembles also identified biomarkers that could be important for further research and the development of new treatments for ADHD.

Keywords: *Ensembles, Convolutional Neural Networks, biomarkers, Random Forest, Automatic Diagnostic.*

LISTA DE ILUSTRAÇÕES

Figura 1 –	Mapa de citações de autores.	28
Figura 2 –	Mapa de co-citações de autores de EEG.	31
Figura 3 –	Mapa de co-citações de autores de fMRI.	32
Figura 4 –	Mapa de citações de palavras-chave.	33
Figura 5 –	Exemplo de Árvore de Decisão.	48
Figura 6 –	Funcionamento de uma SVM.	49
Figura 7 –	Resumo gráfico da metodologia da Tese na parte de EEG.	63
Figura 8 –	Resumo gráfico da metodologia da Tese na parte de fMRI.	64
Figura 9 –	Paradigma de classificação das amostras.	64
Figura 10 –	Diagrama de blocos dos estágios de pré-processamento. Com exceção do primeiro bloco, todas as etapas são digitais.	66
Figura 11 –	Sinais brutos (<i>raw</i>) de EEG de um participante com TDAH hiperativo (NDARAM873GAC).	67
Figura 12 –	Diagrama de blocos com o algoritmo ASR para remoção de ruído.	68
Figura 13 –	Diagrama de blocos com a parte inicial do pré-processamento: filtros e algoritmos de remoção de ruído.	68
Figura 14 –	Histogramas (a) da proporção de segmentos e (b) da proporção de canais remanescentes por participante.	71
Figura 15 –	Bordas de Transição e seus efeitos em um mesmo trecho do conjunto de sinais de EEG em cada faixa de frequência, após a aplicação do ASR.	74
Figura 16 –	Processo de aleatorização e de separação das amostras de EEG em conjuntos de treinamento e de teste.	75
Figura 17 –	Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Delta.	76
Figura 18 –	Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Beta.	77
Figura 19 –	Otimização AG da RF para classificação TDAH/Controle para banda Delta.	78
Figura 20 –	Otimização AG da SVM para classificação TDAH/Controle para banda Teta.	82
Figura 21 –	Fusão dos resultados de classificação baseados nas bandas de EEG.	83
Figura 22 –	Otimização AG da Fusão das bandas de EEG por <i>XGBoost</i> para classificação TDAH/Controle.	87

Figura 23 – Curva de Otimização da fusão <i>XGBoost</i> para classificação TDAH-C/TDAH-I.	87
Figura 24 – Fluxograma do processamento de fMRI desde o dado bruto até a classificação final.	88
Figura 25 – Histogramas do controle de qualidade de fMRI com base no FDJ.	90
Figura 26 – Representação do parcelamento em 200 regiões de interesse conforme (CRADDOCK <i>et al.</i> , 2012). Código de cores arbitrariamente escolhidas por para destacar os limites entre cada região.	92
Figura 27 – Mapa de Yeo de 7 Segmentos. Legenda: Rede Visual (roxo), Rede Somatomotor (azul), Rede de Atenção Dorsal (verde), Rede de Atenção Ventral (Violeta), Rede Límbica (amarelo), Rede Frontoparietal (laranja), Rede de modo padrão (vermelho), demais <i>voxels</i> (preto).	93
Figura 28 – Redes de conectividade e sua representação na matriz de correlações.	95
Figura 29 – Topologia da CNN projetada.	96
Figura 30 – Fluxograma da fusão dos classificadores baseados nas redes de conectividade de fMRI com redes neurais convolucionais (CNN) como classificadores primários.	109
Figura 31 – Matrizes de confusão com as <i>accuracies</i> obtidas pelas RF para as bandas de EEG na classificação entre participantes com TDAH e do grupo controle no grupo de teste.	112
Figura 32 – Matrizes de confusão com as <i>accuracies</i> obtidas pelas RF para as bandas de EEG na classificação de subtipos (TDAH-C <i>versus</i> TDAH-I) no conjunto de teste.	113
Figura 33 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pelas SVM para as bandas de EEG na classificação TDAH <i>versus</i> Controle.	115
Figura 34 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pelas SVM para as bandas de EEG na classificação TDAH-C <i>versus</i> TDAH-I.	116
Figura 35 – Matrizes de confusão do conjunto teste "separado" da RF para classificação de TDAH.	117
Figura 36 – Matrizes de confusão do conjunto teste "separado" da RF para classificação de subtipos.	118
Figura 37 – Matrizes de confusão do conjunto teste "separado" da SVM para classificação de TDAH.	119
Figura 38 – Matrizes de confusão do conjunto teste "separado" da SVM para classificação de subtipos.	119
Figura 39 – Pesos usados em cada método de fusão para cada banda de EEG para obtenção do resultado de classificação de TDAH <i>versus</i> Controle.	121
Figura 40 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pelos métodos de fusão de EEG na classificação TDAH <i>versus</i> Controle.	123
Figura 41 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pelos métodos de fusão de EEG na classificação TDAH-C <i>versus</i> TDAH-I.	125
Figura 42 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pelo <i>XGBoost</i> para amostras reais nas classificações binária e de subtipos.	127
Figura 43 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pela CNN a partir das matrizes de correlação de fMRI nas classificações entre TDAH e grupo controle e de subtipos.	128

Figura 44 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pela CNN a partir das redes de conectividade de fMRI na classificação de TDAH <i>versus</i> Controle.	129
Figura 45 – Matrizes de confusão com as <i>accuracies</i> de teste obtidas pela CNN a partir das redes de conectividade de fMRI na classificação de TDAH-C <i>versus</i> TDAH-I.	131
Figura 46 – Pesos determinados por cada método de fusão para cada rede de conectividade para classificações de TDAH/Controle e TDAH-C/TDAH-I. A curva vermelha indica a <i>accuracy</i> total obtida por cada meta-classificador.	134
Figura 47 – Matrizes de confusão obtidas pelas técnicas de fusão por <i>ensemble</i> na classificação de TDAH/Controle, no conjunto de teste, a partir das imagens de fMRI.	135
Figura 48 – Matrizes de confusão obtidas pelas técnicas de fusão por <i>ensemble</i> na classificação de subtipos, no conjunto de teste, a partir das imagens de fMRI.	137

LISTA DE TABELAS

Tabela 1 –	Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com EEG.	38
Tabela 2 –	Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com fMRI.	44
Tabela 3 –	Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com fMRI.	47
Tabela 4 –	Tempos das instruções do paradigma de coleta de dados de estado de repouso de EEG.	59
Tabela 5 –	Quantidades de indivíduos por classe antes das exclusões de amostras devido ao controle de qualidade.	61
Tabela 6 –	Quantidade de amostras com EEG e fMRI aprovados em ambos os controles de qualidade.	61
Tabela 7 –	Proporção de sujeitos aprovados no controle de qualidade de EEG de acordo com a classe dos participantes. Legenda: 0 - sexo masculino 1 - sexo feminino. A idade e o gênero se referem às amostras após o controle de qualidade.	71
Tabela 8 –	Valores de projeto dos filtros digitais Chebyshev tipo II.	73
Tabela 9 –	Intervalos investigados por AG para otimizar o desempenho do classificador RF baseado em sinais de EEG.	78
Tabela 10 –	Hiperparâmetros otimizados pelo AG para classificar TDAH/Controle através de RFs.	79
Tabela 11 –	Hiperparâmetros otimizados pelo AG para classificar TDAH-C/TDAH-I através de RFs.	80
Tabela 12 –	Intervalos e aleatorização investigados por AG para otimizar o desempenho do classificador SVM baseado em sinais de EEG.	81
Tabela 13 –	Hiperparâmetros otimizados pelo AG para classificar TDAH/Controle através de SVMs.	81
Tabela 14 –	Hiperparâmetros otimizados pelo AG para classificar TDAH-C/TDAH-I através de SVMs.	81
Tabela 15 –	Pesos utilizados no Voto Majoritário Ponderado das predições obtidas por classificadores baseados nas bandas de EEG.	84
Tabela 16 –	Intervalos investigados por AG para otimizar o desempenho do meta-classificador AdaBoost baseado em sinais de EEG.	84

Tabela 17 –	Intervalos investigados por AG para otimizar o desempenho do meta-classificador <i>XGBoost</i> baseado em sinais de EEG.	86
Tabela 18 –	Proporção de sujeitos aprovados no controle de qualidade de fMRI de acordo com a classe dos participantes. Legenda: 0 - sexo masculino 1 - sexo feminino. A idade e o gênero se referem às amostras após o controle de qualidade.	91
Tabela 19 –	Intervalos de hiperparâmetros testados pelo AG para a CNN projetada.	98
Tabela 20 –	Hiperparâmetros da CNN selecionados por AG para classificação entre participantes dos grupos TDAH e controle.	100
Tabela 21 –	Hiperparâmetros da CNN selecionados por AG para classificação entre participantes das classe TDAH-C e TDAH-I.	102
Tabela 22 –	Hiperparâmetros da CNN LeNET-5 selecionados por AG para classificação TDAH/Controle.	104
Tabela 23 –	Hiperparâmetros da CNN LeNET-5 selecionados por AG para classificação TDAH-C/TDAH-I.	105
Tabela 24 –	Hiperparâmetros das CNNs VGG-16 e ResNet-50, ambas com aprendizado por transferência selecionados por AG usando a matriz de correlações.	107
Tabela 25 –	Intervalos investigados por AG para otimizar o desempenho do meta-classificador <i>XGBoost</i> baseado em fMRI.	110
Tabela 26 –	Métricas obtidas por cada modelo de fusão a partir de bandas de EEG para diagnóstico TDAH/Controle.	124
Tabela 27 –	Métricas obtidas por cada modelo de fusão a partir de bandas de EEG para diagnóstico de subtipos (TDAH-C/TDAH-I).	126
Tabela 28 –	<i>Overfitting</i> em cada modelo de fusão de bandas de frequência de EEG.	126
Tabela 29 –	Pesos para cada rede de conectividade funcional usados em cada meta-classificador na classificação de TDAH/Controle. Os dois pesos com maiores valores conforme cada meta-classificador foram marcados em negrito na tabela.	133
Tabela 30 –	Pesos para cada rede de conectividade funcional usados em cada meta-classificador na classificação de subtipos. Os pesos que receberam valores positivos foram marcados em negrito na tabela.	133
Tabela 31 –	<i>Overfitting</i> em cada modelo de fusão de redes de conectividade de fMRI.	138
Tabela 32 –	Métricas obtidas por cada modelo de fusão de fMRI para diagnóstico TDAH/Controle.	138
Tabela 33 –	Métricas obtidas por cada modelo de fusão de fMRI para diagnóstico de subtipos (TDAH-C/TDAH-I).	139
Tabela 34 –	Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com EEG.	151
Tabela 35 –	Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com fMRI.	156

LISTA DE ABREVIATURAS

AdaBoost	Impulso Adaptativo (do inglês, <i>Adaptative Boosting</i>)
Adadelta	Delta Adaptativo (do inglês, <i>Adaptive Delta</i>)
Adam	Estimativa de Momento Adaptativo (do inglês, <i>Adaptive Moment Estimation</i>)
ADHD	TDAH em inglês: <i>Attention-deficit hyperactivity disorder</i>
AG	Algoritmo Genético
ANOVA	Análise de variância (do inglês, <i>Analysis of variance</i>)
ARS	Escala de avaliação de TDAH (do inglês, <i>ADHD rating scale</i>)
ASR	Reconstrução do subespaço de artefatos (do inglês, <i>Artifact Subspace Reconstruction</i>)
BSS	Soma de quadrados entre classes (do inglês, <i>Between class sum of squares</i>)
BOLD	Dependente do nível de oxigênio no sangue (do inglês, <i>Blood-oxygen-level dependent</i>)
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CART	Árvores de classificação e regressão (do inglês, <i>Classification and regression trees</i>)
CNN	Rede Neural Convolutacional (do inglês, <i>Convolutional Neural Network</i>)
C-PAC	<i>Pipeline</i> configurável para análise de conectomas (do inglês, <i>Configurable Pipeline for the Analysis of Connectomes</i>)
CPR	Correlação entre probabilidades de recorrências (do inglês, <i>Correlation between probabilities of recurrences</i>)
CPU	Unidade de Processamento Central (do inglês, <i>Central Processing Unit</i>)
DART	Abandonos Adequados às Árvores de Regressão Aditiva Múltipla (do inglês, <i>Dropouts Meet Multiple Additive Regression Trees</i>)
DPARSFA	Assistente de Processamento de Dados para o Estado de Repouso (do inglês, <i>Data Processing Assistant for Resting-State</i>)
DSM	Manual Diagnóstico e Estatístico de Transtornos Mentais (do inglês, <i>Diagnostic and Statistical Manual of Mental Disorders</i>)
DT	Árvores de Decisão (do inglês, <i>do inglês, Decision Trees - DT</i>)

ECG	Eletrocardiografia
EEG	Eletroencefalografia
ELM	Máquina de Aprendizado Extremo (do inglês, <i>extreme learning machine - ELM</i>)
ELU	Unidade Linear Exponencial (do inglês, <i>Exponential Linear Unit</i>)
EMG	Eletromiografia
EOG	Eletro-oculografia
ERP	Potenciais Relacionados a Eventos (do inglês, <i>event-related potentials</i>)
ERD	Dessincronização relacionada à eventos (do inglês <i>event-related desynchronization - ERD</i>)
fALFF	Amplitude fracionária de flutuações de baixa frequência (do inglês, <i>Fractional amplitude of low-frequency fluctuations</i>)
FDA	Administração de Alimentos e Medicamentos (do inglês, <i>Food and Drug Administration</i>)
FDJ	Deslocamento relativo à quadros Jenkinson (do inglês, <i>Framewise Displacement Jenkinson</i>)
fMRI	Imagens de Ressonância Magnética funcional (do inglês, <i>Functional Magnetic Resonance Imaging</i>)
gbtree	Árvore de aumento de gradiente (do inglês, <i>gradient boosting tree</i>)
GPU	Unidade de Processamento Gráfico (do inglês, <i>Graphics Processing Unit</i>)
GRU	Unidade recorrente com portas (do inglês, <i>Gated recurrent unit</i>)
HBN	Rede do Cérebro Saudável (do inglês, <i>Healthy Brain Network</i>)
HOG	Histogramas Orientados a Gradientes
HIV	Vírus da Imunodeficiência Humana (do inglês, <i>Human Immunodeficiency Virus</i>)
IEE	Laboratório de Instrumentação Eletroeletrônica e Inteligência Artificial
IC	Inteligência Computacional
ICA	Análise de Componentes Independentes (do inglês, <i>Independent Component Analysis</i>)
k-NN	k-Vizinhos Mais Próximos (do inglês, <i>K-Nearest Neighbors</i>)
KGC	Causalidade de Núcleo Granger (do inglês, <i>Kernel Granger Casuality</i>)
L1	<i>Lasso</i>
L2	<i>Ridge</i>
L1L2	<i>ElasticNet</i>
LSTM	Memória de longo prazo (do inglês, <i>Long short-term memory</i>)
MDS	Escalonamento Multi-Dimensional (do inglês, <i>Multi-Dimensional Scaling</i>)

MKL	Aprendizado de Múltiplos Núcleos (<i>multi-kernel learning</i>)
MRI	Imagens de Ressonância Magnética (do inglês, <i>Magnetic Resonance Imaging</i>)
NaN	Não-número (do inglês, <i>Not a Number</i>)
NBN	Neurônio por neurônio (do inglês, <i>Neuron-by-neuron</i>)
NRCDAE	Auto-codificador de Eliminação de Ruído Convolutivo Residual Aninhado (do inglês, <i>Nested Residual Convolutional Denoising Autoencoder</i>)
PCA	Análise de Componentes Principais (do inglês, <i>Principal Component Analysis</i>)
PPGEE	Programa de Pós-Graduação em Engenharia Elétrica
QI	Quociente de Inteligência
RAM	Memória de Acesso Aleatório (do inglês, <i>Random Access Memory</i>)
RBF	Função de Base Radial (do inglês, <i>Radius Basis Function</i>)
ResNet	Rede residual (do inglês, <i>Residual Network</i>)
ReLU	Unidade Linear Retificada (do inglês, <i>Rectified Linear Unit</i>)
RF	Floresta Aleatória (do inglês, <i>Random Forest</i>)
RFE	Eliminação Recursiva de Características (do inglês, <i>Recursive Feature Elimination</i>)
RNA	Redes Neurais Artificiais
RNN	Redes Neurais Recorrentes (do inglês, <i>Recurrent Neural Network</i>)
RMS	Raiz Quadrada Média (do inglês, <i>Root Mean Square</i>)
ROC	Curva Característica de Operação do Receptor (do inglês, <i>Receiver Operating Characteristic Curve receiver operating characteristic - ROC</i>)
ROI	Regiões de Interesse (do inglês, <i>Region Of Interest</i>)
rs-EEG	EEG em Estado de Repouso (do inglês, <i>Resting-State-EEG</i>)
rs-fMRI	fMRI em Estado de Repouso (do inglês, <i>Resting-State-fMRI</i>)
RSS	Soma Residual dos Quadrados (do inglês, <i>Residual Sum of Squares</i>)
SBC	Correlação Baseada em Semente (do inglês, <i>Seed-Based Correlation</i>)
SeLU	Unidade Linear Exponencial Escalonada (do inglês, <i>Scaled Exponential Linear Unit</i>)
SF	Fator de Forma (do inglês, <i>Shape Factor</i>)
SFDR	Faixa Dinâmica Livre Espúria (do inglês, <i>Spurious Free Dynamic Range</i>)
SGD	Descida do gradiente estocástico (do inglês, <i>Stochastic gradient descent</i>)
SINAD	Relação sinal-ruído e distorção (do inglês, <i>Signal to Noise and Distortion Ratio</i>)
SNR	Relação Sinal-Ruído (do inglês, <i>Signal to Noise Ratio</i>)
SMOTE	Técnica de sobreamostragem Minoritária Sintética (do inglês, <i>Synthetic Minority Over-sampling Technique</i>)

sMRI	Imagens de Ressonância Magnética estrutural (do inglês, <i>Structural Magnetic Resonance Imaging</i>)
SVM	Máquina de Vetor de Suporte (do inglês, <i>Support Vector Machine</i>)
SWA	Média da Janela Espacial (do inglês, <i>Spatial Window Averaging</i>)
SWAN	Pontos Fortes e Fracos dos Sintomas de TDAH e Comportamento Normal (do inglês, <i>Strengths and Weakness of ADHD-Symptoms and Normal-behavior</i>)
tanh	Tangente hiperbólica
TDAH	Transtorno de Déficit de Atenção com Hiperatividade
TDAH-C	Transtorno de Déficit de Atenção com Hiperatividade Combinado
TDAH-H	Transtorno de Déficit de Atenção com Hiperatividade Hiperativo
TDAH-I	Transtorno de Déficit de Atenção com Hiperatividade Desatento
UFRGS	Universidade Federal do Rio Grande do Sul
VHERS	Modo Variacional e Separação de Ritmo de EEG baseada na Transformada de Hilbert (do inglês, <i>variational mode and Hilbert transform-based EEG rhythm separation</i>)
VGG	Grupo Geométrico Visual (do inglês, <i>Visual Geometry Group</i>)
WISC-IV	Escala Wechsler de Inteligência para Crianças
WSS	Soma dos Quadrados Dentro da Classe (do inglês, <i>Within-class sum of squares</i>)
WURS	Escala de Classificação de Wender Utah (do inglês, <i>Wender Utah Rating Scale</i>)
XGBoost	Alavancamento Extremo de Gradiente (do inglês, <i>Extreme Gradient Boosting</i>)

LISTA DE SÍMBOLOS

'	Símbolo para identificar a transposição de um vetor ou de uma matriz
α_t	Importância do classificador no AdaBoost
α_{Reg}	Força da regularização imposta ao modelo no classificador <i>Ridge</i>
b	Viés (do inglês, <i>bias</i>)
C	Constante positiva da SVM determinada experimentalmente
c_i	Coefficientes real na SVM
c_j	Coefficientes real na SVM
d	Distância entre um vetor de suporte e o hiperplano de decisão
Dim	Dimensionalidade da matriz
dim_x	Dimensão horizontal de uma matriz antes da convolução
dim_y	Dimensão vertical de uma matriz antes da convolução
den	Coefficiente do denominador
\sum	Somatório
ϵ_{rip}	Fator de <i>Ripple</i> em filtros
ϵ	Erro no meta-classificador <i>AdaBoost</i>
$E()$	Matriz de entrada da convolução
$\varphi()$	Função não linear
$f()$	Função
fil_x	Dimensão horizontal do filtro convolucional (<i>kernel</i>)
fil_y	Dimensão vertical do filtro convolucional (<i>kernel</i>)
f_k	Estrutura independente de árvore
FN	Falso Negativo
FP	Falso Positivo
γ	Mínima redução na função custo para aprovação de uma divisão de galhos no <i>XGBoost</i>
γ_{SVM}	hiperparâmetro que controla a largura do <i>kernel</i>

g_i	Gradiente de primeira ordem do <i>XGBoost</i>
$G_n()$	Função de transferência
η	Fator de encolhimento (em inglês, <i>shrinkage</i>)
H()	Saída da convolução
H(s)	Função de transferência
h_i	Gradiente de segunda ordem do <i>XGBoost</i>
h_t	hipótese predita de um modelo com índice t
I	União das instâncias esquerda (I_L) e direita (I_R) de um nó da árvore depois da divisão dos galhos.
i	Índice
i_a	Índice do somatório horizontal da convolução
i_b	Índice do somatório vertical da convolução
I_L	Instâncias esquerda de um nó da árvore depois da divisão dos galhos.
I_R	Instâncias direita de um nó da árvore depois da divisão dos galhos.
j	Índice
$K()$	Filtro (<i>Kernel</i>)
k	Número de conjuntos de dados (<i>folds</i>)
$L(t)$	Objetivo de Aprendizado Regularizado
$l()$	Objetivo de Aprendizado
L_1	Taxa de regularização <i>Lasso</i>
L_2	Taxa de regularização <i>Ridge</i>
λ	Taxa de regularização
m	Índice máximo
μ	Média
n	Índice máximo
num	Coefficiente do numerador
ξ	Variável de folga
$poli$	Grau do polinômio
p_{pad}	Espessura da borda da matriz em <i>pixels</i>
Q	Superfície não-linear para a decisão
r	Coefficiente que pode ser configurado como um hiperparâmetro na SVM
σ	Desvio Padrão
s	passo <i>stride</i>
t	Passo

T	Número de folhas da árvore
T_n	Polinomial de Chebyshev de ordem n .
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
W_{p1}	Frequências limite inferior da banda de rejeição
W_{p2}	Frequências limite superior da banda de rejeição
x	Característica de entrada
x_i	Características na <i>AdaBoost</i> e no <i>Ridge</i> ou Variáveis de folga na SVM
x_1	Classe 1 na SVM
x_2	Classe 2 na SVM
y	Variável de saída
y_i	Saída Esperada na SVM
y_{obj}	Objetivo de saída
y_{pred}	Saída predita
z	Valor normalizado
Z_t	Fator de Normalização
ω	Frequência (no pré-processamento) ou Peso (nos métodos de <i>Machine learning</i>)
ω_0	Frequência de corte
ω_j^*	Peso ótimo de cada folha no <i>XGBoost</i>
Ω	Unidade de resistência
Ω_X	Termo de regularização do <i>XGBoost</i>

SUMÁRIO

1	INTRODUÇÃO	22
1.1	Justificativas	24
1.2	Objetivos	25
1.3	Estrutura da Tese	26
2	ESTADO DA ARTE	27
2.1	Análise Bibliométrica dos Trabalhos Publicados	27
2.2	Análise dos Trabalhos do Estado da Arte	34
2.2.1	Referencial Teórico dos principais trabalhos de classificação de TDAH a partir de sinais de EEG	35
2.2.2	Referencial Teórico dos principais trabalhos de classificação de TDAH a partir de fMRI	39
2.2.3	Trabalhos com EEG e fMRI	45
2.3	Fundamentação Teórica	47
2.3.1	Árvores de Decisão e <i>Random Forest</i> (RF)	47
2.3.2	<i>Support Vector Machine</i> (SVM)	48
2.3.3	Rede Neural Convolutacional (CNN)	50
2.3.4	Classificador Linear <i>Ridge</i>	52
2.3.5	<i>Adaptive Boosting</i> (AdaBoost)	53
2.3.6	<i>Extreme Gradient Boosting</i> (XGBoost)	54
3	MATERIAIS E MÉTODOS	57
3.1	Base de Dados: <i>Healthy Brain Network</i> (HBN)	57
3.2	Resumo da Metodologia	62
3.3	Eletroencefalografia (EEG)	65
3.3.1	Pré-processamento de EEG	65
3.3.2	Processamento de EEG: <i>Random Forest</i> (RF)	77
3.3.3	Processamento de EEG: <i>Support Vector Machine</i> (SVM)	80
3.3.4	Processamento de EEG: Fusão de Classificações a partir das Bandas de EEG	82
3.4	Classificação de TDAH a partir de Imagens de Ressonância Magnética Funcional (fMRI)	87
3.4.1	Controle de Qualidade de fMRI	89
3.4.2	Pré-processamento	91
3.4.3	Extração de características de fMRI	92
3.4.4	Processamento de fMRI: Rede Neural Convolutacional (CNN) Projetada	96
3.4.5	Processamento de fMRI: Modelos Clássicos	103

3.4.6	Processamento de fMRI: Fusão de Classificações a partir das Redes de Conectividade de fMRI	108
4	RESULTADOS	111
4.1	Resultados da Classificação de TDAH baseada em EEG	111
4.1.1	Resultados da Classificação de TDAH baseada em EEG: RF	111
4.1.2	Resultados de classificação de TDAH e de subtipos com SVM a partir de EEG	114
4.1.3	Resultados de classificação de de TDAH e de subtipos com RF e SVM a partir de EEG no grupo separado	116
4.1.4	Resultados da Classificação de TDAH baseada em EEG: Fusão de Bandas	120
4.2	Resultados da Classificação de TDAH baseada em fMRI	127
4.2.1	Resultados da Classificação de TDAH baseada em fMRI: Arquitetura CNN Projetada	127
4.2.2	Resultados: Fusão de Classificadores de fMRI baseados nas Redes de Conectividade	132
5	DISCUSSÃO	140
5.1	Análise do Significado dos Biomarcadores	142
5.2	Cenários de Classificação	149
5.3	Medida de desempenho dos classificadores	149
5.4	Comparação com os principais trabalhos de classificação de TDAH . . .	150
5.4.1	Comparação com os principais trabalhos de classificação de TDAH com base em EEG	151
5.4.2	Comparação com os principais trabalhos de classificação de TDAH com base em fMRI	155
5.5	Limitações	163
6	CONCLUSÃO	165
6.1	Trabalhos Futuros	167
	REFERÊNCIAS	168
	APÊNDICE A GRÁFICOS DAS PCAS DAS BANDAS DE FREQUÊNCIA DE EEG	188
A.1	Gráficos das respostas em frequência dos filtros de EEG	188
A.2	Gráficos das PCAs das Bandas de frequência de EEG para o diagnóstico entre grupos TDAH e controle	190
A.3	Gráficos das PCAs das Bandas de frequência de EEG diagnóstico TDAH-C/TDAH-I	193
	APÊNDICE B CURVAS DE OTIMIZAÇÃO DOS ALGORITMOS GENÉTICOS	197
B.1	Curvas de Otimização das RF	197
B.2	Curvas de Otimização das SVM	203
B.3	Curvas de Otimização dos XGBoost	209

APÊNDICE C	RESULTADOS DE CLASSIFICAÇÃO DE TDAH E DE SUBTIPOS COM OS MODELOS CLÁSSICOS DE CNN A PARTIR DE fMRI	210
C.1	Resultados da Classificação de TDAH baseada em fMRI: Modelos clássicos de CNN em suas arquiteturas originais	210
C.2	Resultados da Classificação de TDAH baseada em fMRI: Modelos Clássicos de CNN Otimizados	211
APÊNDICE D	CONCEITOS BÁSICOS SOBRE TDAH	218
D.1	Diagnóstico do TDAH	219
D.2	Causas e consequências do TDAH	220
APÊNDICE E	LISTA DE INDIVÍDUOS CUJOS DADOS FORAM UTILIZADOS NA CONFECÇÃO DESTA TESE	223
E.1	Indivíduos que forneceram dados de EEG	223
E.2	Indivíduos que forneceram dados de fMRI aprovados no Controle de Qualidade	229

1 INTRODUÇÃO

A engenharia médica tem como objetivo melhorar o bem-estar dos indivíduos da sociedade. Na engenharia biomédica, a instrumentação biomédica tem por finalidade obter informações que possibilitem melhorar o conhecimento humano sobre os corpos biológicos. O aprofundamento deste entendimento poderá facilitar diagnósticos, como os do Transtorno de Deficit de Atenção com Hiperatividade (TDAH ou em inglês *Attention-deficit hyperactivity disorder* - ADHD) e auxiliar no desenvolvimento e monitoramento da evolução de tratamentos médicos. O TDAH aflige ao menos 5% da população mundial (VOS *et al.*, 2016) e afeta cerca de 5 a 7% das crianças de acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais (do inglês, *Diagnostic and Statistical Manual of Mental Disorders*, DSM-V) (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013) ou 5% segundo (FARAONE *et al.*, 2015). Dentre 30-50% destas crianças permanecem com os sintomas mesmo na vida adulta. Conforme o estudo de (FARAONE *et al.*, 2015) 2,5% dos adultos possuem este transtorno. De acordo com (WOLRAICH *et al.*, 2019) o TDAH é a condição comportamental mais comum na infância, atingindo cerca de 8,4% da população norte-americana, e a segunda doença crônica mais comum em crianças. Segundo (PALOYELIS *et al.*, 2007), o TDAH é o mais comum transtorno psiquiátrico na infância. O trabalho de (WOLRAICH *et al.*, 2019) também afirma que o TDAH é uma condição vitalícia que requer tratamento contínuo dos sintomas, pois não há uma cura definitiva, assim como outras doenças crônicas. Os sintomas do TDAH, usualmente, são os seguintes: dificuldades para perceber detalhes e manter atenção e para permanecer sentados (ADLER *et al.*, 2006); impulsividade e facilidade para se distrair, parecendo esquecidos em atividades diárias; dificuldades para organizar tarefas e atividades, às vezes perdendo objetos necessários para a sua realização, bem como para completar tarefas longas ou tediosas (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013).

No entanto, o diagnóstico clínico do TDAH baseado nos sintomas do paciente apresenta dificuldades, pois os sistemas classificatórios baseados em questionários não abrangem a complexidade de quadros observada na prática clínica (DE SOUZA *et al.*, 2007). Ademais, (WASSERSTEIN, 2005) destaca a dificuldade de adaptar os critérios do DSM-

IV (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 1994) para o diagnóstico de adultos e de adolescentes, uma vez que o manual cita critérios como "o paciente escalar ou correr com muita frequência". O manuscrito de (PHAM & RIVIERE, 2015) ressalta que nas últimas décadas houveram muitas discussões sobre como deveria ser realizado o diagnóstico do TDAH. Por isso, há divergências entre os psiquiatras clínicos sobre os métodos de diagnóstico do TDAH. Por exemplo, o trabalho de (MUSSO & GOUVIER, 2014) mostrou uma pesquisa onde o uso dos critérios da escala de classificação de TDAH (do inglês, *ADHD rating scale* - ARS) e da escala de classificação de Wender Utah (do inglês, *Wender Utah rating scale* - WURS) produziram diagnósticos diferentes para no mínimo 1,7% de indivíduos da mesma amostra de 770 estudantes de ensino superior. O trabalho de (HINSHAW, 2018) mostrou que há controvérsias sobre o aumento dos diagnósticos de TDAH na última década e neste aumento ser percebido de forma diferente em países distintos. O trabalho de (WEIS *et al.*, 2019) destacou que apenas 23,4% dos médicos, os quais avaliaram 214 pacientes, diagnosticam TDAH considerando todos os sintomas descritos no DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013). Por outro lado, o estudo de (MUSSO & GOUVIER, 2014) relatou que uma quantidade significativa de estudantes colegiais finge possuir sintomas de TDAH para obter a medicação. Estes estudantes conseguem iludir os testes baseados em questionários, como a Escala de classificação de Conners para adultos com TDAH (*Conners' Adult ADHD Rating Scale*), a WURS e a ARS.

A eletroencefalografia (EEG) vem sendo utilizada há 75 anos para o estudo de TDAH (LENARTOWICZ & LOO, 2014). A técnica de EEG possibilita a medição de conjuntos de impulsos elétricos cerebrais captados por eletrodos posicionados em torno do couro cabeludo. Em EEG, uma maior medida de potenciais elétricos está associada a uma maior atividade cerebral na região correspondente. Os ritmos cerebrais, em adultos saudáveis, estão associados a estados mentais, como, por exemplo, os estados de sono, relaxamento ou de alerta (?). Porém, não se chegou a uma medida definitiva de EEG ligada ao TDAH. A literatura explica que isto ocorre em razão da existência de múltiplas alterações cerebrais que conduzem a ocorrência dos sintomas do TDAH (LENARTOWICZ & LOO, 2014).

O uso das imagens de ressonância magnética funcional (fMRI) (OGAWA *et al.*, 1990) tem se destacado como uma ferramenta para a interpretação e entendimento do funcionamento do cérebro e de seus possíveis transtornos, tais como o TDAH (PALOYELIS *et al.*, 2007). Para tanto, a fMRI detecta variações no fluxo sanguíneo, em resposta a uma atividade neural, o que permite o mapeamento das funções cerebrais em resposta à estímulos. O funcionamento mais intenso de uma área do cérebro corresponde a uma maior irrigação sanguínea nesta área, a qual é mensurável através das técnicas de fMRI (OGAWA *et al.*, 1990). Esta medida indireta da atividade neural por meio do fluxo sanguíneo é chamada de nível de dependência do oxigênio no sangue (do inglês, *blood-oxygen-level dependent*

- BOLD).

A integração multimodal de métodos de imagem cerebral possibilita um entendimento mais completo da atividade cerebral (MULERT & LEMIEUX, 2023). A técnica de EEG possui melhor resolução temporal que fMRI, a qual em contrapartida possui melhor resolução espacial do que EEG (JATOI & KAMEL, 2017) (MULERT & LEMIEUX, 2023). Por conta desta complementariedade, o uso das técnicas de EEG e de fMRI em conjunto pode auxiliar tanto na obtenção de um diagnóstico automático mais preciso como em aumentar o entendimento do TDAH ou estabelecer novas relações, como demonstra (DRON *et al.*, 2019) para o caso da epilepsia.

Os sinais de EEG são de natureza estocástica, e, por isso, seu entendimento adequado requer o uso de técnicas complexas, tais quais aquelas pertencentes à área da Inteligência Computacional (IC), cujas aplicabilidades tem aumentado graças ao desenvolvimento da computação nos últimos anos. Assim como EEG, fMRI também produz informações complexas com alta quantidade de dados, cuja interpretação pode se beneficiar do avanço da computação e das técnicas de IC e da instrumentação inteligente.

1.1 Justificativas

Os diagnósticos clínicos de TDAH apresentam imprecisões (BISWAS *et al.*, 2020) e são realizados analisando os sintomas apresentados pelos pacientes. O TDAH pode ser dividido entre as subclasses TDAH hiperativo e TDAH desatento, nas quais apenas os sintomas de hiperatividade ou de desatenção são predominantes, respectivamente. Além destes, há o subtipo TDAH combinado, onde uma mescla de sintomas de hiperatividade e de desatenção aflige os pacientes (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 1994). Embora todos os subtipos de TDAH apresentem sintomas que provocam dificuldades sociais e acadêmicas, estes sintomas e suas consequências são diferentes entre si (SAAD *et al.*, 2020). Por exemplo, o subtipo TDAH combinado costuma estar mais associado com comorbidades como o transtorno de conduta opositiva e desafiadora, além de provocar maior dificuldade no controle de impulsos (SAAD *et al.*, 2020). Já o subtipo TDAH desatento apresenta a tendência de possuir maior timidez, passividade na vida social, baixa autoestima, ansiedade e depressão (SAAD *et al.*, 2020). O subtipo hiperativo está mais associado ao transtorno de ansiedade e à pior qualidade de vida (SALVI *et al.*, 2019). O estudo de (SALVI *et al.*, 2019) ressaltou que estudos em adultos sugerem que os pacientes não transitam entre os subtipos ao longo do tempo. Constataram-se diferenças entre subtipos do TDAH por EEG e os efeitos de medicações estimulantes em cada subtipo (CLARKE *et al.*, 2020). Por terem alterações diferentes nos seus cérebros, as medicações necessárias para tratar estes subtipos podem divergir (MIRANDA *et al.*, 2020). O uso de bancos de dados com grande quantidade de dados possibilita que sejam encontrados biomarcadores que levem a uma identificação e diagnóstico mais preciso des-

tes subtipos. A combinação multimodal de diferentes técnicas tem potencial para elevar o entendimento destas diferenças no cérebro assim como produzir ferramentas para auxiliar o diagnóstico (ANGRIMAN *et al.*, 2014).

Esta proposta de Tese tem por objetivo construir soluções confiáveis através de algoritmos de classificação de EEG e fMRI que possam proporcionar um ganho de conhecimento, além de desenvolver um classificador com o máximo de *accuracy* possível, minimizando, assim, erros de diagnóstico.

A fusão de EEG e de fMRI não foi realizada em função da baixa quantidade de participantes, cujos exames foram aprovados no controle de qualidade para ambas as modalidades.

1.2 Objetivos

De forma geral, esta proposta de Tese tem como objetivo pesquisar a construção de classificadores que utilizem dados de fMRI e de EEG para diagnosticar TDAH. Em função das considerações iniciais, o objetivo principal consiste em pesquisar o desenvolvimento de um método de Inteligência Computacional que possa auxiliar os profissionais da área da saúde na identificação da ocorrência do TDAH e de seus subtipos através de meta-classificadores baseados em dados de EEG e de fMRI, bem como verificar se é possível com estas técnicas identificar possíveis biomarcadores. Os subtipo de TDAH analisados são o Transtorno de Déficit de Atenção com Hiperatividade Combinado (TDAH-C) e o Transtorno de Déficit de Atenção com Hiperatividade Desatento (TDAH-I). Os objetivos específicos propostos são:

- desenvolver algoritmos para classificar TDAH e subtipos (TDAH-C e TDAH-I) a partir de sinais de EEG com *Random Forest* (RF) e *Support Vector Machine* (SVM),
- pesquisar meta-classificadores que diferenciem indivíduos com TDAH de indivíduos controle com base em classificadores de faixas de frequência de EEG e em redes de conectividade de fMRI,
- pesquisar meta-classificadores que identifiquem os subtipos de TDAH (TDAH-C e TDAH-I) com base em classificadores de faixas de frequência de EEG e em redes de conectividade de fMRI,
- testar os pesos designados pelos otimizadores dos métodos de *ensemble* para cada possível biomarcador para serem utilizados como biomarcadores de EEG e de fMRI,
- investigar o uso da combinação de predições a partir de possíveis biomarcadores para aumentar o desempenho de classificadores de TDAH e de seus subtipos,

- pesquisar o uso de uma arquitetura de redes neurais convolucionais para a classificação de indivíduos com TDAH e de subtipos (TDAH-C e TDAH-I) a partir de dados de fMRI.

1.3 Estrutura da Tese

A Tese está estruturada em cinco capítulos. A revisão do estado da arte está compreendida no Capítulo 2, destacando as produções mais relevantes na área nos últimos anos, os desafios a serem superados e as técnicas utilizadas na atualidade pela comunidade científica.

O Capítulo 3 apresenta a metodologia utilizada no desenvolvimento desta Tese. Este capítulo inicia com a descrição da base de dados considerada, a *Healthy Brain Network*® (HBN) (ALEXANDER *et al.*, 2017). Na sequência, expõe-se a metodologia de pré-processamento dos sinais de EEG, os métodos de *machine learning* testados como classificadores de TDAH base e as técnicas de fusão de bandas de frequência de EEG. A seguir, apresenta-se a metodologia de pré-processamento e os classificadores de TDAH utilizados para as imagens de fMRI, bem como as técnicas de fusão de redes de conectividade de fMRI. Então, expõe-se a metodologia da fusão de EEG com fMRI aplicada a classificação de TDAH.

Os resultados obtidos estão expostos e discutidos no Capítulo 4. A Discussão formulada a partir do estudo desenvolvido está no Capítulo 5. O Capítulo 4 se divide em três partes: EEG e fMRI. Por fim, apresentam-se as conclusões e as possibilidades de trabalhos futuros no Capítulo 6.

2 ESTADO DA ARTE

2.1 Análise Bibliométrica dos Trabalhos Publicados

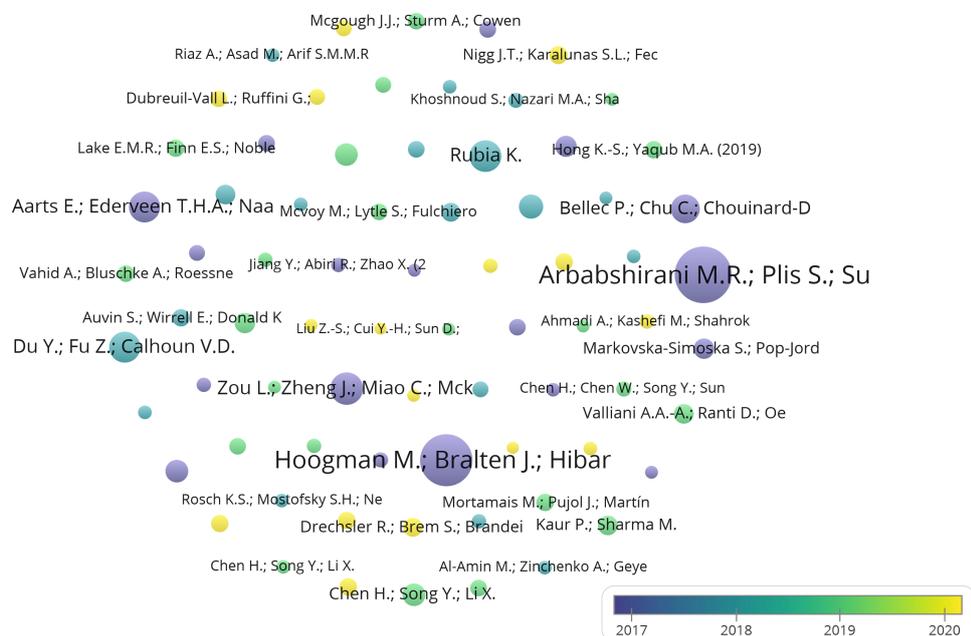
Para avaliar o estado da arte no desenvolvimento de um sistema de classificação do TDAH realizou-se uma análise bibliométrica do assunto na base Scopus abrangendo esta área de pesquisa nos últimos 11 anos. Utilizou-se na pesquisa bibliométrica a ferramenta Vosviewer®.

Na base bibliométrica fornecida pela Scopus, procurou-se por ("*Attention deficit hyperactivity disorder*" OR ADHD) AND (classifi* OR "computer-aided" OR automat* OR computational OR diagnosis OR "machine learning" OR "deep learning") AND ((Electroencephalo* OR EEG) OR ("Magnetic Ressonance Image" OR fMRI OR MRI)) em listas de palavra-chave, resumos (*abstracts*) e títulos. Desta forma, encontraram-se 1.536 artigos. Filtrou-se os resultados desta pesquisa para o período entre 2017 e 2023. Neste período, 938 artigos foram publicados neste tema. Limitando a pesquisa bibliométrica para apenas os artigos em inglês (916) e em espanhol (9), a Scopus listou 923 trabalhos. Limitou-se ainda esta pesquisa bibliométrica em artigos com até 25 autores. Esta pesquisa bibliométrica foi realizada no dia 24 de Setembro de 2023.

O mapa de citações entre publicações, apresentado na Figura 1, mostra as publicações que receberam maior número de citações do escopo no qual se situa esta pesquisa bibliométrica. Este mapa foi organizado de forma que quanto maior o diâmetro de um nó, maior o número de citações de um determinado trabalho. A variação das cores corresponde à variação no período de publicação, de modo que as esferas roxas, conforme a legenda, correspondem aos trabalhos mais antigos, os quais possuem a tendência a possuírem um maior número de citações justamente por estarem disponíveis para leitura há mais tempo. A posição espacial agrupa os autores de forma que aqueles citados pelas mesmas fontes sejam colocados em posições mais próximas. Limitou-se o grafo da Figura 1 a mostrar apenas os trabalhos com ao menos 30 citações. Nestas condições, a pesquisa bibliométrica revelou 74 publicações, as quais são apresentadas na correspondente figura. Os trabalhos que recebem o maior número de citações costumam ser aqueles que já foram publicados há mais tempo, por isto, os trabalhos de 2017 a 2018 recebem maior destaque nestes

mapas. O mapa mostra que as publicações mais citadas no período foram as de (AR-BABSHIRANI *et al.*, 2017), que foi citado 550 vezes nestes 11 anos. Este estudo recebeu esta quantidade de citações por ser um trabalho de revisão de 200 artigos de classificação com *machine learning* de diversos transtornos mentais, atraindo citações de trabalhos em diferentes áreas. Os autores de (DU *et al.*, 2018) também se destacaram com seu artigo de revisão que aborda o pré-processamento da conectividade funcional de fMRI de diversas formas e o uso de diversas técnicas de classificação de dados para diversos transtornos e condições neurológicas.

Figura 1 – Mapa de citações de autores.



Fonte: O Autor

O estudo de (HOOGMAN *et al.*, 2017) realizou uma mega-análise de MRI estrutural (sMRI) de um grupo de 1713 pessoas com TDAH e 1529 voluntários saudáveis com um modelo linear misto. Com esta abordagem, (HOOGMAN *et al.*, 2017) observou reduções de volume na amígdala bilateral (região do sistema límbico associada aos processos emocionais), no núcleo *accubens* (componente da região ventral e ligado ao aprendizado por reforço e à motivação) e no *hippocampus* (estrutura conectada ao aprendizado e à memória) em indivíduos com TDAH. O trabalho de (ZOU *et al.*, 2017) foi citado 188 vezes neste período e apresenta uma Rede Neural Convolutiva (CNN) para o diagnóstico do TDAH a partir de fMRI e sMRI da base de dados da competição ADHD-200 (CONSORTIUM, 2012). Os autores geraram mapas de características que foram enviadas separadamente para 2 CNNs com 4 camadas convolucionais e uma camada totalmente conectada de neurônios. As CNNs das duas modalidades foram unidas através de uma última camada de neurônios. Com esta abordagem, os autores de (ZOU *et al.*, 2017) obtiveram

69.15% de *accuracy* média utilizando uma amostra de 94 indivíduos do grupo controle e 77 com TDAH. Os autores observaram a complementariedade entre as características de sMRI e de fMRI: matéria cinza e da amplitude das flutuações de baixa frequência (0,01Hz-0,1Hz) (fALFF), respectivamente, para a detecção de TDAH.

Outros trabalhos com mais de 140 citações foram: (DU *et al.*, 2018), (RUBIA, 2018), (?), e (AARTS *et al.*, 2017). O artigo de (?) foi citado mais de 140 vezes e relatou o pré-processamento de fMRI e sMRI da base de dados da ADHD-200 seguindo diferentes *pipelines* com o objetivo de facilitar trabalhos futuros de *machine learning* com a base. (AARTS *et al.*, 2017) observou uma relação entre o aumento da quantidade de bactérias no intestino com uma redução nas ativações cerebrais (mensuradas por fMRI) da região ventral durante a antecipação por recompensas em sujeitos com TDAH.

O manuscrito de (RUBIA, 2018) apresenta uma revisão e uma meta-análise de trabalhos de classificação. Através desta meta-análise, (RUBIA, 2018) observou que os pacientes de TDAH possuem alterações nas redes de conectividade dorsal, ventral, frontoparietal, límbica que estão envolvidas nos processos de controle cognitivo, atenção, memória de trabalho e motivação. (RUBIA, 2018) também notou uma dificuldade dos TDAH em desativar a rede de conectividade de modo *Default*, a qual costuma estar ativa quando as pessoas estão acordadas, porém em devaneios ou sonhos acordados.

Observou-se uma tendência dos artigos conectados à fMRI receberem um maior número de citações no período observado do que os artigos de EEG. E, também, há um menor número de artigos de EEG (436) sobre a temática considerando as limitações relatadas sobre o período e idiomas pesquisados. Destaca-se que os trabalhos mais citados de EEG são: (CHEN *et al.*, 2019a), (MARKOVSKA-SIMOSKA & POP-JORDANOVA, 2017), (SAAD *et al.*, 2018), (VAHID *et al.*, 2019), (DUBREUIL-VALL *et al.*, 2020), (TOR *et al.*, 2021), todos com mais de 50 citações. O trabalho de (CHEN *et al.*, 2019a), que recebeu 95 citações, avaliou dados de EEG de 50 pessoas com TDAH e 51 sem o transtorno coletados por um sistema de 32 eletrodos utilizando como características medidas de sincronização e de ciência de redes. Estas características foram organizadas em uma matriz de conectividade de EEG para cada indivíduo e classificadas com uma CNN. Esta abordagem obteve 94.67% de *accuracy* na classificação binária. Os mesmos autores de (CHEN *et al.*, 2019a), publicaram outro estudo (CHEN *et al.*, 2019b), no qual utilizaram uma CNN para avaliar sinais de 10 minutos de EEG em estado de descanso (rs-EEG) de 50 indivíduos com TDAH e 57 do grupo controle. Neste trabalho, estes autores obtiveram *accuracy* de 90,29% \pm 0,58% ao investigar anormalidades no espaço na frequência espacial. No entanto, em nenhum destes trabalhos, (CHEN *et al.*, 2019a) e (CHEN *et al.*, 2019b), realizaram a classificação de subtipos.

O estudo de (MARKOVSKA-SIMOSKA & POP-JORDANOVA, 2017), citado 128 vezes, avaliou a razão entre a potência espectral do espectro Teta e a potência espectral do espectro Beta (razão Teta/Beta), além da potência espectral absoluta e relativa de rs-EEG.

Esta pesquisa utilizou como base de dados um grupo de 60 sujeitos com TDAHs e 60 sujeitos formando o grupo denominado de controle, dos quais metade de cada grupo era formada por crianças e a outra metade por adultos. Apenas as crianças com TDAH apresentaram uma maior razão Teta/Beta e uma maior potência espectral absoluta tanto para a faixa Delta como para a faixa Teta. Este trabalho se utilizou de análises estatísticas de Análise de variância (do inglês, Analysis of variance - ANOVA) e da Curva Característica de Operação do Receptor (do inglês, Receiver Operating Characteristic Curve *receiver operating characteristic* - ROC). O estudo obteve 84,3% na métrica da área abaixo da curva ROC (do inglês, Area Under the Curve - AUC) com as crianças, e, porém, somente uma AUC de 68,3% com os adultos.

O artigo de revisão de (SAAD *et al.*, 2018) revelou as inconsistências na literatura do uso da razão Teta/Beta aprovada pela organização *Food and Drug Administration* (FDA) para o diagnóstico de TDAH e propõe o uso de diferenças em uma frequência de transição entre as faixas alfa e teta como um biomarcador de TDAH.

A pesquisa de (TOR *et al.*, 2021) trata de um sistema de classificação que diferencia pessoas com TDAH de pessoas com o transtorno de conduta (uma comorbidade de TDAH) e de pessoas que possuem os dois transtornos. No entanto, neste estudo não há uma comparação entre TDAH e indivíduos do grupo controle ou entre subtipos de TDAH.

O estudo de classificação de TDAH apresentado em (DUBREUIL-VALL *et al.*, 2020) se utiliza de uma CNN para discriminar 20 indivíduos com TDAH de 20 indivíduos do grupo controle. Os artefatos de EEG foram retirados utilizando *Independent Component Analysis* (ICA). Os autores transformaram os sinais de EEG em espectrogramas que foram fornecidos a CNN. Com esta técnica, os autores alcançaram *accuracy* de $88,0 \pm 1,12\%$ para EEG coletados durante a execução de tarefas, nos quais analisaram *event-related potentials* (ERPs). Contudo, na avaliação de EEG de estado de descanso, este método alcançou uma *accuracy* de apenas $66,0 \pm 1,13\%$.

No trabalho de (VAHID *et al.*, 2019), os sinais de EEG foram coletados ao longo de uma tarefa de estímulo visual, na qual os participantes estimavam a passagem do tempo. Sua base de dados contou com 18 sujeitos com TDAH-C, 15 com TDAH-I e 25 Controle. Para a classificação, utilizaram uma CNN com quatro camadas convolucionais, duas camadas de *pooling* e uma camada de neurônios (densa) chamada de EEGNet. Nessas condições, este grupo obteve $83,0 \pm 23\%$ de *accuracy* na classificação entre TDAH-I e grupo controle. E, uma *accuracy* de $80 \pm 21\%$ entre TDAH-C e grupo controle. Quando foram consideradas as três classes simultaneamente, a *accuracy* foi reduzida para $55,5\%$. Através deste estudo, observou-se diferenças entre os subtipos no córtex superior parietal.

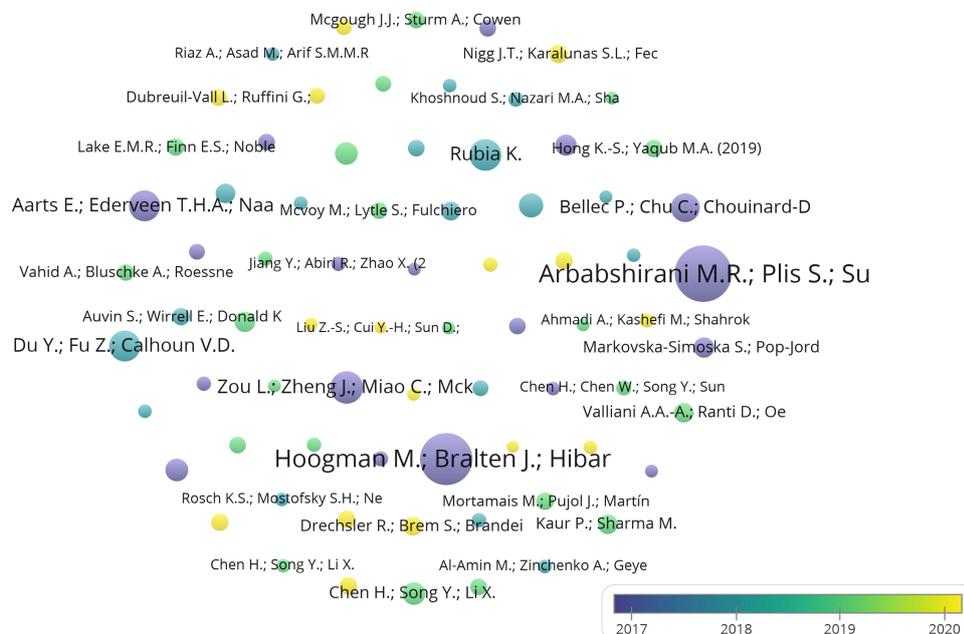
A Figura 2 mostra o mapa de co-citações entre autores limitado a estudos envolvendo EEG. Neste mapa destacam-se os autores já comentados (CHEN *et al.*, 2019b), (VAHID *et al.*, 2019), (MARKOVSKA-SIMOSKA & POP-JORDANOVA, 2017) e (DUBREUIL-VALL *et al.*, 2020), além das pesquisas de (NIGG *et al.*, 2020), (LOO *et al.*, 2013) e

(DRECHSLER *et al.*, 2020). Em seu trabalho de revisão, (NIGG *et al.*, 2020) discutiu diversos estudos de classificação de TDAH com métodos computacionais baseados em comportamento, traços emocionais, EEG e MRI.

A revisão de (DRECHSLER *et al.*, 2020), contudo, mostra que apesar de uma ampla gama de pesquisas sobre TDAH ter sido realizada entre 10 e 20 anos antes desta revisão, nenhum dos trabalhos anteriores havia apresentado marcadores biológicos que pudessem ser utilizados para classificar indivíduos com TDAH de forma a não cometer equívocos. Esta revisão discute o conceito e definições de TDAH, assim como controvérsias nesta área.

O estudo de (LOO *et al.*, 2013) se tornou importante por constatar que a razão entre as bandas Teta e Beta difere significativamente em adultos, mas não em crianças com TDAH em relação ao grupo controle. A razão teta/beta (como costuma ser chamada) foi apresentada por (LUBAR, 1991) como um possível biomarcador de TDAH. No entanto, diversos estudos subsequentes não conseguiram identificar diferenças significativas na razão teta/beta em pessoas com TDAH, o que (LOO *et al.*, 2013) atribui a uma série de diferenças metodológicas relativas às amostras, aos parâmetros de EEG e às diferenças na forma da análise.

Figura 2 – Mapa de co-citações de autores de EEG.



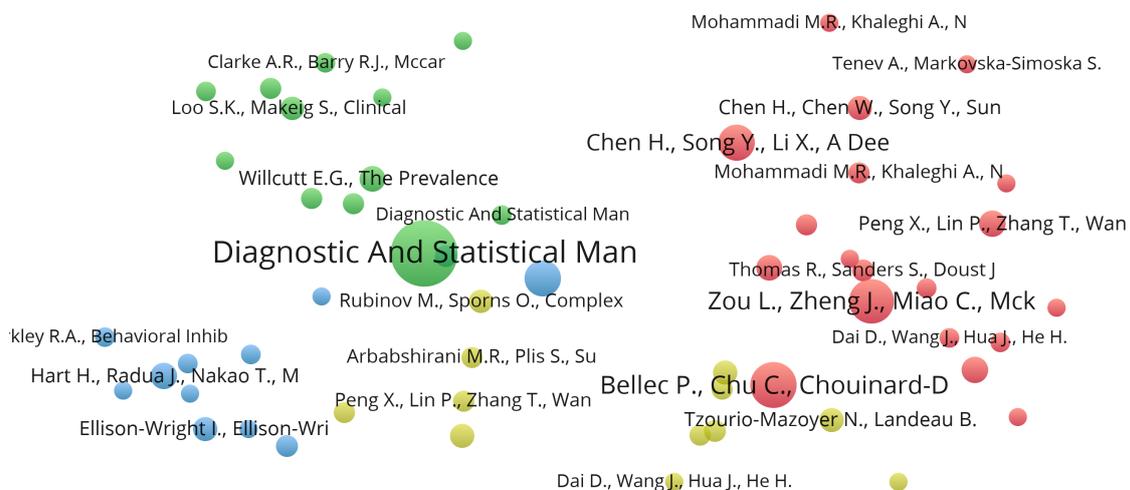
Fonte: O Autor

Na Figura 3, apresenta-se o mapa de co-citações de fMRI, as quais ocorrem quando um determinado trabalho cita ambos os trabalhos conectados no mapa. Neste mapa, cada nó representa um trabalho e a dimensão de seu círculo é proporcional ao seu número de citações. A espessura das linhas aumenta à medida que um determinado par de autores re-

cebe uma maior quantidade de citações conjuntas por um terceiro autor. As cores dos nós variam de acordo com agrupamentos organizados pelo *software Vosviewer*. Neste mapa, o maior número de co-citações pertence ao manual DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013). Outras duas edições do manual DSM-V também aparecem no mapa dos mais citados. Esta é a publicação mais citada porque a maioria dos psicólogos e psiquiatras se baseia nos princípios do manual DSM-V para o diagnóstico do TDAH e de seus subtipos. Uma vez que esta avaliação determina as classes a qual pertencem os sujeitos em bancos de dados, o manual recebe citações em grande parte das publicações sobre o tema.

No mapa da Figura 3 também se destacaram os já citados trabalhos de (BELLEC *et al.*, 2017) e (ZOU *et al.*, 2017), além das meta-análises de (HART *et al.*, 2013) e de (HART *et al.*, 2012).

Figura 3 – Mapa de co-citações de autores de fMRI.



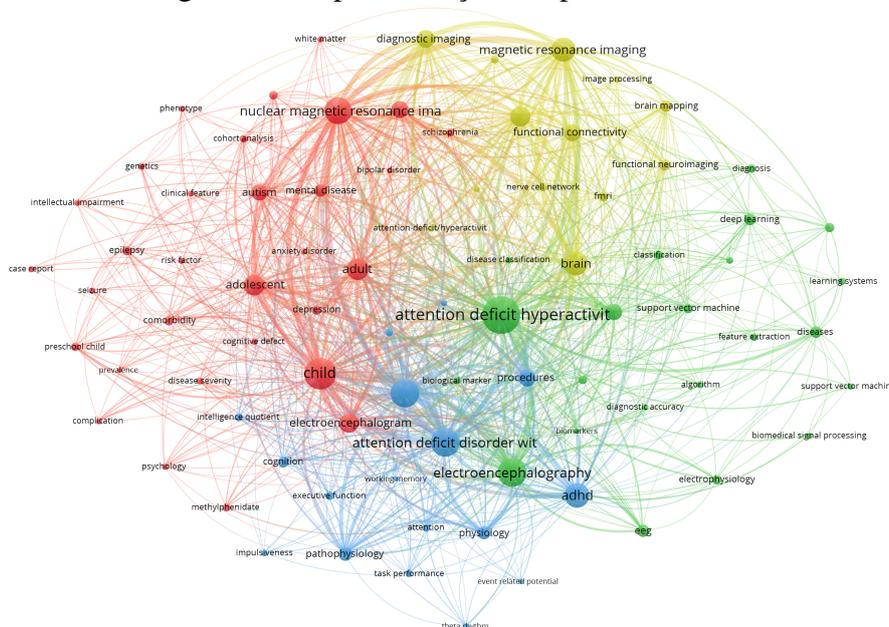
Fonte: O Autor

(HART *et al.*, 2012) realizou uma meta-análise envolvendo 11 estudos de fMRI sobre TDAH envolvendo ao todo 150 pacientes de TDAH e 145 participantes do grupo controle. Nestes estudos, os indivíduos realizavam tarefas ao longo dos exames. Com esta metodologia constataram que funções de temporização cerebrais sofrem alterações em pessoas com TDAH. Os pacientes apresentam diferenças nas partes frontal e parietal do cérebro, além do cerebelo. Porém, a longo prazo, medicação estimulante pode normalizar as ativações na região frontal.

Em seu estudo seguinte, (HART *et al.*, 2013), estes autores realizaram uma nova meta-análise envolvendo 21 estudos sobre fMRI coletados em tarefas de inibição de impulsos e 13 adquiridos em tarefas de atenção. Nas tarefas de inibição, os pacientes apresentaram uma menor ativação no córtex frontal, na área motora suplementar, no córtex cingulado anterior e no tálamo. Já nas tarefas de atenção, os participantes com TDAH apresentaram baixa ativação córtex pré-frontal dorsolateral, gânglios basais posteriores e nas regiões parietal e do tálamo.

Na Figura 4, apresenta-se o mapa de citações de palavras-chave. Nesta figura, os resultados foram filtrados, de modo a somente representar palavras-chave com mais de 40 ocorrências. Neste mapa, a palavra de maior destaque é *attention deficit hyperactivity disorder*, que é a grafia por extenso em inglês de TDAH.

Figura 4 – Mapa de citações de palavras-chave.



Fonte: O Autor

O *cluster* vermelho contém palavras relacionadas a outros transtornos, como epilepsia, autismo, bipolaridade, depressão, ansiedade, deficiência intelectual e esquizofrenia, além de convulsão. Algumas destas são comorbidades de TDAH. Neste *cluster*, também se destacam palavras referentes às faixas etárias (pré-escolar, criança, adolescente e adulto). O medicamento metilfenidato utilizado no tratamento de TDAH é apresentado neste agrupamento. Outras palavras que recebem destaque são a matéria branca do cérebro (normalmente medida em sMRI), fator de risco, prevalência, complicação, defeito cognitivo, fenótipo, característica clínica, convulsão, relatório de caso, genética e característica clínica, psicologia, ressonância magnética nuclear e eletroencefalograma.

Apesar de que palavras de maior destaque no *cluster* verde são TDAH e eletroencefalografia, a maioria das demais palavras chave deste grupo são referentes aos algoritmos de

classificação: SVM, *deep learning*, algoritmo, classificação, sistemas de aprendizado, extração de características. Além destas, o mapa mostra neste agrupamento palavras como eletrofisiologia, biomarcadores, processamento de sinal biomédico, *accuracy* de diagnóstico, doenças e diagnóstico.

O *cluster* amarelo relaciona palavras utilizadas em estudos de fMRI, como fMRI, conectividade funcional, MRI, diagnóstico por imagem, mapeamento cerebral, neuroimagem funcional, cérebro e processamento de imagens. Além destas, as palavras TDAH e rede de células nervosas foram ressaltadas neste *cluster*. Rede de células nervosas é uma forma de referir-se às redes neurais naturais do cérebro.

No *cluster* azul destacam-se o ritmo de EEG Teta, a memória de trabalho (*working memory*), cognição, impulsividade, quociente de inteligência, atenção, desempenho em tarefas, *event-related potential* (ERP), fisiologia, função executiva, procedimentos e patologia. Ou seja, neste *cluster* são mostradas palavras relacionadas a características percebidas como diferentes em pessoas com TDAH. Observa-se o destaque, ainda que pequeno da banda de frequência Teta associada ao TDAH.

2.2 Análise dos Trabalhos do Estado da Arte

Optou-se por analisar os trabalhos na amplitude temporal a partir de 2012, pois há neste ano um marco temporal: a realização da competição de classificadores de dados de pessoas com TDAH conhecida como ADHD-200 (CONSORTIUM, 2012) que impulsionou a pesquisa nesta área, especialmente, na parte de fMRI.

Observou-se tanto para os trabalhos de EEG como para os de fMRI que estes não seguem um padrão único. Desta forma, apresentam diferenças relevantes no tamanho e balanço das bases de dados; no protocolo de coleta de dados; nas técnicas de extração de características e pré-processamento e nas técnicas de classificação. Por isso, todas as comparações entre trabalhos são indiretas. Muitos trabalhos apenas realizam classificações entre TDAH e grupo controle, ignorando a questão das diferenças entre os subtipos de TDAH.

Notoriamente, as bases de dados tem diferentes quantidades de indivíduos, variando entre menos de uma dezena até centenas. Observa-se que muitos trabalhos com *accuracies* elevadas estão associados a uma quantidade pequena de indivíduos amostrados. Isto reduz a robustez e a generalização das análises, pois os sujeitos mensurados podem formar uma amostra mais homogênea do que a população real ou podem possuir alguma característica com valor anômalo para determinada classe.

As bases de dados apresentam diferenças nas distribuições de gênero e de idade dos participantes, podendo trabalhar apenas com adultos, somente com crianças, ou misturar crianças, adolescentes e adultos.

Os paradigmas de coleta se dividem principalmente em duas categorias: coleta em

estado de repouso e durante uma tarefa. As tarefas variam de observação de figuras, reação (às imagens, aos vídeos ou aos sons), apertar botões, jogar *videogame*, entre outras que podem durar de 30 segundos até 20 minutos. A coleta em estado de repouso ocorre com o participante parado. No entanto, há paradigmas onde os participantes permanecem com olhos abertos, fechados ou variam em abrir e fechar os olhos (ALEXANDER *et al.*, 2017).

Também, existem variações nos protocolos de coleta em relação aos equipamentos utilizados, especialmente em EEG, onde os estudos possuem variações de números de canais, frequência de amostragem, janelas de segmentação e tempo de amostragem. Em fMRI, os tempos de amostragem costumam ser de 10 minutos (ALEXANDER *et al.*, 2017).

Em EEG, por vezes, a extração de características é relatada como parte do pré-processamento. Além disso, o pré-processamento de EEG inclui a remoção da interferência da rede, separação de fonte, exclusão de artefatos, segmentação e seleção de bandas e métodos de normalização.

2.2.1 Referencial Teórico dos principais trabalhos de classificação de TDAH a partir de sinais de EEG

Em seu estudo (ALCHALABI *et al.*, 2018), citado por 94 artigos, utilizou uma SVM com um *kernel* de Função de Base Radial (RBF) para classificar segmentos de EEGs pertencentes a 4 pacientes com TDAH e a 5 indivíduos do grupo controle. Os dados destes indivíduos foram segmentados de forma os 9 sujeitos passaram a ser representados por 7545 amostras. Nestas condições, os autores obtiveram uma *accuracy* média de $98,2 \pm 0,3\%$ utilizando a validação cruzada *k-fold*. A inovação deste trabalho está na coleta de EEG enquanto crianças jogavam *videogame* durante 20 minutos. Neste trabalho, não se investigou rs-EEG e nem a classificação de subtipos.

A pesquisa de (PEREDA *et al.*, 2018) (34 citações) trabalhou com classificação de sinais de rs-EEG de 14 sujeitos da classe controle e 19 da classe TDAH-C. Estes autores constataram uma menor conectividade entre as regiões cerebrais em sujeitos com TDAH-C em relação ao grupo controle. Estes autores obtiveram uma *accuracy* 95,0% com sua rede *baeysiana* (BN) na classificação binária entre as classes TDAH e grupo controle. Porém, os autores observaram que suas métricas podem ter sido elevadas em função de seu banco de dados ser composto de uma pequena quantidade de sujeitos, o que pode ter contribuído para a homogeneidade das amostras.

O estudo de (ZHANG & LI, 2019) utilizou 50 amostras de EEG de crianças com TDAH com idades na faixa de $10,44 \pm 0,75$ e 57 crianças saudáveis com a mesma faixa etária. Neste estudo, as amostras de EEG foram gravadas por 10 minutos utilizando o sistema *HydroCel Geodesic Sensor Net* de 128 canais com uma frequência de amostragem de 1000 Hz. Os autores selecionaram apenas 32 dos 128 canais com segmentação de 4

segundos. Então, removeram os artefatos de interferência da rede, de atividade muscular e de movimentos oculares e filtraram os sinais restantes de EEG entre 0,5 e 40 Hz. Calcularam o índice de atraso de fase entre cada par de eletrodos. A média destes índices de atraso de fase através dos períodos de 4 segundos produziu matrizes de conectividade, as quais foram enviadas para uma *Visual Geometry Group Network* (VGG) de (SIMONYAN & ZISSERMAN, 2014), com 16 camadas de aprendizagem, a VGG-16. Estes autores utilizaram de aprendizado por transferência, aproveitando o treinamento prévio da arquitetura VGG-16 para a clássica base de dados Cifar10. Estes autores usaram as cinco camadas iniciais já treinadas da VGG-16 e adicionaram uma camada convolucional e outra de *pooling* para realizar a classificação das matrizes de conectividade de EEG e obter 94,2% de *accuracy* no diagnóstico de TDAH. Contudo, o estudo relatado considerou apenas classificação binária entre os grupos Controle e TDAH.

O trabalho de (MOGHADDARI *et al.*, 2020) utilizou uma CNN (de três camadas convolucionais e duas camadas densas) para analisar imagens formadas por sinais de EEG coletados em tarefas de 31 indivíduos com TDAH e 30 sujeitos do grupo controle. Os artefatos de piscadas de olho, movimentos musculares e batimento cardíaco foram retirados através da *toolbox* do MATLAB conhecida como EEGLab (DELORME & MAKEIG, 2004). Depois, os pesquisadores dividiram cada sinal de 30 segundos em 10 segmentos. A seguir, filtraram as bandas de frequência teta, alfa e beta, associando cada banda com uma cor. Então, produziram uma imagem para cada indivíduo onde os eixos das linhas (x) e colunas (y) correspondiam aos canais de EEG e a quantidade de amostras ao longo do tempo, respectivamente. Os valores resultantes da soma das cores formaram os *pixels* da imagem, depois foi normalizada e disponibilizada como entrada à CNN. Através deste método, os autores obtiveram uma *accuracy* de 97,47% na classificação binária de entre classes TDAH e Controle.

O estudo de (AHMADI *et al.*, 2021) (62 citações) utilizou uma CNN especialmente projetada para classificar subtipos de TDAH. Para tanto, utilizaram sinais de rs-EEG coletados com olhos permanentemente abertos de uma base de dados com 14 voluntários do grupo controle, 13 do grupo TDAH-C e 12 do grupo TDAH-I. Os autores obtiveram uma *accuracy* de 99,7% nesta classificação. Através deste experimento, constataram potências menores na faixa Alfa e maiores nas faixas Delta e Teta em indivíduos com TDAH.

A pesquisa de (TOSUN, 2021) utilizou uma rede neural recorrente (RNN) do tipo *Long Short-Term Memory* (LSTM) para classificar uma base de 16 indivíduos igualmente dividida entre o grupo controle e o TDAH. Os sinais de EEG foram coletados em um protocolo onde os olhos eram abertos e fechados repetidamente com presença de estímulos fotográficos, depois o participante era orientado à respirar fundo e descansar. O autor obteve uma *accuracy* de 92,2% ao classificar 17 segmentos de 3 segundos, com *overlap* de 50%, dos indivíduos entre estas duas classes durante períodos com os olhos fechados. A pesquisa identificou uma diferença entre os eletrodos das regiões pré-frontal e frontal

esquerdas do cérebro em pacientes com TDAH.

O trabalho de (HOLKER & SUSAN, 2022) (2 citações) obteve *accuracy* de 81,8% com uma RF para classificar dados coletados em EEG de tarefa de 60 e 61 participantes dos grupos controle e TDAH, respectivamente, entre estas classes. A partir desta metodologia, os autores constataram uma relação entre a presença de TDAH e alterações nas regiões Frontal, Parietal e Occipital. O estudo de (ABBAS *et al.*, 2021) realizou uma análise estatística e a curva ROC na mesma base de dados, obtendo uma *accuracy* de 75,4%. Através deste estudo, observou que há maior discriminação entre as classes na banda Beta considerando como características a entropia e a transferência de informação.

A pesquisa de (KHARE *et al.*, 2022) (21 citações) utilizou uma base de dados formada por sinais de EEG em tarefa de 61 e 60 participantes dos grupos TDAH e controle, respectivamente, com idades entre 7 a 12 anos. Este trabalho apresentou um método de pré-processamento chamado *variational mode and Hilbert transform-based EEG rhythm separation* (VHERS). Este método estima de forma adaptativa os intervalos relevantes das bandas de frequência: Delta (0–4 Hz), Teta (4–8 Hz), Alfa (8–12 Hz), Beta (12–30 Hz) e Gama (30–60 Hz) para a classificação. Como características, este trabalho utilizou diferentes formas de calcular a entropia, a média e o desvio padrão. Para o classificador, este grupo empregou uma Máquina de Aprendizado Extremo (do inglês, *extreme learning machine* - ELM) e a comparou com Árvores de Decisão (do inglês, *do inglês, Decision Trees* - DT), uma SVM e um k-vizinhos mais próximos (do inglês, *K-nearest neighbors* (k-NN)). Em seu melhor resultado, obteve uma *accuracy* de 99,95% para a classificação binária TDAH/Controle analisando a banda Delta. As *accuracies* obtidas com as bandas Teta, Alfa, Beta e Gama foram, respectivamente, de 97,49%, 97,89%, 96,96% e 93,44%. No entanto, este grupo não estudou a classificação de subtipos de TDAH.

O trabalho de (BAKHTYARI & MIRZAEI, 2022) desenvolveu uma rede LSTM com uma camada convolucional (ConvLSTM) e um módulo de atenção para classificar sinais de EEG coletados em tarefas de 30 segundos. Este grupo utilizou uma base de dados formada por 46 sujeitos com TDAH e 45 do grupo controle. Contudo, cada sujeito desta base corresponde a um número diferente de amostras. Por isso, há um desbalanço no número de amostras: 233 de TDAH e 167 do grupo de controle. Os autores não descreveram o uso de nenhuma técnica de balanceamento de amostras. Com esta abordagem, os autores obtiveram uma *accuracy* média de 99,34%, na classificação binária entre TDAH e grupo controle. Para tanto, usaram como biomarcadores padrões de conectividade dinâmica da combinação das bandas Teta, Alfa, Beta e Gama. Os autores relatam terem obtido *recall* e Precisão de 99,7% e de 99,7%, respectivamente, entretanto estes são valores máximos e não médios, os quais não foram relatados.

A Tabela 1 compara os tipos de dados, técnicas, bancos de dados, *accuracies* e contribuições obtidas, além de especificar se foi realizada a classificação entre TDAH e grupo controle ou entre subtipos e o número de citações recebido por cada publicação. Esta

tabela apresenta os principais trabalhos recentes, publicados nos últimos 5 anos.

Tabela 1 – Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com EEG.

Referência (Número de citações)	Tipo de dado	Técnica	Banco de dados	Accuracies	Contribuição
(PEREDA <i>et al.</i> , 2018) 34 citações	rs-EEG	BN	14 Controle 19 TDAH-C	95,0% (binária)	Conectividade menor em TDAH
(VAHID <i>et al.</i> , 2019) 72 citações	EEG em tarefa	EEGNet	25 Controle 18 TDAH-C 15 TDAH-I	83,0% (subtipos)	Observou diferenças no córtex superior parietal
(CHEN <i>et al.</i> , 2019a) 130 citações	rs-EEG	CNN proj.	51 Controle 50 TDAH	94,7% (binária)	Matrizes de conectividade funcional para EEG
(DUBREUIL-VALL <i>et al.</i> , 2020) 83 citações	EEG em tarefa rs-EEG	CNN	20 Controle 20 TDAH	88,0% (binária)	Diferença entre classificar EGG em tarefa e rs-EEG
(AHMADI <i>et al.</i> , 2021) 62 citações	rs-EEG	CNN proj.	14 Controle 13 TDAH-C 12 TDAH-I	99,5% (subtipos)	Alterações de potência na faixa Delta, Teta e Alfa para TDAH
(TOSUN, 2021) 23 citações	rs-EEG	LSTM	8 Controle 8 TDAH	92,2% (binária)	Diferença entre pré-frontal e frontal no esquerdo
(KHARE <i>et al.</i> , 2022) 21 citações	EEG em tarefa	VHERS ELM	60 Controle 61 TDAH	97,9% (binária)	Relevância da análise da banda Delta

O estudo de (JOY *et al.*, 2022) (23 citações) utilizou uma rede neural artificial (RNA) do tipo *Multi-Layer Perceptron* (MLP) com apenas uma camada oculta e comparou o desempenho do classificador com diferentes formas de se calcular a entropia. Com a entropia de permutação, obteve 99,8% de *accuracy* média em uma base de dados com

apenas 10 crianças com idades entre 7 a 12 anos (metade com TDAH e metade sem). Os dados de cada indivíduo foram transformados em 5 amostras de 60 segundos através da segmentação dos sinais de EEG. Encontrou diferenças significativas em TDAH para a região frontal e pré-frontal em sinais de EEG coletados com olhos fechados com testes *t* de *Student*.

2.2.2 Referencial Teórico dos principais trabalhos de classificação de TDAH a partir de fMRI

Em 2011, o consórcio ADHD-200 (CONSORTIUM, 2012) promoveu a competição de mesmo nome, onde imagens de fMRI e de MRI estrutural (sMRI) de 947 indivíduos (362 com TDAH e 585 controles), com idades entre 7 e 21 anos, foram disponibilizadas para a comunidade científica (BELLEC *et al.*, 2017). Os dados da competição foram coletados em oito locais diferentes. Nesta competição, os pesquisadores foram incentivados a desenvolver métodos de classificação de dados para distinguir indivíduos com TDAH do grupo controle. A competição ADHD-200 contou com 21 equipes (DAI *et al.*, 2012) e resultou em 84 publicações em revistas e congressos pelas equipes participantes segundo (BELLEC *et al.*, 2017). Ao longo dos anos seguintes a 2011 os dados da competição ADHD-200 (CONSORTIUM, 2012) seguiram sendo utilizados por diversos pesquisadores para o teste de métodos de diagnóstico de TDAH a partir de fMRI. Estas publicações apresentaram classificadores com accuracies entre 55,0 e 70,7%. Na classificação de subtipos, a máxima *accuracy* obtida na competição foi de 64,7%. Esta competição representou um marco temporal para a classificação automatizada de TDAH, impulsionando o interesse e aumentando a quantidade de trabalhos na área a partir de sua realização.

O trabalho de (COLBY *et al.*, 2012) extraiu as seguintes características de fMRI da competição ADHD-200: correlação de Pearson das matrizes de conectividade, medidas teóricas de gráficos nodais e globais, espectros de potência nodal, conectividade global *voxel-wise* e homogeneidade regional *voxel-wise*. Estas características ranqueadas através do uso de uma SVM para eliminação recursiva de características (RFE), as quais foram selecionadas priorizando a generalização em resultados de testes obtidos por uma SVM com *kernel* baseado em função radial (RBF). Com esta abordagem, o estudo de (COLBY *et al.*, 2012) obteve 55,0% de *accuracy*, 33,0% de sensibilidade e 80,0% de especificidade na classificação binária de TDAH.

O método de (DAI *et al.*, 2012) obteve a sexta colocação dentre as 21 equipes participantes da competição ADHD-200 utilizando *multi-kernel learning* (MKL) para integrar características multimodais das imagens. Esta equipe utilizou características de conectividade funcional. Este trabalho utilizou uma técnica de seleção de características, após constatar que a dimensionalidade das características utilizadas aumentava a complexidade computacional e elevava o *overfitting*. Para tanto, criaram uma lista ordenada com a capacidade discriminativa das características que eram avaliadas segundo critérios de

within-class sum of squares (WSS) e *between class sum of squares* (BSS). Assim como o já mencionado trabalho de (COLBY *et al.*, 2012), o estudo de (DAI *et al.*, 2012) empregou uma SVM com RFE. Esta SVM com RFE elimina iterativamente as características que pouco contribuem para a discriminação das classes. Consequentemente, esta técnica otimiza o desempenho da SVM com *kernel* RBF utilizada para a classificação. Com esta abordagem, estes autores obtiveram 61,54% de *accuracy* na classificação binária e 54,1% da *accuracy* na classificação entre controle, TDAH-C e TDAH-I. Este trabalho contou com 776 participantes da base ADHD-200, dos quais 195 foram utilizados para o teste.

O estudo de (BROWN *et al.*, 2012) alcançou o primeiro lugar na competição ADHD-200, obtendo uma *accuracy* de 62,52% na classificação multi-classe (controle, TDAH-C e TDAH-I) através do uso de um classificador logístico baseado nas características pessoais (local da coleta dos dados, idade, gênero, lateralidade e testes de quociente de inteligência) e ignorando os dados de fMRI. Este estudo também avaliou dados de fMRI, porém com menor sucesso. Este trabalho utilizou fMRIs de 839 participantes no treinamento e 171 no teste, o que totaliza 839 indivíduos. Este estudo utilizou uma técnica de redução de dimensionalidade (para reduzir o *overfitting*) e uma máscara binária para excluir os *voxels* que não pertencem ao cérebro. Para a redução de dimensionalidade (70.651.500 características por participante), o trabalho de (BROWN *et al.*, 2012) empregou a técnica de *spatial window averaging* (SWA). Esta técnica utiliza médias móveis nas três dimensões espaciais (altura, largura e comprimento) para reduzir a quantidade de *voxels*. Em seguida, estes dados foram submetidos a dois processos de normalização espacial. A dimensionalidade temporal foi reduzida usando Análise de Componentes Principais (*Principal Component Analysis* – PCA) de 370 para apenas 1 componente. Como classificador empregaram uma SVM linear e obtiveram uma *accuracy* de $70,7 \pm 6,2\%$ para classificação TDAH/Controle. Para a classificação de subtipos os autores alcançaram seus melhores resultados (*accuracy* de $64,7 \pm 4,3\%$) substituindo a normalização espacial por uma normalização temporal e utilizando uma SVM linear.

O trabalho de (DEY *et al.*, 2014) constatou uma diferença considerável para a classificação quando se separa os indivíduos de acordo com o gênero antes de se realizar a classificação dos fMRI. Neste trabalho, os autores trabalharam com 307 voluntários do grupo controle e 180 pacientes com TDAH, sendo que 25% de cada grupo foi destinado ao teste. Os autores utilizaram escalonamento multi-dimensional (do inglês, Multi-Dimensional Scaling - MDS) para reduzir a dimensionalidade dos dados. Para a classificação aplicaram uma SVM com *kernel* polinomial. A separação por gênero elevou a *accuracy* do classificador de 62,8% para 73,6% na classificação entre controle e TDAH.

O trabalho de (DESHPANDE *et al.*, 2015) (157 citações) utilizou medidas de conectividade como características para a classificação estimadas através de *Kernel Granger Causality* (KGC) e de correlações entre probabilidades de recorrência (*Correlation between probabilities of recurrences* - CPR). Através das correlações CPR, os autores obtiveram

as sincronizações de fase entre pares de sinais. A KGC estabelece relação de causalidade entre duas séries temporais. Os autores selecionaram as características através da PCA. Como classificador, utilizaram uma RNA completamente conectada em cascata, treinada com um algoritmo *neuron-by-neuron*. As RNAs completamente conectadas em cascata são RNAs de múltiplas camadas, nas quais há apenas um neurônio por camada. Este trabalho obteve *accuracies* máximas de quase 90% para distinguir pacientes com TDAH (433) de voluntários saudáveis (744) e de quase 95% na distinção entre os subtipos TDAH-C (260) e TDAH-I (173) considerando a amostra desbalanceada. Contudo, as *accuracies* médias obtidas através deste método são consideravelmente menores, em especial, quando esta base é balanceada: em torno de 70% e de 60% para TDAH/Controle e Subtipos, respectivamente. Este trabalho mostrou que as características de conectividade tem maior capacidade de discriminação do que o uso do dado bruto.

O estudo de (QURESHI *et al.*, 2016) utilizou dados de sMRI de 159 indivíduos, sendo um terço de cada grupo (TDAH combinado, TDAH desatento e grupo de controle) da base de dados da ADHD-200 e obteve uma *accuracy* de 60,8% utilizando uma ELM hierárquica e uma RFE. A ELM hierárquica, foi comparada com uma ELM clássica e SVM com *kernels* linear e RBF avaliando as características do volume, espessura média cortical, área da superfície, índice de curvatura médio e índices dobráveis. Na classificação binária entre as três possíveis combinações destes grupos, os autores avaliaram estes resultados tanto por validação cruzada como por validação tradicional.

Já (GHIASSIAN *et al.*, 2016) (105 citações) combinou dados de fMRI, sMRI e informações pessoais de 490 indivíduos do grupo controle e 279 do grupo TDAH da base ADHD-200. Contudo, neste estudo foram utilizados 27,6% dos indivíduos de TDAH e apenas 19,2% dos participantes saudáveis para o teste dos classificadores. Este trabalho usou a técnica histogramas orientados a gradientes (HOG) para a extração de características e uma SVM com *kernel* RBF para a classificação. Com estas técnicas, os autores obtiveram 69,6% de *accuracy* para a classificação binária entre TDAH e grupo controle. A intenção dos autores deste trabalho era produzir um classificador que pudesse ser utilizado para bancos de dados com a intenção de classificar transtornos diferentes. Por isso, também trabalharam com uma base de dados voltada ao estudo do autismo.

O trabalho de (RIAZ *et al.*, 2017) analisou os sinais brutos de fMRI com uma CNN em busca de encontrar características latentes dos sinais brutos. Este trabalho emprega uma CNN para extrair características das séries temporais e uma RNA que verifica a similaridade entre estas características, constituindo uma rede siamesa. Uma *Elastic Net* seleciona as características, eliminando aquelas altamente correlacionadas. As características selecionadas foram combinadas com dados pessoais. Separando os indivíduos de cada local de coleta da base ADHD-200 e utilizando uma SVM como classificador, os autores obtiveram *accuracies* de 64.0%, 68.6% e 63.4% para cada centro de coleta de fMRI. Em seu trabalho seguinte, (RIAZ *et al.*, 2018), os autores substituíram o uso da rede siamesa

por uma técnica de agrupamento por propagação de afinidade (FREY & DUECK, 2005) em conjunto com o algoritmo de densidade de picos (RODRIGUEZ & LAIO, 2014) para estimar a conectividade funcional. Como em seu trabalho anterior, (RIAZ *et al.*, 2018) usou a *Elastic Net* para selecionar características de fMRI que, depois, foram unidas com os dados pessoais. Os conjuntos de características que compõem as amostras enviadas à SVM com *kernel* linear foram amostrados utilizando *synthetic minority oversampling technique* (SMOTE). A SVM destes autores obteve *accuracies* de 86,7%, 52,7% e 85,8% incluindo dados pessoais e 67,4%, 25,4% e 85,3% utilizando apenas fMRI para cada um dos três locais de coleta da ADHD-200.

O estudo de (JHA *et al.*, 2020) utilizou um *automated anatomical labeling* atlas para dividir cada cérebro em 116 ROI. E, calculou-se a média de todos os sinais captados dos *voxels* de rs-fMRI para cada ROI. Os sinais correspondentes às ROI variam ao longo do tempo. Os autores segmentaram estes sinais em 171 janelas. Então, um grafo foi construído como uma matriz 116x171 relacionando cada região com o passo temporal. Ao aplicar a autocorrelação ao vetor de médias, produziu-se uma matriz de correlações. Através de uma função limiar aplicada sobre a matriz de correlações, os autores geraram uma matriz de grafo adjacente. Os grafos gerados são entregues como características para 3 CNNs siamesas trigêmeas que compartilham dos mesmos pesos. Cada uma das CNNs calcula uma incorporação (*embedding*) que enfatiza as similaridades ou diferenças entre os grafos. Estas, por sua vez, servem para minimizar a função custo e reduzir o *overfitting*. Os autores utilizaram 311 indivíduos da base ADHD-200, dos quais 270 foram usados no treinamento do modelo. O conjunto de teste foi formado por 12 voluntários controle e 29 pacientes, o que constitui 13% do total de participantes. Com esta abordagem, obtiveram uma *accuracy* de 68,29% na classificação binária de TDAH.

A pesquisa de (MENG *et al.*, 2022) investigou a base ADHD-200 utilizando uma abordagem com uma rede neural com aprendizado por transferência. Este estudo utilizou 142 pacientes com TDAH e 102 indivíduos do grupo controle de parte da base ADHD-200 coletada na *Peking University Child Study Center*. Os autores optaram por utilizar somente os dados deste local, supondo que possíveis diferenças na coleta em cada local poderiam influenciar o classificador. Para o pré-processamento de fMRI, os autores utilizaram o programa *Data Processing Assistant for Resting-State fMRI* (DPARSFA) na plataforma MATLAB (R2016a). Neste *software*, aplicaram uma correção temporal, excluíram sujeitos com movimentação excessiva de cabeça, aplicaram uma normalização espacial e utilizaram um *kernel* Gaussiano para suavizar as imagens, e, desta forma reduzir o ruído. Calcularam os coeficientes de correlação parcial e de Pearson para obter duas matrizes de conectividade funcional por participante. Estas serviram como entrada para os classificadores. Este trabalho, então, comparou a aplicação de duas arquiteturas tipicamente usadas para aprendizado por transferência: a VGG (SIMONYAN & ZISSERMAN, 2014) e a *Residual Neural Network* (ResNet-50) de (HE *et al.*, 2016). Os modelos VGG

e ResNet obtiveram, respectivamente, 82,0% e 80,0% de *accuracy* para a classificação binária entre grupos TDAH e controle.

Em seu estudo, (QURESHI *et al.*, 2017) analisaram tanto a classificação binária como a classificação de subtipos de TDAH. Os autores restringiram seu escopo à análise dos subtipos TDAH-C e TDAH-I, o que é o mais comum na literatura, uma vez que o subtipo TDAH-H é muito raro nos bancos de dados. Este trabalho avaliou características de sMRI e fMRI em estado de repouso (rs-fMRI) com uma ELM, obtendo *accuracies* de 92,9% e de 76,2% para os diagnósticos de TDAH/Controle e de subtipos, respectivamente. Suas análises mostraram que a espessura das redes de atenção dorsal e ventral diminuem com a idade no grupo Controle, mas não no grupo TDAH.

A pesquisa de (MAO *et al.*, 2019) utilizou uma base de dados maior do que as demais, contando com exames de fMRI de 439 voluntários do grupo controle e 359 pacientes com TDAH. Este grupo empregou uma CNN com a extração de características espaço-temporais realizada somente pelas suas camadas convolucionais e, assim, obteve uma *accuracy* de 71,3% na classificação TDAH/Controle. Este trabalho utilizou técnicas de aumento de dados para aumentar em 20 vezes seu banco de dados. Os autores deste trabalho consideraram que havia um balanço relativo entre as classes TDAH e controle, as quais compreendiam 45% e 55% da base de dados usada, respectivamente. Por isto, avaliaram que não era necessário aplicar técnicas de subamostragem ou de sobreamostragem dos dados.

O trabalho de (DE SILVA *et al.*, 2021) constatou que pessoas com TDAH apresentam diferenças em características de correlação baseada em semente (SBC) empregando uma CNN. Seu estudo contou com exames de rs-fMRI de 20 indivíduos de cada classe, Controle e TDAH e obteve uma *accuracy* de 85,0%.

O estudo de (SAHA & SARKAR, 2022) utilizou DTs para investigar regras determinísticas que poderiam ser usadas para o diagnóstico de TDAH. Esta pesquisa pretendia desenvolver um método que facilitasse a interpretabilidade dos resultados. Utilizando como regras, as diferenças entre as redes de conectividade *Default* e *Límbica*, os autores obtiveram uma *accuracy* de 77,1% na classificação de subtipos. Este trabalho se baseou em uma base de dados composta por 86 indivíduos do grupo controle e 86 pacientes com TDAH, sendo metade com TDAH-C e a outra metade com TDAH-I.

A pesquisa de (LIU *et al.*, 2022) investigou características espaço-temporais extraídas de exames de fMRI de 289 pacientes com TDAH e 389 do grupo Controle. Os autores não mencionaram o uso de técnicas para o balanceamento das amostras. A classificação entre TDAH e grupo controle obteve 72,4% de *accuracy*. A técnica *nested residual convolutional denoising autoencoder* (NRCDAE) foi utilizada para retirar ruído dos sinais de fMRI e extrair as características espaço-temporais. Uma rede neural recorrente (RNN) do tipo *gated recurrent unit* (GRU) com uma camada convolucional foi aplicada para analisar os sinais de fMRI que foram, então, enviados a um classificador sigmoide.

Produziu-se uma tabela bibliométrica somente com trabalhos de classificação de TDAH (similar à Tabela 1). Assim, através da Tabela 2, destacam-se os trabalhos mais relevantes de fMRI. Nesta tabela, também observa-se a utilização (ou não) de um extrator de características antes da aplicação da técnica de *machine learning* empregada para a classificação.

Tabela 2 – Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com fMRI.

Referência (citações)	Tipo de dado	Técnica	Banco de dados	Accuracies	Extrator	Contribuição
(BROWN <i>et al.</i> , 2012) 199 citações	fMRI Fenótipo	SVM	523 Controle 192 TDAH-C 124 TDAH-I	70,7% (binária) 64,75 (subtipos)	Sim	Accuracy ampliada por uso de fenótipo
(QURESHI <i>et al.</i> , 2017) 68 citações	sMRI e rs-fMRI	ELM	53 Controle 53 TDAH-C 53 TDAH-I	92,9% (binária) 76,2% (subtipos)	Sim	Espessura Dorsal e Ventral não reduz com idade em TDAH
(MAO <i>et al.</i> , 2019) 145 citações	rs-fMRI	CNN	439 Controle 359 TDAH	71,3% (binária)	Não	Diferenças espaço-temporais
(DE SILVA <i>et al.</i> , 2021) 23 citações	rs-fMRI	CNN	20 Controle 20 TDAH	85,0% (binária)	Sim	Diferenças em SBC
(SAHA & SARKAR, 2022) 3 citações	rs-fMRI	DT	86 Controle 43 TDAH-C 43 TDAH-I	77,1% (subtipos)	Sim	Diferenças em <i>Default</i> e Límbica
(LIU <i>et al.</i> , 2022) 16 citações	rs-fMRI	NRCDA RNN- GRU	389 Controle 289 TDAH	72.4% (binária)	Sim	Diferenças espaço-temporais
(SHARMA <i>et al.</i> , 2023) 0 citações	rs-fMRI Fenótipo	Voto maj. RNA SVM	243 Controle 157 TDAH	86,7% (binária)	Sim	accuracy ampliada por uso de fenótipo

O trabalho de (ARIYARATHNE *et al.*, n.d.) utilizou correlações entre *voxels* e segmentos para distinguir 20 indivíduos com TDAH de 20 indivíduos do grupo controle na base de dados da competição ADHD-200 (CONSORTIUM, 2012). O estudo utilizou uma CNN de 7 camadas e pesquisou seu desempenho na classificação dos dados de 4 regiões diferentes do cérebro da rede de conectividade *Default*, obtendo no melhor caso 85,1% de *accuracy*.

O estudo multi-modal de (SHARMA *et al.*, 2023) combinou exames de rs-fMRI com dados fenotípicos de 243 voluntários do grupo controle e 157 indivíduos com TDAH da base ADHD-200. Em seu estudo (SHARMA *et al.*, 2023) utilizou uma *Independent Component Analysis* (ICA) para extrair as regiões de interesse (*Region of Interest* - ROI). O trabalho desses autores construiu matrizes de correlação para serem as características avaliadas pela RNA para classificar os exames de rs-fMRI. Este trabalho também apresenta um estudo multi-modal que combina as características de rs-fMRI com informações pessoais. Os dados fenotípicos foram avaliados com uma SVM. A fusão entre os dois métodos foi realizada através do voto majoritário com peso 0,5 para cada método. Através desta fusão, os autores aumentaram a *accuracy* de 86,7% obtida pelo classificador de fMRI para 92,5% na classificação TDAH/Controle. Assim, verificaram que usando dados fenotípicos é possível melhorar a performance dos classificadores, corroborando o que havia sido constatado por (BROWN *et al.*, 2012).

Porém, esta abordagem de (SHARMA *et al.*, 2023) reduziu a *accuracy* em relação ao uso somente dos dados pessoais, com os quais a SVM conseguia obter uma *accuracy* de 99,2%. As características fenotípicas avaliadas por estes autores incluem o número de identificação do sujeito, o local da coleta, a idade, o gênero, metodologia da medida do quociente de inteligência (QI), QI verbal, QI de desempenho, QI completo, medidas de controle de qualidade dos *pipelines* de pré-processamento da base ADHD-200. Além destas, os autores também usaram como características o diagnóstico de TDAH, o método de realização deste diagnóstico, a presença de déficit de atenção e, separadamente, a presença de comportamento hiperativo. O uso do diagnóstico como entrada e como saída pretendida em um sistema de aprendizagem supervisionado com RNA explica a *accuracy* de 99,2% obtida pelo classificador fenotípico. No entanto, não comprova a eficácia e utilidade deste método para classificar sujeitos que não tenham sido, previamente, diagnosticados. Mesmo que não houvessem usado o diagnóstico de TDAH tanto como entrada como saída desejada, um modelo que avalia apenas os dados pessoais não contribui para a análise de possíveis biomarcadores cerebrais relacionados ao transtorno.

2.2.3 Trabalhos com EEG e fMRI

O estudo multi-modal de (KARCH *et al.*, 2010) (69 citações) utilizou dados de EEG-fMRI simultâneos para investigar pacientes com TDAH utilizando apenas 8 pacientes entre 26 e 47 anos, além de 8 indivíduos sem TDAH com aproximadamente a mesma

faixa etária. Neste estudo, foi encontrada baixa correlação entre as medidas de EEG e de fMRI, contrariando outros estudos, provavelmente, devido à pequena quantidade de participantes do estudo. Os dados de EEG-fMRI foram coletados durante tarefas de apertar botão, na qual os pacientes com TDAH se saíram melhor, e em tarefas de escolha voluntária onde os pacientes com TDAH tiveram maior dificuldade. Observou-se que as resposta hemodinâmicas frontais e laterais foram menores em pacientes com TDAH.

O trabalho de (LENARTOWICZ *et al.*, 2016) (48 citações) utilizou a abordagem da medição simultânea de EEG e de fMRI para estudar a conectividade fronto-parietal e a dessincronização relacionada à eventos (do inglês *event-related desynchronization* - ERD) de ondas na faixa de frequências alfa em pacientes de TDAH para estudar sua memória de trabalho. A memória de trabalho corresponde a habilidade de armazenar e manipular informações transitoriamente na memória, a qual é uma das principais funções afetadas pelo TDAH conforme (LENARTOWICZ *et al.*, 2016). A pesquisa destes autores utilizou dados coletados de 30 meninos em idades entre 12 e 16 anos, metade deles com TDAH. Este estudo demonstrou que as ativações na região occipital são menores em participantes do tipo TDAH desatento, os quais possuem maior conectividade fronto-parieto-occipital, o que pode ser uma resposta do cérebro para compensar a menor atividade na região occipital.

O trabalho de (WIRSICH *et al.*, 2020) (16 citações) utilizou medições de EEG e de fMRI obtidas simultaneamente de 26 pacientes saudáveis para investigar a independência destas redes de conectividade funcional. De acordo com os autores, a conectividade funcional de EEG pode ser obtida através do casamento de fase (*phase-coupling*) dos sinais ou pelo casamento da amplitude utilizando correlação de Pearson. Ambas as medições de EEG e de fMRI foram segmentadas no tempo. Os autores então juntaram a parte superior das matrizes de correlação de fMRI com as matrizes de correlação de EEG para cada faixa de frequência de EEG (delta, teta, alfa, beta, gama) e utilizaram PCA para reduzir a dimensionalidade dos componentes que explicam de 75 a 90% dos dados. A seguir, aplicaram o método ICA para selecionar os componentes mais robustos. O estudo também encontrou diferenças nas medições entre EEG e fMRI e seus autores afirmam que isto pode estar relacionado a estas técnicas medirem preferencialmente diferentes tipos de atividade neuronal, pois EEG mediria mais fortemente as vias de condução rápida e fMRI os caminhos de condução lenta.

O trabalho de (NENTWICH *et al.*, 2020) (63 citações) comparou a conectividade funcional em EEG com a conectividade funcional observada por fMRI no banco de dados HBN. Para tanto coletaram EEG de 1330 e fMRI de 776 indivíduos com 10 diferentes fenótipos e variáveis demográficas (idade e gênero). Utilizou-se o método de (NOLTE, G., et al., 2004) para extrair a conectividade funcional dos sinais de EEG, e o método de (YEO *et al.*, 2011), tipicamente utilizado em estudos de fMRI, para formar *clusters* de EEG e de fMRI.

A Tabela 3 apresenta a tabela bibliométrica dos trabalhos que utilizaram fusão de EEG e de fMRI para estudar TDAH. Uma vez que os autores destes trabalhos não se utilizaram de classificadores em seus estudos, não se apresenta a coluna de *accuracies*.

Tabela 3 – Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com fMRI.

Referência (citações)	Tipo de dado	Técnica	Banco de dados	Contribuição
(KARCH <i>et al.</i> , 2010) 69 citações	EEG-fMRI em tarefa	Análise estatística	8 Controle 8 TDAH	Associação entre as dificuldades dos TDAH em realizar decisões e respostas reduzidas na região frontal
(LENARTOWICZ <i>et al.</i> , 2016) 48 citações	EEG-fMRI em tarefa	Análise estatística	15 Controle 15 TDAH	As ativações occipitais são reduzidas e a conectividade fronto-parieto-occipital é maior no sub-tipo TDAH-I

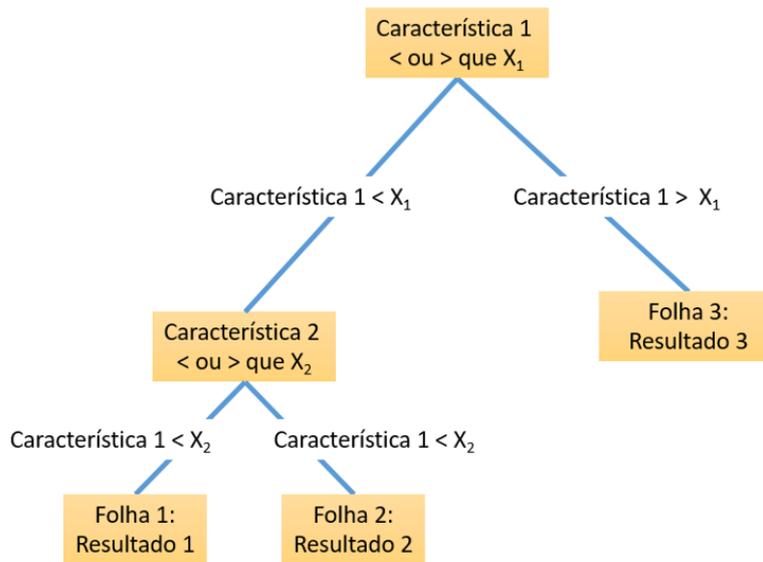
2.3 Fundamentação Teórica

2.3.1 Árvores de Decisão e *Random Forest* (RF)

As árvores de decisão são algoritmos supervisionados de tomada de decisão, nos quais executa-se uma sequência de divisões de dados baseadas em suas características. Os possíveis resultados finais são chamados de folhas da árvore. Já os caminhos são conhecidos como ramos. O desempenho das árvore de decisão depende de se saber quais divisões devem ser realizadas e qual profundidade a árvore deve ter. No algoritmo de árvores de classificação e regressão (do inglês - *classification and regression trees* - CART) para cada folha de uma árvore se designa uma pontuação. Para formar uma floresta, soma-se as pontuações das folhas correspondentes de diferentes árvores. A Figura 8 apresenta um exemplo de árvore de decisão.

A *Random Forest* (RF) é uma técnica proposta por (HO, 1995), na qual uma floresta de árvores não correlacionadas era criada projetando os dados de treinamento em subespaços aleatórios. Esta técnica foi aprimorada por (BREIMAN, 2001), o qual introduziu o CART para a criação de florestas de árvores não-correlacionadas, combinadas com *Bagging* e a otimização da aleatorização dos nós. O algoritmo de *Bagging*, também conhecido como agregação de *bootstraps*, seleciona aleatoriamente com reposição um grupo de amostras

Figura 5 – Exemplo de Árvore de Decisão.



Fonte: O Autor.

de treinamento para gerar árvores. O uso do *bootstrapping* reduz a variância do modelo sem elevar o *bias*, aumentando sua capacidade de generalização e reduzindo o *overfitting*. Os grupos de características usadas para a decisão são selecionados de forma aleatória com reposição e as árvores resultantes são avaliadas de acordo com os critérios do índice Gini ou da entropia. Ao fim, as previsões das RFs são obtidas a partir da média das previsões de árvores individuais (para árvores de regressão) ou do voto majoritário (para árvores de classificação).

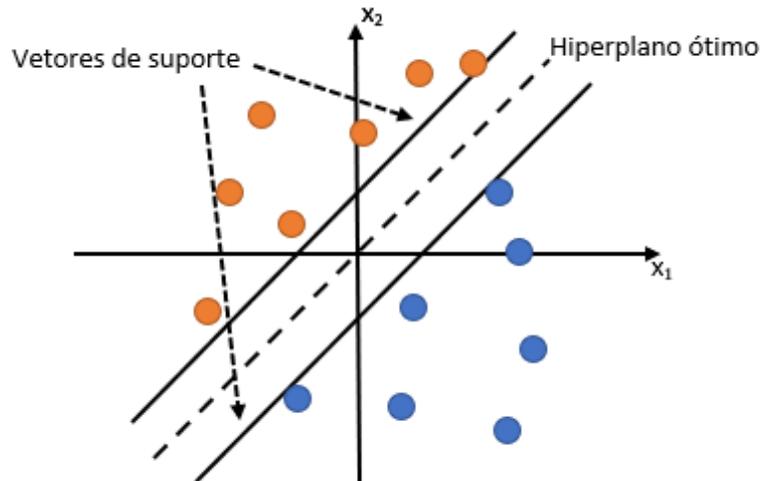
2.3.2 Support Vector Machine (SVM)

As máquinas de vetor de suporte (do inglês, *Support Vector Machine* - SVM) foram propostas por (BOSER *et al.*, 1992). As SVM constroem um hiperplano (ou um conjunto de hiperplanos) em um espaço dimensional para separar as classes. Este espaço dimensional é construído através das características apresentadas pelos elementos, as quais são fornecidas à entrada da SVM.

Os algoritmos de SVM buscam obter a maior distância possível entre os hiperplanos e os vizinhos mais próximos de cada uma das classe separadas por estes planos.

Nesta técnica, um algoritmo busca hiperplanos que possam separar corretamente duas ou mais classes de acordo com as características das classes avaliadas. Além de calcular um hiperplano ótimo, a SVM também infere vetores de suporte, que formam uma margem em torno do hiperplano de decisão. O algoritmo visa maximizar a distância entre os vetores de suporte e, portanto, aumentar a margem entre cada uma das classes. A Figura 9 apresenta um exemplo onde duas classes podem ser separadas em um plano (x_1, x_2) por um hiperplano ótimo linear acompanhado de seus vetores de suporte.

Figura 6 – Funcionamento de uma SVM.



Fonte: Adaptado de (HAYKIN, 2000).

A margem total entre duas classes pode ser expressa conforme a Equação 1.

$$Margem_{total} = \frac{2d}{\|w\|}, \quad (1)$$

onde d é a distância entre um vetor de suporte e o hiperplano de decisão (também chamado de ótimo); w é o vetor normalizado de pesos. Portanto, para maximizar a margem total entre os vetores de suporte, é necessário minimizar w ou maximizar a distância d . O vetor de saída y pode ser expresso por um somatório de funções não-lineares ($\varphi(x)$), de acordo com a Equação 2.

$$y = \sum_{n=1}^m w_n \varphi_n(x), \quad (2)$$

onde w_n é o peso correspondente ao valor n . Já, o hiperplano de separação que minimiza o erro de classificação de uma SVM pode ser encontrado através da minimização do vetor de pesos da Equação 3 (HAYKIN, 2000).

$$\varphi(w, \xi, \xi_i) = \frac{w'w}{2} + C \sum_{i=1}^N \xi_i, \quad (3)$$

onde C é uma constante positiva determinada experimentalmente e ξ são variáveis de folga. $'$ indica a transposição do vetor de pesos. As variáveis de folga são a distância até a qual é permitido que algumas amostras estejam do limite correto da margem. O hiperparâmetro C atua como na regularização do erro de classificação de uma SVM. Quanto menor o parâmetro C , maior é a regularização inversa promovida através deste parâmetro.

No entanto, muitos problemas de classificação não possuem classes linearmente separáveis. Nestas situações, pode-se empregar *kernels* não-lineares para construir hiperplanos adequados à classificação. Nesta Tese, experimentou-se o uso de *kernels* polinomiais,

sigmoide ou RBF, além dos lineares. O produto interno é a soma dos produtos dos elementos de dois vetores. Esta operação é realizada nos *kernels* para medir a similaridade entre pares de características. Esta operação é representada por $\langle \rangle$ nas equações.

O *kernel* linear usa apenas o produto interno entre os vetores (x e x_i) de características que representam elementos (participantes da base de dados) e é representado, respectivamente, pela Equação 4.

$$K(x, x_i) = \langle x, x_i \rangle, \quad (4)$$

O *kernel* sigmoide modifica o produto interno entre os vetores de acordo com a Equação 5,

$$K(x, x_i) = \tanh(\gamma_{SVM} \langle x, x_i \rangle + r), \quad (5)$$

onde r é um coeficiente que pode ser configurado como um hiperparâmetro e γ_{SVM} é o hiperparâmetro que controla a largura do *kernel*. Utilizou-se o valor padrão para o coeficiente r que é zero.

O *kernel* de função polinomial trabalha realizando alterações no produto interno entre os vetores de característica. Estas modificações permitem que os hiperplanos e vetores de suportes possam assumir formas mais complexas para separar amostras distribuídas em um espaço dimensional. O *kernel* de função polinomial pode ser expresso pela Equação 6, de acordo com a biblioteca *sklearn*,

$$K(x, x_i) = (\gamma_{SVM} \langle x, x_i \rangle + r)^{poli}, \quad (6)$$

onde *poli* é o grau do polinômio. Nos experimentos, utilizou-se polinômios de grau 4.

O *kernel* RBF é calculado na biblioteca *sklearn* através da Equação 7

$$K(x, x_i) = e^{-\gamma_{SVM} \|x - x_i\|^2}, \quad (7)$$

2.3.3 Rede Neural Convolutacional (CNN)

As CNNs (LECUN *et al.*, 1989) se popularizaram pelos bons resultados que obtiveram, especialmente, em tarefas de classificação e reconhecimento de imagens. O modelo de CNN consiste em uma série de camadas que realizam a operação de convolução intercaladas por camadas de *pooling* e seguidas de camadas de neurônios artificiais. As camadas convolucionais e de *pooling* extraem características espaciais das matrizes de conectividade. A seguir, são posicionadas as camadas formadas por neurônios artificiais. Estas camadas neuronais também são chamadas de densas ou de totalmente conectadas. Quanto mais camadas convolucionais e de *pooling*, menor será a complexidade computacional com a qual as camadas totalmente conectadas terão que lidar. No entanto, uma

redução demasiada pode sub-representar a complexidade das amostras. Embora as matrizes de correlação de fMRI não sejam imagens e não possuam contornos que possam ser destacados por camadas de convolução, os neurônios das camadas densas ainda se beneficiam da redução de complexidade.

Tanto as camadas de convolução, como as de *pooling* reduzem a dimensionalidade da matriz de entrada. No entanto, estas operações são diferentes. A operação de convolução consiste em multiplicar (ponto por ponto) uma área da matriz de entrada (esta área é chamada de janela) por um filtro (*kernel*), e, em seguida, somar todos os valores da matriz resultante. Esta operação pode ser descrita através da Equação 8.

$$H(x_{im}, y_{im}) = \sum_{i_a=0}^{fil_x-1} \left(\sum_{i_b=0}^{fil_y-1} (E(x_{im} - i_a, y_{im} - i_b)K(i_a, i_b)) \right), \quad (8)$$

onde E representa a entrada e K o filtro (*kernel*) de dimensão $fil_x \times fil_y$; fil_x e fil_y são as dimensões horizontal e vertical dos filtros, respectivamente; i_a e i_b são os parâmetros espaciais que são incrementados ao longo da convolução para percorrer-se toda a imagem. Então, aplica-se a função de ativação. Após, desliza-se a área da matriz da imagem um passo (geralmente um *pixel*) adiante e repete-se a operação, sucessivamente, até percorrer a matriz completa. Quando o passo utilizado entre uma área de convolução e a seguinte é maior que um *pixel*, chama-se o procedimento de convolução a passos largos (*stride convolution*), por isso o número de *pixels* pulados é, comumente, designado como parâmetro *stride*.

A operação de convolução resulta na redução das dimensões da matriz original. A matriz resultante é chamada de mapa de características. A redução na dimensionalidade da matriz (Dim) pode ser expressa segundo a Equação 9.

$$Dim = \left(\frac{dim_x - fil_x + 1}{s} \right) \left(\frac{dim_y - fil_y + 1}{s} \right), \quad (9)$$

onde dim_x e dim_y são as dimensões horizontal e vertical, respectivamente, da uma matriz original; o parâmetro s é o valor do passo *stride* utilizado.

Uma camada de convolução de uma CNN costuma possuir diversos filtros que são testados pela rede, ao longo da etapa de treinamento, com o objetivo de a rede aprender qual o detector de características que permite obter o melhor resultado. Desta forma, as camadas de convolução geram diferentes mapas de característica paralelos, um para cada filtro aplicado.

O procedimento de *pooling* agrupa os dados de uma região dos mapas e compacta estes dados segundo um critério. Nesta Tese, adotou-se o método de *max pooling* que seleciona o valor máximo de uma janela. Após as camadas de convolução e de *pooling*, a camada *flatten* transforma os mapas de características em um vetor eu é enviado para ser processado pelas camadas de neurônios.

Ao longo da operação de convolução, o filtro é deslizado sobre a matriz original. Contudo, os dados nos cantos das matrizes participam apenas uma vez das operações, enquanto os demais dados são utilizados múltiplas vezes. Além disto, cada convolução provoca a redução da imagem. E, uma sequência de convoluções pode provocar uma redução demasiada da imagem que implique na, indesejável, perda de informação. Isto pode ser evitado através do uso da técnica de *padding* ou *zero-padding* (WANG *et al.*, 2017). Nesta técnica zeros são adicionados nas bordas das matrizes. Isto eleva a utilização dos dados próximos às bordas da matriz da imagem original, tornando-os tão relevantes quanto as células centrais. Este procedimento possibilita o aumento da quantidade de camadas de convolução, evitando a redução da dimensionalidade causada pela convolução. Utilizando a técnica de *padding*, a Equação 9 é reformulada conforme a Equação 10.

$$Dim = \left(\frac{dim_x + 2p_{pad} - fil_x + 1}{s} \right) \left(\frac{dim_y + 2p_{pad} - fil_y + 1}{s} \right), \quad (10)$$

onde p_{pad} corresponde ao número da espessura da borda de zeros (medida em pixels) da técnica de *padding*.

2.3.4 Classificador Linear *Ridge*

O classificador *Ridge* escala os valores das características de entrada (as classificações dos modelos primários) para valores entre -1 e 1 e então trata a classificação como uma tarefa de regressão. O classificador *Ridge* é um modelo linear, na forma da Equação 11.

$$y_i = b + \omega_1 x_{i1} + \dots + \omega_n x_{in}, \quad (11)$$

onde ω_1 e ω_n representam os pesos, x_{i1} , x_{in} são as características utilizadas como entradas do estágio de fusão (as predições dos modelos RF) e b é uma constante aditiva chamada de *bias*. O modelo linear busca otimizar seus pesos de forma a minimizar a função custo da soma residual dos quadrados (do inglês, *residual sum of squares* - RSS), conforme a Equação 12.

$$RSS = \sum_{i=1}^n (y_i - (\omega x_i + b))^2 \quad (12)$$

Contudo, a minimização desta função custo pode elevar a variância do modelo e, conseqüentemente, aumentar o *overfitting*. Para evitar esta condição, pode-se utilizar as regularizações *Lasso* (L1), *Ridge* (L2) ou *ElasticNet* (L1L2). A regularização L1 tende a selecionar uma das características correlacionadas e reduzir a zero as demais para minimizar a penalização L1. Ambas as regularizações adicionam novos termos à equação da função custo (Equação 12). Na regularização L1, este termo adicionado à função custo corresponde ao somatório do valor absoluto dos pesos multiplicados por uma constante

de regularização (α_{Reg}). Desta forma, a equação da função custo pode ser reformulada conforme a Equação 13.

$$RSS_{Lasso} = \sum_{i=1}^n (y_i - (\omega x_i + b))^2 + \alpha_{Reg} \sum_{j=1}^m |\omega_j| \quad (13)$$

Na regularização linear L2, o termo de penalização da função custo é o somatório do quadrado dos pesos multiplicados por uma constante α_{Reg} , de acordo com a Equação 14.

$$RSS_{Ridge} = \sum_{i=1}^n (y_i - (\omega x_i + b))^2 + \alpha_{Reg} \sum_{j=1}^m \omega_j^2 \quad (14)$$

O classificador *Ridge* possui, como hiperparâmetro, o alfa (α_{Reg}) que representa a força da regularização imposta ao modelo e é equivalente à $1/2C$, onde C é o parâmetro usado na SVM linear. Quanto maior o valor de alfa, maior vai ser o efeito da regularização. Este hiperparâmetro foi encontrado através de uma otimização manual extensiva, resultando igual a 0,0001.

2.3.5 Adaptive Boosting (AdaBoost)

Os métodos de alavancagem (*boosting*) de resultados, como o *XGBoost* e o *AdaBoost*, treinam modelos sequencialmente. Nestes métodos, o erro encontrado por cada modelo na classificação de cada amostra é registrado. Com base nestes erros, calculam-se novos pesos para cada amostra. Desta forma, as amostras correspondentes aos maiores erros do classificador anterior corresponderão as amostras que receberão maior peso no novo modelo. Os métodos de alavancagem (*boosting*) se baseiam no princípio de que: Um modelo capaz de classificar adequadamente amostras mais difíceis também poderá classificar corretamente as demais amostras, mas o oposto não é verdadeiro. Por isso, as técnica de (*boosting*) priorizam que o modelo aprenda a classificar amostras mais difíceis.

No *AdaBoost*, todas as amostras recebem inicialmente o mesmo peso, tipicamente $1/n$, onde n é o número de amostras. Então, geram-se árvores de decisão para as características de entrada. A árvore com menor índice Gini é selecionada. Então, calcula-se a importância (α_t) deste classificador, de acordo com a Equação 15 (SCHAPIRE, 2013).

$$\alpha_t = \frac{1}{2} \log\left(\frac{1 - \epsilon}{\epsilon}\right), \quad (15)$$

onde ϵ é o erro da árvore de decisão selecionada.

Os novos pesos das amostras são determinados de acordo com a Equação 16 (SCHAPIRE, 2013).

$$w_t = w_{t-1} \frac{e^{-\alpha_t y_i h_{pred}(x_i)}}{Z_t}, \quad (16)$$

onde y_i é a saída esperada para determinada amostra (i), $h_t(x_i)$ é a hipótese predita pelo modelo para a classificação da amostra com base nas características x_i e Z_t é um fator

de normalização. Quando a predição é incorreta, a exponencial da Equação 16 resulta em um número maior e, conseqüentemente, o novo peso será maior para as amostras classificadas incorretamente. Então, o algoritmo usa os novos pesos para determinar a probabilidade do algoritmo sortear as amostras: Quanto maior o peso de uma amostra, maior a chance dessa ser selecionada (podendo ser escolhida a mesma amostra mais de uma vez). Assim, o *AdaBoost* gera um novo conjunto de dados (*bucket*), para o qual novas árvores de decisão são geradas. O processo descrito é repetido minimizando o erro progressivamente.

2.3.6 *Extreme Gradient Boosting (XGBoost)*

O método *XGboost* (?) agrega algumas ferramentas para otimizar a técnica de *boosting*: a regularização do objetivo de aprendizado (*regularizing learning objective*), *gradient tree boosting* e a busca por melhores divisões de galho (*split finding algorithms*). O *gradient tree boosting* é usado para calcular os novos pesos de cada amostra para a formação de um novo conjunto de dados. O *XGBoost* incorpora neste cálculo a regularização do objetivo de aprendizado (*regularizing learning objective*). Já as técnicas de otimização da divisão de galhos são utilizadas na formação de árvores de regressão que funcionam como estimadores base do *ensemble*.

No presente trabalho, cada uma destas árvores de regressão utiliza como características, em suas entradas, as predições obtidas através dos classificadores RF de cada banda de EEG. O *XGboost* substitui as árvores de decisão (utilizadas no *AdaBoost*) por árvores de regressão, onde designa-se um peso ω para cada folha (i). Para cada característica gera-se uma pequena árvore de regressão e, então, somam-se os pesos de cada folha correspondentes a uma determinada predição.

O objetivo de aprendizado consiste em calcular o erro dado pela diferença entre a predição (y_{pred}) e o objetivo (y_{obj}). O erro é medido com uma função convexa, ou seja, que possua a forma da Equação 17.

$$f(tx + (1 - t)y) \geq t(fx) + (1 - t)f(y). \quad (17)$$

Para evitar o *overfitting*, o *XGboost* suaviza o aprendizado dos pesos, adicionando um termo de regularização ao objetivo de aprendizado. Por isso, o *XGboost* utiliza a Equação 18.

$$L(t) = \sum_{i=1}^n l(y_{pred}, y_{obj}) + \sum_k \Omega_X(f_k), \quad (18)$$

onde t são os passos, f_k representa uma estrutura independente de árvore e Ω_X é o elemento de regularização que penaliza a complexidade do modelo. Este termo é calculado com a Equação 19.

$$\Omega_X(f) = \gamma T + 0.5\lambda w^2, \quad (19)$$

onde T é o número de folhas da árvore, λ é o hiperparâmetro da regularização e γ é o hiperparâmetro que representa a mínima redução na função custo para uma divisão de galhos ser aprovada.

O algoritmo *Gradient Tree Boosting* (abreviado internamente como *gbtree*) utiliza uma abordagem *greedy* para otimizar a equação do *ensemble* (18) através de um treinamento. Isto é realizado adicionando-se uma estrutura independente de árvore (f_t) que minimiza o objetivo de aprendizado regularizado, agora representado pela Equação 20.

$$L(t) = \sum_{i=1}^n l(y_{obj}, y_{pred}^{(t-1)} + f_t(x_i) + \Omega_X(f_t)). \quad (20)$$

A Equação 20 pode ser representada por uma aproximação de segunda ordem, apresentada na Equação 21.

$$L(t) \approx \sum_{i=1}^n (l(y_{obj}, y_{pred}^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2 + \Omega_X(f_t)). \quad (21)$$

onde $g_i = \delta_{y_{pred}^{(t-1)}} l(y_{pred}, y_{obj})$ e $h_i = \delta_{y_{pred}^{(t-1)}}^2 l(y_{pred}, y_{obj})$ são os gradientes de primeira e segunda ordem da função custo. Removendo-se os elementos constantes, é possível representar o objetivo de aprendizado regularizado através da Equação 22.

$$L(t) \approx \sum_{i=1}^n (g_i f_t(x_i) + \frac{1}{2} h_i f_t^2 + \Omega_X(f_t)). \quad (22)$$

O peso ótimo de cada folha é calculado de acordo com a Equação 23.

$$w_j^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \quad (23)$$

O método *XGBoost* utiliza o objetivo de aprendizagem ótimo para medir a qualidade da estrutura de cada árvore. Quanto menor o valor ser calculado com a Equação 24, melhor será a árvore.

$$L(t) = -0.5 \sum_{j=1}^n \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \quad (24)$$

Para formar as estruturas das árvores, o *XGBoost* adiciona novos galhos iterativamente e avalia a redução na função custo após a divisão do galho, através da Equação 25.

$$L_{split} = 0.5 \left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma, \quad (25)$$

onde I é a união das instâncias esquerda (I_L) e direita (I_R) de um nó da árvore depois da referida divisão dos galhos.

O *XGBoost* usa ainda as técnicas de redução de escala dos pesos (*shrinkage*) e a subamostragem das características para reduzir o *overfitting*. Contudo, a subamostragem das características não foi utilizada na presente pesquisa porque deseja-se saber o peso que o algoritmo otimizado designará para cada característica, pois esta informação é relevante para a investigação dos biomarcadores.

A redução de escala dos pesos é realizada normalizando cada peso do algoritmo *gradient boosting* através da sua divisão por um fator de *shrinkage* (η , lê-se eta).

3 MATERIAIS E MÉTODOS

3.1 Base de Dados: *Healthy Brain Network* (HBN)

Na presente pesquisa, usou-se a base de dados HBN apresentada em (ALEXANDER *et al.*, 2017). Esta é uma base de dados multimodal de larga escala lançada pelo *Child Mind Institute*, localizado em Nova York nos Estados Unidos com o objetivo de obter amostras de múltiplos diagnósticos psiquiátricos para possibilitar investigações que vissem descobrir os biomarcadores cerebrais correspondentes aos diagnósticos de transtornos psiquiátricos. Para tanto, o *Child Mind Institute* coletou dados de pessoas do sexo masculino ou feminino com idades entre 5 e 21 anos em comunidades diversas através de Nova York. A HBN conta com dados psiquiátricos, comportamentais, cognitivos, fenótipos de estilo de vida; fMRI, sMRI, EEG, rastreamento ocular, gravação de voz e de vídeo, genética e actigrafia.

Contudo, o escopo do presente trabalho se limita a realizar investigações baseadas em fMRI e EEG. Os ensaios de EEG e de fMRI foram coletados pela base de dados de forma individual não-simultânea, ou seja, as amostras de EEG e de fMRI foram coletadas em momentos diferentes. Os dados da base HBN podem ser acessados através da internet no endereço eletrônico: http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network_ com autorização dos funcionários da base de dados.

No entanto, a equipe de gerenciamento da HBN não disponibiliza os dados identificadores pessoais dos participantes com o objetivo de proteger sua privacidade, com exceção de informações de sexo e de idade. Assim, na base de dados e no estudo, os participantes são identificados somente por um código (ALEXANDER *et al.*, 2017). A lista com os indivíduos identificados por estes códigos, cujos dados foram utilizados na presente pesquisa, está no ANEXO A da presente Tese.

A equipe de coleta da base HBN exclui pacientes com quociente intelectual menor que 66; com doenças neuro-degenerativas; com HIV; histórico de dependência de substâncias (exceto nicotina e maconha); diagnóstico recente (menos de seis meses) de transtornos de bipolaridade, de esquizofrenia, esquizoafetivo ou autismo de baixo funcionamento. A

base também removeu pacientes não-tratados que cometeram tentativas de suicídio ou de homicídio até três meses antes da coleta, ou que tiveram episódios psicóticos ou maníacos até seis meses antes da coleta. A base excluiu pessoas com problemas linfáticos, com encefalite, com problemas metabólicos juvenis ou incapazes de preencher formulários ou assinar a documentação.

A base requer que os participantes sejam fluentes em inglês e os pais das crianças podem ser fluentes em espanhol ou em inglês. Pessoas intoxicadas no momento de alguma das coletas ou com histórico de dependência química também foram retiradas do experimento. Não foram realizadas coletas de sujeitos com dificuldades visuais ou acústicas que impedissem sua participação em experimentos baseados em tarefas (ALEXANDER *et al.*, 2017).

O *Child Mind Institute* coletou, para a HBN, dados de EEG em três experimentos ativos (execução de tarefa) e três passivos (não dependentes de tarefa). Os três experimentos ativos são baseados nos paradigmas de sequência de aprendizagem (5 minutos), percepção visual/tomada de decisão (9 minutos) e procura de símbolos da Escala Wechsler de Inteligência para Crianças (WISC-IV) (2 minutos). Os três experimentos passivos são estado de repouso (5 minutos), excitação/inibição (3,5 minutos) e estímulo naturalístico (11 minutos).

Cada um destes seis experimentos possui um paradigma de coleta diferente. Primeiro, são aplicados os três experimentos ativos e a seguir os passivos, de acordo com a ordem apresentada. Limitou-se o escopo deste trabalho à análise dos dados de EEG medidos em estado de repouso. Ao todo, os seis experimentos têm 35,5 minutos de duração.

Os sinais de EEG de estado de repouso foram coletados ao longo de aproximadamente 5 minutos. Nestes 5 minutos, o participante do experimento é instruído a olhar para uma cruz desenhada no centro de uma tela ao longo do experimento. Esta cruz permanece imóvel durante o experimento. Uma voz gravada e eletronicamente reproduzida em sincronia com a coleta de dados de EEG instrui os participantes a fechar e abrir os olhos alternadamente em intervalos de 20 a 40 segundos ao longo do experimento.

A Tabela 4 mostra os períodos em que se espera que os participantes estejam de olhos abertos ou fechados. Este paradigma de coleta de dados de EEG segue o padrão proposto por (FOX & GREICIUS, 2010).

Tabela 4 – Tempos das instruções do paradigma de coleta de dados de estado de repouso de EEG.

Número do evento	Instrução	Tempo (s) após comando anterior	Instante do Evento(s)
1	Início do experimento	0	0
2	Abrir os olhos	14	14
3	Fechar os olhos	20	34
4	Abrir os olhos	40	74
5	Fechar os olhos	20	94
6	Abrir os olhos	40	134
7	Fechar os olhos	20	154
8	Abrir os olhos	40	194
9	Fechar os olhos	20	214
10	Abrir os olhos	40	254
11	Fechar os olhos	20	274
12	Abrir os olhos	40	314
13	Fim do experimento	4	318

Uma vez que o tempo total dos 6 experimentos de coleta de EEG é de 35,5 minutos e o tempo para preparar a rede de EEG é superior a 30 minutos, a impedância dos eletrodos é verificada no início da gravação e novamente testada a cada 30 minutos. Caso necessário, o supervisor dos experimentos adiciona a substância salina aos eletrodos para manter as impedâncias abaixo de 40 k Ω .

As gravações de EEG são realizadas em uma sala blindada contra som. O equipamento de medida de EEG é um *128-channel EEG geodesic hydrocel system* da *Electrical Geodesics Inc.* Os dados da base foram coletados com uma frequência de amostragem de 500 Hz e uma banda de passagem de 0,1 a 100 Hz. Este sistema utiliza 128 canais com eletrodos bivariados e seu canal de referência fixado no centro da cabeça (Cz no sistema 10-20). Para cada participante, é selecionada uma rede de eletrodos de acordo com tamanho compatível com sua cabeça.

A HBN disponibilizou imagens de fMRI de estado de repouso coletadas ao longo de dois períodos curtos de 5 minutos, os quais foram separados para evitar que a perda de atenção e necessidade de movimento dos pacientes, especialmente os mais jovens, produzisse movimentos que prejudicassem a medição. Ambas as gravações são realizadas em estado de repouso com o paciente fixando o olhar em uma cruz, como no exame de EEG. A HBN também realizou dois experimentos de coleta de fMRI baseadas em um

paradigma naturalístico de visualização.

No primeiro experimento, cada participante é instruído a assistir um vídeo com um fragmento de 10 minutos do filme “Meu Malvado Favorito” (*Despicable Me*, no original). Após 1 minuto de descanso, um filme curta-metragem de 3,47 minutos chamado “O Presente” (*The Present*, no original) é exposto ao participante. Ambos os filmes são animações, sendo que “O Presente” busca evocar emoções e empatia nos participantes. No entanto, o escopo do presente trabalho, se limita a análise das imagens obtidas em estado de repouso. Uma vez que as duas coletas de fMRI de estado de repouso são exames muito similares, na presente pesquisa, considerou-se que cada uma constitui uma amostra diferente e que pode ser utilizada no mesmo experimento de classificação. Deste modo, cada participante é associado a duas amostras de fMRI de estado de repouso.

As imagens de fMRI foram coletadas para a HBN utilizando um *scanner* Siemens Avanto 3 T Tim Trio MRI localizado na Universidade *Rutgers University Brain Imaging Center* em Nova York. Este *scanner* é equipado com gradientes eletromagnéticos de 45 mT/m instalado em um veículo *motor-home* e aperfeiçoado com 32 canais receptores de rádio frequência (Siemens *32-channel*).

Para maximizar a estabilidade do sistema a longo prazo, este veículo *motor-home* foi estacionado sobre apoios de concreto com 10 polegadas de espessura. Ao longo dos ensaios de fMRI, são também coletadas imagens de MRI estrutural. Na etapa de calibração, a equipe da HBN orienta os participantes a fixar o olhar em 27 pontos diferentes para estimar para onde este estará olhando durante os ensaios.

Na Tabela 5, apresentam-se as quantidades iniciais de indivíduos de todas as classes usadas tanto para EEG como para fMRI antes das exclusões devido a controles de qualidade (As quantidades posteriores serão apresentadas nas tabelas 7 e 18 nas páginas 71 e 91, respectivamente). Nesta Tabela, observa-se que há uma maior proporção de indivíduos do grupo TDAH do que do grupo controle, de modo que ambos os subtipos mais populosos de TDAH possuem uma maior quantidade de sujeitos na base HBN do que o grupo controle.

A quantidade de indivíduos com TDAH na base de dados é 2,7 e 4,7 vezes maior que a quantidade de sujeitos da classe controle em EEG e fMRI, respectivamente. Para EEG, a classe TDAH-C é majoritária com 37% dos indivíduos, contra 33% de TDAH-I e 26% da classe controle. Já em fMRI, há um equilíbrio entre a quantidade de sujeitos com TDAH-C (333) e com TDAH-I (334).

Porém, o desbalanço entre os grupos TDAH e controle é mais acentuado em fMRI que representam, respectivamente, 78% e 17% dos sujeitos amostrados. Já, a classe excluída, TDAH-H, apresenta somente 5,0% e 3,6% das amostras de fMRI e EEG, respectivamente. Mesmo antes das exclusões do controle de qualidade, a quantidade de sujeitos com TDAH-H é muito menor se comparada às demais classes com apenas 33 e 43 indivíduos fornecendo dados de EEG e fMRI, respectivamente, à base de dados. Em EEG

e em fMRI, respectivamente, o grupo TDAH-H tem 10 e 7,7 vezes menos sujeitos que a categoria TDAH-C. Tanto para EEG como para fMRI, considerou-se a classe TDAH como a união das classes TDAH-C e TDAH-I.

Tabela 5 – Quantidades de indivíduos por classe antes das exclusões de amostras devido ao controle de qualidade.

Classe	Quantidade de indivíduos EEG (%)	Quantidade de amostras fMRI (%)
Controle	193 (24%)	142 (17%)
TDAH (TDAH-C+TDAH-I)	617 (76%)	667 (78%)
TDAH-C	329 (40%)	333 (39%)
TDAH-I	288 (36%)	334 (39%)
TDAH-H	33 (3,6%)	43 (5,0%)

Como mostrado na Tabela 5, a base de dados dispõe de um grande número de amostras de EEG e de fMRI. Porém, a quantidade de participantes que possuem dados coletados por ambos os exames com qualidade aceitável é muito menor, como mostra a Tabela 6. A fusão dos classificadores utilizando dados de ambos os exames não foi realizada neste trabalho, pois, esta quantidade diminuta de dados que poderia impactar os classificadores. Os referidos controles de qualidade de EEG e de fMRI serão apresentados nas seções 3.3 e 3.4, respectivamente.

Tabela 6 – Quantidade de amostras com EEG e fMRI aprovados em ambos os controles de qualidade.

Classe	Quantidade de indivíduos com exames de EEG e de fMRI
Controle	26
TDAH (TDAH-C+TDAH-I)	94
TDAH-C	31
TDAH-I	63

A menor quantidade de pessoas com TDAH-H na base de dados ocorre porque o grupo TDAH-H tem presença menor na população: apenas 2 a 14% dentre as crianças diagnosticadas com TDAH possuem o transtorno do subtipo TDAH-H (GIBBINS *et al.*, 2010)-(SMALLEY *et al.*, 2007). Em adultos, o TDAH-H é ainda mais raro, uma vez que,

os sintomas de hiperatividade tendem a desaparecer à medida que a pessoa amadurece (GIBBINS *et al.*, 2010). Frequentemente, os pacientes com TDAH-H mudam de subtipo para o TDAH-C ao envelhecerem (LAHEY *et al.*, 2005). Isto ocorre porque as pessoas que sofrem de hiperatividade costumam ter também (ou desenvolver futuramente) os sintomas de desatenção e, por isso, passam a ser classificados como TDAH-C.

A movimentação excessiva produz artefatos tanto nos sinais de EEG como em fMRI. Estes artefatos de movimento, além de produzir maior ruído, o qual dificulta a leitura dos sinais coletados, podem gerar uma situação onde os classificadores decidem a qual classe pertence cada indivíduo de acordo com a quantidade de movimento, o que é indesejável no presente estudo que pretende relacionar os sintomas de cada sujeito com o comportamento cerebral. Por isso, há a necessidade de excluir do estudo sujeitos que apresentaram demasiada movimentação ao longo da coleta.

Alguns participantes não conseguiram completar com sucesso todas as etapas do protocolo de coleta da HBN por motivos diversos. Por exemplo, alguns participantes podem ter abandonado o exame de fMRI devido a sofrerem sintomas de claustrofobia. Participantes com dificuldades sensoriais podem ter tido dificuldades de acompanhar o protocolo de coleta de EEG. Por isso, nem todos os participantes dos experimentos de EEG estão presentes nos exames de fMRI e vice-versa. Nos exames de fMRI, foram excluídos os participantes com 5 anos, por serem a parcela etária com mais movimento de cabeça (ALEXANDER *et al.*, 2017).

Os pacientes com TDAH-H possuem uma maior dificuldade de permanecer parados durante os exames de EEG e de fMRI em razão dos sintomas do próprio transtorno. Por isso, as coletas de dados fisiológicos de pacientes com TDAH-H têm uma maior suscetibilidade a apresentarem sinais com menor qualidade. Isto se apresenta como um problema, uma vez que este subtipo de TDAH é muito mais raro que os demais. Desta forma, na base HBN há sinais de EEG de uma pequena quantidade de indivíduos com TDAH-H ($n = 33$), dos quais apenas 18 foram aprovados no controle de qualidade (Tabela 7, página 71). Já em fMRI, a base HBN tem 43 participantes com TDAH-H e apenas 25 destes possuíam qualidade aceitável segundo os critérios adotados no presente estudo (Tabela 18, página 91). Como a quantidade de amostras de TDAH-H são pouco representativas comparadas às das demais classes, a classe TDAH-H foi excluída dos experimentos desta Tese. Os representantes desta classe também não foram retirados do conjunto de dados que compõem a classe TDAH.

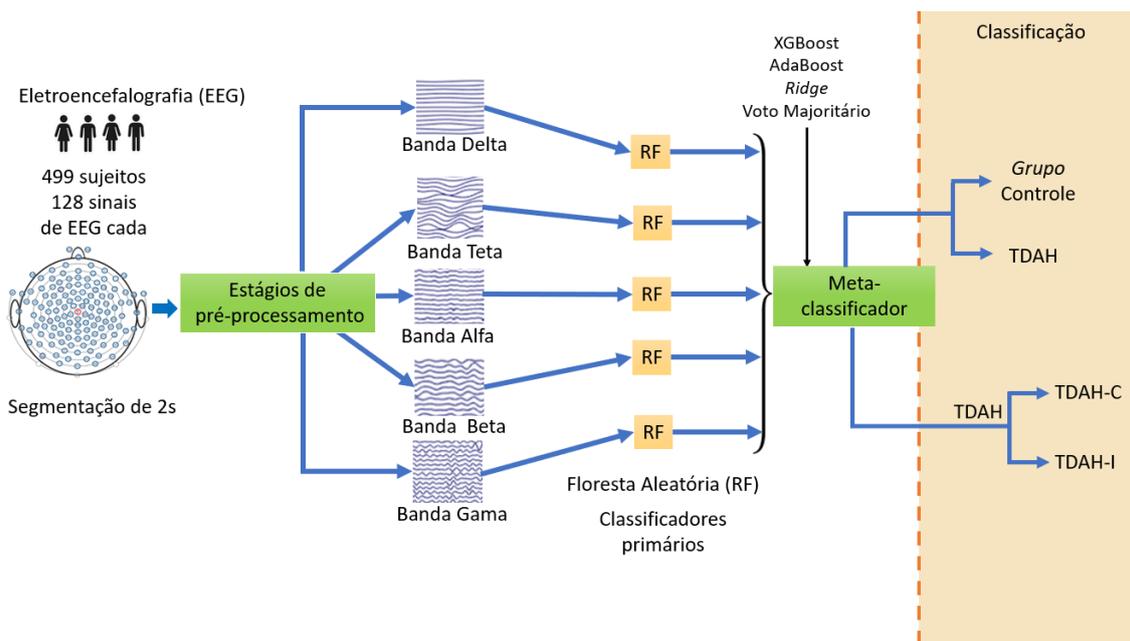
3.2 Resumo da Metodologia

Para facilitar a compreensão da metodologia proposta nessa Tese, consultar as figuras 7 e 8. Estas figuras resumem as metodologias empregadas nesta Tese para o processamento dos sinais de EEG e de fMRI, respectivamente.

Os sinais de EEG coletados de 499 sujeitos são submetidos a uma série de estágios de pré-processamento para garantir a qualidade dos sinais. Utilizam-se filtros para separar nestes sinais as bandas de frequência clássicas de EEG. As características mensuradas dos sinais em cada faixa de frequência são enviadas a uma RF correspondente. Experimentou-se, também, o uso de SVM como classificadores primários, embora estes tenham obtido resultados menos satisfatórios quando utilizados, independentemente, do que aqueles obtidos pelas RF. Por isto, optou-se pelo uso das RF como classificadores primários nos meta-classificadores.

As classificações produzidas pelas RFs servem de características para o meta-classificador que realiza a classificação final, separando os pacientes com TDAH dos voluntários da classe controle. O mesmo método também é usado, em paralelo, para classificar os pacientes de TDAH entre os subtipos: TDAH-C e TDAH-I. Tanto para EEG como para fMRI, considerou-se quatro possibilidades de meta-classificadores: Voto Majoritário, classificador *Ridge*, *AdaBoost* e *XGBoost*; cujo desempenho foi comparado.

Figura 7 – Resumo gráfico da metodologia da Tese na parte de EEG.

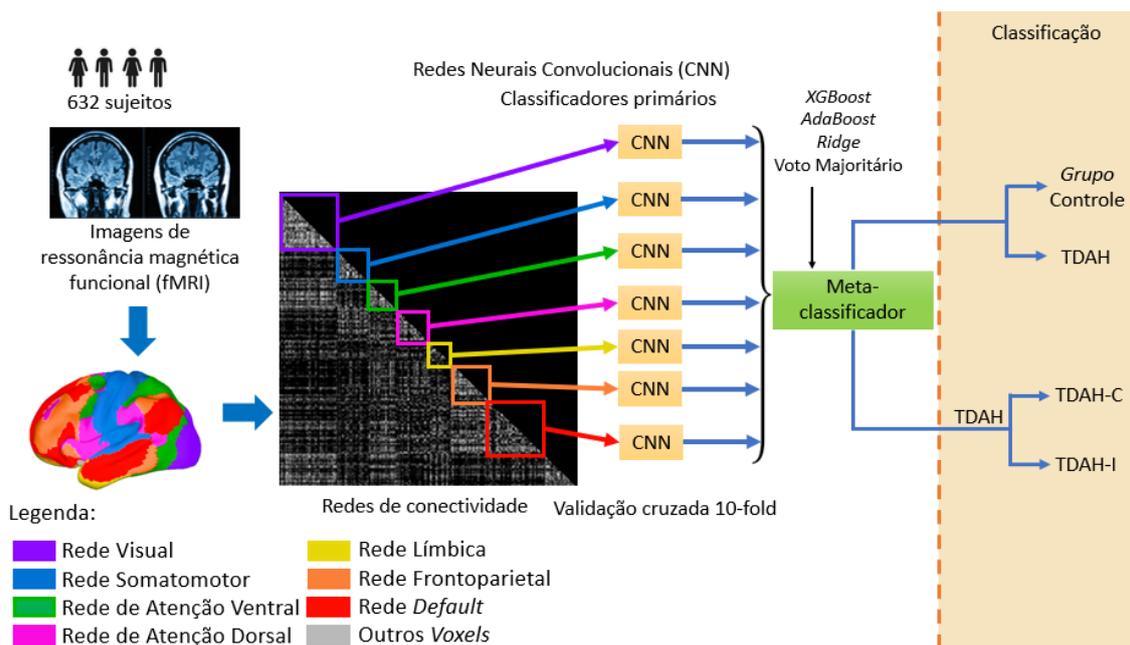


Fonte: O Autor.

Os sinais de fMRI coletados de 632 sujeitos são submetidos a uma sequência de etapas de pré-processamento, de forma que as médias aferidas são organizadas em um conjunto de 190 regiões de interesse (ROI). Realizou-se a autocorrelação do vetor com as médias das medições em cada uma destas ROI, gerando-se uma matriz de correlações (esta será explicada em detalhes na seção 3.4.3, que inicia na página 92). Esta matriz é organizada de forma que nesta possam ser localizadas as regiões cerebrais cujos neurônios costumam trabalhar em conjunto, conhecidas como redes de conectividade. Cada uma

destas redes de conectividade é enviada a uma CNN correspondente. Cada CNN prediz a classe de cada indivíduo, e o conjunto das predições de todas as CNNs forma o vetor de características apresentado à entrada do meta-classificador. Assim como em EEG, o meta-classificador aponta a classe dos sujeitos, separando os pacientes de TDAH do grupo controle. O mesmo método é aplicado para identificar os subtipos (TDAH-C e TDAH-I) entre os indivíduos já diagnosticados com TDAH.

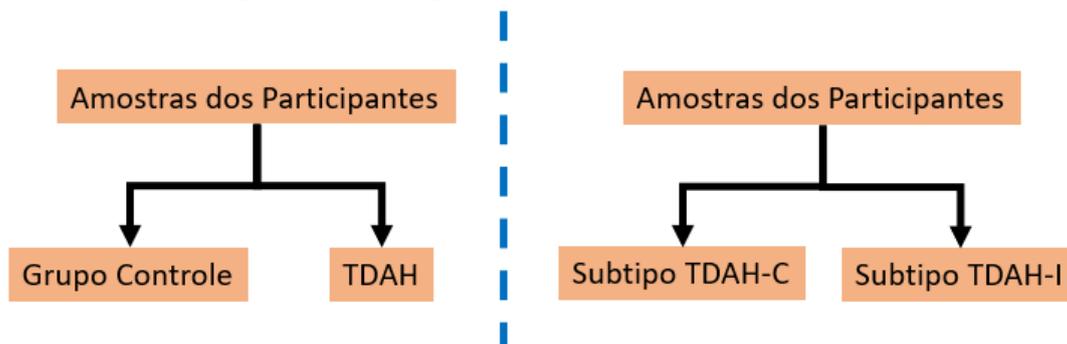
Figura 8 – Resumo gráfico da metodologia da Tese na parte de fMRI.



Fonte: O Autor.

Considerou-se, tanto na parte de EEG como na parte de fMRI, o paradigma de classificação apresentado na Figura 9. Neste, as amostras dos participantes são identificadas como pertencentes a sujeitos com TDAH ou ao grupo controle. Em paralelo, as amostras de TDAH são classificadas como pertencentes aos grupos TDAH-C ou TDAH-I.

Figura 9 – Paradigma de classificação das amostras.



Fonte: O Autor

A arquitetura do computador no qual os experimentos computacionais foram executados é sempre a mesma: Um *Desktop* com Unidade de Processamento Central (do inglês, *Central Processing Unit* - CPU) Intel (R) Core (TM) i9-9900KF) @3.60GHz com 32 GB de Memória de Acesso Aleatório (do inglês, *Random Access Memory* - RAM) instalada. Este computador tem uma Unidade de Processamento Gráfico (do inglês, *Graphics Processing Unit* - GPU) NVIDIA GeForce RTX 3070 Ti e possui um sistema operacional de 64 bits com um processador baseado em x64.

3.3 Eletroencefalografia (EEG)

3.3.1 Pré-processamento de EEG

O pré-processamento tem como objetivo reduzir os efeitos da contaminação dos sinais por ruído e por artefatos fisiológicos, além de preparar as características dos segmentos dos sinais para que sejam, adequadamente, processadas pelos classificadores.

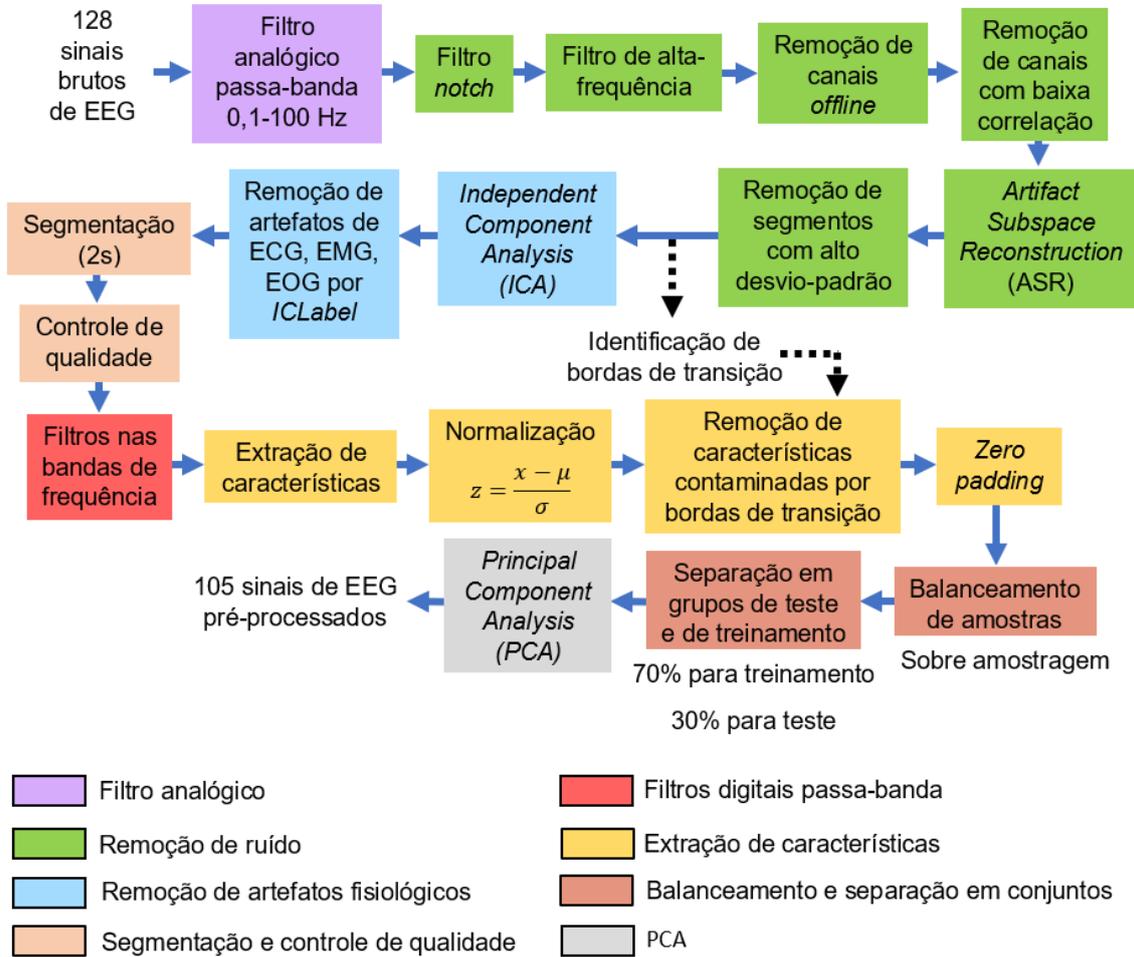
O funcionamento de cada etapa representada do pré-processamento será detalhado nas subseções apresentadas a seguir. Estas etapas foram separadas em subseções de acordo com suas funções: remoção de ruído, exclusão de artefatos fisiológicos, segmentação e separação de grupos de treinamento e de teste, controle de qualidade, filtros passa-faixas e extração de características.

A subseção de remoção de ruído abriga os blocos do filtro *notch*, filtro de alta-frequência, remoção de canais *offline*, remoção de segmentos e *Artifact Subspace Reconstruction* (ASR) e remoção de canais com baixa correlação com os demais.

A subseção de exclusão de artefatos é composta pela *Independent Component Analysis* (ICA) e pela remoção dos artefatos. As subseções de segmentação e de controle de qualidade; e filtros passa-faixas contém apenas as etapas referidas nos títulos destas subseções.

A subseção de extração de características é, também, composta pelas etapas de normalização, de remoção das características afetadas por bordas de transição, de balanceamento de amostras, separação de grupos de treinamento e de teste e de *Principal Component Analysis* (PCA). A Figura 10 apresenta o diagrama de blocos completo de pré-processamento de EEG neste trabalho.

Figura 10 – Diagrama de blocos dos estágios de pré-processamento. Com exceção do primeiro bloco, todas as etapas são digitais.



Fonte: O Autor

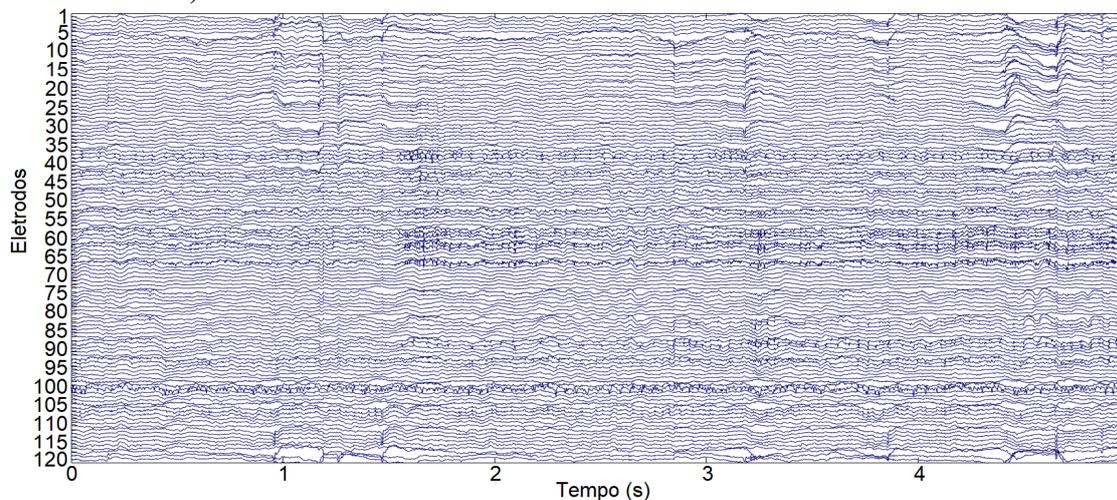
3.3.1.1 Remoção de ruído

Os 128 sinais de EEG de cada indivíduo foram pré-processados utilizando a *toolbox* EEGLAB através da plataforma computacional MATLAB r2016a. Esta subseção apresenta uma descrição dos passos de pré-processamento, os quais foram aplicados na sequencia apresentada, separadamente, para cada participante.

Seguiu-se o protocolo de pré-processamento de dados recomendado pela equipe da base de dados e pelo desenvolvedor da *toolbox* EEGLab (DELORME & MAKEIG, 2004), Arnaud Delorme em (<https://eeglab.org/>).

A Figura 11 apresenta, como exemplo, os sinais brutos (*raw*) de EEG de um dos participantes da subclasse TDAH hiperativo (NDARAM873GAC) antes do pré-processamento.

Figura 11 – Sinais brutos (*raw*) de EEG de um participante com TDAH hiperativo (NDA-RAM873GAC).



Fonte: O Autor

Filtrou-se cada um dos 128 sinais de EEG com um filtro *notch* ajustado na frequência da rede do local de coleta (60 Hz). Este filtro é necessário porque o cabo que transmite os sinais de EEG da touca de eletrodos para o computador foi posicionado muito próximo a uma tomada da rede elétrica e, por isso, está sujeito à interferência eletromagnética.

A seguir, aplica-se um filtro *butterworth* do tipo passa-altas com frequência de corte em 0,5 Hz para eliminar o comportamento não estacionário dos sinais. Então, excluem-se os sinais com mais de 5 segundos contínuos sem variações *offline*, pois este é um indicio de que estes canais estão com problemas de transmissão de sinais. Este comportamento costuma estar relacionado à falta de reposição do gel em torno dos eletrodos que eleva à impedância dos canais.

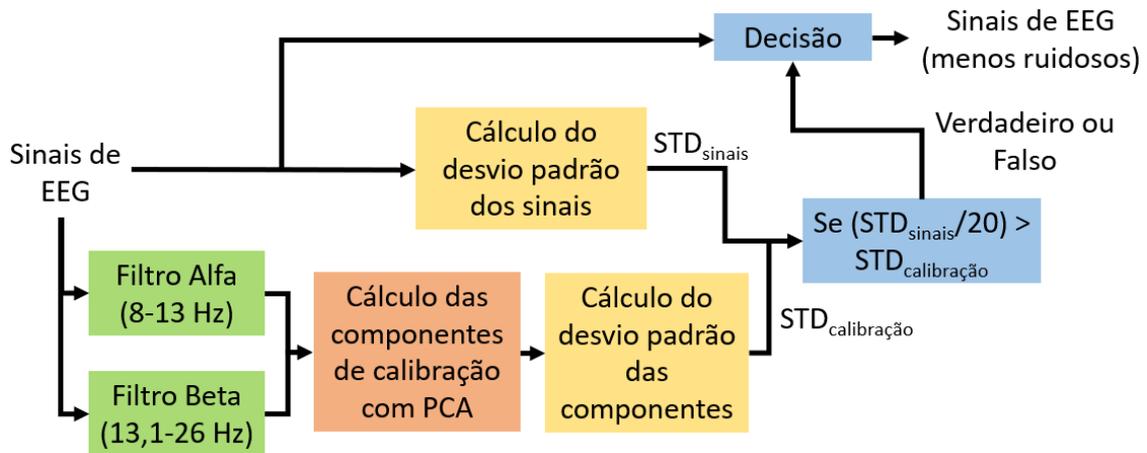
O algoritmo identifica os canais com menos de 70% de correlação com os demais canais (do mesmo indivíduo) como excessivamente ruidosos e os exclui.

A técnica *Artifact Subspace Reconstruction* (ASR), apresentada por (KOTHE & JUNG, 2016), calcula seus componentes de calibração aplicando uma análise de componentes principais (*Principal Component Analysis* - PCA) após retirar as componentes nas frequências Alfa (8–13 Hz) e Beta (13,1–26 Hz) através de filtros rejeita-banda. Os filtros e a PCA utilizados nesta etapa são utilizados apenas para calcular as componentes de calibração da ASR (as componentes PCA) e não alteram diretamente o sinal de saída.

De posse destas componentes, a ASR rejeita os segmentos cujo desvio padrão é ao menos 20 vezes superior ao desvio padrão das componentes de calibração.

A Figura 12 apresenta um diagrama de blocos para auxiliar o entendimento do funcionamento do algoritmo ASR.

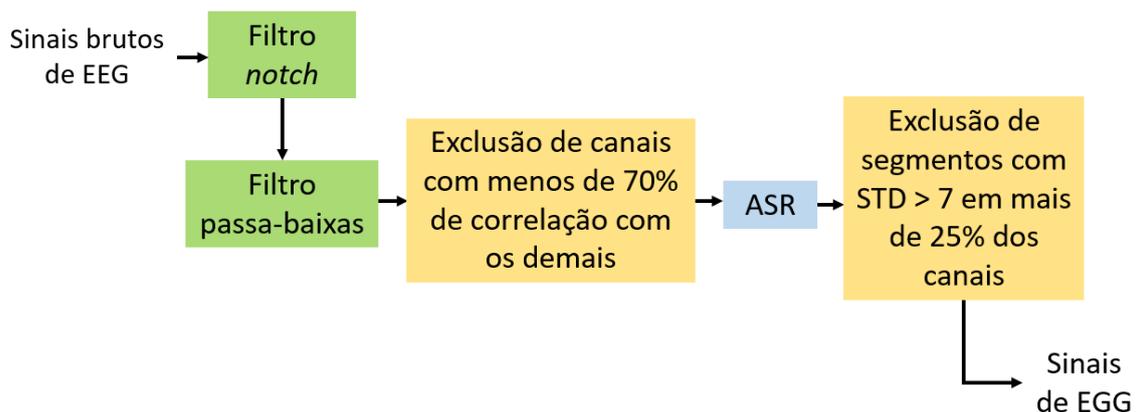
Figura 12 – Diagrama de blocos com o algoritmo ASR para remoção de ruído.



Fonte: O Autor

Mesmo depois do procedimento da ASR, os sinais de EEG ainda podem ter segmentos ruidosos. Por isso, aplicou-se mais um algoritmo de remoção de ruído. Este algoritmo identifica os segmentos que possuem desvio padrão maior que 7 em mais de 25% dos canais, estes são excluídos. Os processos de remoção de ruído são aplicados, separadamente, para cada participante. O diagrama de blocos com as etapas de filtros e de remoção de canais e de segmentos ruidosos é apresentado na Figura 13.

Figura 13 – Diagrama de blocos com a parte inicial do pré-processamento: filtros e algoritmos de remoção de ruído.



Fonte: O Autor

Infelizmente, a remoção de segmentos gera transições bruscas (bordas) entre o início e o final do segmento removido. Estas transições abruptas podem impactar a extração de características temporais. Então, identificou-se as posições temporais nas quais estas transições ocorrem, para a remoção futura das características correspondentes a estes segmentos.

3.3.1.2 Remoção de artefatos fisiológicos

Os sinais de EEG contém informações nas frequências menores que 0-100 Hz (?). Contudo, outros órgãos do corpo também atuam produzindo sinais nesta faixa de frequência, os quais contaminam as medições de EEG. Os principais contaminantes são os sinais de eletrocardiografia (ECG), eletromiografia (EMG) e eletro-oculografia (EOG).

Embora ocorram em bandas de frequências inseridas na faixa onde ocorrem os sinais de EEG, os artefatos possuem um comportamento diferente dos sinais de EEG e entre si.

Os artefatos de ECG são usados para mensurar a atividade elétrica do coração e costumam apresentar um padrão repetitivo, regular, com baixa amplitude (URIGÜEN & GARCIA-ZAPIRAIN, 2015) em uma frequência, principalmente, entre 14 e 33 Hz (JIANG *et al.*, 2007). Artefatos de pulso podem ocorrer quando um eletrodo é colocado sobre um vaso sanguíneo pulsante, como uma artéria do escalpo. Embora os pulsos sejam similares aos sinais de EEG em frequência e em período de duração, os pulsos ocorrem apenas nos eletrodos diretamente sobre o vaso. Porém, estes pulsos são regulares e precedem o ECG em um intervalo constante (URIGÜEN & GARCIA-ZAPIRAIN, 2015). Os movimentos oculares produzem sinais de EOG que contaminam os canais pré-frontais e frontais de EEG. O movimento lateral dos olhos, provoca a elevação do potencial elétrico em um hemisfério cerebral e a redução no outro. As piscadas oculares são medidas como sinais de alta frequência com uma amplitude muito maior do que a dos sinais de EEG (URIGÜEN & GARCIA-ZAPIRAIN, 2015).

Os sinais de EMG são gerados por contrações musculares. Os artefatos em EEG podem se originar em músculos da face e do pescoço. Os artefatos de EMG possuem uma ampla distribuição espectral e podem contaminar quaisquer das faixas de frequências clássicas de EEG, especialmente, Alfa e Beta (URIGÜEN & GARCIA-ZAPIRAIN, 2015). Contudo, os artefatos de EMG costumam ter curta duração, frequências mais elevadas e uma maior potência que os sinais de EEG (URIGÜEN & GARCIA-ZAPIRAIN, 2015).

O procedimento de *Independent Component Analysis* (ICA) (HERAULT & JUTTEN, 1986) foi usado para a separação de fonte de cada eletrodo. Isto permite separar e distinguir os artefatos de ECG, EMG e EOG dos sinais de EEG. A ICA é realizada separadamente para os dados de cada participante.

O teorema do limite central determina que a misturar linear entre infinitos sinais independentes entre si tende a se comportar como uma distribuição normal. A ICA rotaciona os eixos dos sinais de forma a obter distribuições dos sinais com a menor semelhança possível a uma distribuição normal. Este processo é iterativo: a cada época o algoritmo realiza uma rotação e verifica se houve a redução da semelhança até atingir um limiar pré-determinado. A ICA foi realizada através da *toolbox* EEGLab (DELORME & MAKEIG, 2004). Nesta ferramenta, selecionou-se o uso do Infomax, por este ser o mais usado. O Infomax promove um procedimento recursivo que modifica cada sinal e mantém esta modificação caso esta tenha aprimorado a independência estatística entre os canais. Isto

equivale a maximizar entropia entre os sinais.

Ao separar as fontes, a ICA também separa os contaminantes de ECG, EMG e EOG. Estes passam a se concentrar em apenas poucos canais. O algoritmo ICLabel, do EEGlab (DELORME & MAKEIG, 2004), busca os canais nos quais se concentram estes artefatos. O ICLabel verifica a semelhança de cada sinal com o comportamento típico de cada contaminante e exclui os canais com 90% de probabilidade de serem artefatos.

No entanto, este método possui a desvantagem de que a quantidade de canais excluídos não é a mesma para todos os indivíduos. Isto origina a necessidade do controle de qualidade, apresentado na subseção 3.2.1.3.

3.3.1.3 Segmentação e Controle de Qualidade

Os sinais de cada indivíduo são segmentados em intervalos de 2 segundos sem sobreposição (*overlapping*). Este intervalo foi utilizado por ser um período de segmentação comum a outros trabalhos como, por exemplo, por (ALCHALABI *et al.*, 2018), (WOLTERING *et al.*, 2012), (SANDER *et al.*, 2010), (CETIN *et al.*, 2022) e (AYDN *et al.*, 2022).

Após a aplicação da ICA e das técnicas de rejeição de segmentos e de canais, observou-se que o número de segmentos e de canais restantes difere entre os participantes. Consequentemente, sem um controle de qualidade, analisariam-se com os classificadores indivíduos com números de características distintos. Por isso, foi necessário estabelecer-se limiares. Os indivíduos com quantidade de segmentos ou de canais abaixo dos limiares foram excluídos.

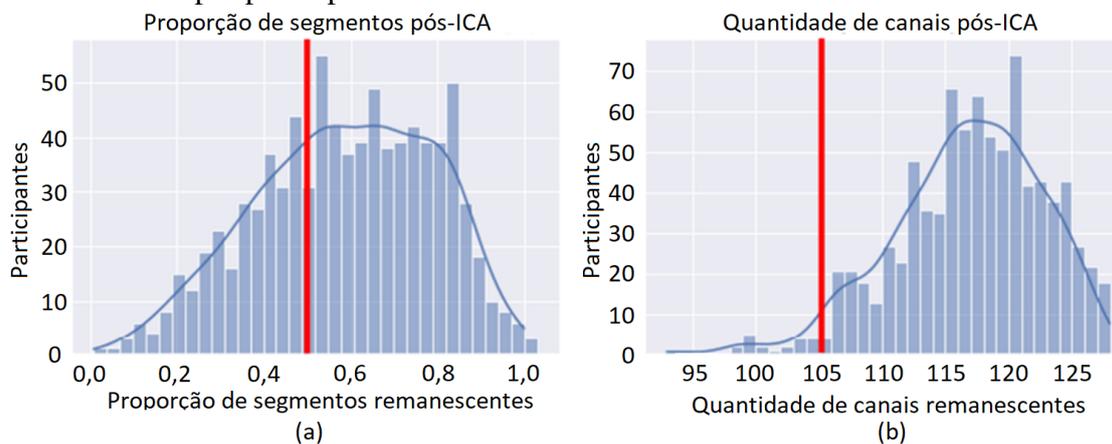
Para definir o número mínimo de segmentos e de canais a serem analisados de cada participante, analisou-se as distribuições dos indivíduos quanto a estes parâmetros. Para tanto, traçaram-se histogramas apresentados na Figura 14. Na Figura 14 (a) mostra-se a proporção de segmentos restantes por participante e na Figura 14 (b) apresenta-se a proporção de canais restantes por participante. A forma destas distribuições se assemelha à forma da distribuição normal com cauda para a esquerda.

No histograma da Figura 14 (a), uma parte considerável dos indivíduos tem menos de 50% de segmentos não excluídos, o que corresponde a menos de 2,5 minutos de dados. Uma parte considerável da comunidade científica considera que são necessários no mínimo 2 minutos de sinais de EEG para a análise de dados e diversos autores adotam este tempo mínimo de gravação de EEG em seus trabalhos (KIM *et al.*, 2023), (ANJUM *et al.*, 2020), (YANG *et al.*, 2022), (PAN *et al.*, 2024), (STRIK *et al.*, 1995). Para trabalhar-se com uma margem de tempo um pouco maior do que a mínima, estabeleceu-se 2,5 minutos de gravação dos sinais de EEG como o limiar de corte. Portanto, excluiu-se os participantes que após as exclusões de segmentos mantiveram menos de 2,5 minutos de gravação de EEG.

Através do histograma da Figura 14 (b), observou-se que os procedimentos elimina-

ram uma pequena quantidade de canais na maioria dos participantes. Uma vez que esta maioria possui mais de 105 canais, estabeleceu-se 105 canais como o limite mínimo de canais aceitável para um indivíduo participar dos experimentos de classificação.

Figura 14 – Histogramas (a) da proporção de segmentos e (b) da proporção de canais remanescentes por participante.



Fonte: O Autor

A Tabela 7 apresenta a proporção de sujeitos aprovados no controle de qualidade, cujos dados, portanto, são usados ao longo dos experimentos de classificação.

Tabela 7 – Proporção de sujeitos aprovados no controle de qualidade de EEG de acordo com a classe dos participantes. Legenda: 0 - sexo masculino 1 - sexo feminino. A idade e o gênero se referem às amostras após o controle de qualidade.

Classe	Total de Sujeitos	Aprov. (amostra)	Aprov. (%)	Proporção aprovados	Idade (anos)	Sexo
Controle	193	149	77,2%	30%	9,6±3,7	0,47±0,50
TDAH (TDAH-C +TDAH-I)	617	350	51,4%	70%	9,7±3,2	0,28±0,44
Subtipos de TDAH						
TDAH-C	329	169	52,8%	48%	9,0±3,0	0,23±0,42
TDAH-I	288	181	62,9%	52%	10,5±3,4	0,33±0,47

3.3.1.4 Filtros passa-faixas

Os sinais restantes foram filtrados nas bandas de frequência Delta (0,5-3,9 Hz), Teta (4-7,5 Hz), Alfa (8-13 Hz), Beta (13,1-26 Hz) e Gama (32-45 Hz), cujas faixas foram

definidas de acordo com (?). Além disto, manteve-se os sinais antes da aplicação destes filtros, pois esta representa o espectro completo dos sinais de EEG, o que é conhecido como banda larga (em inglês *broad band* – 0,5 - 100 Hz).

Projetou-se filtros passa-banda Chebyshev do tipo II (filtro Chebyshev inverso) para extrair os espectros Delta, Teta, Alfa, Beta, Gama. Utilizaram-se filtros Chebyshev do tipo II porque estes possuem uma declividade de atenuação mais íngreme e, por isso, uma menor região de transição entre a faixa de passagem e a faixa de rejeição (*roll-off* mais rápido) do que os filtros Butterworth. Diferentemente dos filtros Chebyshev do tipo I, os filtros do tipo II possuem uma faixa de passagem plana, comparável a banda de passagem dos filtros Butterworth (PAARMANN, 2005). Nos filtros Chebyshev do tipo II, a ondulação (mais conhecida como *ripple*) ocorre apenas na faixa de rejeição. A função de transferência do filtro Chebyshev do tipo II é dada pela Equação 26.

$$G_n(\omega) = \sqrt{\frac{\epsilon_{rip}^2 T_n^2\left(\frac{\omega_0}{\omega}\right)}{1 + \epsilon_{rip}^2 T_n^2\left(\frac{\omega_0}{\omega}\right)}}, \quad (26)$$

onde ϵ_{rip} é o fator de *ripple*, ω é a frequência de corte e C_n representa o polinômio de Chebyshev de ordem n . O termo T_n é um polinomial de Chebyshev de ordem n .

A ordem foi obtida aplicando-se a função `cheby2ord` no *software* MATLAB 2020b. A função de transferência dos filtros Chebbychev do tipo II também pode ser escrita na forma apresentada na Equação 27.

$$H(s) = \frac{\sum_{j=0}^m num_j s^j}{\sum_{i=0}^n den_i s^i}, \quad (27)$$

onde os coeficientes do denominador *den* e do numerador *num* foram determinados em função da ordem, do *ripple* máximo determinado para a faixa de rejeição, da banda de passagem, da frequência de Nyquist, e dos coeficientes W_{p1} e W_{p2} . Os coeficientes W_{p1} e W_{p2} representam as frequências limite da banda de rejeição, obtidos através da função `cheby2` do MATLAB. A frequência de Nyquist é 250 Hz, metade da frequência de amostragem. Durante a criação da base de dados, os pesquisadores da HBN, aplicaram um filtro passa-banda na faixa de 0,1 e 100Hz.

Os filtros das faixas de frequência de EEG foram projetados para apresentarem *ripples* máximos de 10 e de 40 na faixa de passagem e na faixa de rejeição, respectivamente. Entretanto, uma vez que o filtro utilizado é o Chebyshev do tipo II, o *ripple* na banda de passagem será nulo.

Os coeficientes W_{p1} e W_{p2} foram pré-ajustados como a razão entre a frequência de corte pretendida para a faixa de passagem e de rejeição, respectivamente. Estes valores foram, então, ajustados empiricamente com ajuda das funções para projetos de filtros (`freqz`) no MATLAB. A Tabela 8 apresenta a ordem n e os coeficientes das frequências de corte W_{p1} e W_{p2} dos filtros passa-faixas usados para cada banda de EEG. As magnitudes das respostas em frequência dos filtros projetados são apresentadas no APÊNDICE A.

Tabela 8 – Valores de projeto dos filtros digitais Chebyshev tipo II.

Faixa	Ordem (n)	W_{p1}	W_{p2}
Delta	3	0,48	2,10
Teta	2	0,27	3,80
Alfa	5	0,86	1,13
Beta	5	0,83	1,21
Gama	4	0,80	1,18

3.3.1.5 Extração de características

As seguintes características foram extraídas: Entropia, relação sinal-ruído (*signal to noise ratio* - SNR), desvio padrão, curtose, amplitude média, frequência média, frequência mediana, potência da banda (*band power*), distorção harmônica total, relação sinal-ruído e distorção (*signal to noise and distortion ratio* - SINAD), densidade de potência, *spurious free dynamic range* (SFDR), fator de impulso, 12 parâmetros autoregressivos e seus máximos atrasos negativo e positivo, *root mean square* (RMS), fator de folga (*clearance/clearing fator*), fator de crista (*crest fator*) e fator de forma (*shape factor* - SF).

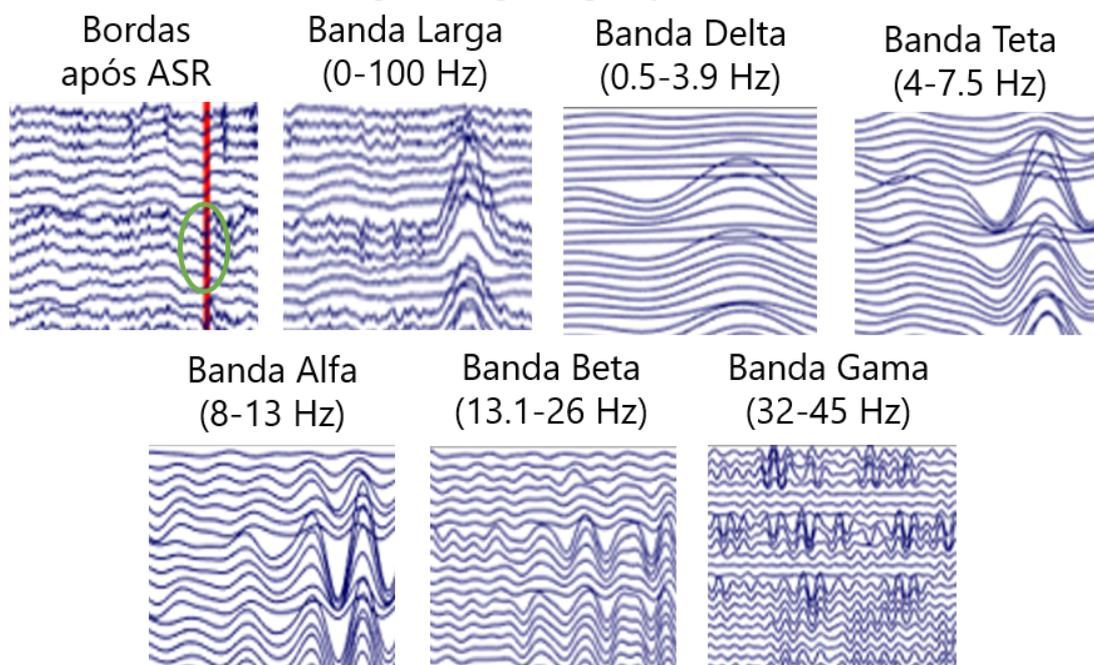
Estas características foram extraídas de cada segmento de cada canal de cada participante formando vetores de 32 características. Os vetores de 32 características de cada um dos 105 sinais são enfileirados na ordem dos sinais organizados após o ICA para cada amostra (cada uma constituída por um segmento), formando um vetor maior com 3.360 características.

Na sequência, aplicou-se a normalização z-score, também conhecida como *zero mean variance-unit* devido à sua equação característica. Esta normalização foi selecionada por ser uma das mais comuns e conhecidas na área de aplicação e foi utilizada também, por exemplo, por (MWAMBA *et al.*, 2019), (TAGHI BEYGLOU *et al.*, 2020), (TOR *et al.*, 2021). Nesta técnica de normalização, subtrai-se a média (μ) de cada elemento do vetor de características e divide-se o resultado pelo desvio padrão (σ), conforme a Equação 28.

$$z = \frac{x - \mu}{\sigma} \quad (28)$$

Após o *Artifact Subspace Reconstruction* (ASR), foram identificados os segmentos onde estavam localizadas as bordas de transição provocadas pela retirada de segmentos. A Figura 15 mostra os efeitos destas bordas em um determinado trecho de um segmento do indivíduo NDARAC853DTE (classe TDAH-C) em todas as bandas de frequência analisadas. Para evitar os efeitos que estas transições poderiam provocar na extração das características, optou-se pela remoção das características correspondentes aos segmentos onde as bordas foram identificadas.

Figura 15 – Bordas de Transição e seus efeitos em um mesmo trecho do conjunto de sinais de EEG em cada faixa de frequência, após a aplicação do ASR.



Fonte: O Autor

3.3.1.6 Balanceamento de amostras

Para o balanceamento das amostras de EEG utilizou-se a técnica de sobreamostragem (*oversampling*) conhecida como *Synthetic Minority Over-sampling Technique* (SMOTE) (CHAWLA *et al.*, 2002). O desempenho dos modelos com esta técnica foi comparado ao desempenho dos modelos usando a técnica de sub amostragem (*undersampling*) e à não utilização de métodos de balanceamento de amostras. Resultados iniciais apontaram que os modelos obtinham melhores *accuracies* usando SMOTE do que com a aplicação da técnica de sobreamostragem aleatória, na qual as amostras são selecionadas aleatoriamente para serem repetidas até obter-se um número de amostras igual para todas as classes.

No balanceamento das amostras, a técnica SMOTE foi aplicada de forma a igualar o número de amostras das classes minoritárias com o número de amostras da classe majoritária. O algoritmo do SMOTE seleciona aleatoriamente uma amostra x da classe minoritária, e, então, calcula seus cinco vizinhos mais próximos. Dentre estes cinco vizinhos, o algoritmo seleciona duas e gera uma nova amostra sintética com características selecionadas em um ponto aleatório entre as características destas duas amostras.

O algoritmo SMOTE é executado e estas amostras são selecionadas aleatoriamente a cada vez que cada classificador que trabalha com EEG é executado. Portanto, a cada execução dos modelos, as amostras sintéticas são geradas novamente de forma aleatória.

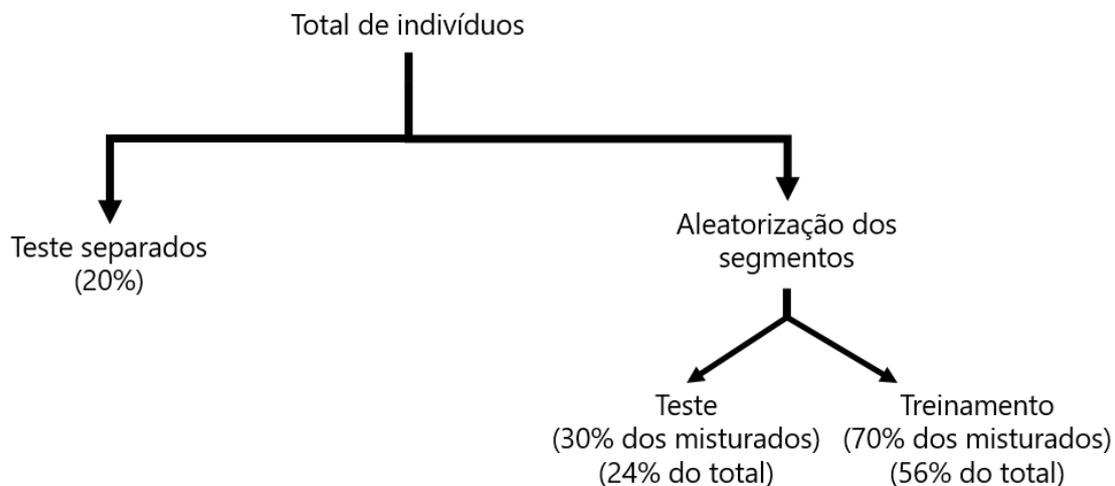
3.3.1.7 Separação em grupos de teste e de treinamento

Considerou-se como uma amostra cada conjunto de segmentos de EEG de dois segundos de todos os canais (sempre ordenados da mesma forma). Destas amostras, 20% do total de amostras após o controle de qualidade foram aleatoriamente separadas em um conjunto de teste “separado”. Para os modelos de classificação de sinais de EEG desta Tese, embaralharam-se os indivíduos e as amostras de EEG aleatoriamente. Embaralhou-se, aleatoriamente, os participantes de forma que os indivíduos de diferentes classes estão misturados nos grupos de teste e de treinamento, mantendo seus rótulos de classe. Então, 70% e 30% destas amostras são separadas em conjuntos de treinamento e de teste, respectivamente. Os segmentos do grupo "separado" são aleatorizados, separadamente. O grupo "separado" consiste em um conjunto de dados, cujos segmentos nunca foram vistos no treinamento.

A separação do grupo de teste “separado” tem o objetivo de testar o desempenho dos classificadores com dados de indivíduos que o classificador nunca viu nenhuma parte de seus sinais de EEG antes. Os resultados obtidos com o conjunto de teste "misturados" são apresentados na seção de resultados e os resultados obtidos para o conjunto de teste "separados" são relatados no APÊNDICE C.

O fluxograma da Figura 16 resume este procedimento de separação das amostras em conjuntos de treinamento e de teste.

Figura 16 – Processo de aleatorização e de separação das amostras de EEG em conjuntos de treinamento e de teste.



Fonte: O Autor

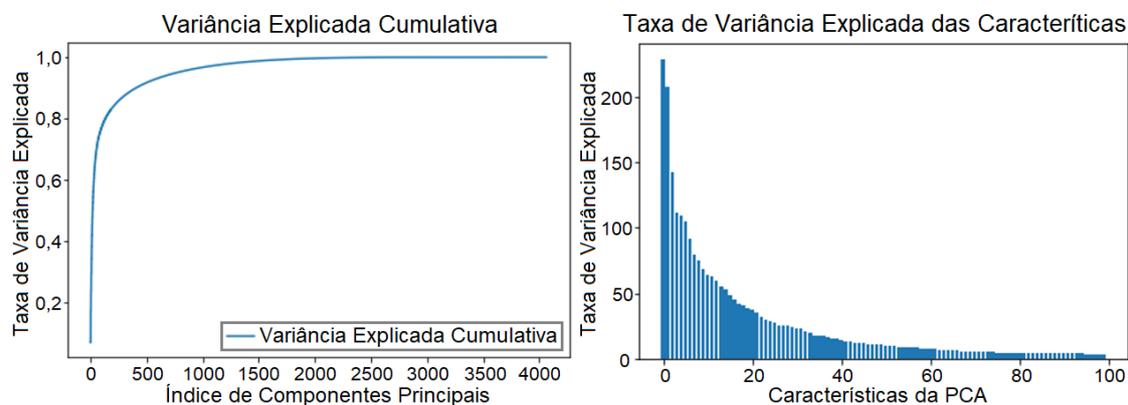
3.3.1.8 Principal Component Analysis (PCA)

Realizou-se uma PCA das características para cada faixa de frequência para cada diagnóstico de TDAH e de subtipos (TDAH-C e TDAH-I). A PCA foi utilizada com o objetivo de reduzir possíveis redundâncias de características. Estas redundâncias poderiam confundir os classificadores. A PCA maximiza a informação relevante e reduz a complexidade de processamento dos classificadores.

Para definir a quantidade de componentes principais a serem utilizadas como entradas pelos classificadores, traçaram-se os gráficos da curva da variância explicada cumulativa e da taxa de variância explicada das características da PCA. Estes gráficos foram muito similares entre as bandas larga (0-100 Hz), Delta, Teta e Alfa. As curvas das bandas Beta e Gama diferem das demais, porém se assemelham entre si. Com o objetivo de manter-se um alto nível de taxa de variância explicada pela PCA (97%) para todos os casos, decidiu-se trabalhar com as primeiras mil componentes da PCA para cada faixa de frequência separadamente tanto nas classificações entre grupos TDAH e controle como nas classificações entre subtipos (TDAH-C e TDAH-I).

A Figura 17 apresenta a curva da variância explicada cumulativa e a taxa de variância explicada das características da PCA para a banda Delta considerando a classificação TDAH/Controle. E, a Figura 18 apresenta as mesmas curvas, porém obtidas considerando a faixa Beta.

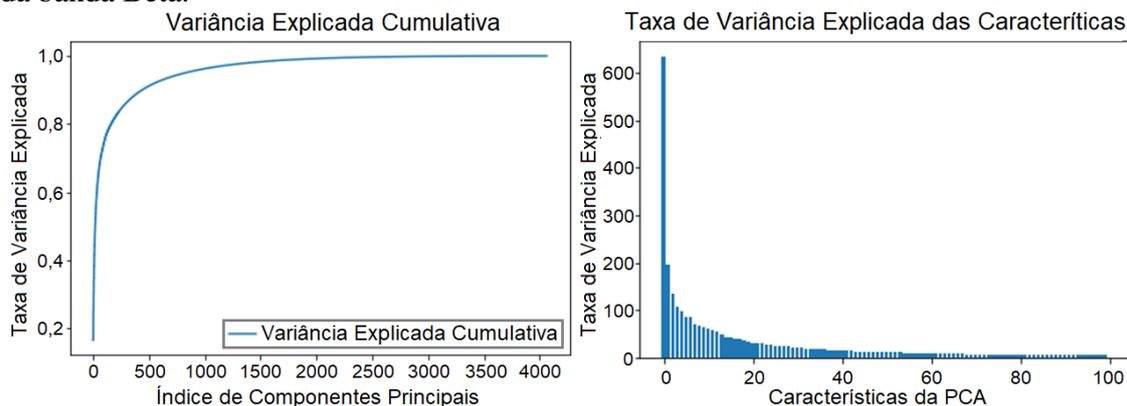
Figura 17 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Delta.



Fonte: O Autor

Os gráficos da variância explicada cumulativa e a taxa de variância explicada das características da PCA para as demais faixas de frequência se encontram no APÊNDICE A.

Figura 18 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Beta.



Fonte: O Autor

3.3.2 Processamento de EEG: *Random Forest* (RF)

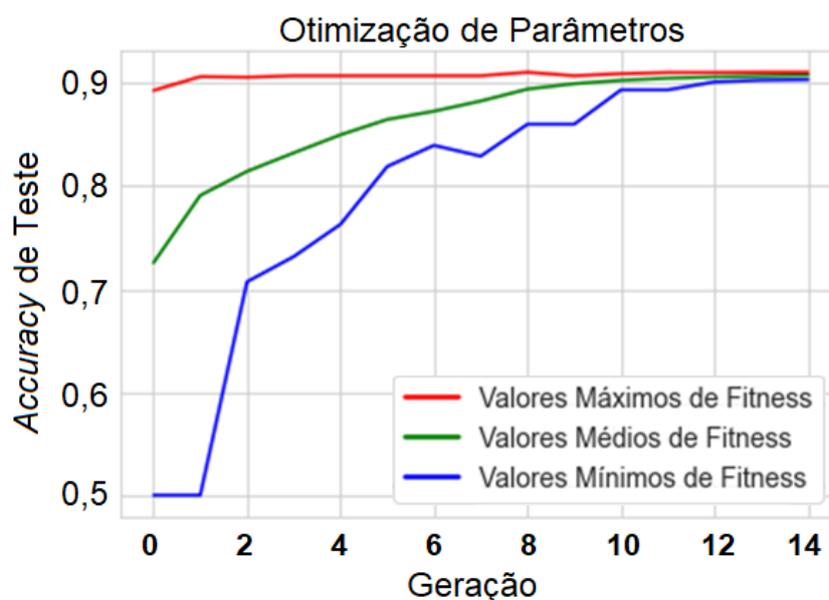
Para a otimização da RF utilizou-se um algoritmo genético (AG), pesquisando-se hiperparâmetros em uma população de 100 indivíduos (onde cada indivíduo é uma combinação de hiperparâmetros) através de 15 gerações. Estabeleceu-se uma probabilidade de mutação de 0,002 e uma probabilidade de cruzamento de 0,95. Caso o AG pare de aprimorar o modelo e, por 10 gerações, encontre o mesmo valor mínimo de aptidão (*Fitness*), o procedimento é interrompido e a configuração resultante é aceita como sendo a melhor. Utilizou-se com um conjunto de hiperparâmetros da RF: Número de estimadores, critério, profundidade máxima, divisão mínima de amostras e mínimo de amostras por folha. Os hiperparâmetros da RF foram investigados nos intervalos fornecidos na Tabela 9. Estes intervalos foram pré-selecionados em uma busca exaustiva anterior ao uso do AG. Uma vez que otimizações meta-heurísticas são processos aleatórios, cada AG foi executado 10 vezes afim de garantir a melhor combinação de hiperparâmetros.

A Figura 19 mostra as curvas de otimização obtidas pelo AG para o diagnóstico TDAH/Controle com as características observadas na banda Delta. Os valores máximos, médios e mínimos de *Fitness* (*accuracy*) apresentados na figura são referentes aos indivíduos da população até a geração apresentada no eixo 'Geração' do gráfico. A curva apresentada no gráfico corresponde a uma execução do AG e não a média das 10 vezes que o AG foi executado para obter-se a melhor combinação de hiperparâmetros.

Tabela 9 – Intervalos investigados por AG para otimizar o desempenho do classificador RF baseado em sinais de EEG.

Hiperparâmetros	Intervalo	Aleatorização
Número de Estimadores	1 a 200	Uniforme
Critério	Gini ou Entropia	Uniforme
Profundidade Máxima	1 a 30	Uniforme
Divisão mínima de amostras	2 a 30	Uniforme
Mínimo de amostras por folha	0 a 5	Uniforme

Figura 19 – Otimização AG da RF para classificação TDAH/Controle para banda Delta.



Fonte: O Autor.

Estas curvas, da Figura 19, mostram a evolução dos valores de *Fitness*. Estas *Fitness* correspondem as *accuracies* de teste de cada membro da população de possibilidades de soluções do AG. Assim, este gráfico mostra os valores máximos, médios e mínimos das *accuracies* de teste da RF convergindo para um alto valor ao longo das gerações.

As curvas de otimização dos classificadores RF referentes as demais bandas se encontram no APÊNDICE B. Observou-se que na maioria destas curvas, assim como mostra a Figura 19, os valores máximos de *Fitness* iniciam com valores altos e aumentam um pouco ao longo das gerações, convergindo ao valor final. Já os valores mínimos de *Fitness* sofrem uma alteração maior: no caso demonstrado na figura, o incremento de *accuracy* dos valores mínimos é de 0,4.

A Tabela 10 apresenta os hiperparâmetros selecionados através da otimização promovida pelo AG. Nesta tabela observa-se que os AGs escolheram quantidades de estimadores próximas ao máximo investigado (200) para todas as bandas, exceto para a banda larga. Como critério de avaliação, a entropia foi selecionada para todas as faixas de frequência com exceção da Delta, na qual o AG apontou para o uso do índice Gini. todos os AGs apontaram para profundidade máxima = 5 e número mínimo de amostras por folha = 0. Já a divisão mínima de amostras esteve entre 20 e 29 para a maioria das bandas (a única exceção foi a banda Beta).

Tabela 10 – Hiperparâmetros otimizados pelo AG para classificar TDAH/Controle através de RFs.

Parâmetros Selecionados	Banda Larga	Delta	Teta	Alfa	Beta	Gama
Número de Estimadores	130	186	196	194	173	185
Critério	Entropia	Gini	Entropia	Entropia	Entropia	Entropia
Profundidade Máxima	5	5	5	5	5	3
Divisão mínima de amostras	28	20	29	25	11	2
Mínimo de amostras por folha	0	0	0	0	0	0

A Tabela 11 mostra os valores dos hiperparâmetros das RFs otimizados pelo AG para classificar TDAH-C/TDAH-I para cada banda de EEG. Nesta tabela, notou-se que o número de estimadores para realizar a classificação TDAH-C/TDAH-I é consideravelmente menor do que o escolhido para classificar TDAH/Controle para as bandas Delta e Gama. Para as faixas Teta, Alfa e Beta foi escolhido um valor similar para classificar TDAH-C/TDAH-I ao que havia sido selecionado para TDAH/Controle. No entanto, o número de estimadores foi consideravelmente maior na classificação de subtipos com a banda larga. Este aumento ocorreu devido a maior dificuldade de cada estimador em obter uma *accuracy* alta na classificação de subtipos.

Na seleção do critério que mede a qualidade das divisões dos galhos, os AGs escolheram a Entropia para as bandas larga, Delta e Teta e o índice Gini para as demais. A divisão mínima de amostras para classificar TDAH-C/TDAH-I foi entre 5 (banda larga) e

18 (Beta), logo menor do que o que havia sido obtido para TDAH/Controle. Isto significa que a classificação TDAH-C/TDAH-I requer RFs mais flexíveis.

Tabela 11 – Hiperparâmetros otimizados pelo AG para classificar TDAH-C/TDAH-I através de RFs.

Parâmetros Seleccionados	Banda Larga	Delta	Teta	Alfa	Beta	Gama
Número de Estimadores	161	142	199	192	175	117
Critério	Entropia	Entropia	Entropia	Gini	Gini	Gini
Profundidade Máxima	5	5	5	5	5	5
Divisão mínima de amostras	5	12	14	10	18	11
Mínimo de amostras por folha	0	0	0	0	0	0

3.3.3 Processamento de EEG: *Support Vector Machine* (SVM)

Para a otimização da SVM através de AG utilizou-se o mesmo critério de parada e os mesmos critérios do AG, população de 100 indivíduos, 15 gerações, probabilidade de mutação de 0,002 e probabilidade de cruzamento de 0,95. No entanto, a parada do AG pode ser antecipada caso não haja variação nos valores mínimos de *Fitness* por 10 gerações. Otimizou-se os hiperparâmetros da SVM: *kernel*, Gama e C. Selecionou-se estes intervalos com base em resultados de uma busca exaustiva realizada, de forma preliminar, manualmente antes de se empregar o AG.

Na Tabela 12, apresentam-se os intervalos investigados e a distribuição utilizada na aleatorização tanto da seleção inicial das sequências genéticas avaliadas como na mutação dos hiperparâmetros.

Tabela 12 – Intervalos e aleatorização investigados por AG para otimizar o desempenho do classificador SVM baseado em sinais de EEG.

Hiperparâmetros	Intervalo	Aleatorização
<i>Kernels</i>	Linear, sigmoide, polinomial ou RBF	Uniforme
C	0,01 a 1	Gaussiana
γ_{SVM}	1 a 1000	Uniforme

A Tabela 13 e a Tabela 14 apresentam os hiperparâmetros otimizados obtidos através dos AGs, respectivamente, para as classificações TDAH/Controle e TDAH-C/TDAH-I. Na maioria das bandas, os GAs optaram pelo uso do *kernel* RBF.

Tabela 13 – Hiperparâmetros otimizados pelo AG para classificar TDAH/Controle através de SVMs.

Parâmetros Seleccionados	Banda Larga	Delta	Teta	Alfa	Beta	Gama
<i>Kernel</i>	Linear	RBF	RBF	RBF	Poli. (grau 4)	RBF
C	41	267	87	218	934	958
Gama	0,66	0,042	0,87	0,90	0,38	0,15

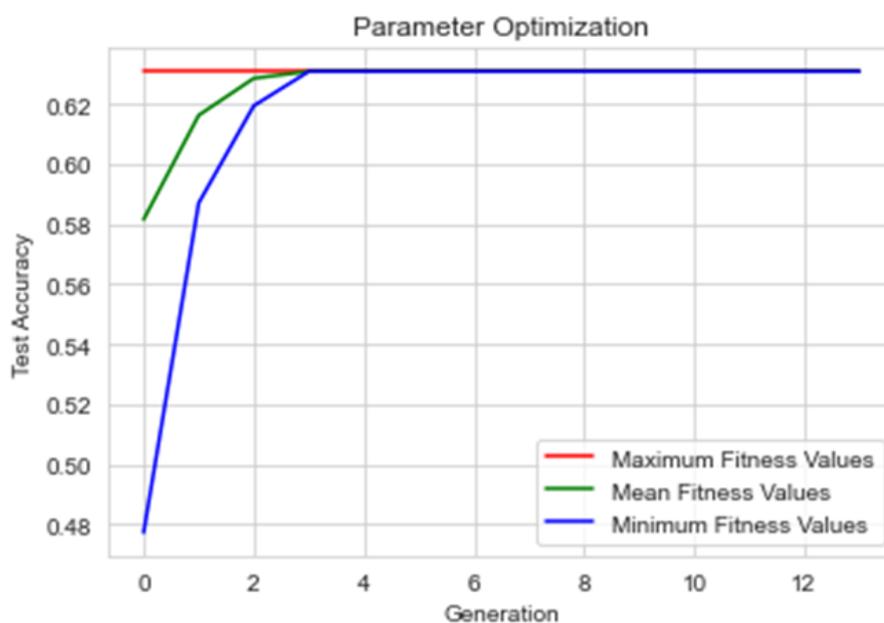
Tabela 14 – Hiperparâmetros otimizados pelo AG para classificar TDAH-C/TDAH-I através de SVMs.

Parâmetros Seleccionados	Banda Larga	Delta	Teta	Alfa	Beta	Gama
<i>Kernel</i>	Poli (grau 4)	RBF	Linear	RBF	Poli (grau 4)	Poli (grau 4)
C	44	29	198	62	145	268
Gama	0,294	0,420	0,940	0,450	0,796	0,085

Assim como realizado para as RFs, os AGs de otimização das SVMs foram executados 10 vezes afim de garantir a escolha dos melhores hiperparâmetros.

A Figura 20 apresenta, como exemplo, a curva de otimização dos hiperparâmetros obtida para o caso da classificação de participantes dos grupos TDAH e controle a partir dos dados da banda Teta. As demais curvas de otimização para classificação TDAH/Controle e TDAH-C/TDAH-I se encontram no APÊNDICE B. Observou-se nas curvas de otimização, que os AGs encontraram os valores ótimos dos hiperparâmetros das SVM com um menor número de gerações do que necessitaram para as RFs. Isto, possivelmente, se deve ao uso de um menor número de hiperparâmetros nas SVMs. Assim como para as curvas apresentadas na RF, o gráfico da Figura 20 mostra as curvas referente a apenas uma das 10 execuções do AG de otimização do modelo.

Figura 20 – Otimização AG da SVM para classificação TDAH/Controle para banda Teta.



Fonte: O Autor.

Os resultados referentes às classificações de TDAH e de seus subtipos através da análise de sinais de EEG com SVM encontram-se no APÊNDICE D.

3.3.4 Processamento de EEG: Fusão de Classificações a partir das Bandas de EEG

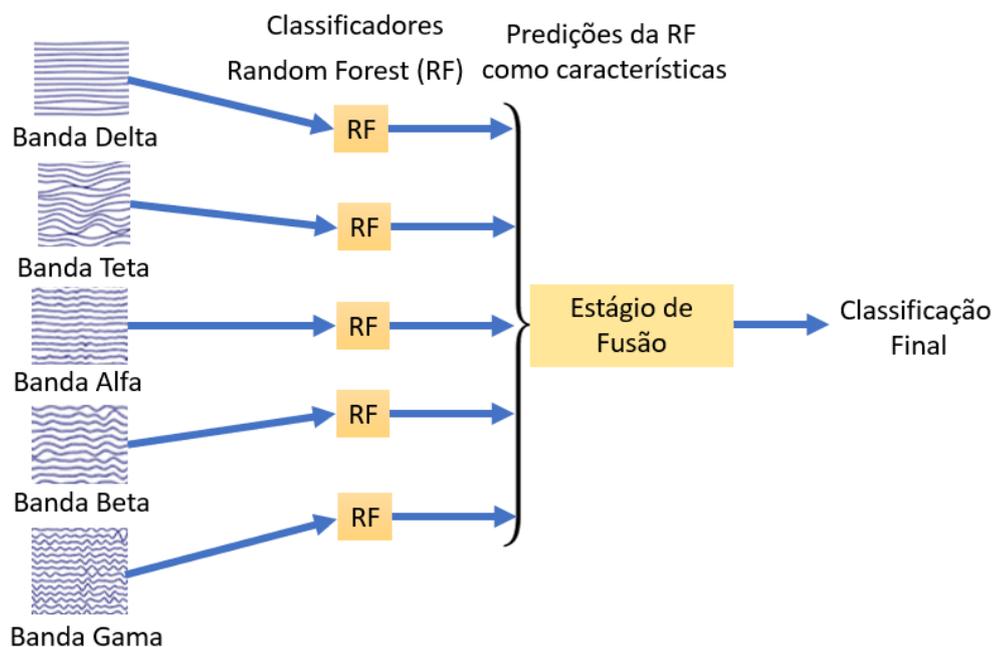
Através das matrizes de confusão obtidas para cada banda de EEG (apresentadas no capítulo de Resultados), constatou-se que os classificadores de cada uma das bandas Delta, Teta, Alfa, Beta e Gama conseguiram distinguir as classes. Isto foi observado tanto para o diagnóstico do TDAH como para a distinção entre subtipos.

Contudo, as *accuracies* obtidas analisando cada faixa de frequência de EEG são diferentes entre si. Por isto, elaborou-se a hipótese de que com a união das informações obtidas pelos classificadores seria possível obter uma classificação com maior *accuracy* total e manter-se o balanço entre as *accuracies* obtidas para cada classe.

Para realizar a união das informações extraídas de cada faixa de EEG, optou-se por uma abordagem de *ensemble*, pois esta abordagem permite aproveitar as previsões já obtidas por cada classificador de cada banda. Dessa forma, as classificações obtidas através dos modelos primários são enviadas como características para o meta-classificador, o qual produz em sua saída a classificação final do modelo.

Uma vez que as RFs obtiveram maiores *accuracies* que as SVMs nas classificações individuais, as previsões das RFs foram selecionadas como características para o meta-classificador. Este modelo de fusão dos resultados dos classificadores das bandas de EEG com meta-classificadores é resumido através da Figura 21.

Figura 21 – Fusão dos resultados de classificação baseados nas bandas de EEG.



Fonte: O Autor.

Para o estágio de fusão, consideraram-se os métodos: Voto Majoritário Ponderado, classificador linear *Ridge*, *AdaBoost* (FREUND & SCHAPIRE, 1996) e *XGBoost* (CHEN & GUESTRIN, 2016).

O Voto Majoritário consiste em somar as previsões binárias das classes ('0' ou '1') obtidas por cada classificador para cada indivíduo e compará-las a um limiar. A comparação com o limiar define a conclusão do meta-classificador.

No âmbito desta Tese, o Voto Majoritário utilizado é ponderado. Nesta técnica, atribui-se pesos para cada classificador primário de acordo com a importância percebida de cada modelo. Estes pesos multiplicam a classe predita ('0' ou '1') antes da soma dos votos.

A Tabela 15 apresenta os pesos utilizados no Voto Majoritário Ponderado para as classificações entre TDAH e grupo controle e de subtipos (TDAH-C *versus* TDAH-I).

Tabela 15 – Pesos utilizados no Voto Majoritário Ponderado das predições obtidas por classificadores baseados nas bandas de EEG.

Banda de Frequência	TDAH vs Controle	Subtipos
Delta	$0,9105^2 = 0,8290$	$0,8605^2 = 0,7405$
Teta	$0,8874^2 = 0,7875$	$0,8655^2 = 0,7491$
Alfa	$0,8529^2 = 0,7274$	$0,8605^2 = 0,7405$
Beta	$0,7625^2 = 0,5814$	$0,7925^2 = 0,6281$
Gama	$0,6679^2 = 0,4461$	$0,8075^2 = 0,6521$

No presente estudo, o peso atribuído à classificação de um determinado classificador foi o quadrado da *accuracy* obtida por este classificador quando utilizado individualmente. O peso de cada *accuracy* foi elevado ao quadrado para aumentar a relevância das bandas apontadas como mais relevantes pela análise dos classificadores baseados nas RF. O meta-classificador apresentou maiores *accuracies* nos ensaios quando estes pesos foram elevados ao quadrado.

No voto majoritário realiza-se a soma dos votos multiplicados pela classe predita ('0' ou '1'). Esta soma é comparada à um valor limiar pré-determinado para decidir a qual classe pertence o indivíduo avaliado. O valor limiar foi selecionado para ser a metade do valor máximo que poderia ser obtido pela soma dos votos quando todos os classificadores primários apontarem para a classe '1'.

Os hiperparâmetros do *AdaBoost* utilizado para a fusão das predições dos classificadores das bandas de EEG são o número de estimadores e a taxa de aprendizado. Para determinar a melhor combinação de hiperparâmetros, utilizou-se um AG. O *AdaBoost* otimizado possui 1039 estimadores e taxa de aprendizado de 0,65568698.

A Tabela 16 apresenta os hiperparâmetros e seus intervalos investigados pelo GA.

Tabela 16 – Intervalos investigados por AG para otimizar o desempenho do meta-classificador *AdaBoost* baseado em sinais de EEG.

Hiperparâmetros	Intervalo	Aleatorização
Número de Estimadores	50 a 5000	Uniforme
Taxa de Aprendizado	0,001 a 1	Uniforme

Os valores dos hiperparâmetros do *XGBoost* foram obtidos através de um AG que investigou os intervalos de hiperparâmetros apresentados na Tabela 17. O *Minimum child weight* é a soma mínima de peso que um nó deve ter para que o processo de produção de

novos galhos continue. O AG apontou para o uso de 2047 estimadores base, $\eta = 1,5715$, $\gamma = 6,0514$, Máxima profundidade das árvores = 5, *Minimum child weight* = 5, taxa de regularização L1 = 3 e taxa de regularização L2 = 10.

Utilizou-se o método clássico de *Gradient Tree Boosting*, conhecido como *gbtree*, para otimizar o *XGBoost*, o qual apresentou maiores *accuracies* e maior velocidade de cálculo do que o método DART de (VINAYAK & GILAD-BACHRACH, 2015) em comparações nos experimentos iniciais realizadas de forma manual. O método para a otimização da divisão de galhos utilizado é o *Basic Exact Greedy (?)*, no qual todas as possíveis divisões de galhos são avaliadas para todas as características.

Tanto os AGs utilizados para otimização do *Adaboost* como do *XGBoost* foram executados 10 vezes cada para reduzir os efeitos da aleatoriedade da seleção de populações iniciais, cruzamento e mutação inerentes ao AG sobre a seleção da melhor combinação de hiperparâmetros. Contudo, cada um dos gráficos apresentados abaixo mostram curvas de otimização referentes a apenas uma das execuções do AG.

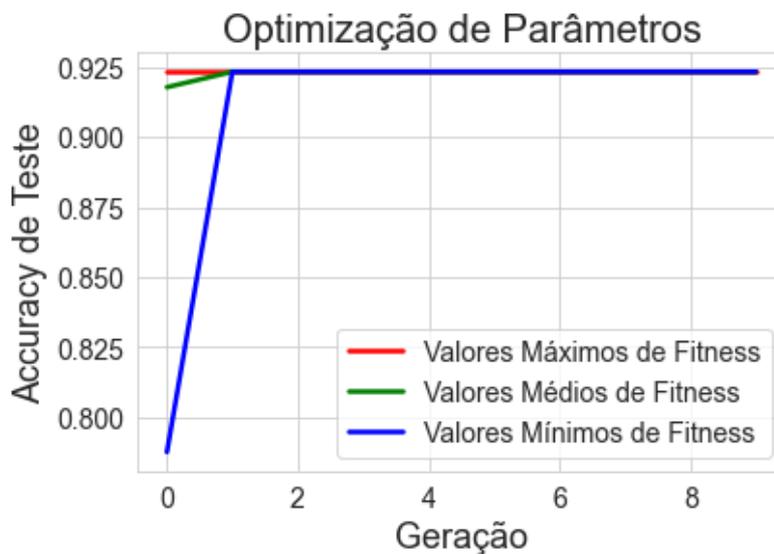
A execução do AG do *XGBoost* com os parâmetros apresentados na Tabela 17, usaram tamanho de população igual a 70, probabilidade de cruzamento igual a 0,5, probabilidade de mutação de 0,0001 por 10 gerações, tanto para a classificação entre grupos TDAH e controle como para a identificação do subtipo ao qual o paciente com TDAH pertence. A otimização da distinção das classes TDAH e Controle por AG transcorreu conforme apresentado na Figura 22. Nesta figura, os valores de *fitness*, correspondem a *accuracy*. O valor máximo de *accuracy* obtido pela população desde a primeira geração é de 0,925. Os valores médios e mínimos de *accuracy* convergem para 0,925 a partir da segunda geração (geração 1 no gráfico).

Tabela 17 – Intervalos investigados por AG para otimizar o desempenho do meta-classificador *XGBoost* baseado em sinais de EEG.

Hiperparâmetros	Intervalo	Aleatorização
Número de Estimadores	1 a 10000	Uniforme
η	0,001 a 10	Uniforme
γ	0,001 a 10	Uniforme
Máxima profundidade das árvores	1 a 6	Uniforme
<i>Minimum child weight</i>	1 a 10	Uniforme
Taxa de L1	0 a 10	Uniforme
Taxa de L2	0 a 10	Uniforme

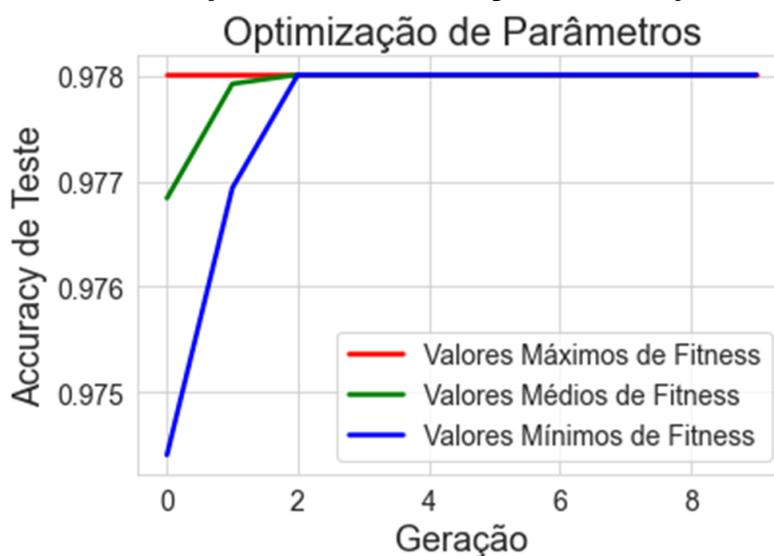
A Figura 23 mostra a otimização por AG da classificação dos subtipos de TDAH. Esta figura, é bastante similar a Figura 22. No entanto, na otimização para subtipos, o AG do meta-classificador *XGBoost* demorou uma geração a mais para convergir ao valor ótimo da *accuracy* (0,978), o que ocorreu na terceira geração (geração 2 na figura). Porém, para a identificação de subtipos o *XGBoost* otimizado alcançou uma *accuracy* maior. Os hiperparâmetros otimizados são: 9367 Estimadores, $\eta = 0,144$, $\gamma = 4,128$, *Minimum child weight* = 1, Profundidade de árvore = 6, taxa de L1 = 1 e taxa de L2 = 5.

Figura 22 – Otimização AG da Fusão das bandas de EEG por *XGBoost* para classificação TDAH/Controle.



Fonte: O Autor.

Figura 23 – Curva de Otimização da fusão *XGBoost* para classificação TDAH-C/TDAH-I.



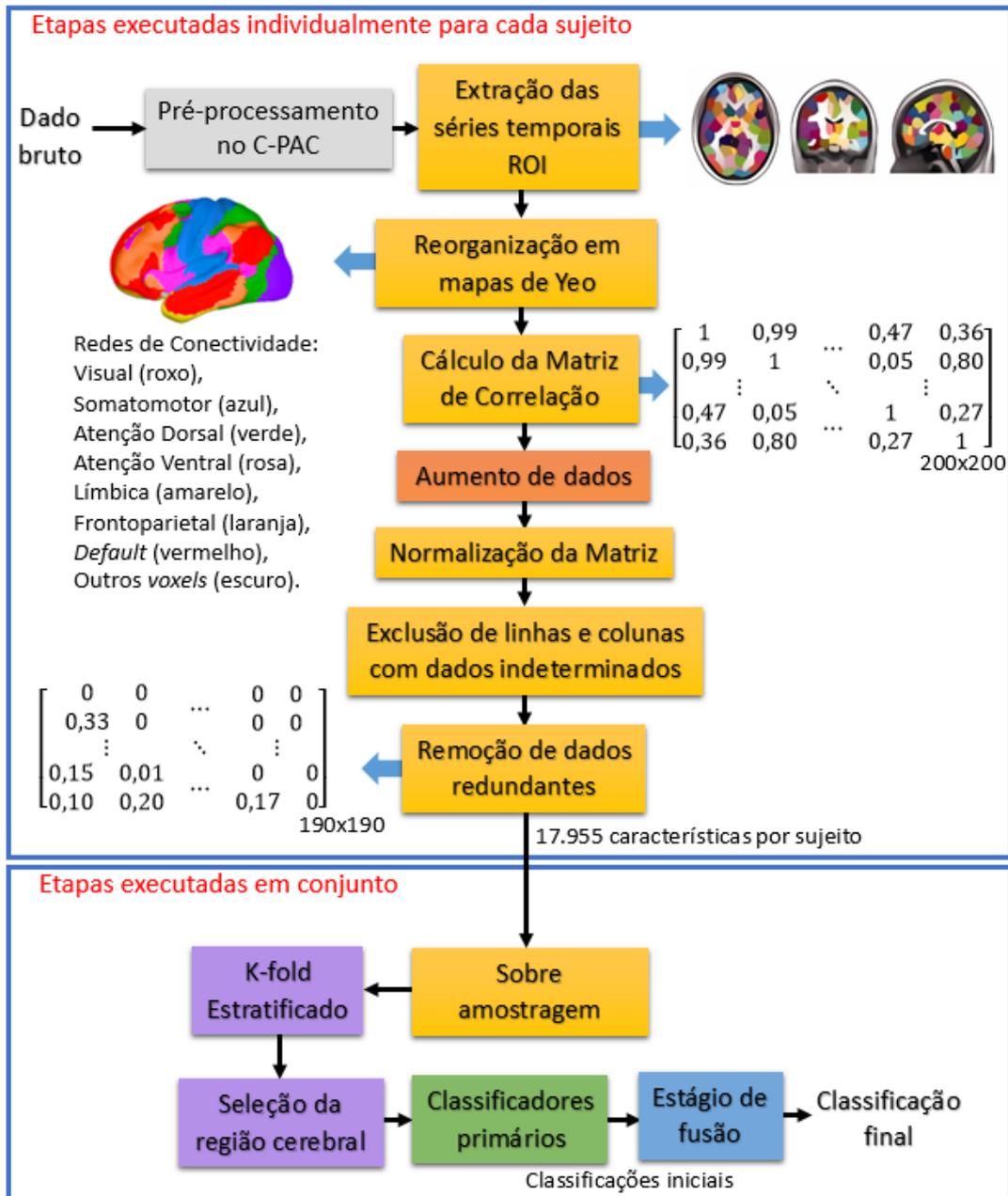
Fonte: O Autor

3.4 Classificação de TDAH a partir de Imagens de Ressonância Magnética Funcional (fMRI)

A classificação de TDAH (e de seus subtipos) a partir das imagens de fMRI foi executada de forma análoga ao processamento dos sinais de EEG, no que diz respeito ao uso dos meta-classificadores. No caso de fMRI, os biomarcadores investigados são as redes de conectividade funcional.

A Figura 24 mostra o fluxograma completo de processamento dos dados, desde o dado bruto até a classificação final.

Figura 24 – Fluxograma do processamento de fMRI desde o dado bruto até a classificação final.



Fonte: O Autor

As etapas realizadas ao longo do processamento de fMRI são: o pré-processamento de fMRI; a extração de características formada por uma sequência algoritmos que culminam no cálculo das matrizes de correlação funcional (serão explicadas em detalhes na seção 3.4.3 que inicia na página 92).; o aumento de dados; a normalização; a exclusão de ROIs com dados indeterminados e de dados redundantes. Após agrupar os indivíduos, realiza-se a sobreamostragem para o balanceamento de amostras; a validação cruzada *k-fold*; a

seleção da região cerebral que será a entrada de cada classificador primário; e, por fim, o meta-classificador que realiza a classificação final. Estas etapas apresentadas serão detalhadas nas subseções seguintes.

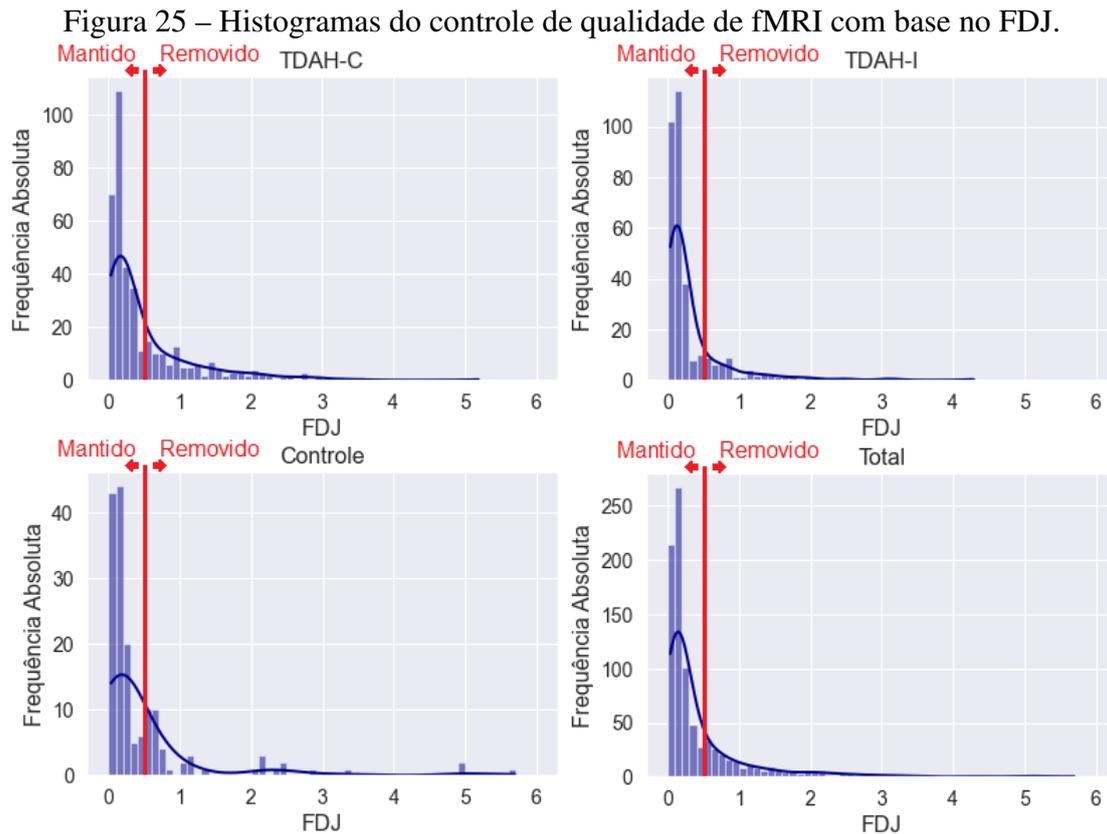
3.4.1 Controle de Qualidade de fMRI

Sabe-se que, em fMRI, o movimento dos pacientes é a principal causa de ruído e de distorções nos sinais eletromagnéticos medidos (ALEXANDER *et al.*, 2017). Os autores de (JENKINSON *et al.*, 2002) desenvolveram uma forma de medir estes movimentos, conhecida como *frame-wise displacement Jenkinson* (FDJ). O FDJ é uma série temporal que quantifica ao longo do exame o comportamento dinâmico da cabeça de um paciente. O FDJ foi aferido pela base HBN ao longo dos exames de fMRI para cada indivíduo. O FDJ de cada amostra é um sinal avaliado segundo uma medida estocástica.

Considerou-se a possibilidade de utilizar-se a média ou a mediana como medida de avaliação do FDJ. Para esta escolha, utilizou-se como critério o fato de que um único evento de movimento de maior amplitude prejudica menos a leitura dos sinais de fMRI do que uma quantidade expressiva de movimentos corporais frequentes (ALEXANDER *et al.*, 2017).

Na média aritmética calcula-se a soma de todos os pontos do sinal e se divide esta soma pela quantidade de pontos do sinal. Por isto, um evento de alta amplitude no sinal pode elevar muito a sua média. Já, para calcular a mediana ordenam-se os pontos do sinal em ordem crescente. Se o número de elementos do sinal (n) for ímpar, o valor da mediana é o valor do elemento na posição $(n+1)/2$. E, se n for par, a mediana é o valor da média aritmética entre os valores nas posições $(n/2)$ e $(n+1)/2$. Portanto, a mediana não é influenciada por eventos únicos de alta amplitude, cujos valores das amplitudes seriam posicionados em uma das extremidades do vetor ordenado. Por outro lado, elementos de movimentos frequentes produzem uma maior quantidade de elementos com amplitude elevada que mais provavelmente influenciaram o valor da mediana. Por isso, preferiu-se avaliar o sinal FDJ com a mediana ao invés da média.

A fim de avaliar qual o valor máximo aceitável neste estudo para o FDJ, traçaram-se histogramas para verificar a relação entre os níveis de FDJ e a quantidade de indivíduos. Esta análise foi necessária para equilibrar a qualidade e a quantidade dos dados remanescentes. Estes histogramas são apresentados na Figura 25.



Fonte: O Autor

Como mostra a Figura 25, além de analisar as quantidades de dados em função do FDJ para o conjunto total de dados, avaliou-se o comportamento dos histogramas do FDJ para cada classe separadamente. Através desta separação verificou-se que os FDJ das classes eram similares e que nenhuma classe seria excessivamente penalizada com as exclusões de dados. Observando-se estes histogramas, concluiu-se que este equilíbrio poderia ser melhor atingido estabelecendo-se o FDJ máximo como 0,5.

A Tabela 18 mostra o total de amostras por classe, antes e depois do controle de qualidade, bem como a a percentagem de indivíduos aprovados por classe, a proporção da classe no total de indivíduos aprovados. Além disso, apresentam-se a idade e o gênero médio e o correspondente desvio padrão das amostras.

Tabela 18 – Proporção de sujeitos aprovados no controle de qualidade de fMRI de acordo com a classe dos participantes. Legenda: 0 - sexo masculino 1 - sexo feminino. A idade e o gênero se referem às amostras após o controle de qualidade.

Classe	Total de amostras	Aprov. (amostra)	% de Aprov.	Proporção aprovados	Idade (anos)	Gênero
Controle	142	109	76,8%	17%	10,9±3,2	0,45±0,48
TDAH (TDAH-C +TDAH-I)	667	523	78,4%	83%	10,8±3,3	0,32±0,46
Subtipos de TDAH						
TDAH-C	333	244	73,3%	39%	9,51±2,7	0,27±0,44
TDAH-I	334	279	83,5%	44%	12,0±3,3	0,36±0,48

3.4.2 Pré-processamento

O pré-processamento dos sinais de fMRI foi realizado através do uso do *software Configurable Pipeline for the Analysis of Connectomes (C-PAC)* pela equipe da HBN (O'CONNOR *et al.*, 2017). O pré-processamento tem como objetivo remover ao máximo interferências que possam prejudicar a análise dos dados. Para tanto, realizaram a correção de movimentos dos pacientes, em seguida alinharam as imagens a partir de registros de limite (*boundary-based registration*) (GREVE & FISCHL, 2009).

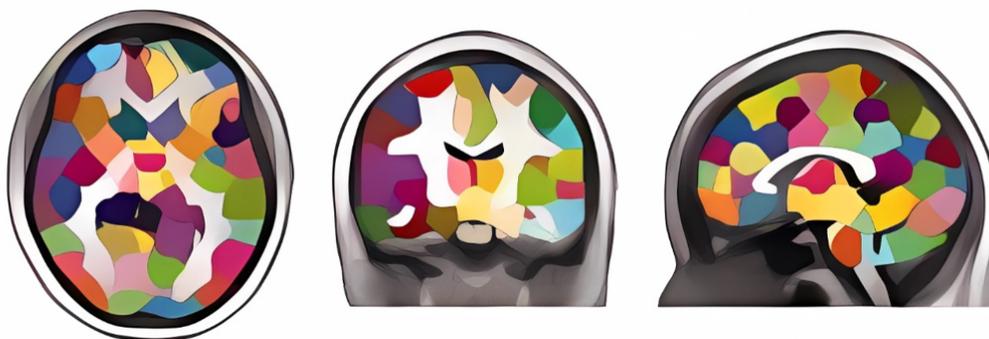
Depois, aplicaram uma série de técnicas de regressão para a redução de ruído, incluindo regressões com polinômios de primeira e de segunda ordem, regressor de modelo de movimento (FRISTON *et al.*, 1996), sinal médio da máscara da matéria branca e sinal médio do líquido cefalorraquidiano. E, então, extraíram as séries temporais correspondentes a cada ROI do atlas de (CRADDOCK *et al.*, 2012) para cada indivíduo.

O pré-processamento dos sinais de fMRI foi realizado através do uso do *software Configurable Pipeline for the Analysis of Connectomes (C-PAC)* pela equipe da HBN (O'CONNOR *et al.*, 2017). O pré-processamento tem como objetivo remover ao máximo interferências que possam prejudicar a análise dos dados. Para tanto, realizaram a correção de movimentos dos pacientes, seguida do alinhamento das imagens para um espaço padrão (MNI152 - (MAZZIOTTA *et al.*, 1995)) usando a técnica de registros de limite (*boundary-based registration*) (GREVE & FISCHL, 2009).

Depois, aplicaram uma série de técnicas de regressão para a redução de ruído, incluindo regressões com polinômios de primeira e de segunda ordem, regressor de modelo de movimento (FRISTON *et al.*, 1996), sinal médio do líquido cefalorraquidiano e substância branca. E, então, extraíram as séries temporais correspondentes a cada ROI do

atlas de (CRADDOCK *et al.*, 2012) para cada indivíduo. A Figura 26 apresenta uma representação a divisão das regiões cerebrais conforme o atlas de Craddock.

Figura 26 – Representação do parcelamento em 200 regiões de interesse conforme (CRADDOCK *et al.*, 2012). Código de cores arbitrariamente escolhidas por para destacar os limites entre cada região.



Fonte: Figura adaptada de Craddock (CRADDOCK *et al.*, 2012) com uma ferramenta de super resolução.

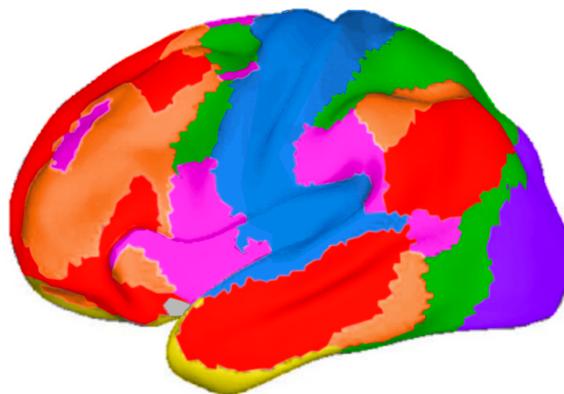
Na seção seguinte, é apresentado o restante do pré-processamento (as etapas chamadas de extração de características), o qual foi executado no âmbito da presente Tese.

3.4.3 Extração de características de fMRI

As medidas tridimensionais de fMRI são agrupadas em 200 ROI. A variação das médias espaciais das medidas eletromagnéticas ao longo do exame em cada ROI forma uma série temporal. O conjunto de séries temporais para cada participante se apresenta em uma matriz, cujas linhas distinguem as ROI e as colunas os segmentos temporais. Cada séries temporal possui 375 pontos amostrados ao longo de 5 minutos, o que corresponde à frequência de amostragem de 1,25 Hz. As dimensões do nível do sinal BOLD destes dados costumam ser pequenas ($-0,072 \pm 20,9$ a. u.), porém podem chegar a valores altos (máximos de ± 300 a. u.) em eventos. Os sinais BOLD costumam ser apresentados em termos de unidades arbitrárias (a. u.).

As médias das ROI ao longo do tempo geram um vetor. Este vetor é, então, reordenado seguindo uma ordem pré-estabelecida para representar as redes de conectividade: visual, somatomotor, de atenção ventral, de atenção dorsal, límbica, frontoparietal *default* e 'outros *voxels*', representadas no mapa de Yeo de 7 segmentos (YEO *et al.*, 2011). Este mapa divide o cérebro em sete regiões, também conhecidas como redes de conectividade, cujas sub-regiões trabalham em conjunto. Ainda, há a área onde se localiza um grupo de pequenas redes, a qual é chamada de 'outros *voxels*'. Estas estão localizadas perto do cerebelo, onde é mais difícil para sensores identificarem o comportamento dos neurônios. A Figura 27 mostra as localizações destas regiões no cérebro.

Figura 27 – Mapa de Yeo de 7 Segmentos. Legenda: Rede Visual (roxo), Rede Somato-motor (azul), Rede de Atenção Dorsal (verde), Rede de Atenção Ventral (Violeta), Rede Límbica (amarelo), Rede Frontoparietal (laranja), Rede de modo padrão (vermelho), demais *voxels* (preto).



Fonte: (YEO *et al.*, 2011)

O vetor de reordenamento é o seguinte: [1, 137, 2, 79, 138, 80, 117, 39, 180, 181, 139, 118, 128, 140, 102, 60, 61, 182, 3, 82, 40, 141, 119, 41, 120, 4, 103, 42, 142, 183, 62, 104, 83, 121, 43, 63, 5, 122, 123, 143, 6, 124, 105, 7, 184, 144, 106, 8, 145, 64, 146, 185, 147, 9, 84, 148, 107, 149, 85, 44, 150, 10, 65, 66, 45, 46, 86, 186, 87, 11, 132, 72, 67, 156, 125, 88, 187, 108, 89, 188, 12, 154, 90, 189, 13, 190, 109, 47, 14, 48, 155, 110, 68, 191, 126, 49, 15, 50, 127, 69, 156, 16, 17, 157, 18, 158, 159, 19, 160, 111, 51, 112, 128, 70, 129, 91, 161, 192, 92, 193, 52, 20, 53, 130, 93, 21, 131, 94, 162, 194, 22, 71, 163, 54, 195, 72, 95, 23, 164, 165, 166, 24, 73, 167, 133, 55, 25, 196, 132, 26, 133, 27, 168, 56, 114, 74, 57, 28, 29, 115, 96, 197, 75, 134, 97, 169, 170, 98, 135, 30, 76, 31, 171, 172, 32, 198, 33, 199, 34, 99, 173, 100, 136, 101, 58, 174, 175, 77, 35, 36, 176, 37, 177, 200, 38, 178, 78, 116, 179, 59]. Os valores dos elementos correspondem às posições das médias das ROI antes do reordenamento e as posições destes valores são as novas posições que as ROI assumem no vetor reorganizado.

Após a reorganização, foi realizada a média de cada série temporal e formou-se um vetor destas médias. Então realizou-se a autocorrelação destas médias para formar uma matriz de correlações conhecida como matriz de conectividade funcional. A posição de cada valor na matriz corresponde a uma posição espacial ou a uma relação entre duas posições espaciais.

Cada matriz é formada por um exame de 5 minutos de um sujeito em repouso, o que corresponde a uma amostra. Estas matrizes foram normalizadas usando o método *z-score* para garantir a igual relevância de características distintas. A função *z-score*, também, foi usada para normalizar matrizes de correlação em (SU *et al.*, 2023), (BU *et al.*, 2020) e (DUFFY *et al.*, 2021).

Observa-se que, nos exames de fMRI, há uma caixa chamada de campo de visão (*field*

of view) que é mantido fixa para todos os participantes. Apenas as partes do cérebro que estiverem neste campo são medidas pelo exame de fMRI. O campo de visão é pré-configurado para obter medições adequadas na maioria dos indivíduos.

Contudo, alguns sujeitos possuem um cérebro um pouco maior, de modo que uma pequena parte de seus cérebros, geralmente, localizada no cerebelo, não é medida no exame. Reconfigurar o campo de visão modificaria a resolução e a distância de onde são adquiridas as imagens. Isto prejudicaria a padronização com as imagens adquiridas anteriormente dos demais sujeitos e poderia gerar *bias* indesejados. Por isso, a HBN não mede as partes do cérebro que ficam fora do campo de visão. Os dados correspondentes a estas ROI são identificados como *Not a Number* (NaN) no pré-processamento no C-PAC.

Uma vez que descartar as amostras onde ocorreram dados NaN reduziria a representatividade dos dados optou-se pela exclusão das ROIs onde este problema ocorre. Por isto, trabalhou-se com matrizes de 190x190. A extração de características nesta Tese foi realizada com o *software* MATLAB versão 2020b.

Embora na ciência de dados se saiba que imagens são matrizes, o oposto nem sempre é verdade. Todas as matrizes de correlação possuem uma diagonal principal onde todos os os valores são '1', uma vez que este valor representa a correlação de um valor com o próprio valor. Os valores acima da diagonal principal são espelhados embaixo desta diagonal. Enviar dados redundantes aos classificadores poderia prejudicar seu desempenho, o que foi observado em testes iniciais. Além disto, a velocidade da execução das CNNs aumenta, consideravelmente, o que reduz o tempo de execução e, assim, viabiliza o uso dos algoritmos genéticos correspondentes.

Por isso, a diagonal principal e o espelhamento inferior foram removidos de cada matriz. Contudo, para evitar a perda da informação das posições dos valores da matriz, estes dados foram substituídos por '0'. Por isso, cada amostra é formada por 17.955 valores de características e 18.145 (17.955+190) zeros.

Caso as posições dos valores nas matrizes fossem alteradas, estas mudanças poderiam alterar a identificação dos locais de cada rede de conectividade e, assim, alterar a interpretabilidade dos resultados e prejudicar a análise de biomarcadores. Por isso, não se utilizou as técnicas de aumento de dados usadas comumente com imagens: rotações, translações e escalonamento de partes das matrizes. Embora a HBN possua uma quantidade de amostras grande para os padrões de fMRI, este número ainda é considerado pequeno na ciência de dados. Portanto, o aumento de dados é desejável. Como as técnicas de transformação, neste caso, poderiam prejudicar o estudo, foi necessário utilizar outro método. Este procedimento consiste em realizar uma média de cada uma das características de cada uma destas 10 matrizes reais selecionadas aleatoriamente (com substituição) para formar uma nova matriz generativa que será uma nova amostra.

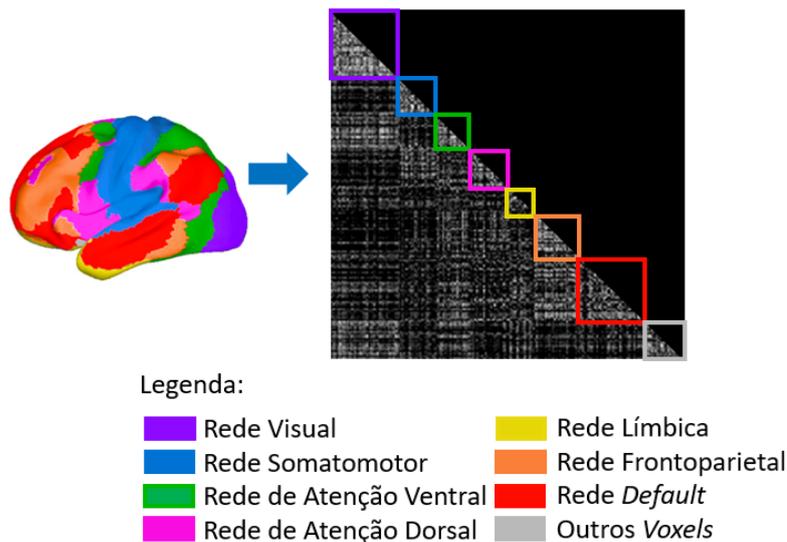
O procedimento de aumento de dados foi realizado para cada classe separadamente, aumentando-se a base de dados em 20%. A seguir, aplicou-se a técnica de sobreamos-

tragem *Synthetic Minority Oversampling Technique* (SMOTE) (CHAWLA *et al.*, 2002) para aumentar o número de amostras das classes minoritárias e assim garantir um balanço igual entre as classes.

A técnica de validação cruzada de *k-fold* estratificado (com $k=10$) foi usada para avaliar os modelos preservando as porcentagens de amostras de cada classe. Evitou-se utilizar um maior número de conjuntos de dados (*folds*) para evitar que estes se tornassem muito homogêneos, prejudicando a generalização dos modelos. O uso da validação cruzada aumenta a generalização dos modelos, e, desta forma, reduz o *overfitting*. Assim como foi feito em EEG, destinou-se o uso de 70% dos dados para o treinamento e 30% para o teste. Os subconjuntos (*folds*) são divisões utilizadas para validação no conjunto de treinamento. O conjunto de teste foi guardado para a avaliação final dos modelos.

Então, selecionou-se a região cerebral a ser analisada na matriz de conectividade. A área correspondente à rede selecionada é recortada da matriz completa e enviada ao classificador. A Figura 28 apresenta a matriz completa de conectividade funcional com as localizações de cada rede de conectividade nesta matriz.

Figura 28 – Redes de conectividade e sua representação na matriz de correlações.



Fonte: O Autor

Na Figura 28, observa-se que as áreas das correlações internas das redes de conectividade estão em torno da diagonal principal. As áreas diretamente à esquerda de determinada rede possuem as correlações entre esta rede e a rede acima na figura. As correlações entre redes foram analisadas nesta Tese somente quando se observa a matriz de correlações completa. Nesta figura, observa-se que cada rede de conectividade possui um tamanho diferente em *voxels*, o qual corresponde as diferentes dimensões de área na matriz de correlações. As dimensões das regiões visual, somatomotor, de atenção dorsal, de atenção ventral, límbica, frontoparietal, *default* e outros *voxels* na matriz de correlações

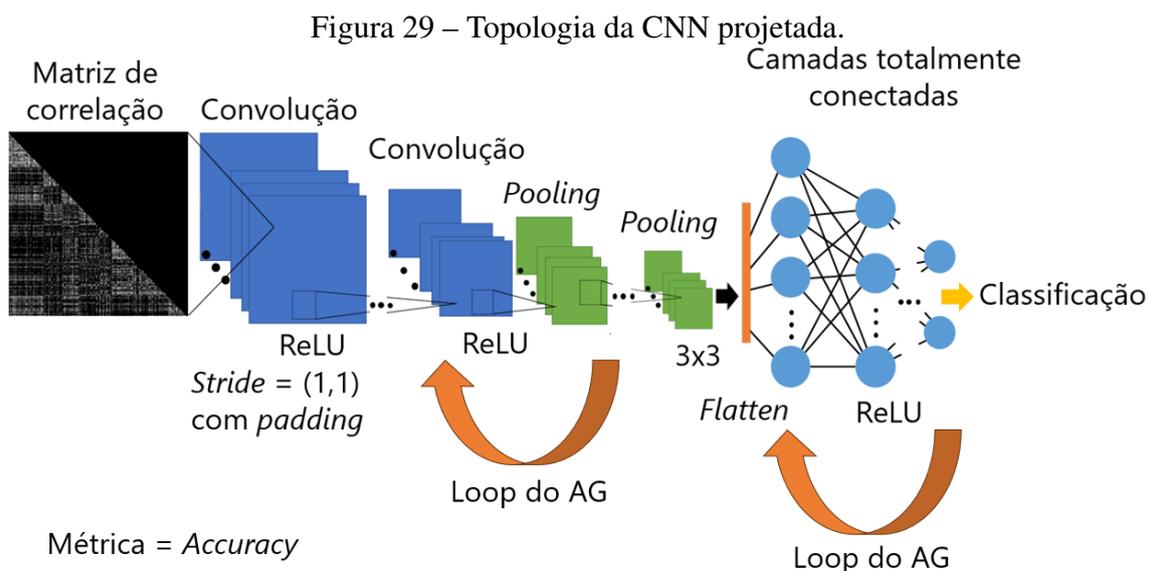
são 38x38, 21x21, 19x19, 23x23, 10x10, 19x19, 43x43 e 17x17, respectivamente.

3.4.4 Processamento de fMRI: Rede Neural Convolutacional (CNN) Projetada

As matrizes de correlação não são imagens e, por isso, têm pouco em comum com as imagens com que costumam ser treinadas os modelos clássicos. Portanto, decidiu-se por projetar uma CNN, especialmente, para trabalhar com estas matrizes. A CNN foi, então, otimizada, independentemente, para cada matriz de entrada (matriz completa ou cada rede de conectividade) para cada diagnóstico (TDAH/Controle ou TDAH-C/TDAH-I). A otimização foi realizada com um algoritmo genético (AG) implementado no *Python* v. 3.6. através da biblioteca *deap*. Para programar a CNN, utilizou-se a biblioteca *keras*.

Na topologia da CNN projetada, aplica-se uma convolução à matriz de características da entrada com diversos filtros (ou *kernels*). Este processo gera um mapa de características para cada filtro. Então, aplica-se uma sequência de convoluções intercaladas por camadas de *pooling* a estes mapas. A quantidade de camadas nesta sequência varia de acordo com a matriz de entrada (matriz completa ou cada rede de conectividade) e foi definida através do AG.

A Figura 29 apresenta a topologia da CNN projetada. Todos os AGs indicaram que as topologias obtêm melhores *accuracies* quando a normalização de *batch* não é realizada independentemente de qual for a matriz de entrada e se a classificação é TDAH/Controle ou TDAH-C/TDAH-I. Por este motivo, a normalização de *batch* não foi representada na Figura 29.



Fonte: O Autor

Após a sequência de convoluções, experimentou-se a aplicação de uma normalização de *batch*. Esta normalização é seguida de uma camada de *pooling* com janela 3x3. Então, as matrizes com os mapas de características são reorganizadas na forma de um

vetor através do processo de *flatten*. Este vetor é enviado aos neurônios da camada de entrada da RNA. As quantidades de neurônios e de camadas ocultas desta RNA foram otimizadas através do AG para cada matriz de entrada em cada um dos tipos de diagnóstico (TDAH/Controle ou TDAH-C/TDAH-I). Para a camada de saída utilizou-se a função de ativação *softmax* e a função custo de *binary crossentropy*.

Os parâmetros que foram testados pelos AGs para as CNN projetadas foram o tamanho de *batch*; a quantidade de filtros convolucionais; o tamanho das janelas de convolução; o tamanho das janelas de *pooling*; a quantidade de neurônios por camada totalmente conectada (as camadas de neurônios convencionais após a camada de *flatten*); o *drop out* das camadas neuronais; o otimizador; a quantidade de camadas convolucionais; a quantidade de camadas ocultas na RNA; a taxa de regularização L1; e o uso da normalização de *batch*. O *drop out* foi otimizado para todas as camadas simultaneamente. Em trabalhos futuros pode-se melhorar os classificadores otimizando-se o *drop out* de forma separada para cada camada.

Para a classificação de subtipos, foi necessário também otimizar a taxa de aprendizado inicial e o número de épocas em função da maior complexidade desta tarefa. A classificação entre os subtipos é mais difícil porque há maior similaridade no funcionamento dos cérebros das classes TDAH-C e TDAH-I do que entre os grupos TDAH e Controle.

Para a classificação entre pacientes com TDAH e voluntários do grupo controle, as épocas e a taxa de aprendizado inicial foram fixadas em 100 e 10^{-3} , respectivamente. Embora a taxa de aprendizado seja um parâmetro importante para as RNAs, os otimizadores Estimativa de Momento Adaptativo (do inglês, *Adaptive Moment Estimation* - Adam) e Delta Adaptativo (do inglês, *Adaptive Delta* - Adadelta) empregam métodos adaptativos para ajustar esta taxa ao longo do treinamento. Configurou-se o número de neurônios para ser reduzido pela metade para cada nova camada. Nas convoluções, utilizou-se um *stride* com janela 1x1 e *padding*.

O AG possui seus próprios hiperparâmetros, os quais foram selecionados através de testes manuais extensivos. Estes hiperparâmetros são: tamanho da população (150), probabilidade de cruzamento (0,5), probabilidade de mutação (0,0001) e número de gerações (até 20). Cada AG foi executado 10 vezes para garantir que as configurações ótimas foram obtidas. Todas as avaliações dos AGs foram realizadas com base no critério da *accuracy* no conjunto de teste. Os AGs indicaram configurações ótimas das CNNs para cada rede de conectividade (ou para a matriz completa) e para cada classificação (entre classes TDAH e Controle ou entre classes TDAH-C e TDAH-I).

Os ensaios mostraram que CNNs com menores quantidades de camadas de convoluções e de *pooling* obtiveram melhores resultados (*accuracies*) em classificações de TDAH (tanto para distinguir pacientes com TDAH e indivíduos do grupo controle como entre os subtipos de TDAH). Isto possivelmente ocorreu porque a RNA precisa de informações que um maior número de convoluções esconde.

A Tabela 19 apresenta os intervalos nos quais os hiperparâmetros foram testados pelo AG para a CNN. Os intervalos foram selecionados de acordo com testes manuais extensivos, os quais foram realizados de forma preliminar. Parte dos intervalos de pesquisa do AG são distintos quando usa-se o mapa cerebral completo (matriz completa) e quando emprega-se individualmente as redes de conectividade que compõem o mapa. Isto foi realizado desta forma porque as redes de conectividade possuem dimensões menores do que a matriz completa e porque, através de experimentos, constatou-se que as redes de conectividade necessitam de uma investigação mais profunda de seus hiperparâmetros.

Tabela 19 – Intervalos de hiperparâmetros testados pelo AG para a CNN projetada.

Hiperparâmetro	Matriz Completa	Redes de Conectividade
Tamanho de <i>Batch</i>	2 a 36	2 a 36
Quantidade de Filtros Convolucionais	8 a 48	8 a 100
Dimensão da Janela da Convolução	1x1 a 3x3	1x1 a 3x3
Dimensão da Janela de <i>Pooling</i>	2x2 a 3x3	2x2 a 3x3
Neurônios por camada	128 a 512	64 a 1024
<i>Drop out</i>	0,15 a 0,95	0,15 a 0,95
Optimizador	Adadelta, SGD ou Adam	Adadelta, SGD ou Adam
Camadas Convolucionais	1 a 7	1 a 7
Camadas Ocultas	1 a 7	1 a 7
L1 ($\times 10^{-5}$)	0.00001 a 0.01	0.00001 a 0.01
Normalização de <i>Batch</i>	Sim ou Não	Sim ou Não
Taxa de aprendizado inicial	10^{-3}	10^{-5} a 10^{-2}
Épocas	100	20 a 200

3.4.4.1 Hiperparâmetros determinados para as CNNs usadas na distinção das classes TDAH e Controle

Para a classificação entre indivíduos das classes TDAH e Controle, os AGs preferiram o uso de apenas uma camada convolucional, além da camada convolucional inicial (apresentada, anteriormente, na Figura 29).

As dimensões das janelas das convoluções e de *pooling* foram selecionadas como 3x3 e 1x1, respectivamente, para todas as redes de conectividade. Para a rede de atenção dorsal, a dimensão da janela convolucional é 2x2. Porém, para o caso da matriz completa, o AG preferiu reduzir a dimensão da janela da camada convolucional para 1x1. Portanto, esta se tornou uma convolução do tipo *pointwise* (LIN *et al.*, 2014), na qual apenas se aplica a função de ativação unidade linear retificada (do inglês, *rectified linear unit - ReLU*) e se multiplica cada ponto da matriz pelo valor do filtro convolucional. Desta forma, geram-se novos mapas convolucionais com a mesma dimensão dos anteriores da convolução *pointwise*. Para compensar a menor janela na segunda camada convolucional, o AG aumentou as dimensões da janela de *pooling* para 2x2.

Para o otimizador, na classificação entre pacientes com TDAH e participantes da classe Controle, os AGs optaram pela Descida do Gradiente Estocástico (do inglês, *Stochastic gradient descent*) (SGD) na maioria dos casos (redes de atenção dorsal e ventral, límbica, frontoparietal e outros *voxels*). Já para as redes visual e *default*, o otimizador selecionado foi o Adam. Na matriz completa, o AG preferiu o uso do Adadelta, combinado ao uso do maior *dropout* e um menor número de neurônios por camada. Os *dropouts* selecionados para as redes de conectividade são relativamente baixos (todos inferiores a 35%), o que indica baixos *overfittings*. As CNNs que usaram o SGD precisaram de uma maior taxa de regularização L1. Nos casos em que o otimizador escolhido foi o Adam, o AG optou também pelo uso de duas camadas ocultas.

Uma vez que as redes visual e *default* são maiores, estas correspondem a matrizes de entrada com maiores dimensões na CNN. Ou seja, a tarefa de extrair as características mais relevantes das matrizes das redes visual e *default* podem exigir o uso de uma rede com mais camadas que operam melhor com um otimizador mais complexo (Adam). No entanto, para a rede de atenção ventral, o AG optou por usar o SGD em conjunto com três camadas ocultas.

Na distinção entre indivíduos da classe TDAH e do grupo controle, o AG decidiu pelo uso de 24 filtros por camada convolucional para quatro redes de conectividade (de atenção ventral, frontoparietal, *default* e outros *voxels*). Esta quantidade de filtros foi dobrada para a rede límbica e quase dobrada para rede de atenção dorsal. O número de filtros convolucionais foi 13 Para a matriz completa e apenas 3 para a rede visual.

O maior tamanho de *Batch*, 103, foi selecionado para a rede de atenção dorsal que é uma das redes de menor dimensão. Já a matriz completa, cujas dimensões (190x190) são a soma das dimensões de todas as redes de conectividade, corresponderam ao uso

do menor tamanho de *Batch*, 3. Apesar, de sugerir uma proporcionalidade inversa, isto não se confirma em outros casos, uma vez que a maior das redes é a *default*, cujo tamanho de *Batch* é de 60 valores. Os hiperparâmetros otimizados para TDAH/Controle são apresentados na Tabela 20.

Tabela 20 – Hiperparâmetros da CNN selecionados por AG para classificação entre participantes dos grupos TDAH e controle.

Param.	Matriz Completa	Visual	Soma.	Dorsal	Ventral	Límb.	Fronto.	<i>Default</i>	Outros Voxels
Tamanho de <i>Batch</i>	3	27	21	103	43	28	43	60	43
Filtros Conv.	13	3	79	47	24	48	24	24	24
Dim. Janelas Conv.	1x1	3x3	1x1	2x2	3x3	3x3	3x3	3x3	3x3
Dim. Janelas <i>Pooling</i>	2x2	1x1	2x2	1x1	1x1	1x1	1x1	1x1	1x1
Neur. por Camada	152	162	452	582	288	288	582	288	582
<i>Dropout</i>	0,592	0,214	0,396	0,283	0,342	0,342	0,283	0,342	0,283
Otimiz.	Adadelta	Adam	SGD	SGD	SGD	SGD	SGD	Adam	SGD
Camadas Conv.	1	1	2	1	1	1	1	1	1
Camadas Ocultas	1	2	3	1	3	1	1	2	1
L1 (10^{-5})	62,0	5,00	62,1	484	303	140	298	180	487

3.4.4.2 Hiperparâmetros determinados para as CNNs usadas na distinção das classes TDAH-C e TDAH-I

Para a classificação entre participantes dos subtipos TDAH-C e TDAH-I, os AGs inferiram a necessidade de um maior número de camadas convolucionais do que para a classificação entre indivíduos com TDAH e do grupo controle. As camadas de *pooling* também possuem janelas com maiores dimensões. Na maioria dos casos, o tamanho de *batch* foi menor e o número de filtros convolucionais foi maior para a classificação de subtipos, especialmente, para as redes visual, somatomotor, frontoparietal e *default*. O aumento do número das camadas convolucionais, camadas ocultas, dimensão das camadas de *pooling* e filtros convolucionais ocorreu como consequência da maior dificuldade inerente à classificação de subtipos.

Para todos os casos de classificação de subtipos, os AGs apontaram o uso de janelas convolucionais e janelas de *pooling* com dimensões 3x3 e 2x2, respectivamente, com exceção da rede de atenção ventral. Nesta rede, o AG indicou o uso de uma convolução *pointwise* e a realização da redução de dimensionalidade através de uma camada de *pooling* em janelas 3x3.

Os AGs apontaram que a maioria das CNNs melhoraram seu desempenho quando o número de épocas é menor do que 70, com exceção das CNNs correspondentes às redes límbicas e frontoparietais, cujo desempenho ótimo ocorreu em 150 e 100 épocas, respectivamente.

Em geral, os *dropouts* indicados pelos AGs foram baixos, menores que 37% para todos os casos menos para a área dos outros *voxels*, a qual precisou de um *dropout* de 47,6%. O menor *dropout* foi apontado para CNN da rede de atenção dorsal. A CNN com o maior *dropout* foi também aquela configurada com a menor taxa de regularização L1 ($4,6 \times 10^{-5}$) e um dos menores números de épocas (40). Isto pode representar uma tentativa dos AGs de encontrar um equilíbrio entre estes dois hiperparâmetros. Isto, porque tanto a taxa de regularização como o *dropout* são ferramentas para reduzir a ocorrência de *overfitting* e aumentar a capacidade de generalização do modelo, assim como a redução do número de épocas.

As CNNs cujos AGs indicaram o uso de uma maior quantidade de camadas ocultas (3) foram também aquelas para as quais os AGs recomendaram o uso dos otimizadores Adam ou Adadelta. No entanto, a quantidade de três camadas ocultas também foi requisitada para a matriz completa, mesmo os AGs tendo apontado o uso do SGD como otimizador neste caso.

A Tabela 21 mostra os hiperparâmetros selecionados para as CNN na classificação entre indivíduos das classe TDAH-C e TDAH-I.

Tabela 21 – Hiperparâmetros da CNN selecionados por AG para classificação entre participantes das classe TDAH-C e TDAH-I.

Param.	Matriz Completa	Visual	Soma.	Dorsal	Ventral	Límb.	Fronto	Default	Outros Voxels
Tamanho de <i>Batch</i>	29	21	30	23	10	23	27	27	12
Filtros Conv.	28	84	88	59	64	59	79	88	27
Dim. Janelas Conv.	3x3	3x3	3x3	3x3	1x1	3x3	3x3	3x3	3x3
Dim. Janelas <i>Pooling</i>	2x2	2x2	2x2	2x2	3x3	2x2	2x2	2x2	2x2
Neur. por Camada	120	281	433	218	184	296	64	281	346
<i>Dropout</i>	0,361	0,242	0,367	0,161	0,367	0,220	0,248	0,248	0,476
Otimiz.	SGD	SGD	SGD	Adam	SGD	SGD	Adam	Ada-delta	Ada-delta
Camadas Conv.	2	5	5	2	5	2	2	2	7
Camadas Ocultas	3	1	1	3	1	1	3	3	3
L1 (10^{-5})	229	1384	1385	69	1385	69	62	69	4,6
Épocas	35	47	65	40	69	150	100	50	40
Taxa de Aprendiz. ($x10^{-4}$)	3	40	70	0,2	0,147	10	30	30	70

3.4.5 Processamento de fMRI: Modelos Clássicos

Para fins de comparação de resultados obtidos e confirmação dos biomarcadores, utilizaram-se modelos clássicos usualmente empregados para a análise de imagens. As topologias selecionadas para este fim foram: LeNet-5, ResNet-50 e VGG-16. Experimentou-se o uso de AGs para otimizar alguns hiperparâmetros destas redes.

Os hiperparâmetros configurados através de AG e os respectivos intervalos buscados para a LeNet-5 são: Tamanho de *batch* (entre 12 e 20), filtros da primeira camada convolucional (entre 2 e 12), filtros da segunda camada convolucional (entre 12 e 24), janela do filtro da primeira camada convolucional (entre 1 e 5), janela do filtro da segunda camada convolucional (entre 1 e 5), neurônios na primeira camada (de 60 a 240), neurônios na segunda camada (de 32 a 168) e otimizador (Adam, SGD ou Adadelta).

Como função de ativação, experimentaram-se a ReLU (FUKUSHIMA, 1969), a tangente hiperbólica (*tanh*), a unidade linear exponencial (do inglês, *exponential linear unit* - ELU) (CLEVERT *et al.*, 2015) e a unidade linear exponencial escalonada (do inglês, *scaled exponential linear unit* - SeLU) (KLAMBAUER *et al.*, 2017). Este AG possui população de 200 combinações genéticas, probabilidade de cruzamento de 0,50, probabilidade de mutação de 0,002 e 10 gerações, parando antecipadamente caso o valor mínimo de aptidão não fosse alterado por 5 gerações. Utilizou-se o operador de cruzamento de dois pontos (*two point*) e o operador de seleção baseado em um torneio entre 10 indivíduos.

Para todas as redes de conectividade, o AG designou o otimizador Adam para trabalhar com a LeNet-5. Entretanto o AG preferiu o uso do otimizador Adadelta para a LeNet que buscou discrepâncias entre as características com a matriz de correlações funcionais completa. A função de ativação ELU foi selecionada para as entradas com maior número de características: matriz completa e rede *default*. Para estas regiões, também, o AG selecionou o uso de um maior número de filtros convolucionais da segunda camada 22 e 21, respectivamente. As regiões que usaram a ELU precisaram de poucos neurônios na segunda camada densa, conforme o AG apontou. Já, as regiões nas quais a ReLU foi empregada necessitaram do uso de uma maior quantidade de neurônios.

O tamanho de *batch* selecionado está mais próximo do limite superior pesquisado em 5 das redes de conectividade, e próximo ao limite inferior e para a matriz completa e para a rede visual.

Na classificação TDAH/Controle com a LeNet-5, o número de filtros da primeira camada foi selecionado como sendo 9 para as regiões dorsal, ventral, límbica e 'outros voxels'. Estas foram as mesmas regiões para as quais o AG decidiu por empregar a função SeLU.

Para as dimensões das janelas dos filtros convolucionais, o AG selecionou apenas valores 2x2 ou 3x3 para trabalhar com todas as redes de entrada, com exceção das redes frontoparietal e *default*. Nestas regiões, uma janela de convolução 1x1 foi designada para

as camadas convolucionais 1 e 2, respectivamente. A Tabela 22 apresenta os hiperparâmetros selecionados para a CNN do tipo LeNet-5, através do AG descrito nesta seção para a classificação entre os grupos TDAH e Controle.

Tabela 22 – Hiperparâmetros da CNN LeNET-5 selecionados por AG para classificação TDAH/Controle.

Param.	Matriz Completa	Visual	Soma.	Dorsal	Ventral	Límb.	Fronto	<i>Default</i>	Outros Voxels
Tamanho de <i>Batch</i>	13	12	19	15	18	19	20	19	16
Filtros conv. 1	5	7	8	9	9	9	6	7	9
Filtros conv. 2	22	22	13	15	15	15	16	21	17
Janelas conv. 1	2x2	3x3	2x2	3x3	3x3	3x3	1x1	2x2	2x2
Janelas conv. 2	3x3	3x3	2x2	2x2	3x3	3x3	2x2	1x1	2x2
Neur. cam. 1	98	74	222	187	206	194	136	182	79
Neur. cam. 2	52	148	156	155	43	46	100	69	115
Função ativação	ELU	ReLU	ReLU	SeLU	SeLU	SeLU	ReLU	ELU	SeLU
Otimiz.	Adadelta	Adam	Adam	Adam	Adam	Adam	Adam	Adam	Adam

A Tabela 23 mostra os hiperparâmetros selecionados pelo AG para a CNN do tipo LeNet-5, considerando a classificação entre os grupos TDAH-C e TDAH-I. O AG procedeu da mesma forma para a classificação de subtipos em relação à classificação TDAH/Controle no que diz respeito as escolhas de otimizadores. Para os tamanhos de *Batch* houve a tendência a de o AG preferir o uso de um tamanho reduzido para a classificação de subtipos em comparação com a classificação TDAH/Controle para as redes para as redes dorsal, ventral, límbica, frontoparietal e e *default*.

Para o diagnóstico de subtipos, novamente o AG apontou o uso do otimizador Adadelta somente para a matriz completa e o Adam para as redes de conectividade. Como a matriz completa possui uma quantidade maior de características, o processamento desta

é mais complexo. Um processamento mais complexo requer um treinamento mais longo, não apenas em tempo real como em épocas. Este aspecto foi agravado pela restrição da arquitetura LeNET-5 ter apenas 2 camadas convolucionais, pois a quantidade de características que são reduzidas pelas camadas convolucionais é menor quando há mais características na entrada da CNN. O otimizador Adam precisa de mais tempo de treinamento para convergir do que o Adadelta (YAZAN & TALU, 2017). Como, neste experimento, o número de épocas foi fixado, o uso do Adam na LeNet-5, possivelmente, foi preterido por requerer um maior número de épocas para o treinamento convergir. A Tabela mostra os hiperparâmetros selecionados na arquitetura LeNet-5 pelo AG na classificação de subtipos.

Tabela 23 – Hiperparâmetros da CNN LeNET-5 selecionados por AG para classificação TDAH-C/TDAH-I.

Param.	Matriz Completa	Visual	Soma.	Dorsal	Ventral	Límb.	Fronto	Default	Outros Voxels
Tamanho de <i>Batch</i>	19	14	20	14	14	13	12	13	20
Filtros conv. 1	8	9	7	9	9	4	6	4	4
Filtros conv. 2	17	14	13	20	14	22	19	17	14
Janelas conv. 1	1x1	3x3	2x2	2x2	3x3	2x2	3x3	3x3	2x2
Janelas conv. 2	4x4	2x2	3x3	2x2	3x3	2x2	3x3	2x2	2x2
Neur. cam. 1	143	186	157	136	66	204	152	162	74
Neur. cam. 2	65	145	164	75	158	115	158	58	148
Função ativação	tanh	SeLU	SeLU	SeLU	ELU	SeLU	tanh	SeLU	ReLU
Otimiz.	Adadelta	Adam	Adam	Adam	Adam	Adam	Adam	Adam	Adam

Na classificação de subtipos, bem como ocorreu no diagnóstico TDAH/Controle, a função de ativação escolhida mais frequentemente pelo AG foi a SeLU. A SeLU possui como vantagem a característica de normalizar a saída da camada de RNA associada a esta

função de ativação, desde que a entrada desta camada também tenha sido normalizada (mantendo a média dos valores de saída dos neurônios em zero e sua variância unitária). Conseqüentemente, a entrada da camada posterior estará normalizada. Isto evita os problemas da explosão e desaparecimento do gradiente (KLAMBAUER *et al.*, 2017). A ReLU também evita o desaparecimento do gradiente, porém, para as ativações menores que zero, o gradiente da ReLU será nulo e os pesos correspondentes não são ajustados ao longo da descida. Isso provoca nos neurônios artificiais um estado de falta de resposta a variações na entrada (*dying ReLU problem*).

Para a VGG-16, utilizou-se um AG com os mesmos hiperparâmetros de probabilidade de cruzamento, probabilidade de mutação, gerações, parada antecipada, operador de cruzamento e o operador de seleção que o AG usado para a LeNet-5. Porém, o tamanho da população foi de 100 combinações genéticas. Os hiperparâmetros investigados para a VGG-16 foram o tamanho de *batch* (entre 2 e 20), camadas congeladas (de 0 a 15), taxa de aprendizado (de 0,00001 a 0,01), taxa do regularizador L1 (de 0,0001 a 0,02), *dropout* (de 0,15 a 0,85) e épocas (de 15 a 50). Como otimizador utilizou-se o Adam.

Na ResNet-50 buscou-se com AG encontrar a combinação ótima dos seguintes hiperparâmetros: Tamanho de *batch* (de 10 a 46), camadas congeladas (de 0 a 49), taxa de aprendizado (de 0,00001 a 0,001), taxa do regularizador L1 (de 0,0001 a 0,01), *dropout* (de 0,15 a 0,95) e épocas (de 50 a 200). Utilizou-se as mesmas configurações de hiperparâmetros do AG da VGG-16. Como otimizador utilizou-se o Adam.

A técnica de aprendizado por transferência consiste em se aproveitar características aprendidas por arquiteturas de RNAs durante um treinamento em uma base de dados (normalmente muito grande) para outra base de dados, usualmente, menor. Contudo, os treinamentos prévios disponibilizados na biblioteca *keras* para estas redes se baseiam na base *imagenet*, na qual as arquiteturas foram treinadas para classificar imagens de objetos como carros, aviões, pessoas, sapatos, etc. Estas imagens não tem atributos em comum com as matrizes de correlação investigadas na presente Tese.

A Tabela 24 apresenta os hiperparâmetros selecionados para as CNNs do tipo VGG-16 e ResNet-50 com o AG para as classificações de TDAH/Controle e de TDAH-C/TDAH-I a partir da matriz de correlações funcionais completa. Considerou-se o uso destas redes utilizando a técnica de aprendizado por transferência. Observou-se que com as VGG-16, os AG optaram por congelar todas (16) ou quase todas (14) as camadas. Para a maioria dos casos, o tamanho de *batch* foi escolhido em torno de 32. A ResNet-50 resultante do AG apresentou hiperparâmetros similares para o tamanho de *batch*, camadas congeladas, taxa de aprendizado e *dropout* para as classificações TDAH/Controle e TDAH-C/TDAH-I. Em ambos os cenários, o valor do *dropout* designado para a ResNet-50 foi maior do que o da VGG-16, o que pode estar relacionado à VGG-16 possuir uma menor profundidade do que a ResNet-50.

Tabela 24 – Hiperparâmetros das CNNs VGG-16 e ResNet-50, ambas com aprendizado por transferência selecionados por AG usando a matriz de correlações.

Hiperparam.	VGG-16 TDAH/ Con- trole	VGG-16 TDAH-C/ TDAH-I	ResNet-50 TDAH/ Con- trole	ResNet-50 TDAH-C/ TDAH-I
Tamanho de <i>Batch</i>	6	33	30	34
Camadas congeladas	14	16	39	28
Taxa de aprendizado (10^{-3})	0,861	4,415	1,609	1,470
Taxa reg. L1 (10^{-2})	1,41	0,19	0,68	1,20
<i>Dropout</i>	0,4678	0,6513	0,8013	0,8123
Épocas	37	96	172	103

Também experimentou-se o uso das topologias originais ResNet-50 e VGG-16 sem otimização de hiperparâmetros, sem o uso de aprendizagem por transferência e com o treinamento de todas as suas camadas (sem utilizar o recurso de congelamento de camadas).

Os resultados referentes às classificações de TDAH e de seus subtipos através da análise de sinais de fMRI com os modelos clássicos encontram-se no APÊNDICE E. Por conveniência, para a comodidade do leitor, resumem-se nesta seção tais resultados.

Utilizando as topologias originais, sem aprendizado por transferência e sem o congelamento de camadas, o desempenho da ResNet-50, LeNet-5 e da VGG-16 foi insatisfatório mesmo na classificação binária. Nestas condições, a ResNet-50 apresentou *accuracies* superiores as demais, embora consideravelmente abaixo daquelas obtidas pela CNN projetada. As *accuracies* obtidas com a ResNet-50 foram de 100% para o treinamento e de 68,8% para o teste, o que caracteriza a ocorrência de *overfitting*.

Com a otimização das topologias dos modelos clássicos através de AG e o uso de aprendizado por transferência, as *accuracies* de todas as redes foram aprimoradas. Desta forma, a *accuracy* da ResNet-50 passou a ser de $71,0 \pm 2,2\%$ e a da LeNet-5 de $78,0 \pm 2,3\%$. Embora tenha ocorrido uma evolução também para o desempenho da VGG-16, esta foi pouco relevante, já que sua *accuracy* foi de menos de 52%.

Para a classificação de subtipos, as 3 arquiteturas apresentaram desempenho semelhante entre si. LeNet-5, VGG-16 e ResNet-50 obtiveram *accuracies* de $71,8 \pm 9,8\%$,

68,8±6,5% e 69,6±6,1%. Devido ao melhor desempenho da LeNet-5 nestas condições, esta e a CNN projetada foram utilizadas para avaliar as redes de conectividade. A LeNet-5 obteve *accuracies* superiores a 85% na classificação binária avaliando, independentemente, quaisquer das redes de conectividade. O melhor desempenho da LeNet-5 foi obtido analisando as redes *Default*, Somatomotor, Visual e de Atenção Dorsal: 94,8±2,5%, 92,2±1,7%, 92,1±3,5% e 91,4%, respectivamente.

Na avaliação de subtipos, nota-se uma queda considerável no desempenho da LeNet-5 (assim como ocorreu com a CNN projetada). As melhores *accuracies* da LeNet-5 na classificação de subtipos foram de 70,0±4,9%, 69,5±5,0% e 67,5±6,5% para as redes de Atenção Dorsal, Visual e *Default*, respectivamente.

3.4.6 Processamento de fMRI: Fusão de Classificações a partir das Redes de Conectividade de fMRI

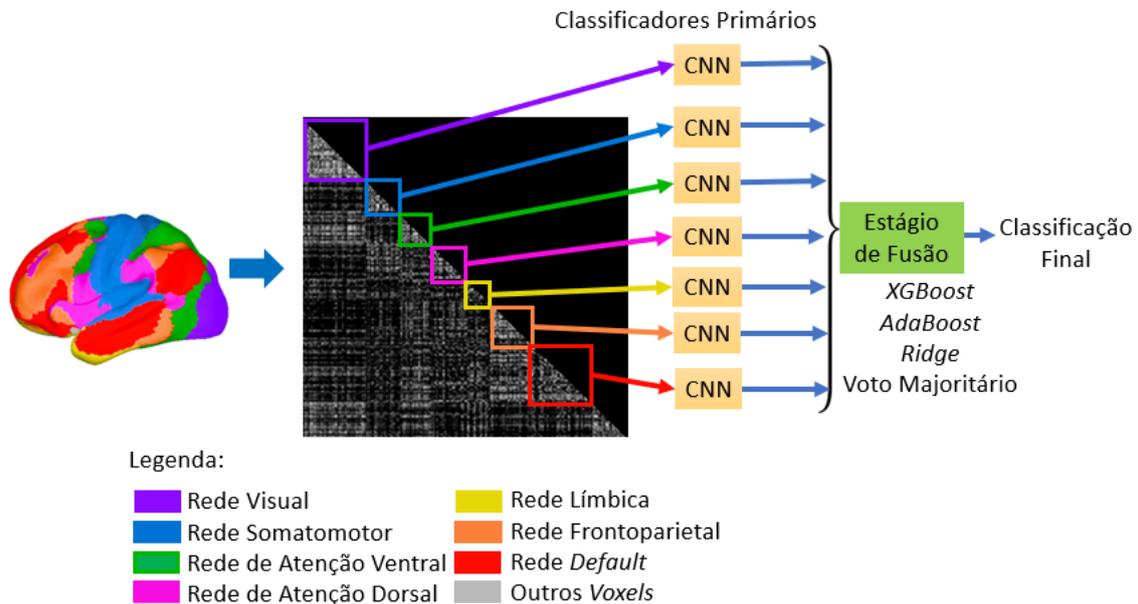
As arquiteturas CNN projetadas com as configurações descritas na seção 3.4.4 foram utilizadas como classificadores primários (também chamados de classificadores-base) em modelos *ensembles*. A fusão das redes de conectividade é realizada de forma análoga à fusão de ritmos de EEG descrita na seção 3.3.4.

Como descrito na seção 3.4.4, cada rede de conectividade (visual, somatomotor, de atenção dorsal, de atenção ventral, límbica, frontoparietal, e *default*) é a entrada de uma CNN configurada, especialmente, para lidar com os dados de cada uma destas redes. A área da matriz de conectividade conhecida como outros *voxels* foi excluída do processo por apresentar *accuracies* individuais consideravelmente menores do que as demais. Os classificadores tendem a apresentar piores desempenhos nesta área porque ela foi formada por correlações entre médias de *voxels* que não costumam operar em conjunto, tipicamente resultando em menores correlações.

O uso de meta-classificadores tem o propósito de possibilitar a obtenção de um classificador capaz de obter *accuracies* mais elevadas e de realizar classificações mais robustas. A combinação dos classificadores base pode endossar diagnósticos com mais confiança do que um único classificador, pois constatou indícios do pertencimento de um determinado indivíduo a uma classe em mais de uma região cerebral. Mesmo a CNN com dados da rede de conectividade que produz classificações com as maiores *accuracies* pode, eventualmente, gerar erros, os quais podem ser corrigidos comparando suas previsões com outras CNNs.

As classificações de cada sujeito por cada uma das CNNs (uma para cada rede de conectividade) são enviadas como características para o estágio de fusão. O estágio de fusão utiliza um modelo de meta-classificador para gerar uma classificação final. O meta-classificador atribui pesos para as classificações iniciais obtidas por cada CNN. A Figura 30 apresenta o método de fusão utilizado para fMRI.

Figura 30 – Fluxograma da fusão dos classificadores baseados nas redes de conectividade de fMRI com redes neurais convolucionais (CNN) como classificadores primários.



Fonte: O Autor

Nesta Tese, para o estágio de fusão, experimentou-se os métodos *XGBoost*, *AdaBoost*, *Ridge* e voto majoritário, previamente descritos na seção 3.3.4.

Para o Voto Majoritário, os pesos das predições com base em cada rede de conectividade foram iguais às *accuracies* obtidas pelas CNNs correspondentes quando estas trabalharam individualmente.

Uma vez que o classificador linear *Ridge* possui apenas um único hiperparâmetro (α), este foi otimizado por buscas manuais extensivas. Seus valores são 617 e 0,0001 para as classificações TDAH/Controle e TDAH-C/TDAH-I, respectivamente.

Para selecionar os hiperparâmetros dos métodos *AdaBoost* e *XGBoost*, empregou-se AGs. Os hiperparâmetros dos AGs foram configurados como: tamanho da população = 200, probabilidade cruzamento = 0,50, probabilidade de mutação = 0,0001 e número de gerações = 15. Cada AG foi executado 10 vezes na busca pela solução ótima.

Para o *AdaBoost*, o AG investigou o número de estimadores no intervalo entre 1 e 10 mil e a taxa de aprendizagem entre 0,00001 e mil. Utilizou-se aleatorização uniforme para a geração das populações e de mutações. Para TDAH/Controle, o AG encontrou a seguinte configuração ótima: 561 estimadores-base e taxa de aprendizado de 5,22. Em TDAH-C/TDAH-I, o AG indicou o uso de 6910 classificadores primários e uma taxa de aprendizado de 1,33. Tanto para o *AdaBoost* como para o *XGBoost*, os AGs selecionaram uma maior quantidade de estimadores-base para a classificação de subtipos. Possivelmente, isto ocorre porque os classificadores primários obtiveram *accuracies* maiores para a classificação TDAH/Controle, facilitando o trabalho dos meta-classificadores.

A busca por AG foi limitada ao uso do algoritmo *gradient boosting tree* (gbtree) descrito, anteriormente, na seção 3.3.4. Utilizou-se a aleatorização uniforme para a geração das populações e mutações de todos os hiperparâmetros da busca. Os hiperparâmetros do *XGBoost* são apresentados na Tabela 25.

Tabela 25 – Intervalos investigados por AG para otimizar o desempenho do meta-classificador *XGBoost* baseado em fMRI.

Hiperparâmetros	Intervalo	Hiperparâmetros selecionado para TDAH/Controle	Hiperparâmetros selecionado para TDAH-C/TDAH-I
Número de Estimadores	1 a 10.000	2047	9367
η	0,0001 a 10.000	1,57	0,144
γ	0,0001 a 10.000	6,05	4,128
Máxima profundidade das árvores	1 a 100	5	6
<i>Minimum child weight</i>	1 a 100	5	1
Taxa de L1	0,0001 a 200	3	1
Taxa de L2	0,0001 a 200	10	5

4 RESULTADOS

Neste capítulo, apresentam-se os resultados obtidos através da metodologia proposta. Dividiu-se o capítulo em duas seções referentes à classificação de TDAH a partir de sinais de EEG e de fMRI. Em ambas as seções, primeiro apresentam-se os resultados dos classificadores primários e, depois, os resultados obtidos através do uso dos meta-classificadores.

Os resultados dos métodos de comparação dos classificadores primários para EEG e para fMRI encontram-se, respectivamente, no APÊNDICE D e no APÊNDICE E. As discussões são apresentadas no Capítulo 5.

A principal métrica utilizada é a *accuracy* que é definida como a relação expressa na Equação 29.

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN}, \quad (29)$$

onde VP são os verdadeiros positivos, VN os verdadeiros negativos, FP os falsos positivos e FN os falsos negativos. Como medida das incertezas, adotou-se o desvio padrão.

4.1 Resultados da Classificação de TDAH baseada em EEG

Os resultados da classificação de TDAH a partir de EEG são apresentados através de subseções correspondentes ao uso das RFs individualmente avaliando as bandas de frequência e ao uso dos meta-classificadores para a fusão da informação destas bandas.

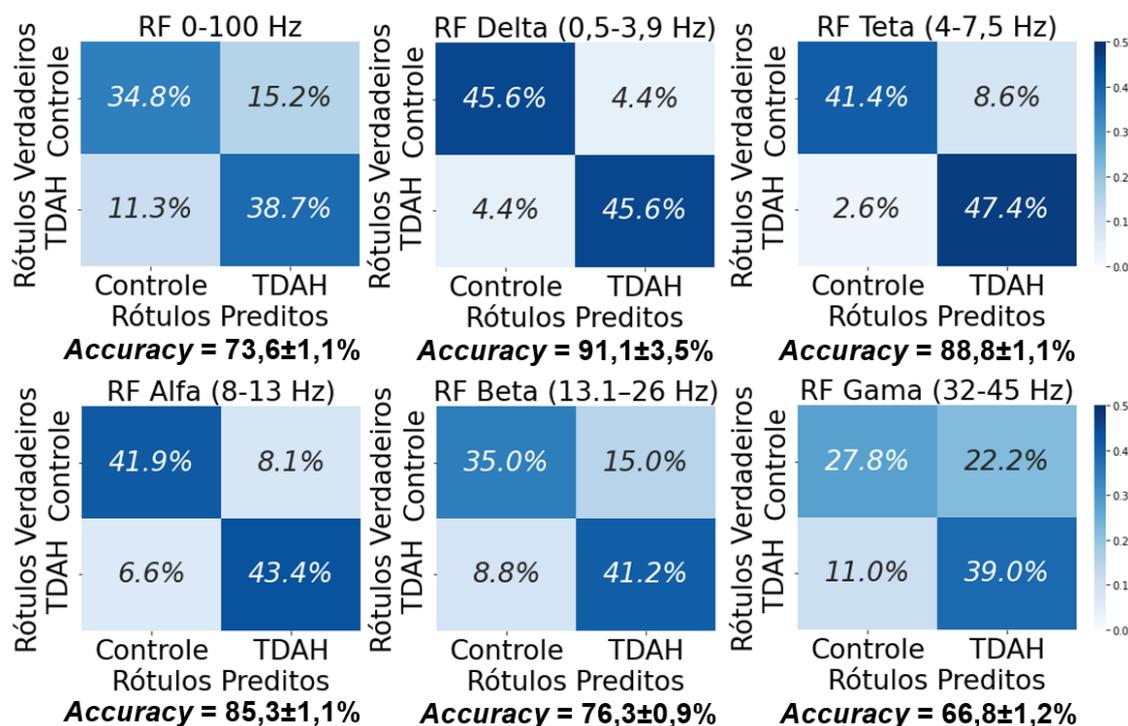
4.1.1 Resultados da Classificação de TDAH baseada em EEG: RF

A subseção de resultados da classificação de TDAH a partir de EEG por cada RF é dividida em duas partes: a classificação binária de TDAH e a classificação de subtipos.

4.1.1.1 RF na classificação entre pacientes com TDAH e voluntários do grupo controle

A Figura 31 apresenta as matrizes de confusão obtidas pela RF para a classificação TDAH/Controle trabalhando com características extraídas de cada faixa de frequência.

Figura 31 – Matrizes de confusão com as *accuracies* obtidas pelas RF para as bandas de EEG na classificação entre participantes com TDAH e do grupo controle no grupo de teste.



Fonte: O Autor.

Como mostra a Figura 31, com a banda larga (0-100 Hz), a RF obteve 73,6% de *accuracy* total de teste na classificação entre pacientes com TDAH e voluntários do grupo controle. Contudo, há um desbalanço entre as *accuracies* obtidas para cada classe: 77,4% para TDAH e 69,7% para Controle. Esta RF apresentou apenas uma pequena diferença de 3% entre a *accuracy* total de treinamento (76,9%) e de teste (73,6%).

Quando diagnosticou participantes a partir da banda Alfa, o classificador RF atingiu uma *accuracy* total de 85,3% com um balanço razoável entre as classes (83,8% de *accuracy* para controles e 86,8% para TDAH). Analisando esta faixa de frequência, o classificador apresentou *overfitting* de 5%, uma vez que a *accuracy* de treinamento foi de 90,3%.

Para a banda Beta, o classificador RF obteve 84,0% e 76,3% de *accuracies*, respectivamente, de treinamento e de teste para a classificação de TDAH/Controle.

O classificador RF alcançou suas melhores *accuracies* de classificação para o diagnóstico TDAH/Controle: 96,1% e 91,1% para os conjuntos de treinamento e de teste, respectivamente, analisando as características da banda Delta. Nesta faixa, para o conjunto de teste, a RF obteve um balanço perfeito entre as *accuracies* de cada classe, como mostra a Figura 31.

Na classificação TDAH/Controle através da análise da banda Gama, a RF obteve *ac-*

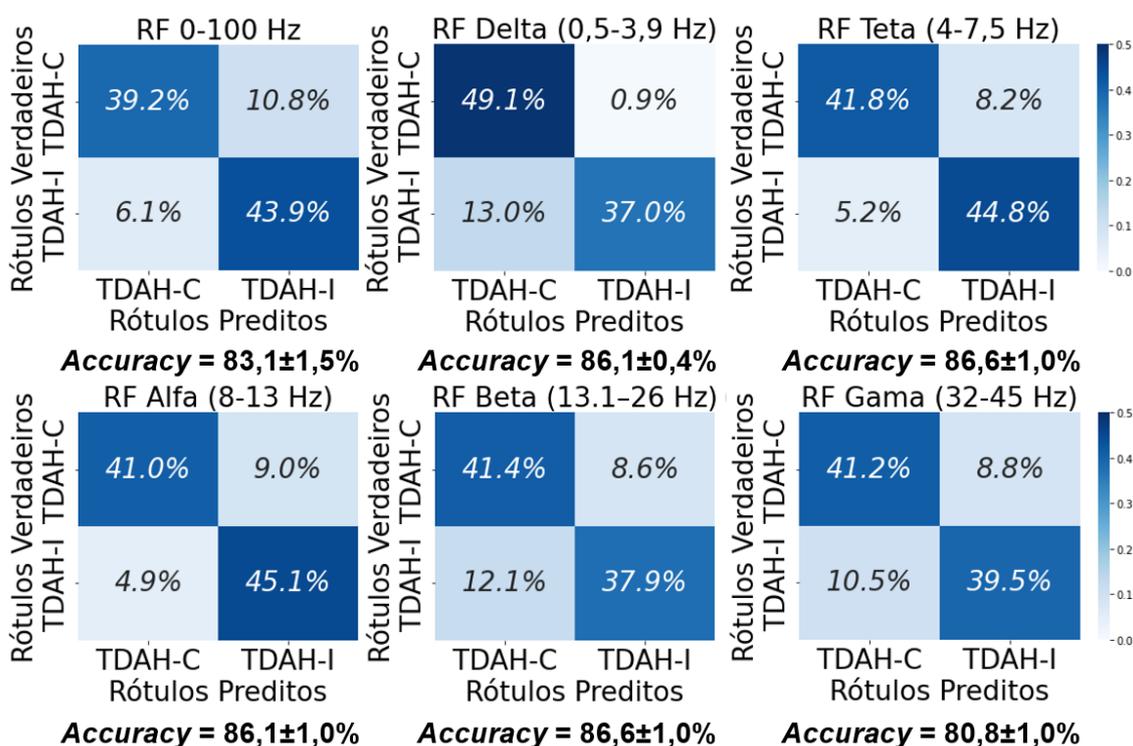
accuracies de 68,9% e 66,8% para o treinamento e para o teste. No conjunto de teste, para a banda Gama, o classificador apresentou um desbalanço marcante entre as classes: 55,6% de *accuracy* para grupo controle e 78,0% para o grupo com TDAH.

Analisando a banda Teta, a RF atingiu *accuracies* de treinamento e de teste de 89,2% e de 88,7%, respectivamente. Este é o menor *overfitting* observado nos experimentos com a RF aplicada ao diagnóstico TDAH/Controle. Além disto, observando a banda Teta, a RF alcançou *accuracies* de 82,7% para controles e 94,8% para TDAH.

4.1.1.2 RF na classificação entre subtipos de TDAH

A Figura 32 apresenta as matrizes de confusão dos classificadores RF para o diagnóstico de TDAH-C/TDAH-I.

Figura 32 – Matrizes de confusão com as *accuracies* obtidas pelas RF para as bandas de EEG na classificação de subtipos (TDAH-C versus TDAH-I) no conjunto de teste.



Fonte: O Autor.

Na classificação de subtipos, o modelo RF apresentou suas melhores *accuracies* para as bandas Delta, Teta e Alfa. Trabalhando com as características extraídas da banda Teta, a RF obteve *accuracies* de 86,8% e de 86,6% para treinamento e teste, respectivamente. Por isto, este foi, o menor *overfitting* obtido na classificação de subtipos. Com a banda Teta, as *accuracies* de cada classe foram de 83,5% e 89,6% para TDAH-C e TDAH-I, respectivamente.

Observando a banda Delta, o classificador RF diagnosticou corretamente 98,2% dos

pacientes com TDAH-C e 73,9% daqueles com TDAH-I no conjunto de teste, o que corresponde a uma *accuracy* de 86,1% (com *accuracy* de 89,3% no treinamento).

A classificação de subtipos com RF apresentou um comportamento similar trabalhando com características extraídas das faixas Beta e Gama. A RF obteve *accuracies* totais de 79,3% and 80,8%, respectivamente, para as bandas Beta e Gama. As *accuracies* das classes TDAH-C e TDAH-I foram similares nestes espectros. Contudo, a RF apresentou uma *accuracy* 0,2% menor para TDAH-C e uma *accuracy* 3,2% maior para TDAH-I trabalhando com as ondas Gama. Este comportamento similar entre as RFs trabalhando com as bandas Beta e Gama pode estar ligado às curvas semelhantes apresentadas na PCA.

Como mostrado na Figura 32, com a faixa Alfa, a RF obteve *accuracies* de 81,9% para TDAH-C, 90,2% para TDAH-I, e, conseqüentemente, 86,1% da *accuracy* total de teste. Este modelo obteve *accuracy* de classificação de 90,3 na etapa de treinamento.

Para a banda larga (0-100 Hz), a RF obteve 84,1% e 77,4% de *accuracies* totais de treinamento e de teste, respectivamente. No teste da banda larga, a RF obteve *accuracies* de 78,4% para TDAH-C e 87,8% para TDAH-I.

4.1.2 Resultados de classificação de TDAH e de subtipos com SVM a partir de EEG

Esta seção detalha os resultados de classificação de TDAH e de subtipos obtidos utilizando o modelo classificador SVM a partir dos sinais pré-processados de EEG. Esta seção é dividida em duas partes: a análise dos resultados de classificação binária e de subtipos.

4.1.2.1 SVM na classificação entre pacientes com TDAH e voluntários do grupo controle

O modelo SVM atingiu suas melhores *accuracies* para diagnóstico TDAH/Controle trabalhando com características extraídas da banda Delta: 87,8%, 76,6% e 99,0% para *accuracy* total, para *accuracy* na classe Controle e para *accuracy* na classe TDAH, respectivamente. No entanto, a SVM alcançou 99,6% de *accuracy* total para o conjunto de treinamento, o que aponta um certo *overfitting* (11,8%) e, conseqüentemente, a uma menor capacidade de generalização.

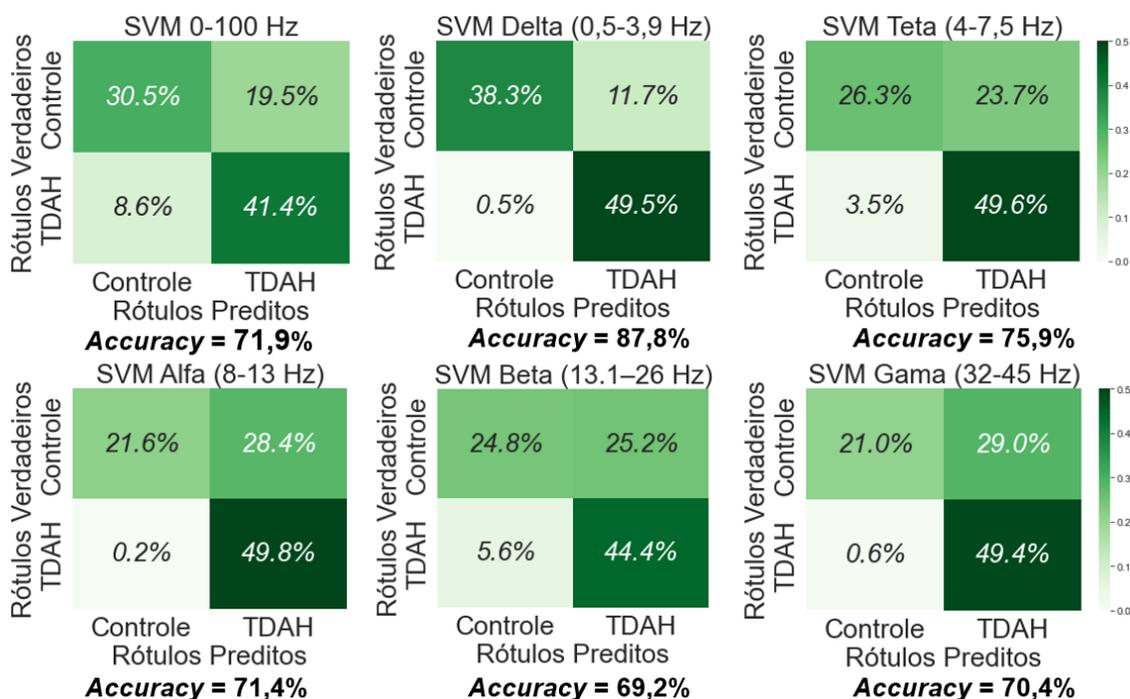
Observou-se que a maioria dos classificadores SVM que avaliam as bandas de EEG obtiveram *accuracies* grandes na classificação de TDAH: 82,8%, 99,0%, 99,3%, 99,5%, 88,7% e 98,9% para as bandas larga (0-100 Hz), Delta, Teta, Alfa, Beta e Gama, respectivamente.

Contudo, os modelos do SVM não apresentaram um balanço equilibrado entre as *accuracies* das classes. As *accuracies* do grupo controle foram 61,0%, 76,6%, 52,6%, 43,2%, 49,7% e 42,0% para as bandas larga (0-100 Hz), Delta, Teta, Alfa, Beta e Gama, respectivamente. Esta falta de balanço nas SVMs provocou menores *accuracies* totais em relação às RFs em todos os espectros, exceto pela banda Gama. Nesta faixa de frequência,

obteve-se 70,5% de *accuracy* total. Para os espectros restantes: larga (0-100 Hz), Teta, Alfa e Beta a SVM atingiu, respectivamente, *accuracies* totais de: 71,9%, 76,0%, 71,4% e 69,2%.

Para melhor visualização dos resultados, apresenta-se, na Figura 33, as matrizes de confusão correspondentes à classificação TDAH/Controle através de SVM.

Figura 33 – Matrizes de confusão com as *accuracies* de teste obtidas pelas SVM para as bandas de EEG na classificação TDAH *versus* Controle.



Fonte: O Autor.

4.1.2.2 SVM na classificação entre subtipos de TDAH

As matrizes de confusão resultantes da aplicação das SVM analisando cada faixa de frequência de EEG para a classificação de subtipos são apresentadas na Figura 34.

No diagnóstico de subtipos, a SVM alcançou sua maior *accuracy* total (65,9%) no conjunto de teste e 67,6% no conjunto de treinamento analisando a banda Delta. Para a banda Delta, a SVM obteve 99,4% de *accuracy* para TDAH-C e apenas 32,3% de *accuracy* para TDAH-I.

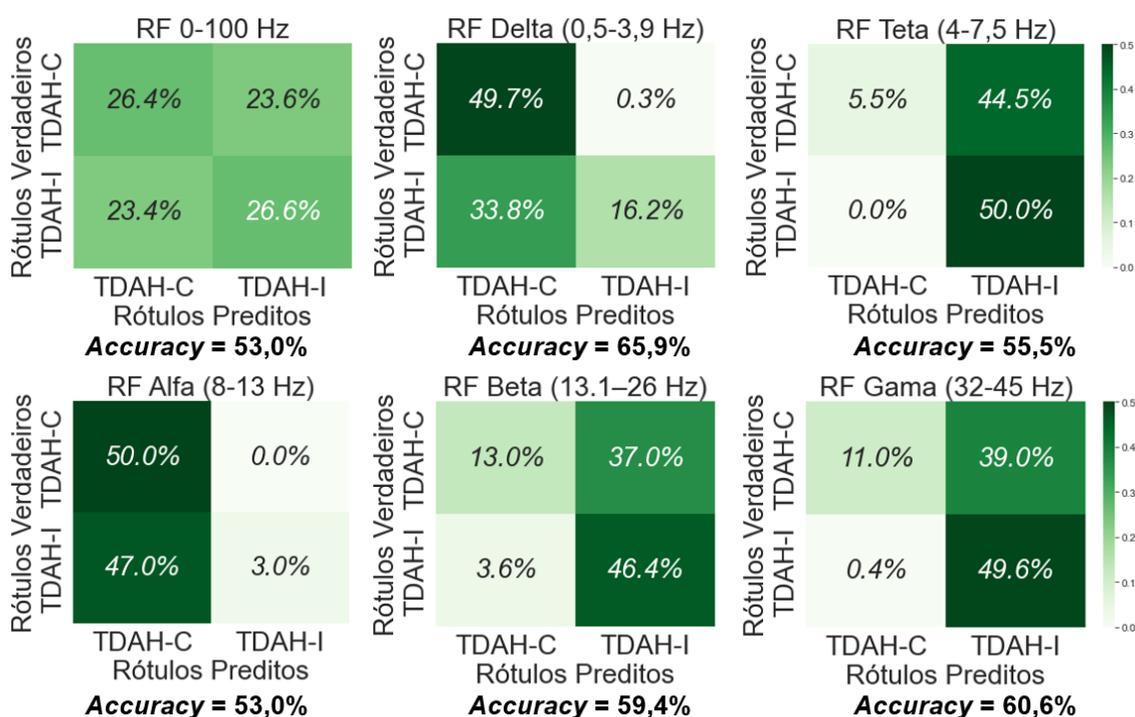
Quando trabalhou com a faixa Alfa, a SVM obteve *accuracies* de 99,9% para TDAH-C e 5,9% para TDAH-I (*accuracy* total de 52,9%), evidenciando um grande desbalanço entre as classificações.

Analisando a banda larga (0-100 Hz), a SVM obteve somente 53,0% de *accuracy*, porém com um balanço mais equilibrado entre as classes: 52,9% e 53,1% para TDAH-C e para TDAH-I, respectivamente.

Com a banda Gamma, a SVM obteve 60.6%, 99.3% e 21.9% de *accuracy* total, *accuracy* de TDAH-I e *accuracy* de TDAH-C, respectivamente. Na análise da banda Beta, a SVM se comportou de forma similar ao reportado para a banda Gama, obtendo: 59,4%, 26,1% e 92,7% de *accuracies* total, de ADHD-C e de ADHD-I.

Na análise da faixa Teta, a SVM acertou 100% dos indivíduos de TDAH-I e somente 11% dos sujeitos com TDAH-C (*accuracy* total de 55,5%).

Figura 34 – Matrizes de confusão com as *accuracies* de teste obtidas pelas SVM para as bandas de EEG na classificação TDAH-C versus TDAH-I.



Fonte: O Autor.

As *accuracies* de treinamento das SVM apresentaram valores menos de 2% maiores do que as do conjunto de teste, o que mostra que não houve um *overfitting* significativo. No entanto, os resultados obtidos pela SVM tanto para a classificação TDAH/Controle como para TDAH-C/TDAH-I possuem *accuracies* consideravelmente inferiores aquelas atingidas pelas RFs, com a única exceção da RF que atua com na classificação TDAH/Controle com as características da faixa Gama.

4.1.3 Resultados de classificação de de TDAH e de subtipos com RF e SVM a partir de EEG no grupo separado

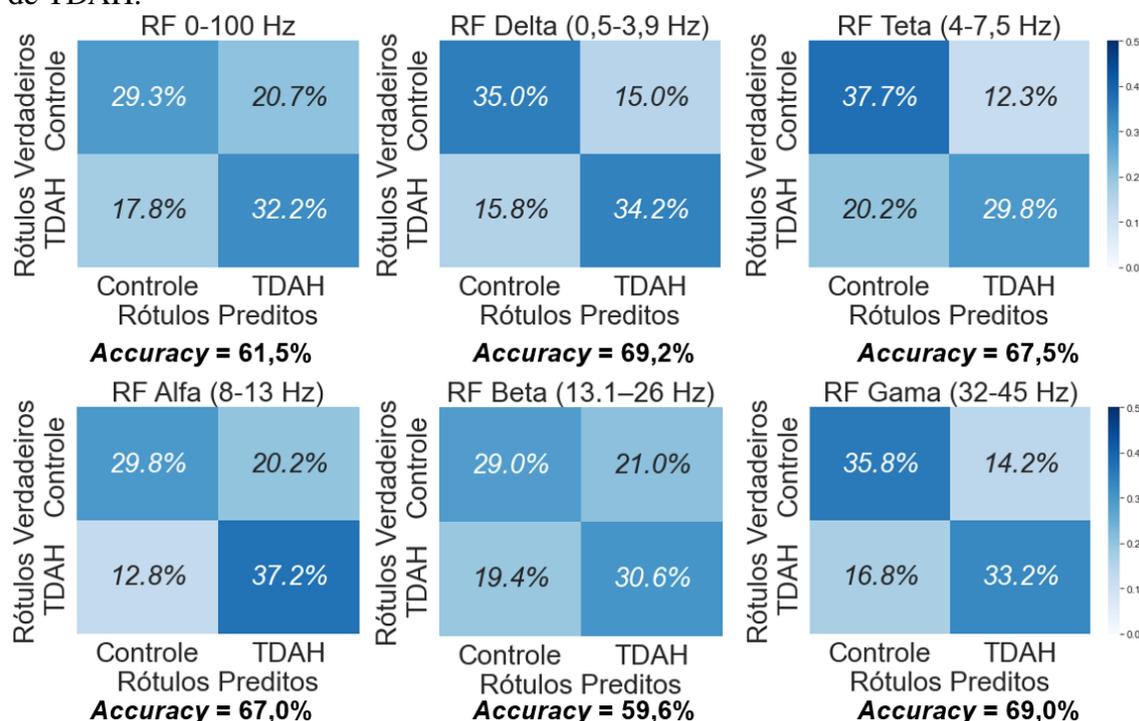
Conforme relatado na seção 3.2.1.7 do Capítulo 3 (Materiais e Métodos), haviam 2 conjuntos de teste para as amostras de EEG. O conjunto de teste "separado" consiste em um grupo de participantes, os quais não tiveram nenhum de seus segmentos participando do treinamento.

Embora as *accuracies* tenham sido consideravelmente menores do que as apresentadas para o conjunto de teste "misturado"(apresentadas no Capítulo 4), as bandas Delta, Teta e Alfa ainda se destacaram para a classificação entre pacientes com TDAH e voluntários do grupo controle. Contudo, neste teste, o uso da banda Gama produziu uma *accuracy* maior do que o uso das bandas Teta e Alfa, mas não que a banda Delta.

Na classificação de subtipos, nas bandas larga (0-100 Hz), Alfa e Beta, os classificadores obtiveram melhores *accuracies* do que para a identificação entre classes TDAH e Controle no conjunto de teste "separado". A RF que atuou com base na banda alfa obteve a melhor *accuracy*, nesta situação: 71,0%. A banda larga obteve uma *accuracy* próxima desta.

As figuras 35 e 36 mostram as matrizes de confusão obtidas no conjunto de teste "separado" com as RFs cujos hiperparâmetros foram otimizados para o conjunto de teste "misturado".

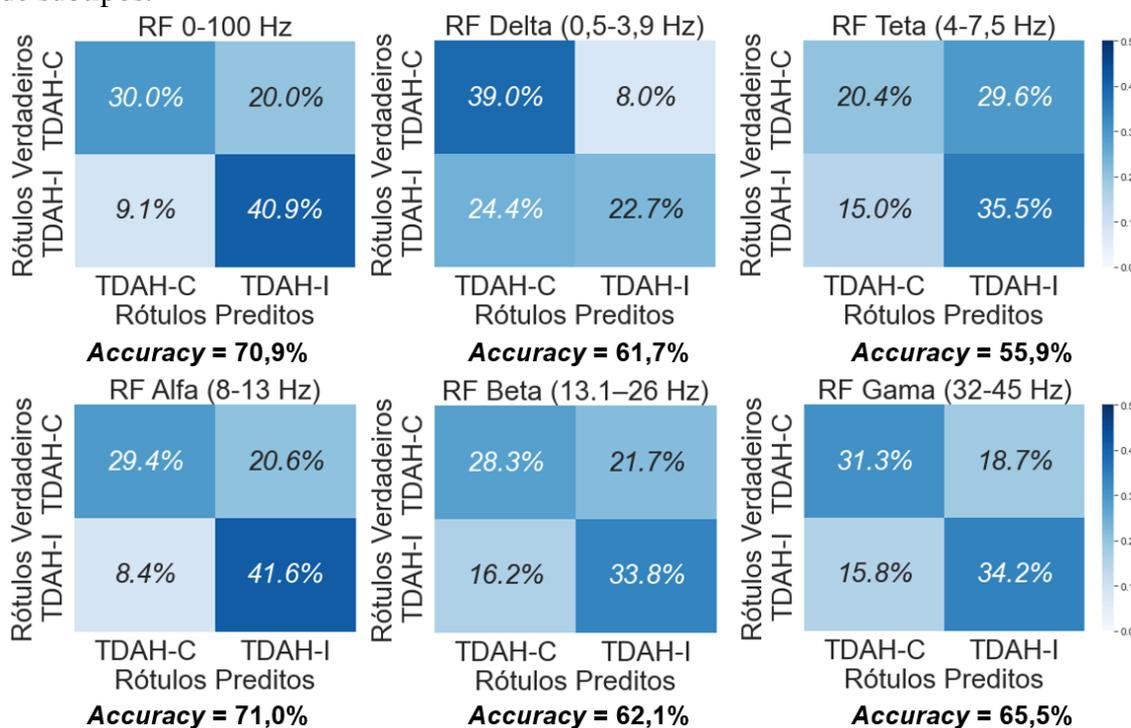
Figura 35 – Matrizes de confusão do conjunto teste "separado" da RF para classificação de TDAH.



Fonte: O Autor

Na classificação de subtipos, a SVM alcançou menos de 55% de *accuracy* para todas as bandas, com exceção da faixa Gama. E, apenas com a banda larga e, especialmente, a banda Teta obteve-se um balanço de acertos entre as classes TDAH-C e TDAH-I.

Figura 36 – Matrizes de confusão do conjunto teste "separado" da RF para classificação de subtipos.



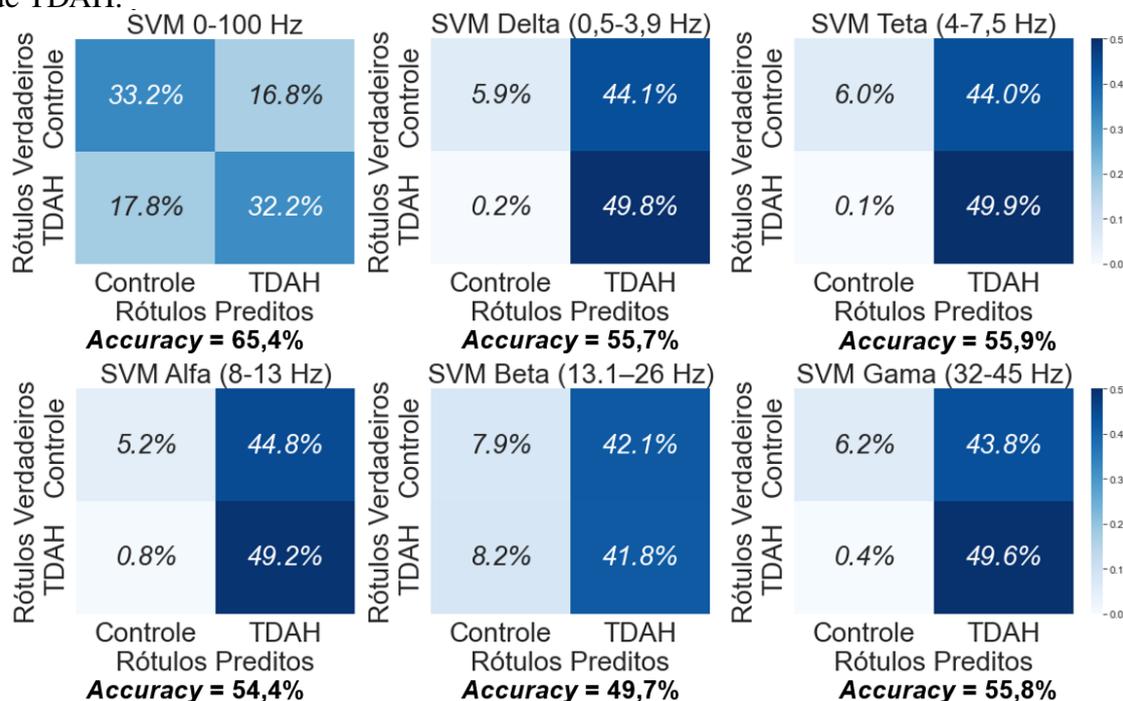
Fonte: O Autor

Os resultados obtidos mostram que as RFs sofrem uma considerável perda nas suas *accuracies* quando estes modelos tem de classificar amostras de indivíduos totalmente novos. Isto demonstra uma menor capacidade de generalização dos modelos nesta situação. Isto pode indicar que há fatores que auxiliam as RFs na classificação dos segmentos quando há outros segmentos dos indivíduos do conjunto teste no treinamento. Contudo, é importante ressaltar que estes modelos foram otimizados somente para classificar indivíduos do conjunto "misturados". Uma otimização realizada considerando um hipotético segundo grupo de indivíduos separados, possivelmente, poderia apresentar melhores resultados no conjunto separado.

Para a distinção entre grupos TDAH e Controle, as SVM obtiveram *accuracies* de apenas cerca de 55% para a maioria das faixas de frequência de EEG neste teste, exceto na banda larga, onde o classificador obteve uma *accuracy* de 65,4%. Outra exceção foi a banda Delta, com a qual a SVM alcançou apenas 49,7% de *accuracy*. A banda larga foi o único caso em que a SVM obteve *accuracies* balanceadas para as duas classes.

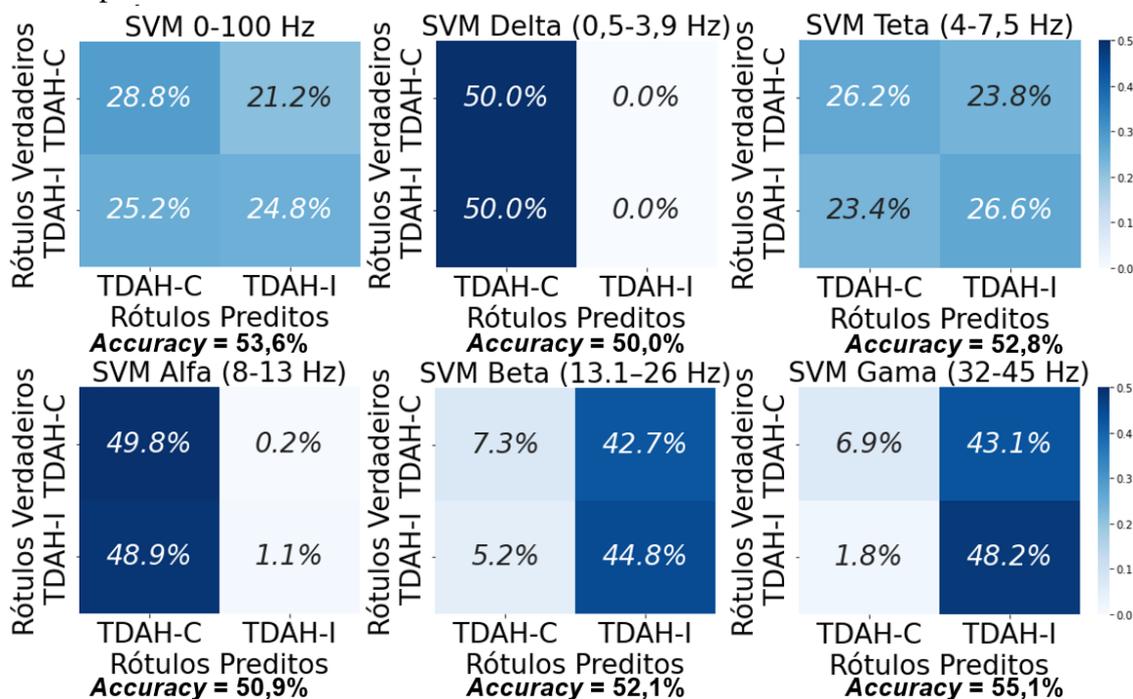
As figuras 37 e 38 apresentam as matrizes de confusão obtidas nas condições mencionadas nesta seção para o conjunto de teste "separado" com as SVMs otimizadas para o outro conjunto de teste.

Figura 37 – Matrizes de confusão do conjunto teste "separado" da SVM para classificação de TDAH.



Fonte: O Autor

Figura 38 – Matrizes de confusão do conjunto teste "separado" da SVM para classificação de subtipos.



Fonte: O Autor

Contudo, ressalta-se que os resultados apresentados nesta seção foram obtidos com

a configuração de classificadores primários otimizada através de AG para classificar o conjunto de teste "misturado". Experimentos posteriores com algoritmos com hiperparâmetros otimizados com AG para classificar o conjunto de teste "separado", obtiveram resultados (*accuracies*, precisão e *recall*) cerca de 5% menores apresentados na seções 4.1.1 e 4.1.2 para o conjunto de teste "misturado". Contudo, estes experimentos não foram concluídos.

4.1.4 Resultados da Classificação de TDAH baseada em EEG: Fusão de Bandas

Com exceção da Votação Majoritária Ponderada, os métodos de fusão utilizados possuem técnicas de otimização que modificam os pesos designados a cada um dos classificadores RF correspondentes as bandas de EEG. Os valores destes pesos quantificam a contribuição de cada banda para a obtenção das melhores *accuracies* de classificação. Portanto, a medição destes pesos indica o potencial de cada banda de EEG como um biomarcador de TDAH e de seus subtipos.

Com a finalidade de realizar uma comparação justa entre os pesos designados por cada meta-classificador, estes foram normalizados de acordo com a Equação 28, reproduzida por conveniência na Equação 30.

$$z = \frac{x - \mu}{\sigma}, \quad (30)$$

onde μ e σ são a média e o desvio padrão de todos os pesos dados por um único método de fusão, separadamente, em cada caso de classificação.

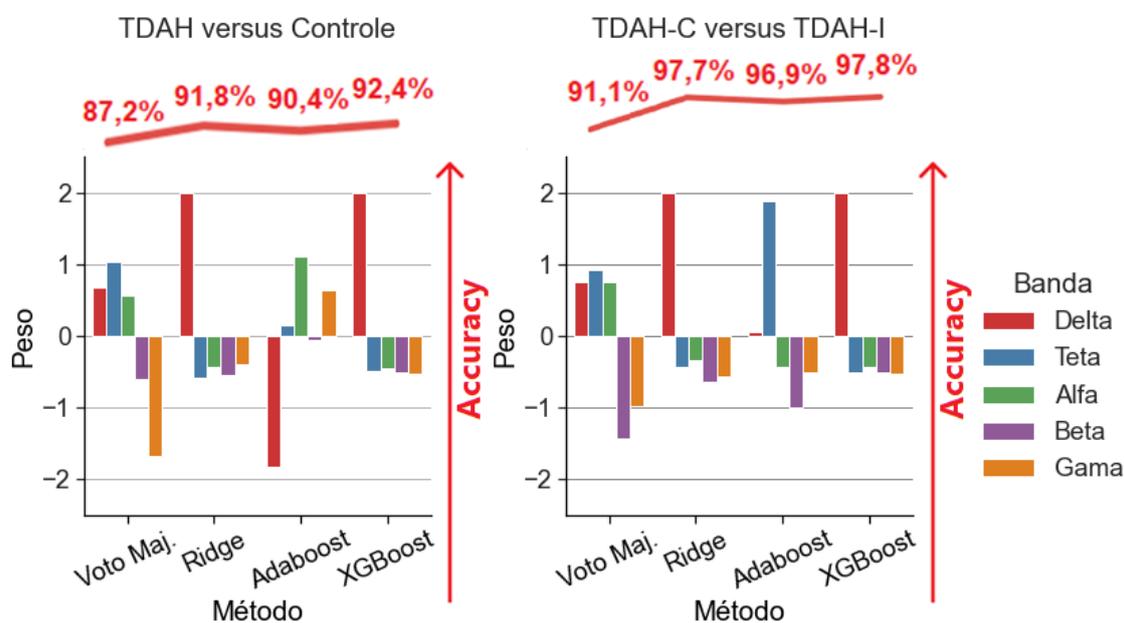
Os pesos atribuídos a cada uma das faixas de frequência na Votação Majoritária Ponderada são as *accuracies* obtidas pela RF correspondente quando aplicadas de forma independente, apresentadas nas figuras 31 e 32, na seção 4.1.1.

O meta-classificador linear *Ridge* aumentou consideravelmente o peso da banda Delta tanto para a classificação entre grupos TDAH e controle como para a identificação de subtipos, de forma similar. Os demais pesos foram reduzidos pela otimização, em ambos os casos. Para o diagnóstico TDAH/Controle, o *Ridge* considerou as ondas Gama como tendo maior importância que Teta, Alfa, e beta, recebendo esta última o menor peso para este caso.

Já para o diagnóstico TDAH-C/TDAH-I, o método *Ridge* aumentou o peso da banda Alfa (em relação a Teta, Beta e Gama). O peso da banda Teta, neste caso, foi maior do que o das faixas Beta e Gama.

A Figura 39 apresenta os pesos normalizados dados para as inferências das RFs de cada banda de EEG de acordo com cada método de fusão para ambos os cenários de classificação. Nestes gráficos, o peso médio é representado por zero devido à normalização.

Figura 39 – Pesos usados em cada método de fusão para cada banda de EEG para obtenção do resultado de classificação de TDAH *versus* Controle.



Fonte: O Autor.

A distribuição de pesos otimizada promovida pelo meta-classificador *XGBoost* foi muito similar para ambos os cenários de classificação (TDAH/Controle e TDAH-C/TDAH-I). Em ambos os casos, o *XGBoost* elevou consideravelmente o peso da banda Delta e reduziu as demais para um nível abaixo da média da normalização. O nível dos pesos dado para as bandas Teta, Alfa, Beta, e Gama foi similar, embora o peso designado para a faixa Alfa seja um pouco maior e o peso dado à Gama um pouco menor. Para o caso do diagnóstico TDAH/Controle, o peso atribuído à banda Beta é um pouco menor do que para o caso da detecção de subtipos de TDAH.

Para a identificação entre classes TDAH e Controle, o meta-classificador *AdaBoost*, favoreceu as predições baseadas nas características extraídas dos sinais nas bandas Alfa e Gama, além de reduzir consideravelmente a importância da banda Delta. Já os pesos das faixas Beta e Teta foram mantidos em valores próximos à média.

No caso da classificação de subtipos, o *AdaBoost* aumentou muito o peso das predições baseadas na faixa Teta e reduziu os demais pesos, especialmente, da banda Beta. Para este caso, o peso normalizado para fins comparativos da faixa Delta é aproximadamente zero.

Através da análise da Figura 39, observou-se que os métodos de fusão que apresentaram as maiores *accuracies*, *XGBoost* e *Ridge* chegaram a conclusões similares sobre os pesos de cada faixa de frequência. Ambos os meta-classificadores aumentaram consideravelmente a influência das predições baseadas na banda Delta e aumentaram levemente o peso das ondas Alfa em relação às demais. Contudo, estas técnicas apresentaram algu-

mas diferenças. Tanto para TDAH/Controle como para TDAH-C/TDAH-I, o classificador *Ridge* atribuiu maiores diferenças entre os pesos de Teta, Beta e Gama. O *XGBoost* atribuiu maior importância para a faixa Teta do que o *Ridge* que preferiu favorecer Gama no caso TDAH/Controle.

Já, o meta-classificador *AdaBoost* que obteve *accuracies* similares ao *XGBoost* e ao *Ridge* optou por caminhos diferentes. O *AdaBoost* reduziu o peso da Delta consideravelmente abaixo da média, na classificação TDAH/Controle, e elevou a relevância de Alfa e de Gama. No diagnóstico TDAH-C/TDAH-I, o *AdaBoost* manteve as ondas Delta com peso médio, elevou consideravelmente o peso da banda Teta e promoveu a maior redução na relevância da Beta.

4.1.4.1 *Meta-classificadores na classificação entre pacientes com TDAH e voluntários do grupo controle a partir das bandas de EEG*

Para diferenciar pacientes com TDAH e membros do grupo controle, o meta-classificador *Ridge* obteve *accuracies* totais de 94,0% e de 91,8% para treinamento e teste, respectivamente, com 95,6% de *accuracy* para controles e 88,0% para TDAH no conjunto de teste.

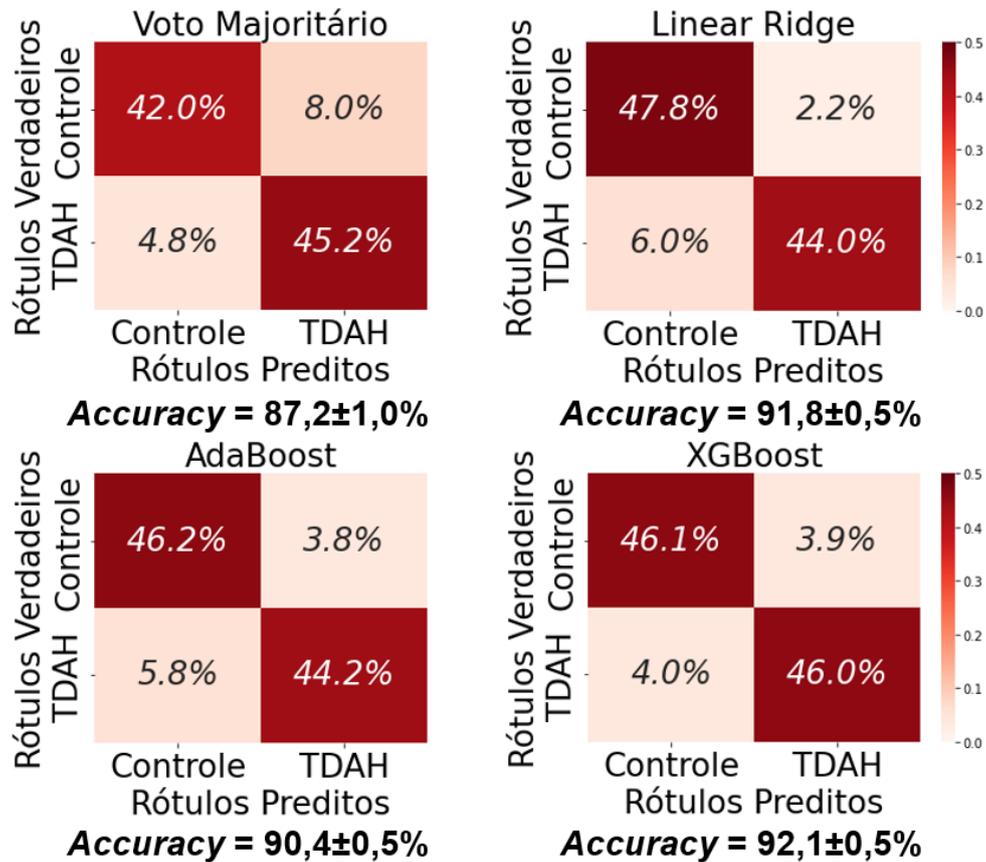
A técnica de fusão *XGBoost* obteve *accuracies* totais de treinamento e de teste iguais à 93,6% e 92,2%, respectivamente, na classificação TDAH/Controle. No conjunto de teste, a aplicação do método *XGBoost* obteve *accuracies* de 92,1% e de 92,1% para as classes controle e TDAH, respectivamente.

O meta-classificador *AdaBoost* obteve *accuracies* de 93,2% e de 90,4% para os conjuntos de treinamento e de teste, respectivamente, na classificação entre participantes com TDAH e do grupo controle. No teste, as *accuracies* nas classes controle e TDAH foram de 92,4% e de 88,4%, respectivamente.

O meta-classificador baseado na Votação Majoritária Ponderada apresentou pior desempenho que os demais métodos *ensemble* para a fusão das bandas de EEG tanto para o diagnóstico de TDAH/Controle como para determinar o subtipo de TDAH. Isto era esperado, pois este meta-classificador não dispõe de um método de otimização de pesos. Na classificação TDAH/Controle, a Votação Majoritária Ponderada obteve uma *accuracy* total de 87,2% e acertou 84,1% das amostras do grupo controle e 90,3% das amostras de TDAH no teste. No treinamento, a *accuracy* total obtida foi de 88,1%.

A Figura 40 apresenta as matrizes de confusão com as *accuracies* resultantes da aplicação dos meta-classificadores (Votação Majoritária Ponderada, Classificador Linear *Ridge*, *AdaBoost* e *XGBoost*) para a classificação de TDAH/Controle.

Figura 40 – Matrizes de confusão com as *accuracies* de teste obtidas pelos métodos de fusão de EEG na classificação TDAH *versus* Controle.



Fonte: O Autor.

A Tabela 26 resume as métricas *accuracy*, a sensibilidade (*Recall*) e Precisão de teste e de treinamento para os modelos de fusão de bandas de EEG testados para o diagnóstico TDAH/Controle. A Precisão e o *Recall* são calculados, respectivamente, de acordo com a Equação 31 e a Equação 32.

$$Precisão = \frac{VP}{VP + FP}, \quad (31)$$

$$Recall = \frac{VP}{VP + FN}, \quad (32)$$

onde VP é o número de Verdadeiros Positivos, FP é o número de Falsos Positivos e FN é o número de Falsos Negativos.

Uma vez que a Votação Majoritária Ponderada não requer treinamento, apresentou-se apenas os resultados obtidos para o conjunto de teste. Notou-se que os demais métodos de *ensemble* apresentaram valores maiores para a Precisão do que para o *Recall*, o que indica que os modelos apresentam mais falsos negativos do que falsos positivos para esta classificação.

Tabela 26 – Métricas obtidas por cada modelo de fusão a partir de bandas de EEG para diagnóstico TDAH/Controle.

Método	Acc. Treino (%)	Acc. Teste (%)	Recall Treino (%)	Recall Teste (%)	Precisão Treino (%)	Precisão Teste (%)
Voto Maj.	——	87,2	——	90,3	——	85,0
<i>Ridge</i>	94,0	91,8	91,2	88,0	96,6	95,3
<i>AdaBoost</i>	93,2	90,4	90,8	88,4	95,3	92,1
<i>XGBoost</i>	93,5	92,2	90,3	92,2	96,6	88,5

4.1.4.2 Meta-classificadores na classificação entre subtipos de TDAH a partir das bandas de EEG

Na determinação dos subtipos de TDAH, a Votação Majoritária Ponderada apresentou uma *accuracy* total de 91,1% (sendo 95,8% para TDAH-C e 86,4% para TDAH-I). No treinamento, a Votação Majoritária Ponderada obteve *accuracy* total um pouco maior: 94,2%.

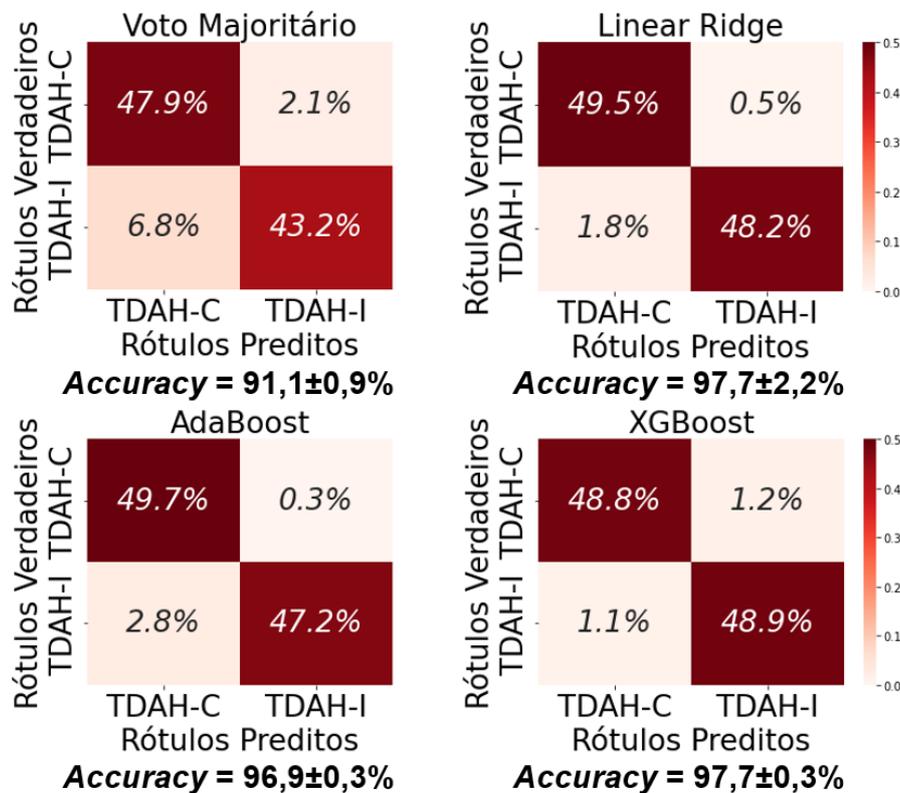
Na classificação de subtipos, o método *Ridge* atingiu 98,2% e 97,7 nas *accuracies* totais de teste e treinamento, respectivamente. No teste, o classificador acertou 99,0% para TDAH-C e 96,4% para TDAH-I.

Para a classificação de subtipos, o *AdaBoost* obteve, no teste, *accuracies* de 99,4% para TDAH-C e 94,4% para TDAH-I, totalizando uma *accuracy* média entre as classes de 96,9%. No treinamento, o valor desta *accuracy* total foi de 97,6, o que evidencia um baixo nível de *overfitting*.

A Figura 41 apresenta as matrizes de confusão com as *accuracies* resultantes da aplicação dos meta-classificadores (Votação Majoritária Ponderada, Classificador Linear *Ridge*, *AdaBoost* e *XGBoost*) para o diagnóstico de subtipos.

Como foi mostrado nas matrizes de confusão da Figura 41, no diagnóstico entre os subtipos TDAH-C e TDAH-I, o *XGBoost* alcançou *accuracies* totais de 98,7% e 97,8% para treinamento e teste. Sendo que, no teste, as *accuracies* para as duas classes são quase idênticas: 97,7% para TDAH-C e 97,8% para TDAH-I. Desta forma, a fusão de bandas de EEG com *XGBoost* alcançou os melhores resultados na presente Tese para as classificações de subtipos.

Figura 41 – Matrizes de confusão com as *accuracies* de teste obtidas pelos métodos de fusão de EEG na classificação TDAH-C *versus* TDAH-I.



Fonte: O Autor.

Ao contrário do que foi observado para a classificação entre grupos TDAH e Controle, para o diagnóstico de subtipos, os meta-classificadores *Ridge* e *AdaBoost* obtiveram valores levemente maiores para o *Recall* do que para a Precisão. Portanto, na classificação de subtipos, há uma maior tendência destes meta-classificadores apresentarem mais falsos positivos do que falsos negativos. Contudo, o método *XGBoost* obteve valores quase iguais para *Recall* e Precisão.

A Tabela 27 resume as métricas *accuracy*, *Recall* e Precisão de teste e de treinamento para os modelos de fusão de bandas de EEG utilizados nesta Tese para o diagnóstico de subtipos (TDAH-C/TDAH-I).

Destaca-se que todos os métodos de fusão de bandas de EEG testados apresentaram baixos níveis de *overfitting* tanto no diagnóstico TDAH/Controle como na determinação do subtipo dos pacientes. O *overfitting* foi avaliado pela diferença entre as *accuracies* totais de treinamento e as de teste. O *overfitting* de cada modelo de fusão em cada cenário de classificação (TDAH/Controle ou subtipos) é resumido na Tabela 28.

Tabela 27 – Métricas obtidas por cada modelo de fusão a partir de bandas de EEG para diagnóstico de subtipos (TDAH-C/TDAH-I).

Método	Acc. Treino (%)	Acc. Teste (%)	Recall Treino (%)	Recall Teste (%)	Precisão Treino (%)	Precisão Teste (%)
Voto Maj.	—	91,1	—	86,4	—	95,4
<i>Ridge</i>	98,2	97,7	97,0	96,4	99,3	99,0
<i>AdaBoost</i>	97,6	96,9	95,7	94,4	99,5	99,4
<i>XGBoost</i>	98,7	97,8	98,4	97,8	99,0	97,7

Tabela 28 – *Overfitting* em cada modelo de fusão de bandas de frequência de EEG.

Método de Fusão	TDAH/Controle (%)	TDAH-C/TDAH-I (%)
Voto Majoritário Ponderado	0,9	3,1
Classificador <i>Ridge</i>	2,2	0,5
<i>AdaBoost</i>	2,8	0,7
<i>XGBoost</i>	2,4	0,9

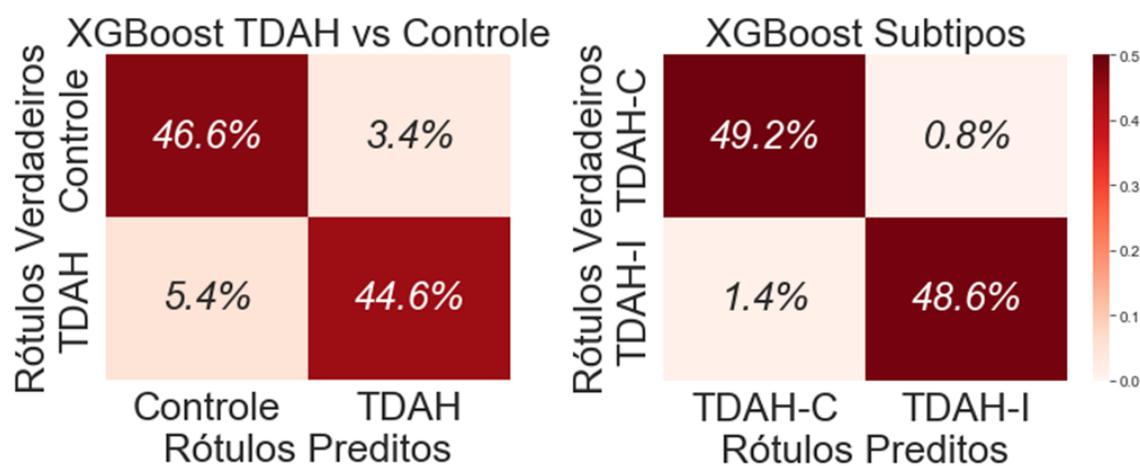
O desempenho dos meta-classificadores pode ser comparado ao uso da RF com a banda larga, pois a banda larga é composta pela combinação de todas as demais bandas de frequência analisadas, e do espectro entre 45 e 100 Hz. Embora também disponha de informações relativas à todas as bandas de frequência, a RF analisando a banda larga produziu uma *accuracy* de somente $73,6 \pm 1,1\%$ para a classificação binária e $83,1 \pm 1,5\%$ para a classificação de subtipos. Estes resultados são inferiores aos obtidos por qualquer um dos meta-classificadores analisados nesta Tese trabalhando com as bandas de EEG. A média do resultado obtido pela RF para a banda larga na classificação binária está quase 20% aquém da *accuracy* de $92,4 \pm 0,5\%$ produzida pelo *XGBoost* para este caso, além de possuir mais que o dobro de desvio padrão apresentado pelo meta-classificador. Na classificação de subtipos, a média dos resultados obtidos pela RF com a banda larga ($83,1 \pm 1,5\%$) foi 14,6% menor do que a média de resultados obtida pelo *XGBoost* ($97,7 \pm 0,3\%$).

4.1.4.3 Meta-classificador *XGBoost* na classificação apenas de amostras reais (não sintéticas) das bandas de EEG

Experimentou, também, testar-se os modelos de classificação de TDAH e de subtipos através do meta-classificador *XGBoost* em um conjunto de teste formado apenas por

amostras reais (sem amostras sintéticas). Nesta situação, a *accuracy* média do modelo XGBoost para identificação de TDAH reduziu de $92,4 \pm 0,2\%$ para $91,2 \pm 0,3\%$. Para a classificação dos subtipos, a *accuracy* média permaneceu em $97,8 \pm 0,3\%$. Isso mostrou que embora o uso de amostras sintéticas tenha auxiliado na fase de treinamento e otimização dos classificadores, esta teve pouca influência nos resultados dos testes destes modelos. A Figura 42 apresenta as matrizes de confusão referentes aos resultados do classificador XGBoost apenas com amostras reais.

Figura 42 – Matrizes de confusão com as *accuracies* de teste obtidas pelo XGBoost para amostras reais nas classificações binária e de subtipos.



Fonte: O Autor.

4.2 Resultados da Classificação de TDAH baseada em fMRI

Esta seção apresenta os resultados de classificação de TDAH obtidos a partir dos sinais de fMRI coletados da base HBN. Esta seção foi dividida em duas partes: os resultados obtidos pelos classificadores primários e os resultados alcançados pelos meta-classificadores.

4.2.1 Resultados da Classificação de TDAH baseada em fMRI: Arquitetura CNN Projetada

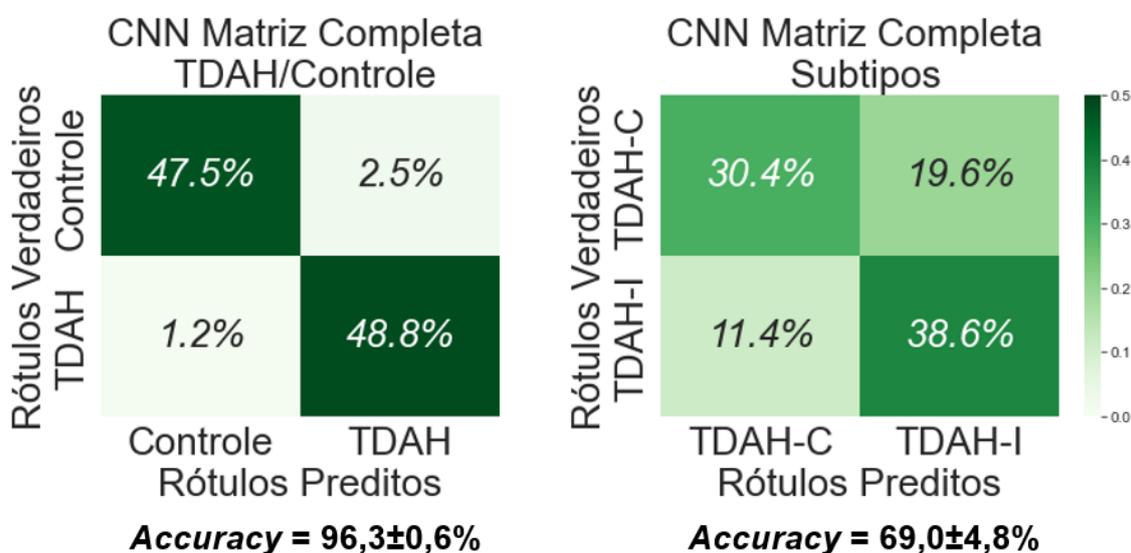
Nesta subseção são apresentados os resultados obtidos pelo modelo classificador baseado na CNN projetada. Primeiramente, apresentam-se o desempenho na classificação de TDAH através da avaliação da matriz de correlações completa. Nas subseções seguintes, considera-se a classificação de TDAH e de subtipos através da análise com os modelos das redes de conectividade.

4.2.1.1 Classificação da matriz de correlações completa através da arquitetura CNN projetada

Primeiramente, apresentam-se os resultados alcançados pela CNN nas classificações utilizando as matrizes de correlação como características de entrada de cada participante. Cada resultado relatado nesta seção é uma média da métrica correspondente para as classificações obtidas na metodologia de validação cruzada *k-fold* com $k=10$. O desvio padrão foi usado como medida da dispersão. O *overfitting* é quantificado pela diferença resultante subtraindo-se a *accuracy* média no teste da *accuracy* média no conjunto de treinamento.

A Figura 43 apresenta as matrizes de confusão alcançadas pelos modelos CNN na classificação de TDAH/Controle e subtipos (TDAH-C/TDAH-I) usando todas as matrizes de correlação como características.

Figura 43 – Matrizes de confusão com as *accuracies* de teste obtidas pela CNN a partir das matrizes de correlação de fMRI nas classificações entre TDAH e grupo controle e de subtipos.



Fonte: O Autor.

A CNN projetada para TDAH versus diagnóstico de controle alcançou $99,9\pm 0,08\%$ e $96,3\pm 0,6\%$ de *accuracy* para conjuntos de treinamento e de teste, respectivamente. No conjunto de testes, a CNN classificou corretamente $95,0\%$ dos voluntários do grupo controle e $97,6\%$ dos pacientes com TDAH. Essa diferença pode ter ocorrido devido à maior presença de pacientes com TDAH na base de dados.

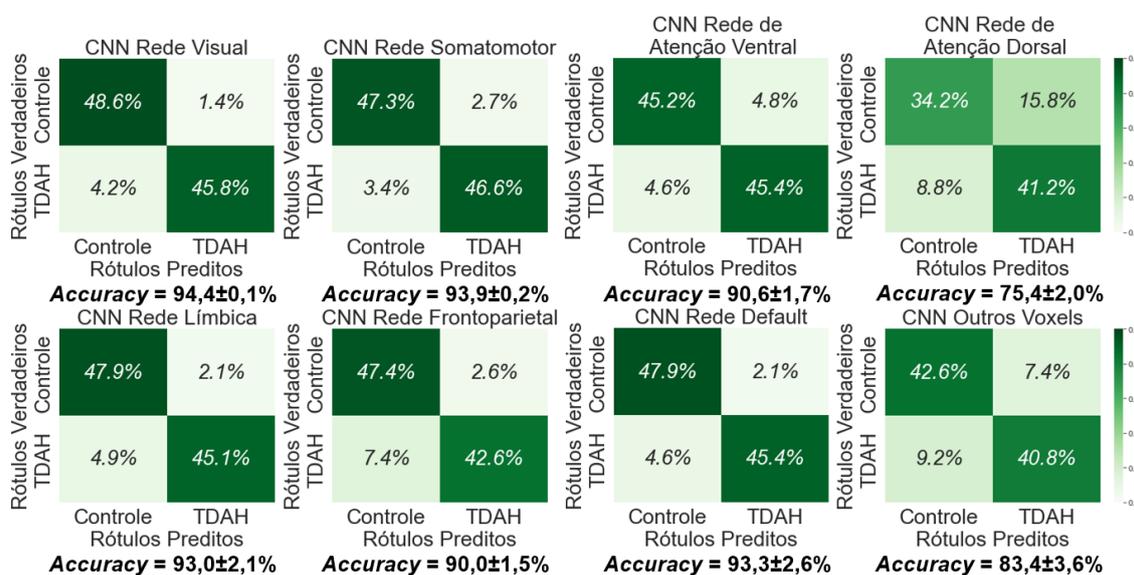
No entanto, o desempenho da CNN usando a matriz completa não foi tão bom para classificação de subtipos. Apesar da otimização do classificador com o algoritmo AG, a *accuracy* média foi de apenas $72,8\pm 4,4\%$ e $69,0\pm 4,8\%$ para os conjuntos de treinamento e de teste, respectivamente. No conjunto de testes, a CNN apresentou maior discrepância

entre os diagnósticos corretos de TDAH-C e de TDAH-I: 60,7% e 77,3%, respectivamente.

4.2.1.2 Redes de Conectividade: Identificação entre classes TDAH e grupo controle com a arquitetura CNN projetada

A Figura 44 apresenta as matrizes de confusão obtidas por cada classificador CNN em cada rede de conectividade para classificação binária de TDAH.

Figura 44 – Matrizes de confusão com as *accuracies* de teste obtidas pela CNN a partir das redes de conectividade de fMRI na classificação de TDAH *versus* Controle.



Fonte: O Autor.

Analisando as redes de conectividade individualmente, as *accuracies* médias totais foram menores, em relação ao uso da matriz de conectividade completa, para identificação de indivíduos com TDAH ou do grupo controle. Contudo, estes resultados podem auxiliar na identificação de marcadores biológicos, uma vez que a *accuracy* será maior quando a CNN receber redes de conectividade que apresentem maiores diferenças entre as classes.

Para a rede visual, o classificador obteve $99,1 \pm 0,1\%$ e $94,5 \pm 0,1\%$ de *accuracies* nos conjuntos de treinamento e teste, respectivamente. No entanto, a *accuracy* foi consideravelmente superior para o grupo controle (97,3%) do que para o grupo TDAH (91,7%).

Trabalhando com a rede somatomotor, a CNN alcançou *accuracy* total ligeiramente inferior nos conjuntos de treinamento e teste ($98,3 \pm 5,1\%$ e $93,8 \pm 2,2\%$, respectivamente). Entretanto, esta CNN obteve maior equilíbrio entre as classes: 94,6% para o grupo controles e 93,1% para pacientes com TDAH.

A CNN que trabalhou com rede de atenção ventral apresentou um maior equilíbrio de *accuracies* entre as classes: 90,5% e 90,9% para os grupos controle e TDAH, respectivamente. No total, esta CNN apresentou *accuracies* de $94,7 \pm 0,3\%$ e de $90,7 \pm 1,7\%$ para os

conjuntos de treinamento e de teste, respectivamente. Portanto, esta CNN apresentou um *overfitting* de 4,0%.

A CNN alimentada pela rede de atenção dorsal obteve as piores *accuracies*: e $77,2 \pm 0,3\%$ no treinamento e $75,5 \pm 2,1\%$ no teste. No entanto, observou-se que o *overfitting* foi de apenas 1,7%. Com esta rede de conectividade, a diferença entre as *accuracies* das classes é agravada. As *accuracies* para os grupos controle e TDAH foram, respectivamente, 68,4% e 82,5%.

Trabalhando com a rede límbica, a respectiva CNN apresentou $98,6 \pm 0,1\%$ e $93,7 \pm 2,1\%$ de *accuracy* para os conjuntos de treinamento e de teste, respectivamente. Esta CNN obteve uma diferença de 5,6% nas *accuracies* de classificação de cada classe, favorecendo a classificação correta dos indivíduos saudáveis.

Atuando com a rede frontoparietal, a CNN correspondente obteve uma *accuracy* de $97,7 \pm 0,1\%$ no treinamento. Já no conjunto de teste, 94,9% de *accuracy* para o grupo controle e 85,1% para o grupo TDAH, totalizando uma *accuracy* de $90,0 \pm 1,5\%$. Consequentemente, o uso da região frontoparietal conduz a um *overfitting* de 7,7%. Observa-se também que ao usar esta rede tem-se quase 10% a mais de chances de classificar corretamente indivíduos do grupo controle do que do grupo TDAH.

A CNN que trabalhou com a rede *default* obteve *accuracies* de $97,8 \pm 1,2\%$ e de $93,2 \pm 2,6\%$ para o treinamento e para o teste, respectivamente. A diferença de *accuracies* entre as classes no teste foi de 5,1%, favorecendo a classe controle: 95,8% e 90,7% na classificação dos indivíduos de controle e de TDAH, respectivamente.

A partir da região dos 'outros voxels', a CNN alcançou *accuracies* de $86,6 \pm 0,6\%$ e $83,5 \pm 3,6\%$ no treinamento e no teste, respectivamente. A discrepância entre as *accuracies* de cada classe foi de apenas 3,7%, uma vez que esta CNN obteve 85,3% de *accuracy* para os voluntários da classe controle e 81,6% para os pacientes com TDAH.

4.2.1.3 *Redes de Conectividade: Identificação entre classes TDAH-C e TDAH-I com a arquitetura CNN projetada*

O diagnóstico dos subtipos (TDAH-C/TDAH-I) com as CNNs não foi tão bem-sucedido quanto a classificação binária de TDAH/Controle.

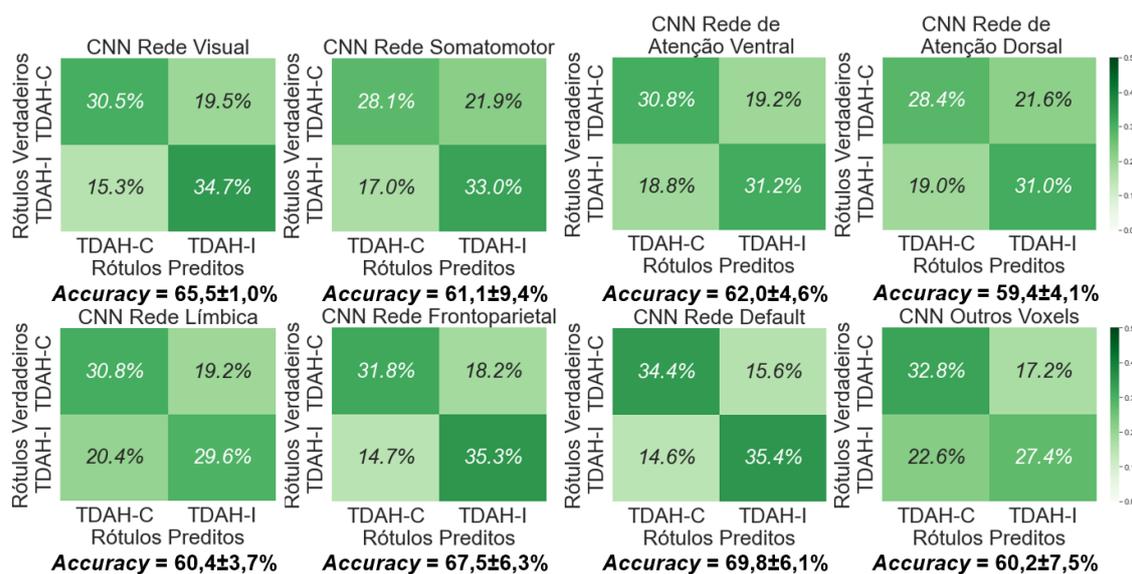
Com a rede visual, a CNN obteve uma *accuracy* de $65,2 \pm 1,0\%$ no conjunto de teste e $69,5 \pm 0,1\%$ no conjunto de treinamento. Desta forma, o *overfitting* é de 4,3%. Para esta rede, a diferença de *accuracy* entre as classes é de 8,4%, favorecendo a classificação correta da classe TDAH-I em detrimento da classificação de TDAH-C. Estas classes apresentam, respectivamente, *accuracies* de 69,4% e de 61,0%. As Precisões de treinamento e de teste são, respectivamente, 69,5% e 63,1%.

Atuando com a rede somatomotor, a CNN alcançou *accuracies* de $65,9 \pm 2,0\%$ e de $61,2 \pm 9,4\%$ no treinamento e no teste. Esta, consequentemente, é a CNN que apresentou maior desvio padrão dos resultados no conjunto de teste e é, portanto, a que apresentou

resultados menos robustos. A partir dos dados da rede somatomotor, esta CNN obteve uma discrepância considerável nas *accuracies* de cada classe: 56,2% para TDAH-C e 66,1% para TDAH-I. As Precisões de treinamento e de teste foram de 65,9% e 61,1%, respectivamente.

A CNN encontrou maiores *accuracies*, e, portanto, discrepâncias entre os subtipos quando investigou as redes *default* e frontoparietal. A Figura 45 apresenta as matrizes de confusão alcançadas pelos classificadores CNN no teste com cada rede de conectividade na classificação de subtipos.

Figura 45 – Matrizes de confusão com as *accuracies* de teste obtidas pela CNN a partir das redes de conectividade de fMRI na classificação de TDAH-C *versus* TDAH-I.



Fonte: O Autor.

Trabalhando com a rede de atenção ventral, a CNN apresentou menores diferenças entre as *accuracies* obtidas para cada classe no teste: 61,6% e 62,5%. Portanto, a *accuracy* do classificador foi de 62,0±4,6% no teste, uma vez que a *accuracy* do classificador é a média entre as *accuracies* que o modelo apresenta para cada classe.

No conjunto de treinamento, esta CNN obteve 67,9±1,0% de *accuracy*. Analisando a rede de atenção dorsal, a CNN obteve seu pior desempenho na classificação de subtipos. Este desempenho foi medido de acordo com as *accuracies* de treinamento: 60,3±2,5%; e de teste: 59,3±4,1%. As *accuracies* das classes no teste foram 56,8% e 61,9% para TDAH-C e TDAH-I, respectivamente. Portanto, a diferença de *accuracies* entre as classes foi de 5,1%.

Ao analisar a rede límbica, a respectiva CNN obteve *accuracies* de treinamento e de teste de 61,9±0,4% e de 60,1±3,7%, respectivamente, para a classificação de subtipos. No teste, a CNN obteve 61,6% e 59,2% de *accuracies* para TDAH-C e TDAH-I, respectivamente. Consequentemente, esta CNN obteve baixo *overfitting* (1,8%) e baixa diferença

entre as *accuracies* de cada classe (2,4%). Na rede frontoparietal, a *accuracy* no conjunto de teste da CNN foi de $67,1 \pm 6,3\%$, sendo 70,6% para TDAH-I e 63,7% para TDAH-C no conjunto de teste. Observou-se um *overfitting* de 8%, já que a *accuracy* de treinamento foi de $75,4 \pm 0,36\%$.

Na rede *default*, o classificador obteve *accuracies* de $74,6 \pm 4,9\%$ e $69,8 \pm 6,1\%$ em conjuntos de treinamento e de teste, respectivamente. Portanto, usando apenas a rede *default*, a CNN alcançou uma *accuracy* maior do a CNN que analisou a matriz de correlações completa. Esta é a maior *accuracy* no conjunto de teste atingida na classificação de subtipos pelas CNNs projetadas. Na rede *default*, as *accuracies* de classe foram de 68,8% para TDAH-C e de 70,9% para TDAH-I.

A CNN que analisou a região dos 'outros voxels' para a classificação de subtipos obteve uma *accuracy* de treinamento de $63,7 \pm 3,2\%$ e uma *accuracy* no conjunto de teste de $60,1 \pm 7,5\%$. A análise da região dos 'outros voxels' conduziu esta CNN a ser a única obter uma maior *accuracy* para a classe TDAH-C (65,5%) do que para a classe TDAH-I (54,7%). Observa-se que a diferença das *accuracies* entre as classes com esta CNN é de 10,8%.

4.2.2 Resultados: Fusão de Classificadores de fMRI baseados nas Redes de Conectividade

A aplicação dos modelos meta-classificadores para a fusão das redes de conectividade funcional permitiu aprofundar as investigações de marcadores biológicos de fMRI e aumentou a *accuracy* dos modelos, o que será mostrado nesta seção.

A análise da distribuição dos pesos e da *accuracy* obtida por cada meta-classificador ajuda a estabelecer conexões entre cada rede de conectividade e a presença de discrepâncias entre os grupos TDAH e controle ou entre os subtipos de TDAH. Para garantir uma comparação justa, aplicou-se a normalização *z-score* aos pesos dos *ensembles*.

O *XGBoost* valorizou a rede somatomotor, embora tenha dado um peso maior para a rede límbica. Os classificadores *Ridge* e *AdaBoost* também colocaram pesos altos na rede límbica, mas atribuíram igual importância para a rede frontoparietal.

A Tabela 29 mostra os valores dos pesos usados por cada meta-classificador após a normalização na classificação TDAH/Controle. Nesta tabela, observa-se que no Voto Majoritário ponderado priorizou-se a classificação conforme as redes visual e somatomotor. Os resultados desta tabela também mostram que, para os métodos de otimização automática, a rede de conectividade mais conectada ao diagnóstico de TDAH/Controle é a límbica, seguida pela frontoparietal e somatomotor. Porém, quanto maior a importância que cada método atribuiu à rede visual, maior foi a *accuracy* alcançada. Todos os métodos, exceto o classificador *Ridge*, apontaram peso mínimo para a rede de atenção dorsal.

Tabela 29 – Pesos para cada rede de conectividade funcional usados em cada meta-classificador na classificação de TDAH/Controle. Os dois pesos com maiores valores conforme cada meta-classificador foram marcados em negrito na tabela.

Redes de Conectividade	Voto Majoritário Ponderado	<i>Ridge</i>	<i>AdaBoost</i>	<i>XGBoost</i>
Visual	0,709009	-0,478132	-0,811149	-0,515711
Somatomotor	0,604934	-0,691162	0,324778	0,206284
Atenção Ventral	0,097570	-0,825860	-0,811149	-0,515711
Atenção Dorsal	-2,374205	-0,648094	-0,811149	-0,515711
Límbica	0,474841	1,571818	1,459909	2,372269
Frontoparietal	-0,021140	1,571818	1,459909	-0,515711
<i>Default</i>	0,508990	-0,500389	-0,811149	-0,515711

A Tabela 30 apresenta os valores dos pesos usados por cada meta-classificador após a normalização para a classificação entre subtipos de TDAH (TDAH-C e TDAH-I).

Tabela 30 – Pesos para cada rede de conectividade funcional usados em cada meta-classificador na classificação de subtipos. Os pesos que receberam valores positivos foram marcados em negrito na tabela.

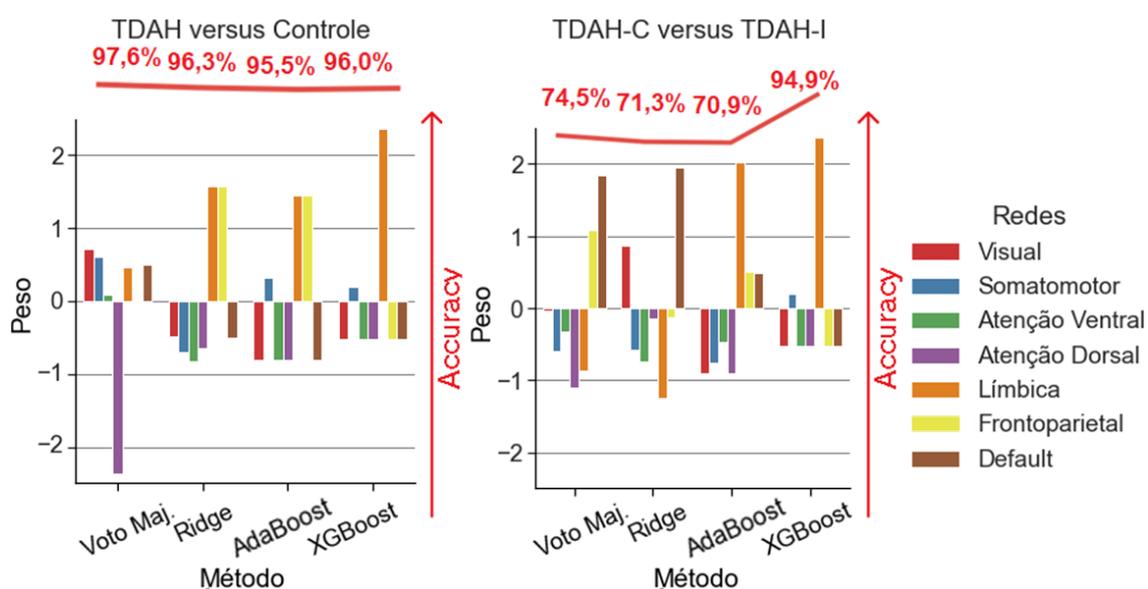
Redes de Conectividade	Voto Majoritário Ponderado	<i>Ridge</i>	<i>AdaBoost</i>	<i>XGBoost</i>
Visual	-0,041752	0,871874	-0,893455	-0,515711
Somatomotor	-0,590093	-0,579928	-0,759616	0,206284
Atenção Ventral	-0,331232	-0,732859	-0,476301	-0,515711
Atenção Dorsal	-1,093900	-0,137893	-0,893455	-0,515711
Límbica	-0,874006	-1,241932	2,025374	2,372269
Frontoparietal	1,088333	-0,125847	0,499352	-0,515711
<i>Default</i>	1,842650	1,946584	0,498101	-0,515711

Na classificação de subtipos, os meta-classificadores Voto Majoritário Ponderado, *Ridge* e, em certa medida, o *AdaBoost* selecionaram pesos altos para a rede *Default*. En-

quanto, no Voto Majoritário a rede frontoparietal foi a segunda em ordem de importância, o classificador *Ridge* designou um peso maior à rede visual. Já, o *AdaBoost* colocou na rede frontoparietal um peso similar ao da rede *Default*, embora quatro vezes menor do que o peso escolhido para a rede límbica, que recebeu a maior prioridade. O *XGBoost* concordou com o *AdaBoost* que a rede límbica é o principal biomarcador. Contudo, o *XGBoost* percebeu uma importância maior em observar diferenças entre as classes através da CNN que trabalha com a rede somatomotor. Considerando que o *XGBoost* obteve uma *accuracy* cerca de 20% maior que os demais *ensembles* na classificação de subtipos, isto indica que as redes límbica e somatomotor são biomarcadores mais robustos na identificação de distinções cerebrais entre as classes TDAH-C e TDAH-I.

Os pesos das redes de conectividade para cada modelo de *ensemble* são apresentados na Figura 46 tanto para a classificação de TDAH/Controle quanto para classificar TDAH-C/ADHD-I. A altura das barras representa a importância que cada classificador deu a cada rede de conectividade. Como os pesos foram normalizados, um peso negativo corresponde a uma rede de conectividade com menor importância para as classificações dos sujeitos. A Figura 46 e as tabelas 29 e 30 mostram que a distribuição de pesos pode variar mesmo para os métodos que alcançaram *accuracies* semelhantes, porque a discrepância entre a as classes é similar em redes de conectividade distintas. A curva vermelha indica as *accuracies* correspondentes alcançadas por cada metodologia de conjunto.

Figura 46 – Pesos determinados por cada método de fusão para cada rede de conectividade para classificações de TDAH/Controle e TDAH-C/TDAH-I. A curva vermelha indica a *accuracy* total obtida por cada meta-classificador.



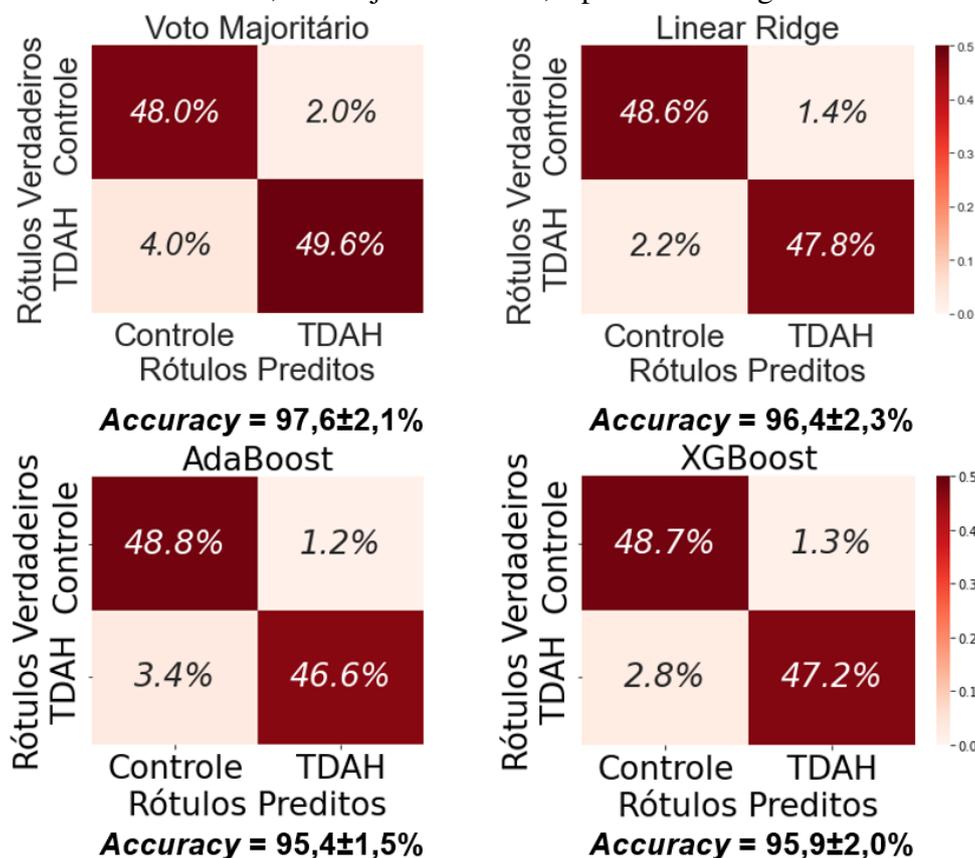
Fonte: O Autor.

A Votação Majoritária Ponderada alcançou $100 \pm 0,05\%$ de *accuracy* no conjunto de

treinamento e $97,6 \pm 2,1\%$ de *accuracy* no conjunto de teste para a classificação entre pacientes com TDAH e sujeitos do grupo controle. A *accuracy* da Votação Majoritária foi um pouco maior para o grupo com TDAH (99,2%) do que para o grupo controle (96,0%) no conjunto de teste.

A Figura 47 mostra as matrizes de confusão obtidas pela aplicação das técnicas de *ensemble* para fundir as informações obtidas pelas CNNs de cada rede de conectividade na classificação entre grupos TDAH e Controle. Observa-se que haviam 186 amostras da classe TDAH e 48 da classe Controle no conjunto de teste, antes das amostras de Controle serem aumentadas para 186 através da técnica de sobreamostragem, SMOTE.

Figura 47 – Matrizes de confusão obtidas pelas técnicas de fusão por *ensemble* na classificação de TDAH/Controle, no conjunto de teste, a partir das imagens de fMRI.



Fonte: O Autor.

Para a classificação TDAH/Controle, o classificador *Ridge* obteve *accuracies* de $96,3 \pm 2,3\%$ no conjunto de teste e $100 \pm 0,04\%$ no conjunto de treinamento. A *accuracy* da classificação dos controles *Ridge* foi de 97,1%. Já a *accuracy* na classe TDAH foi de 95,5%. Por isto, o *Ridge* foi o meta-classificador de fMRI que obteve a menor diferença de *accuracies* entre as classes para a classificação TDAH/Controle.

O *AdaBoost* alcançou *accuracies* de $100 \pm 0,05\%$ no treinamento e $95,5 \pm 1,5\%$ no conjunto de testes para classificação de participantes com TDAH e do grupo controle. O

AdaBoost obteve 97,7% de *accuracy* da classe de controle e 93,3% de *accuracy* da classe de TDAH. Portanto, dentre os meta-classificadores de fMRI, o *AdaBoost* apresentou a maior diferença de *accuracies* conforme as classes para distinguir os grupos TDAH e Controle: 4,4%.

O *XGBoost* obteve uma *accuracy* de treinamento de $100 \pm 0,06\%$ e uma *accuracy* no conjunto de teste de $96,0 \pm 2,0\%$ na classificação entre indivíduos com TDAH e do grupo controle. As *accuracies* das classes *XGBoost* foram de 97,4% e 94,5% para as classes Controle e TDAH, respectivamente.

A *accuracy* total da Votação Majoritária foi maior que aquelas atingidas pelos demais meta-classificadores, graças à obtenção da maior *accuracy* da classe de TDAH (99,2%), apesar de obter a menor *accuracy* no grupo controle: 96,0%. A maior *accuracy* do grupo controle foi obtida pelo *AdaBoost*, 97,7%. O meta-classificador *Ridge* alcançou o maior equilíbrio entre as *accuracies* das classes.

Apesar da Votação Majoritária Ponderada ser a técnica de fusão mais simples empregada nesta Tese, obteve-se com este método a maior *accuracy* na classificação entre pacientes com TDAH e indivíduos do grupo controle. Isto pode ter ocorrido porque as técnicas de *boosting* (*AdaBoost* e *XGBoost*) foram desenvolvidas para aumentar o desempenho dos classificadores-primários considerados como alunos fracos, ou seja, aqueles que apresentam baixos índices de classificações corretas. Entretanto, no caso da classificação entre os grupos TDAH e Controle, quase todos os classificadores-primários são alunos fortes com *accuracies* acima de 90,0% (Figura 44). O trabalho de (WICKRAMARATNA *et al.*, 2001) demonstrou que o uso de técnicas de *boosting* com alunos fortes pode provocar degradação no desempenho. O uso de *XGBoost* e *AdaBoost* foi investigado porque, na classificação de subtipos, temos o cenário oposto, onde os estimadores primários são todos alunos fracos, como foi, anteriormente, mostrado na Figura 45.

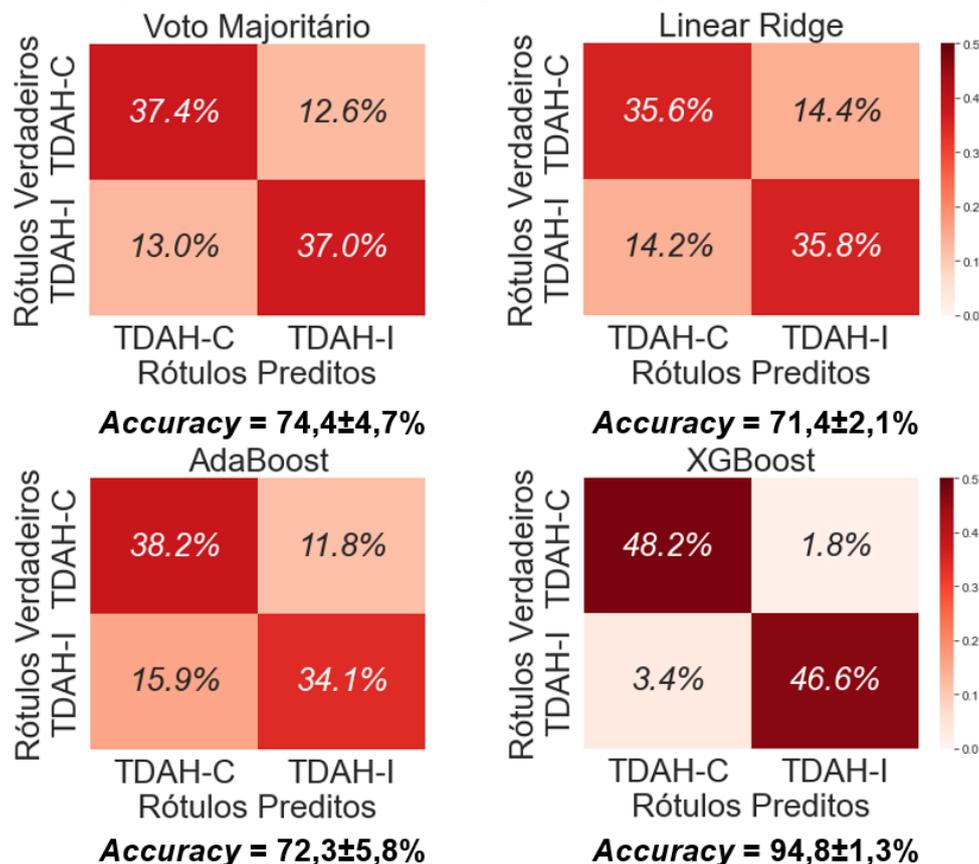
Na classificação de subtipos, as *accuracies* de treinamento e de teste da Votação Majoritária Ponderada foram de $99,9 \pm 0,34\%$ e $74,5 \pm 4,7\%$, respectivamente. Embora a Votação Majoritária não possua uma etapa de treino, esta obteve maior sucesso quando aplicada ao conjunto de treinamento. A diferença de *accuracy* entre as classes TDAH-C e TDAH-I foi de apenas 0,9%, favorecendo a predição de TDAH-C.

O classificador linear *Ridge* obteve uma *accuracy* de treinamento de $92,0 \pm 0,86\%$ e uma *accuracy* no conjunto de teste de $71,3 \pm 2,1\%$ na classificação de subtipos. Portanto, este meta-classificador apresentou um *overfitting* considerável. Este *ensemble* apresenta uma pequena discrepância entre as *accuracies* das classes TDAH-C e TDAH-I que favorece o grupo TDAH-I em apenas 0,3%.

A Figura 48 apresenta as matrizes de confusão obtidas pela fusão das informações extraídas pelas CNNs de cada rede de conectividade na classificação de subtipos. O conjunto de teste utilizou 118 amostras de TDAH-I e 92 amostras de TDAH-C (30% do total de amostras), antes da aplicação da técnica de sobreamostragem SMOTE. Esta técnica

aumentou o número de amostras da classe minoritária gerando novas amostras, de forma a igualar o número de amostras da classe majoritária ($n = 118$).

Figura 48 – Matrizes de confusão obtidas pelas técnicas de fusão por *ensemble* na classificação de subtipos, no conjunto de teste, a partir das imagens de fMRI.



Fonte: O Autor.

Para a classificação de subtipos, *AdaBoost* alcançou $100 \pm 0, 0\%$ de *accuracy* no treinamento e $70, 9 \pm 5, 8\%$ de *accuracy* no teste. O *AdaBoost* apresentou a maior discrepância nas *accuracies* das classes: 11,2%, beneficiando a classificação TDAH-C. O *AdaBoost* obteve apenas 65,3% de *accuracy* na classificação de TDAH-I. Esta foi a pior *accuracy* para a classe TDAH-I, dentre todas as *accuracies* obtidas pelos meta-classificadores. No entanto, a *accuracy* do *AdaBoost* para o grupo TDAH-C foi de 76,5%, ou seja, um pouco maior do que as *accuracies* atingidas pelo *Ridge* e pela Votação Majoritária Ponderada para esta classe.

O desempenho do *XGBoost* superou os demais meta-classificadores na classificação de subtipos, atingindo $94, 9 \pm 1, 3\%$ no teste e $100 \pm 0, 06\%$ nos conjuntos de treinamento. As ferramentas de otimização do *XGBoost* conseguiram evitar o *overfitting*, que prejudicava a *accuracy* dos outros meta-classificadores. No que diz respeito ao balanceamento de *accuracy* entre as classes, o *XGBoost* tem uma tendência ligeiramente maior para classificar corretamente o grupo TDAH-C (96,5% de *accuracy*) do que o grupo TDAH-I (93,2%

de *accuracy*). Mesmo assim, as *accuracies* obtidas pelo *XGBoost* para qualquer das classes superou as métricas correspondentes alcançadas pelos demais meta-classificadores.

Os desvios padrões apresentados pelos meta-classificadores para a classificação entre os grupos TDAH e Controle é no máximo de 2,3% para o *Ridge* e no mínimo de 1,5% para o *AdaBoost*. Para a classificação de subtipos, o *XGBoost* possui um desvio padrão de apenas 1,3%, o que indica uma maior robustez do que o *AdaBoost*, cujo desvio padrão é de 5,8%.

O *overfitting* para cada modelo de fusão em cada cenário de classificação (TDAH/Controle ou subtipo) é retomado na Tabela 31. Nenhuma das técnicas foi superajustada para a classificação TDAH/Controle; entretanto, no diagnóstico de subtipo, apenas o *XGBoost* evitou o *overfitting* e alcançou uma métrica robusta. A Tabela 31 mostra uma comparação das medidas de *overfitting* para os conjuntos em cada cenário.

Tabela 31 – *Overfitting* em cada modelo de fusão de redes de conectividade de fMRI.

Método de Fusão	TDAH/Controle (%)	TDAH-C/TDAH-I (%)
Voto Majoritário Ponderado	2,40	25,4
Classificador <i>Ridge</i>	3,71	20,6
<i>AdaBoost</i>	4,83	21,1
<i>XGBoost</i>	0,05	5,14

A Tabela 32 apresenta as métricas de *accuracy*, precisão e *recall*, para a classificação TDAH/Controle, atingidas nos conjuntos de teste e de treinamento para cada um dos meta-classificadores.

Tabela 32 – Métricas obtidas por cada modelo de fusão de fMRI para diagnóstico TDAH/Controle.

Método	Acc. Treino (%)	Acc. Teste (%)	Precisão Treino (%)	Precisão Teste (%)	Recall Treino (%)	Recall Teste (%)
Voto Maj.	99,98	97,59	99,98	96,17	99,98	99,19
<i>Ridge</i>	100,0	96,29	100,0	97,06	100,0	95,48
<i>AdaBoost</i>	100,0	95,52	100,0	97,66	100,0	93,34
<i>XGBoost</i>	100,0	95,99	100,0	97,48	100,0	94,52

Na Tabela 32, as métricas obtidas no conjunto de treinamento pelos diferentes modelos são quase idênticas, porém ocorreram algumas diferenças quando as técnicas foram

aplicadas no conjunto de teste. Observou-se que embora o Voto Majoritário tenha obtido a maior *accuracy*, a precisão obtida por este modelo é menor do que aquelas alcançadas pelos demais. O *AdaBoost* obteve a maior precisão na classificação TDAH/Controle, embora tenha este método tenha apresentado o menor *recall*. O maior *recall* de teste foi obtido através do Voto Majoritário.

A Tabela 33 mostra as métricas de *accuracy*, precisão e *recall* atingidas nos conjuntos de teste e de treinamento para cada um dos meta-classificadores para a classificação dos subtipos. Neste caso, o *Ridge* apresentou menores valores também para as métricas de treinamento. Este meta-classificador apresentou valores similares para *accuracy*, precisão e *recall* de teste.

Tabela 33 – Métricas obtidas por cada modelo de fusão de fMRI para diagnóstico de subtipos (TDAH-C/TDAH-I).

Método	Acc. Treino (%)	Acc. Teste (%)	Precisão Treino (%)	Precisão Teste (%)	Recall Treino (%)	Recall Teste (%)
Voto Maj.	99,87	74,50	99,90	75,43	99,83	74,10
<i>Ridge</i>	91,96	71,32	90,88	71,45	93,29	71,52
<i>AdaBoost</i>	100,0	70,88	100,0	73,77	100,0	67,27
<i>XGBoost</i>	100,0	94,86	100,0	96,46	100,0	93,23

O *AdaBoost* apresentou os mais baixos *recall* e *accuracy* no conjunto de teste devido ao *overfitting*, porém com uma precisão um pouco acima da encontrada pelo *Ridge*.

O *XGBoost* apresentou valores superiores aos demais em todas as métricas no teste, embora seja notável que a sua precisão (96,46%) esteja acima de sua *accuracy* (94,86%) e de seu *recall* (93,23%) no teste. Isto ocorre porque a precisão depende apenas dos VP e FP, enquanto o *recall* depende dos VP e FN, conforme expresso, anteriormente, nas Equações 31 e 32, respectivamente. E, a *accuracy* depende tanto dos VP e FP como dos VN e FN, como descrito na Equação 29.

Desta forma, a obtenção de uma precisão maior do que o *recall* e a *accuracy* indica que o modelo produz uma menor quantidade de FP do que de FN.

5 DISCUSSÃO

Primeiramente, nesta seção, discute-se o conteúdo comum aos experimentos de classificação de TDAH baseados em EEG e em fMRI.

Em EEG, as etapas de pré-processamento e de controle de qualidade reduziram a quantidade de participantes da base de dados *Healthy Brain Network* utilizado no presente estudo de 856 para 499 indivíduos. Após as exclusões, o estudo passou a contar com as seguintes quantidades de participantes para cada classe: 149 para o grupo controle, 169 para a classe TDAH-C e 181 para o grupo TDAH-I. A classe TDAH consiste na união dos subgrupos TDAH-C e TDAH-I, os quais somam 350 participantes.

Em fMRI, o controle de qualidade foi responsável por reduzir a quantidade de sujeitos de 809 para 632. Após as remoções de participantes, esta base de dados contou com 109 indivíduos do grupo controle e 523 com TDAH, dos quais 244 possuem TDAH-C e 279 tem TDAH-I.

Para a classificação baseada nos sinais de EEG, experimentou-se como classificadores primários as RFs e as SVMs para cada uma das faixa de frequência: banda larga (0-100 Hz), Delta (0,5-3,9 Hz), Teta (4-7,5 Hz), Alfa (8-13 Hz), Beta (13,1-26 Hz) e Gama (32-45 Hz). Cada uma destas bandas de frequência foi analisada por uma RF e por uma SVM como um possível biomarcador tanto para a classificação entre TDAH e grupo controle como para identificar o subtipos.

Antes de enviar os dados de EEG aos classificadores primários, aplicou-se uma sequência de etapas de pré-processamento. Nestas etapas, removeu-se a interferência da rede; canais e segmentos ruidosos; artefatos de ECG, EMG e EOG. Realizou-se, também, a separação de fonte dos canais com ICA e segmentou-se os sinais de EEG em janelas de 2 segundos. Filtrou-se os sinais, extraiu-se características e normalizou-se o vetor de características resultante. Após o balanceamento de amostras, aplicou-se uma PCA para selecionar as componentes mais relevantes, as quais foram enviadas aos classificadores.

Em todos os experimentos de classificação de EEG, as amostras foram aleatorizadas e divididas em conjuntos de treinamento e de teste nas proporções de 70% e de 30%, respectivamente. Este procedimento foi realizado novamente cada vez que os códigos foram executados, e estes foram executados 20 vezes para obter as *accuracies* médias.

Para a classificação entre pacientes com TDAH e integrantes do grupo controle, tanto o *XGBoost* como o *Ridge* proporcionaram um aumento de *accuracy*, superando o melhor resultado obtido através do uso de uma única RF (91,1%). O incremento de *accuracy* proporcionado pelos meta-classificadores foi de 0,7% a 1,3%.

Na classificação de subtipos, o uso de meta-classificadores aumentou a *accuracy* de 4,5% a 11,2% em relação ao máximo que havia sido obtido por uma RF (86,6%).

Os meta-classificadores *XGBoost* alcançaram as maiores *accuracies* a partir de sinais de EEG para as classificações de TDAH/Controle ($92,4 \pm 0,5\%$) e de subtipos ($97,8 \pm 0,3\%$), respectivamente.

De forma análoga, em fMRI, as redes de conectividade visual, somatomotor, de atenção ventral, de atenção dorsal, límbica, frontoparietal, *default* e outros *voxels* foram analisadas como possíveis biomarcadores. Cada uma das redes de conectividade (bem como o mapa de Yeo completo) foi enviada a uma CNN otimizada por AG para lidar especificamente com a área cerebral correspondente. Avaliou-se o uso de uma arquitetura de CNN, projetada, especialmente, para estes dados.

Na análise das redes de conectividade, os resultados obtidos pela arquitetura CNN projetada foram corroborados pela arquitetura LeNet-5 que obteve *accuracies* similares para cada rede. Contudo, na rede de atenção dorsal, a LeNet-5 obteve maior sucesso para classificar os indivíduos entre TDAH e grupo controle. Embora as *accuracies* obtidas através da classificação das redes de conectividade pela LeNet-5 sejam em média um pouco maiores, as CNNs com a arquitetura projetada obtiveram *overfittings* e desvios-padrão menores. Por isto, selecionou-se as CNN com a arquitetura projetada para atuarem como classificadores primários nos *ensembles*.

O desempenho na análise do mapa cerebral completo (matriz de correlações) da CNN, cuja topologia foi projetada nesta Tese, foi comparado com o da LeNet-5, da VGG-16 e da ResNet-50 (APÊNDICE E). Para a classificação entre indivíduos com TDAH e do grupo controle, a CNN projetada obteve $96,3 \pm 0,6\%$ de *accuracy*, superando as *accuracies* de $78,2 \pm 2,3\%$, $51,7 \pm 7,4\%$ e $71,1 \pm 2,2\%$ obtidas, respectivamente, com as arquiteturas LeNet-5, VGG-16 e ResNet-50. Além disso, a incerteza da métrica (desvio padrão) foi menor para a CNN projetada.

Na classificação de subtipos a partir do mapa cerebral completo, a CNN projetada obteve apenas $69,0 \pm 4,8\%$, uma *accuracy* que está aquém da obtida pelas arquiteturas LeNet-5 e ResNet-50: $71,8 \pm 9,8\%$ e $69,7 \pm 6,1\%$, respectivamente, e um pouco maior que a da VGG-16 ($68,3 \pm 6,5\%$). Entretanto, a medida de incerteza foi menor usando a a CNN projetada.

Para unir as informações das redes de conectividade medidas por fMRI, utilizaram-se os mesmos meta-classificadores aplicados para a fusão das bandas de EEG: o Voto Majoritário ponderado, o classificador linear *Ridge*, o *AdaBoost* e o *XGBoost*. Novamente, comparou-se o uso destas técnicas. Como em EEG, os meta-classificadores, também,

foram otimizados.

Em todos os experimentos de fMRI, utilizou-se a validação cruzada *k-fold* com $k = 10$. O *k-fold* foi estratificado de forma a manter as proporções entre as classes nos conjuntos de treinamento e de teste que receberam 70% e 30%, respectivamente, das amostras sorteadas aleatoriamente.

Os meta-classificadores (Voto majoritário e *XGBoost*) apresentaram as melhores *accuracies*: $97,6 \pm 2,1\%$ e $94,9 \pm 1,3\%$ para o diagnóstico entre participantes com TDAH e do grupo controle e para a identificação de subtipos a partir da fusão das redes de conectividade de fMRI, respectivamente.

Na maioria dos casos, considerando tanto os experimentos de EEG como de fMRI, o *XGBoost* superou os demais meta-classificadores, obtendo maiores *accuracies*. O *XGBoost* costuma ter melhor desempenho por ser um método mais complexo e possuir uma maior gama de ferramentas de otimização.

Na classificação TDAH/Controle com fMRI, o Voto Majoritário ponderado superou os demais alcançando precisão e *recall* de 96,2% e de 99,2%, respectivamente, além da *accuracy* de $97,6 \pm 2,1\%$. Possivelmente, Voto Majoritário teve maior sucesso neste caso porque os classificadores primários já possuíam *accuracies* elevadas, constituindo "alunos fortes". As ferramentas dos métodos de *boosting* (*AdaBoost* e *XGBoost*) foram desenvolvidas para alavancar resultados de "alunos fracos", ou seja, de classificadores primários que produzem baixas *accuracies* (SIPPER & MOORE, 2022).

Para o caso da classificação de subtipos em EEG, o meta-classificador com maior *accuracy* foi o *XGBoost*, o qual obteve *accuracy*, precisão e *recall* de $97,8 \pm 0,3\%$, de 97,7% e de 97,8%.

Os métodos de fusão não apenas aumentaram as *accuracies*, em comparação com os classificadores de banda única, mas também diminuíram bastante o *overfitting* observado. Portanto, o uso de informações conjuntas das previsões dos classificadores primários (as RFs para cada banda de frequência e as CNNs para cada rede de conectividade) aumentou a capacidade de generalização dos modelos.

5.1 Análise do Significado dos Biomarcadores

Nesta seção discutem-se as associações entre os biomarcadores encontrados pelos classificadores e sua relação com os sintomas de TDAH, os quais foram relatados no APÊNDICE F desta Tese.

Os meta-classificadores indicaram uma alta relevância da rede de conectividade límbica tanto para identificar a presença de TDAH como para distinguir seus subtipos. O sistema límbico conecta os córtex cerebrais ao tronco cerebral (BROCA, 1978) (?), ou seja, conecta os pensamentos e emoções ao controle das funções básicas corporais. O sistema límbico é formado por algumas redes menores que trabalham em conjunto embora

possuam funções específicas. Dentre estas estruturas, destacam-se o tálamo, a amígdala, o hipocampo, o hipotálamo e o núcleo *accumbens* (MORGANE *et al.*, 2005) (CATANI *et al.*, 2013) (?). A rede límbica possui ainda uma estrutura para processar o olfato, chamada de córtex olfatório piriforme primário (CATANI *et al.*, 2013). Um estudo recente associou TDAH com a dificuldade em identificar odores (MERCAN ISIK *et al.*, 2023).

O tálamo recebe informações sensoriais do corpo e decide o que merece ou não a atenção do sujeito (KAMALI *et al.*, 2023). Caso o tálamo decida que as informações recebidas são importantes, estas são repassadas às regiões cerebrais correspondentes.

A amígdala é responsável por processar as emoções como medo, afetividade, agressividade, excitação, ansiedade e prazer. A amígdala também processa a função de aprendizado por associação, por exemplo, sensações olfativas de comida com sabores percebidos pelo paladar (?). A amígdala possui funções relacionadas à atenção, interações sociais e tomada de decisões. Esta estrutura está conectada aos lobos temporal e occipital (KAMALI *et al.*, 2023).

O hipotálamo produz hormônios que promovem o equilíbrio corporal controlando funções como a temperatura, o ciclo do sono, a adrenalina e a fome. O hipotálamo conecta as emoções (geradas pela amígdala) às sensações físicas. O hipotálamo está conectado aos lobos occipital, temporal, parietal e frontal, além do cerebelo (KAMALI *et al.*, 2023).

O hipocampo transforma as memórias de curto prazo em memórias de longo prazo e, por isso, está associado ao aprendizado (SQUIRE, 1992). O núcleo *accumbens* está relacionado às recompensas, ao prazer e ao vício.

Então, o sistema límbico pode estar conectado aos sintomas de TDAH em razão de alterações na amígdala que pode estar gerando excitação e ansiedade em demasia. A ansiedade contribui para deixar os pacientes mais agitados e sem paciência para esperar por momentos adequados, o que se conecta com os sintomas de hiperatividade.

O tálamo de pessoas com TDAH pode decidir com uma maior frequência ou intensidade quais estímulos sensoriais são importantes e, portanto, não devem ser ignorados. Então, a amígdala produz a excitação que torna o estímulo mais atraente ao indivíduo. Outra possibilidade é que a amígdala de sujeitos com TDAH receba do tálamo os estímulos com a mesma frequência, mas produza respostas mais intensas. As emoções mais intensas de excitação ou de alerta são mais difíceis de ser ignoradas pelos pacientes que reagem a estas e perdem o foco em suas atividades (sintoma de desatenção).

Os sintomas de perda de objetos necessários para terminar tarefas; esquecimento de atividades diárias; e de cometer erros em atividades devido à falta de atenção em detalhes podem estar conectados às alterações no hipocampo. A dificuldade em transformar memórias de curto prazo em memórias de longo prazo provoca esquecimentos que podem ser o motivo da perda dos objetos. O esquecimento de detalhes em instruções pode ser a razão dos erros atribuídos à falta de atenção ou à dificuldade de organizar tarefas sequenciais.

O paciente TDAH-I pode se sentir ansioso em responder aos estímulos, mas não

sente a necessidade de se mexer constantemente, como os TDAH-C e TDAH-H. Isto pode ocorrer por a amígdala produzir ansiedade em uma maior intensidade em pacientes com TDAH-C e TDAH-I. Outra teoria seria que hipotálamo alterado enviaria uma maior quantidade de hormônios ao corpo ou estes seriam enviados à partes diferentes do corpo, o que produziria a necessidade de reações físicas mais enérgicas para manter o equilíbrio corporal.

Os meta-classificadores Voto Majoritário, *XGBoost* e *AdaBoost* apontaram uma contribuição da rede somatomotor, também conhecida como sensorimotor, na classificação entre TDAH e grupo controle. Apenas o *XGBoost* indicou a relevância da rede somatomotor para distinguir subtipos. Contudo, o *XGBoost* obteve *accuracy* mais de 20% maior que os demais métodos neste experimento. Este resultado evidencia que esta rede de conectividade é relevante tanto para a detecção de TDAH como para a identificação de subtipos.

A rede somatomotor é responsável por ligar o cérebro ao restante do corpo, processando estímulos sensoriais e enviando informações de controle do músculos e órgãos (RIECANSKY & LAMM, 2019). Esta região é composta pelo córtex somatossensorial e o córtex motor. O estudo de (HERMAN *et al.*, 2020) relacionou alterações na rede somatomotor com o comportamento impulsivo. A rede somatomotor possui uma função crucial no sono: quando esta é acionada mais frequentemente ao longo do descanso, a qualidade do sono decai (BAI *et al.*, 2022).

O córtex somatossensorial processa as informações dos receptores sensoriais do corpo, como, por exemplo, vibração, toque, detecção de consistência de materiais, temperatura, dor, coceira e sensação de deslizamento. Há uma central de memória somatossensorial que é conectada à amígdala e ao hipocampo da rede límbica e auxilia a reforçar as memórias.

Além de gerar os impulsos que controlam diretamente os movimentos, o córtex motor possui diversas funções relacionadas ao planejamento do movimento. Este córtex é responsável por decidir se fará um movimento, planejar sua trajetórias, velocidade, extensão, duração. Para isso, esta região precisa calcular a orientação espacial e a coordenação motora do corpo. O córtex motor inclui sistemas de recebimento de *feedback* e de aprendizado com erros em movimentos (KIM *et al.*, 2021).

Alterações no córtex somatossensorial podem estar conectadas aos sintomas de desatenção de TDAH. Pois, se este córtex reage com menor intensidade para os pacientes com TDAH do que para o grupo controle, o tálamo pode decidir que esta informação não é suficientemente importante para que a amígdala reaja e conecte a informação a uma emoção ou para que o hipocampo grave a informação nas memórias de longo prazo.

O córtex motor pode ter apresentado diferenças para os pacientes com TDAH-C por ser mais requisitado em um paciente com hiperatividade do que em um que não tenha estes sintomas. Há a possibilidade de os pacientes desenvolverem alterações nesta rede

de forma a compensar as alterações na rede límbica, uma vez que os meta-classificadores apontam uma menor relevância da rede somatomotor em relação à rede límbica. A CNN projetada e a LeNet-5 obtiveram *accuracies* similares processando dados de cada uma destas duas redes de conectividade.

A rede de conectividade frontoparietal foi indicada como um fator importante para a classificação entre grupos TDAH e controle pelos classificadores *Ridge* e *AdaBoost*.

A rede frontoparietal tem funções vinculadas à atenção sustentada, à resolução de problemas complexos e à memória de trabalho. Esta rede manipula as memórias de trabalho em raciocínios complexos para a tomada de decisões direcionadas a objetivos (MENON, 2011). A memória de trabalho é a capacidade de memorizar informações temporariamente (SHAH & MIYAKE, 1999).

A rede frontoparietal possui conexões com as redes *default* e com a rede de atenção dorsal nas funções de processos introspectivos e de atenção perceptiva, respectivamente (DIXON *et al.*, 2018). A rede frontoparietal influencia na habilidade de coordenar o comportamento de forma flexível e rápida para atingir um determinado objetivo (MAREK & DOSENBACH, 2018). Esta rede possui a habilidade de se reconfigurar para adaptação à determinada tarefa (COLE *et al.*, 2013).

O estudo de (MAREK & DOSENBACH, 2018) já havia detectado a conexão entre a rede frontoparietal e o diagnóstico de TDAH.

Alguns dos sintomas de desatenção são o paciente evitar ou não gostar de executar tarefas que exijam a sustentação de esforço mental; e o paciente ter dificuldade de sustentar atenção. Este sintoma pode estar conectado às dificuldades com a atenção sustentada, que é uma das tarefas da rede frontoparietal. Este aspecto tem influência também sobre os sintomas da suscetibilidade às distrações e à perda de atenção ao longo de uma conversa, na qual a partir de determinado momento o paciente parece não estar ouvindo o seu interlocutor. Esta perda de atenção, por sua vez, pode ter como consequência o sintoma no qual o paciente não atenta a detalhes importantes para executar tarefas.

Alterações na rede frontoparietal provocam mudanças na capacidade de resolução de problemas complexos ligados a objetivos. Portanto, este aspecto pode provocar a dificuldade em organizar e gerenciar tarefas e atividades sequenciais, observada em pacientes com TDAH. A dificuldade que os pacientes com sintomas de desatenção apresentam para gerenciar o tempo estaria vinculada à dificuldade em manter a atenção e executar raciocínios ligados à resolução de problemas objetivos complexos.

Para a distinção dos subtipos, o Voto Majoritário (e portanto, a comparação das CNN projetadas) e o *AdaBoost* apontaram a análise da rede frontoparietal como importante. Porém, estes métodos de fusão obtiveram *accuracies* baixas comparada à obtida pelo *XG-Boost*. Isto indica que ou a rede frontoparietal possui diferenças entre TDAH-C e TDAH-I devido a diferenças na intensidade dos sintomas de desatenção. Outra possibilidade seria que as alterações na rede frontoparietal provocam também os sintomas de hiperatividade.

No entanto, considerando as funções da rede frontoparietal, não parece haver conexão entre esta e os sintomas de hiperatividade.

O método *ensemble* que obteve maior sucesso em classificar indivíduos entre os grupos TDAH e controle foi o Voto Majoritário, que designou maiores pesos para as redes visual (também chamada de córtex visual) e *default*.

Após a aprovação do tálamo, a informação visual observada pelos olhos é enviada à rede de conectividade visual (córtex visual), cuja função é o processamento de informações visuais. O córtex visual realiza tarefas de identificação de padrões e identifica se os objetos estão estáticos ou em movimento. A rede visual também processa informação espacial, como, por exemplo, cores, orientação, frequência de movimento de objetos e detecção de bordas.

A explicação para a rede visual apresentar relevância na detecção de TDAH pode estar na maior suscetibilidade dos pacientes a estímulos que frequentemente são visuais. Esta rede, alterada, poderia estar enviando informações um pouco distorcidas sobre uma pequena presença de movimento em objetos estáticos ou detectando estes objetos como se fossem maiores ou mais coloridos que os demais. Estas informações distorcidas poderiam provocar um aumento na necessidade de atenção que estes objetos necessitam receber, conduzindo à perda do foco na atividade anterior.

Outra possibilidade seria a presença de mecanismos compensatórios no córtex visual desenvolvidos pelos sujeitos com TDAH.

A rede *default*, também chamada de rede de modo padrão (*default mode network*), costuma estar ativa quando um indivíduo está acordado e em descanso. Está associada à recordação de memórias (SPRENG *et al.*, 2009), ao sonhar acordado (*daydreaming*) (PHILIPPI *et al.*, 2021), a devaneios e à imaginação (CARROLL, 2020). A rede *default* processa a teoria da mente (SPRENG *et al.*, 2009) (HUGHES *et al.*, 2019), como são chamadas as informações que possibilitam entender as crenças, desejos, intenções, emoções e pensamentos de outros indivíduos e, portanto, é fundamental para as interações sociais.

As modificações na rede de modo padrão (*default*) de indivíduos com TDAH pode elevar a ocorrência de processos de recordação, *daydreaming*, devaneios e imaginação. Estes processos podem, então, surgir em momentos inadequados, enquanto interlocutores falam ou enquanto o indivíduo está envolvido em outras tarefas, produzindo uma fonte de distrações.

A conexão da rede *default* com o entendimento de outras mentes em relações sociais pode estar vinculada aos sintomas da desatenção e aos problemas sociais relatados por pessoas com TDAH (KOFER *et al.*, 2011).

De forma análoga ao observado em fMRI, em EEG, os classificadores obtiveram *accuracies* superiores quando combinaram informações de mais de uma faixa de frequência. Isto mostra que o TDAH provoca alterações, em relação ao grupo controle, em mais de uma faixa de frequência de EEG. Portanto, observa-se, o TDAH altera múltiplas regiões

do cérebro (o que, também, foi constatado na análise dos exames de fMRI). Portanto, há mais de uma faixa de frequência que serve como biomarcador.

Constatou-se que a combinação do uso de características observadas em múltiplas faixas de frequência, também, elevou as *accuracies* obtidas para o caso da classificação de subtipos, de forma similar ao observado em fMRI. Todos os meta-classificadores aumentaram as suas *accuracies* combinando informações de bandas em relação aos classificadores que realizaram o diagnóstico somente com as informações de uma única banda de frequência (mesmo com a banda larga que possui informações de todas as demais).

Se a identificação dos subtipos é influenciada por mais de uma banda, isto indica que mais de uma banda é alterada pelo transtorno e, por isso, serve como um biomarcador da diferença entre os subtipos de TDAH.

Na distinção entre as classes TDAH e Controle, os meta-classificadores obtiveram *accuracies* similares entre si. Esta semelhança nas métricas de *accuracy* foi observada, especialmente, para o *XGBoost* (92,4%), *Ridge* (91,8%) e *AdaBoost* (90,4%). O uso do Voto Majoritário obteve uma *accuracy* um pouco menor (87,2%). Contudo, os meta-classificadores obtiveram estes resultados semelhantes atribuindo conjuntos de pesos diferentes a cada banda de frequência.

Para a identificação binária de TDAH, os meta-classificadores *XGBoost* e *Ridge* priorizaram a análise da RF baseada na banda Delta e apresentaram pequenas diferenças nas atribuições de pesos para as demais bandas.

O *AdaBoost* obteve um sucesso apenas um pouco menor na classificação entre pacientes com TDAH e voluntários do grupo controle, reduzindo o peso da banda Delta, e, em compensação, aumentando a importância das bandas Alfa e Gama. Já o Voto Majoritário combinou informações principalmente das faixas Teta, Delta e Alfa.

Portanto, as únicas faixas apontadas como importantes para a identificação de TDAH por mais de um meta-classificador foram a banda Delta e a banda Alfa.

A banda Delta foi apontada como um indicador relevante na detecção de TDAH e, também na identificação de subtipos. A banda Delta é associada ao sono profundo, mas pode estar presente em tarefas que requerem atenção sustentada (KIRMIZIALSAN, 2006). A banda Delta está relacionada à região occipital do cérebro (SANEI, 2021).

A ocorrência das ondas na banda Alfa costuma ser detectada quando os olhos estão fechados e é associada ao relaxamento e ao controle inibitório. Estas ondas se originam na parte posterior do cérebro, especialmente, na região occipital (SANEI, 2021), a qual costuma ser associada aos sintomas de hiperatividade.

A banda Gama costuma ser associada com reações do sistema (ou córtex) somatosensorial a estímulos visuais e auditivos (SANEI, 2021). Quanto maior o estímulo, maior a resposta Gama em indivíduos com desenvolvimento típico. Isto corrobora com o que foi constatado em fMRI, onde a rede somatomotor, onde se encontra o córtex somatosensorial, foi apontada pelos meta-classificadores como um biomarcador de TDAH. Portanto,

isso evidencia que a suscetibilidade à distrações apresentada pelos pacientes com TDAH pode estar no nível de estímulo recebido pelo sistema somatossensorial ou no nível da resposta ao estímulo que este córtex envia às demais regiões cerebrais.

A banda Gama também é associada à memória somatossensorial, já que estímulos já conhecidos por um sujeito acionam respostas mais fortes nesta faixa (HERRMANN *et al.*, 2010).

Para a classificação de subtipos, os meta-classificadores destacaram a importância da análise de diferenças nas bandas Delta, Teta e Gama.

A análise da banda Delta foi apontada como a mais importante pelos métodos *Ridge* e *XGBoost*. Estes meta-classificadores apontaram pesos semelhantes para as demais bandas, embora o *Ridge*, notadamente, tenha elevado um pouco os pesos das faixas Alfa e Teta em relação ao *XGBoost*.

O *AdaBoost* elevou a relevância da banda Teta e, em menor medida, da banda Delta em relação às suas demais atribuições de pesos. Para o Voto Majoritário, a análise das faixas Teta e Gama apresentaram maior relevância. As *accuracies* obtidas pelos métodos Voto Majoritário, *Ridge*, *AdaBoost* e *XGBoost* foram de 91,1% 97,7%, 96,9% e 97,8%.

A banda Teta costuma ser observada na região frontal do cérebro e pode estar relacionada aos sonhos acordados, os quais costumam também ser atribuídos à rede *default*, e são comuns àqueles que possuem sintomas de desatenção de TDAH. A banda Teta está conectada ao inconsciente, a inspiração criativa, a meditação profunda e à excitação (SANEI, 2021). A banda Teta também se manifesta no hipocampo (parte da rede límbica) (MORGANE *et al.*, 2005). Portanto, as alterações na banda Teta de pacientes com TDAH, e mesmo entre seus subtipos, estão relacionadas com as diferenças nas redes límbica (principalmente) e com a rede *default*, as quais também haviam sido apontadas como biomarcadores pela análise dos sinais de fMRI.

As diferenças na banda Teta entre as classes TDAH e o grupo controle podem estar vinculadas à intensidade e frequência destes sonhos acordados que contribuem para a perda do foco e em diferenças no processo de transformar as memórias de curto em longo prazo.

Os biomarcadores analisados no presente trabalho poderão auxiliar futuros estudos a pesquisar o processo cerebral que provoca o TDAH e as variações do processo cerebral que culminam nos diferentes subtipos de TDAH. Observar os biomarcadores detectados também poderá auxiliar clínicos a monitorar a evolução dos tratamentos de TDAH, bem como desenvolver novos tratamentos específicos para cada subtipo de TDAH. Ou seja, estes indicadores podem ser úteis para desenvolver novos tratamentos orientados à redução das diferenças de biomarcadores entre TDAH e grupo controle.

5.2 Cenários de Classificação

Tanto em EEG como em fMRI, alguns autores, como (QURESHI *et al.*, 2017), (SAHA & SARKAR, 2022) e (AHMADI *et al.*, 2021), decidiram avaliar a identificação dos subtipos de TDAH, simultaneamente, com o diagnóstico do transtorno. Ou seja, trabalharam com ferramentas de classificação multi-classe que avaliaram a possibilidade de cada sujeito pertencer aos subtipos, TDAH-C ou TDAH-I, ou ao grupo controle. Realizar estas tarefas simultaneamente, aumenta o desafio para os classificadores. Contudo, não há necessidade clínica de diagnosticar-se a presença de TDAH e distinguir-se, simultaneamente, seu subtipo específico.

Por isso, nesta Tese propôs-se uma abordagem diferente, na qual a identificação entre classes TDAH e Grupo Controle (classificação binária) é realizada separadamente da classificação de subtipos. Esta separação influencia na melhora do desempenho dos classificadores primários, fornecendo uma base mais sólida para a classificação final realizada pelos métodos *ensemble*, assim como eleva a interpretabilidade dos biomarcadores encontrados.

A configuração dos hiperparâmetros dos classificadores necessários para otimizar os dois modelos (de classificação binária e de subtipos) é diferente. Isto foi observado tanto para os modelos que trabalharam com dados de EEG como para aqueles que analisaram dados de fMRI, como foi mostrado no capítulo de Materiais e Métodos. Portanto, executar os AGs para melhorar ambos os modelos, simultaneamente, poderia resultar em hiperparâmetros intermediários que não alcançariam o resultado ótimo para as classificações referentes.

A classificação binária e de subtipos simultânea poderia impactar tanto no desempenho dos métodos para a classificação como na busca por biomarcadores. A otimização das *accuracies* totais dos modelos de classificação multi-classe por meio do AG poderia priorizar elevar *accuracies* em apenas uma das comparações entre as classes: TDAH-C *versus* TDAH-I, ou TDAH-C *versus* Controle ou TDAH-I *versus* Controle. Ou, então, o AG poderia priorizar a obtenção de uma *accuracy* que não fosse a melhor em cada caso, mas que permitiria a obtenção de um resultado razoável em cada comparação. Além disso, a presente Tese verificou que diversos marcadores biológicos que distinguem os sujeitos com TDAH do grupo controle não são os mesmos que apontam a qual subtipo um determinado sujeito pertence. Em uma classificação multi-classe, os biomarcadores de cada tarefa (classificação binária e de subtipos) poderiam ofuscar-se mutuamente, dificultando a otimização e a análise dos resultados.

5.3 Medida de desempenho dos classificadores

Os sinais de EEG e de fMRI costumam ser aplicados para diagnóstico de TDAH após suspeita de TDAH pelos pais, professores ou clínicas e, frequentemente, após testes clí-

nicos baseados em perguntas sobre o comportamento. O diagnóstico incorreto de TDAH pode conduzir um indivíduo do grupo controle a utilizar medicamentos desnecessários, os quais podem provocar-lhe efeitos colaterais.

Uma pessoa com TDAH não medicada sofrerá todos os sintomas de TDAH e os riscos associados ao transtorno, caso seja diagnosticada como pertencente ao grupo controle. O diagnóstico incorreto do subtipo (TDAH-C/TDAH-I), também, poderá provocar consequências indesejadas.

Por exemplo, um paciente com TDAH-C que não receber medicação para comportamento hiperativo, por ter sido erroneamente classificado como TDAH-I, continuará exibindo os sintomas de hiperatividade. Uma pessoa com TDAH-I pode exibir uma apatia ou tranquilidade exagerada ao receber medicamentos para uma hiperatividade inexistente em seu quadro.

Portanto, é importante obter classificadores com métricas balanceadas para ambas as classes tanto no diagnóstico de TDAH como na identificação de seus subtipos. Por isso, nesta Tese, utilizou-se a *accuracy* como principal métrica para a otimização dos classificadores primários e meta-classificadores. Contudo, também foram informadas as métricas de precisão e de *recall*, as quais são importantes para os possíveis usuários do sistema desenvolvido.

5.4 Comparação com os principais trabalhos de classificação de TDAH

A seguir, comparam-se os resultados obtidos nesta Tese com trabalhos distintos de outros autores. Contudo, é necessário ressaltar que estas comparações são indiretas. Pois, cada grupo de pesquisa realizou seu estudo com dados de bases diferentes, as quais possuem quantidades diversas de dados em cada classe. Além disso, há grandes variações nos métodos de coleta de dados de cada base.

Cada estudo apresenta, também, variações nas técnicas de balanço de dados. Há autores que nem sequer empregam técnicas de balanço destas bases de dados, mesmo quando existem diferenças consideráveis nos tamanhos amostrais de cada classe. Os trabalhos nesta área, também, apresentam diferenças relevantes no controle de qualidade e nos critérios de exclusões de indivíduos, bem como nas etapas de pré-processamento. Observou-se, na etapa experimental desta Tese, as diferenças no desempenho dos classificadores provocadas por acrescentar passos de pré-processamento e pela otimização através dos AGs. Estes fatores podem impactar o desempenho dos classificadores.

Para fins de organização, separou-se as temáticas do restante das do trabalho apresentado nesta Tese com o estado da arte em seções relativas à EEG e à fMRI.

5.4.1 Comparação com os principais trabalhos de classificação de TDAH com base em EEG

Nos experimentos de classificação com meta-classificadores a partir das características extraídas de EEG, no presente estudo, utilizou-se uma quantidade de 499 sujeitos. Esta quantidade é maior do que todos os principais trabalhos do estado da arte, o que pode ser verificado na Tabela 34. Esta tabela reproduz os principais trabalhos da Tabela 1 mostrada anteriormente no Capítulo 2 e compara estes trabalhos com o trabalho com EEG apresentado nesta Tese.

Tabela 34 – Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com EEG.

Referência (Número de citações)	Tipo de dado	Técnica	Banco de dados	Accuracies	Contribuição
(VAHID <i>et al.</i> , 2019) 72 citações	EEG em tarefa	EEGNet	25 Controle 18 TDAH-C 15 TDAH-I	83,0% (subtipos)	Observou diferenças no córtex superior parietal
(CHEN <i>et al.</i> , 2019a) 130 citações	rs-EEG	CNN proj.	51 Controle 50 TDAH	94,7% (binária)	Matrizes de conectividade funcional para EEG
(AHMADI <i>et al.</i> , 2021) 62 citações	rs-EEG	CNN proj.	14 Controle 13 TDAH-C 12 TDAH-I	99,5% (subtipos)	Alterações de potência na faixa Delta, Teta e Alfa para TDAH
(KHARE <i>et al.</i> , 2022) 21 citações	EEG em tarefa	VHERS ELM	60 Controle 61 TDAH	97,9% (binária)	Relevância da análise da banda Delta
Presente Tese	rs-EEG	<i>Ensemble</i> <i>XGBoost</i> RFs	149 Controle 350 TDAH Subtipos: 169 TDAH-C 181 TDAH-I	92,4±0,5% (binária) 97,8±0,3% (subtipos)	Biomarcadores: Binária: Delta, Teta, Alfa e Gama Subtipos: Delta, Teta e Alfa

basearam seus estudos em menos de 10 indivíduos de cada classe. Os trabalhos de (PEREDA *et al.*, 2018) e (AHMADI *et al.*, 2021) contaram com menos de 20 sujeitos por classe. O estudo de (VAHID *et al.*, 2019) se baseou em 25 amostras do grupo controle e menos de 20 de cada um dos subtipos (TDAH-C e TDAH-I). O uso de menos de 50 amostras também foi observado nos trabalhos de (ZHANG & LI, 2019), (MOGHADDARI *et al.*, 2020) e (BAKHTYARI & MIRZAEI, 2022). Nenhum dos demais estudos destacados na pesquisa bibliométrica trabalhou com mais de 65 indivíduos por classe.

A pesquisa apresentada neste manuscrito é apoiada por dados de mais participantes do que qualquer outro trabalho que identifique diferenças entre os subtipos de TDAH relatados por (SLATER *et al.*, 2022) em sua revisão. A presente pesquisa utilizou dados de mais que o dobro do número de participantes do trabalho de (DUPUY *et al.*, 2013), que é o maior estudo, em quantidade de indivíduos, relatado em (SLATER *et al.*, 2022). Portanto, a presente pesquisa compreende o estudo de classificação do subtipo TDAH com maior número de sujeitos até o presente momento.

Assim como observado para fMRI, a maioria dos trabalhos de EEG na literatura realiza apenas a classificação entre TDAH e grupo controle. Os estudos revelados na pesquisa bibliométrica que realizaram classificação de subtipos de TDAH foram os de (VAHID *et al.*, 2019), (AHMADI *et al.*, 2021). O trabalho de (PEREDA *et al.*, 2018) realizou a classificação entre voluntários do grupo controle e, somente, pacientes do subtipo TDAH-C.

O trabalho de (BAKHTYARI & MIRZAEI, 2022) desconsiderou o uso da banda Delta e combinou dados de padrões de conectividade dinâmica das demais faixas. A pesquisa de (KHARE *et al.*, 2022) considerou, como características, diferentes formas de calcular a entropia, a média e o desvio padrão e avaliou as demais bandas como biomarcadores melhores do que a banda Delta. Isto se contrapõe ao observado no presente trabalho.

Os meta-classificadores, *Ridge* e *XGBoost*, que obtiveram melhor desempenho na presente Tese apontaram a banda Delta, justamente, como a mais relevante para a classificação entre as categorias TDAH e controle, discordando do que foi apresentado por (BAKHTYARI & MIRZAEI, 2022) e (KHARE *et al.*, 2022). Estas diferenças entre os estudos podem estar ligadas às características usadas para alimentar os classificadores por cada grupo. Estas diferenças também podem ter surgido porque estes trabalhos se basearam em dados coletados enquanto os sujeitos executavam tarefas. E, nesta Tese, trabalhou-se com sinais coletados com os indivíduos em repouso.

A *accuracy* do classificador RF na classificação entre grupos TDAH e Controle com a banda Delta ($91,1 \pm 3,5\%$) obtida na presente Tese superou em, aproximadamente, 10% os 81,8% obtidos por (HOLKER & SUSAN, 2022) com sua RF.

A RF apresentada nesta Tese, também, superou a *accuracy* máxima para diagnóstico alcançada pela CNN usada para a classificação TDAH/Controle obtidos em tarefa apresentada no estudo de (DUBREUIL-VALL *et al.*, 2020). Esta CNN obteve 88% de

accuracy em um banco de dados formado por de sinais de EEG de 40 indivíduos. No entanto, sua CNN, também, identificou maiores diferenças entre TDAH e participantes do grupo controle nas frequências Alfa, Delta e Teta, o que corrobora os achados do presente estudo.

Em nosso trabalho anterior, (PEDROLLO *et al.*, 2022), usamos uma *toolbox* de *Spiking Neural Networks* para classificar amostras de 243 sujeitos da HBN. Essa abordagem obteve *accuracies* de 72% e de 68%, respectivamente, para a classificação binária entre TDAH e grupo controle e para a classificação de subtipos (entre TDAH-C, TDAH-I e grupo controle). No entanto, aquela *toolbox* carecia de ferramentas de otimização para alcançar desempenhos mais robustos e, por isso, foi abandonada na presente Tese.

O trabalho de (SHAKUR *et al.*, 2023) coletou dados de EEG e de movimentos oculares de 22 indivíduos (apenas 3 deles com TDAH) realizando atividades e classificou estes dados com uma análise exploratória baseada em regras. Sem balancear os dados, este estudo alcançou uma *accuracy* de 91,1% na classificação binária entre grupos TDAH e Controle com precisão de 79,0%. Seu sistema de aprendizagem baseado em regras comparou o movimento dos olhos e as características do EEG principalmente das bandas Teta, Alfa e Beta de acordo com as seguintes características de EEG: dimensão fractal de Haguchi, densidade espectral de potência, raiz quadrada média e desvio padrão.

A *accuracy* total alcançada pelos classificadores de fusão (92,2±0,5%) para classificação de entre pacientes com TDAH e sujeitos do grupo controle no presente trabalho não é tão alta quanto os 98% alcançados pela SVM de (ALCHALABI *et al.*, 2018). No entanto, seu trabalho utilizou um banco de dados com apenas 9 participantes em uma atividade baseada em tarefa (jogo de *videogame*). Portanto, este foi um paradigma, consideravelmente, diferente do investigado na presente pesquisa que investigou diferenças no TDAH (e em seus subtipos) no EEG em estado de repouso de 499 indivíduos.

Alguns trabalhos conseguiram obter uma *accuracy* de classificação entre classes TDAH e controle maior do que a apresentada na presente pesquisa. Porém, parte desses trabalhos utilizou uma amostra muito pequena de sujeitos, este foi o caso dos estudos de (ALCHALABI *et al.*, 2018), (PEREDA *et al.*, 2018), (JOY *et al.*, 2022) e de (AHMADI *et al.*, 2021) (este último na classificação de subtipos).

Além destes, os estudos de (KHARE *et al.*, 2022), (BAKHTYARI & MIRZAEI, 2022), (MOGHADDARI *et al.*, 2020), (ZHANG & LI, 2019) e (CHEN *et al.*, 2019b) usaram uma amostra de algumas dezenas de sujeitos em cada classe. Os classificadores apresentados nestes trabalhos obtiveram *accuracies* superiores às encontradas na presente Tese, porém, trabalharam apenas com a classificação binária entre sujeitos com TDAH e indivíduos do grupo controle.

Entre estes, (KHARE *et al.*, 2022), (MOGHADDARI *et al.*, 2020) e (BAKHTYARI & MIRZAEI, 2022) basearam seus estudos em sinais de EEG coletados enquanto os pacientes e voluntários praticavam tarefas. Conforme demonstrado por (DUBREUIL-VALL

et al., 2020), os classificadores tendem a obter *accuracies* maiores quando trabalham com EEG coletado ao longo de uma tarefa do que com rs-EEG na classificação binária de TDAH. Isto porque os pacientes com TDAH apresentam maiores diferenças nos sinais de EEG em tarefas.

Os trabalhos de (ZHANG & LI, 2019) e (CHEN *et al.*, 2019b) obtiveram *accuracies* de ordem similar (94,4% e 94,7%, respectivamente), embora um pouco maiores do que a obtida nesta Tese (92,2±0,5%) para a distinção entre classes controle e TDAH com sinais de rs-EEG. Ambos os autores investigaram matrizes de conectividade funcional de EEG com arquiteturas CNN, de forma análoga ao que foi realizado para os sinais de fMRI nesta Tese.

O estudo de (ABBAS *et al.*, 2021) usou um banco de dados composto por 61 participantes com TDAH e 60 do grupo controle para investigar as diferenças de conectividade no TDAH usando a entropia dos sinais de EEG como medida de transferência de informações. Este estudo identificou que o TDAH apresenta maiores diferenças de entropia em Beta, do que em outras bandas.

Em (ROSTAMI *et al.*, 2021), os autores tentaram diferenciar todos os subtipos de TDAH do grupo controle com Análise de Função Discriminante *step-wise* e PCA em amostras de 104 indivíduos. Estas amostras são de 43 voluntários do grupo controle e 61 pacientes com TDAH, dos quais 22, 25 e 14 são, respectivamente, identificados como TDAH-C, TDAH-I e TDAH-H. No entanto, os resultados obtidos por estes autores indicaram que os três subtipos de TDAH não podiam ser distinguidos uns dos outros e do grupo de controle simultaneamente apenas com base na potência absoluta das bandas de frequência. No entanto, na presente pesquisa consideramos um paradigma de classificação binária e de subtipos separada, onde os subtipos são identificados após o TDAH já ter sido detectado. Além disto, na presente Tese mais características, além da potência absoluta das bandas foram analisadas pelos classificadores.

O método de classificação de EEG proposto nesta Tese tem um custo computacional relativamente baixo, pois o ICA é a parte mais demorada de todo o processamento, com exceção dos AGs. Portanto, esta técnica é acessível para a maioria dos laboratórios, assim como a coleta dos sinais de EEG.

Assim como na pesquisa apresentada nesta Tese, os estudos de (AHMADI *et al.*, 2021), (BAKHTYARI & MIRZAEI, 2022) e (KHARE *et al.*, 2022) também analisaram a importância do uso de cada faixa de frequência para a classificação de TDAH. Destes, apenas (AHMADI *et al.*, 2021) considerou a classificação de subtipos, embora tenha se baseado em uma pequena quantidade de indivíduos (menos de 15 por classe).

Os autores de (AHMADI *et al.*, 2021) usaram uma CNN para distinguir os subtipos TDAH-I e TDAH-C do grupo de controle. Para tanto, utilizaram um banco de dados composto por apenas 40 crianças destras, das quais 13 eram do grupo TDAH-C, 12 da classe TDAH-I e 14 pertenciam ao grupo controle. Com esta técnica e este banco de

dados, os autores obtiveram 99,46% de *accuracy* com base em características espaciais das bandas Beta e Gama na região posterior do cérebro. O estudo deste grupo apontou a banda Delta como aquela cujo uso resultava nas piores *accuracies*.

Nesta Tese, alcançou-se uma *accuracy* similar a de (AHMADI *et al.*, 2021) para a classificação de subtipos com base em fMRI ($97,8 \pm 0,5\%$), porém com dados de 499 participantes (após a seleção resultante do procedimento de controle de qualidade). O presente estudo encontrou maiores discrepâncias entre as classes TDAH e Controle na banda Delta do que na banda Beta ou Gama. A diferença nos biomarcadores encontrados, possivelmente, se deve às diferenças na base de dados usada ou as demais diferenças nos procedimentos de coleta, pré-processamento, características extraídas ou nos modelos dos classificadores.

O trabalho de (MARKOVSKA-SIMOSKA & POP-JORDANOVA, 2017) encontrou diferenças entre crianças com TDAH e crianças do grupo controle nas bandas Delta e Teta, obtendo uma AUC de 84,3% na classificação através da análise com a curva ROC. Contudo, as diferenças nessas bandas não foram observadas em adultos, o que conduziu à redução na AUC para 68,3% nesta faixa etária no trabalho desses autores. Assim como o trabalho de (MARKOVSKA-SIMOSKA & POP-JORDANOVA, 2017), o presente estudo apontou distinções nas faixas Delta e Teta, porém, a RF, o Voto Majoritário e o *AdaBoost* adicionaram a contribuição das bandas Alfa e Gama.

O estudo de (SAAD *et al.*, 2018) propôs o uso da frequência de transição entre as faixas Alfa e Teta como um biomarcador de TDAH, o que está em conformidade com a alta relevância encontrada com as RFs para ambas as bandas nesta Tese.

Na classificação entre sujeitos com TDAH e do grupo controle, desta Tese, as idades das classes são de $9,7 \pm 3,2$ e de $9,6 \pm 3,7$, respectivamente. Logo, estas classes são constituídas, principalmente, de crianças e de alguns adolescentes. A CNN projetada nesta pesquisa apontou uma maior relevância da banda Delta, com a qual obteve sua maior *accuracy*: $91,1 \pm 3,5\%$. Os meta-classificadores *Ridge* e *XGBoost* confirmaram a maior importância desta faixa. Contudo, o desvio padrão de 3,5% é maior que os desvios-padrão obtidos analisando as demais faixas de frequência, como mostrado na Figura 31. Esta maior variabilidade da *accuracy*, a partir da análise da banda Delta, pode estar relacionada com a variação das idades dos sujeitos, embora não se constitua em uma diferença tão grande como a analisada no estudo de (MARKOVSKA-SIMOSKA & POP-JORDANOVA, 2017). Este fator pode ser tema de estudos futuros.

5.4.2 Comparação com os principais trabalhos de classificação de TDAH com base em fMRI

A presente Tese obteve maiores *accuracies* de classificação binária entre pacientes com TDAH e grupo controle com os meta-classificadores de fMRI do que os demais trabalhos pertencentes ao estado da arte que trabalharam com dados coletados através

desta modalidade de exame. Isto pode ser observado, comparando os resultados desta Tese com os dos demais trabalhos apresentados, anteriormente. A Tabela 35 reproduz os principais trabalhos apresentados na Tabela 2 no Capítulo 2 (Estado da Arte) e os compara com a presente Tese.

Tabela 35 – Tabela bibliométrica de técnicas, bancos de dados utilizados e resultados obtidos por diversos autores investigando classificação de TDAH com fMRI.

Referência (citações)	Tipo de dado	Técnica	Banco de dados	Accuracies	Contribuição
(BROWN <i>et al.</i> , 2012) 199 citações	fMRI Fenótipo	SVM	523 Controle 192 TDAH-C 124 TDAH-I	70,7% (binária) 64,75 (subtipos)	<i>Accuracy</i> ampliada por uso de fenótipo
(QURESHI <i>et al.</i> , 2017) 68 citações	sMRI e rs-fMRI	ELM	53 Controle 53 TDAH-C 53 TDAH-I	92,9% (binária) 76,2% (subtipos)	Espessura Dorsal e Ventral não reduz com idade em TDAH
(MAO <i>et al.</i> , 2019) 145 citações	rs-fMRI	CNN	439 Controle 359 TDAH	71,3% (binária)	Diferenças espaço-temporais
(DE SILVA <i>et al.</i> , 2021) 23 citações	rs-fMRI	CNN	20 Controle 20 TDAH	85,0% (binária)	Diferenças em SBC
(SAHA & SARKAR, 2022) 3 citações	rs-fMRI	DT	86 Controle 43 TDAH-C 43 TDAH-I	77,1% (subtipos)	Diferenças em <i>Default</i> e Límbica
Presente Tese	rs-fMRI	<i>Ensemble</i> <i>XG-Boost</i> CNNs	109 Controle 523 TDAH Subtipos: 244 TDAH-C 279 TDAH-I	97,6±2,1% (binária) 94,9±1,3% (subtipos)	Biomarcadores: Binária: Visual, Somatomotor, <i>Default</i> , Límbica e Frontoparietal Subtipos: Frontoparietal, <i>Default</i> e Límbica

Na classificação entre grupos TDAH e controle, as *accuracies* alcançadas através da

metodologia proposta na presente Tese superaram as *accuracies* obtidas por (QURESHI *et al.*, 2017), (MAO *et al.*, 2019), (SHARMA *et al.*, 2023), (SAHA & SARKAR, 2022), (DE SILVA *et al.*, 2021), (LIU *et al.*, 2022) e (ARIYARATHNE *et al.*, n.d.). Estes trabalhos obtiveram, respectivamente, *accuracies* de 92,9%, 71,3%, 86,7%, 77,1%, 85,0%, 72,4%, 85,0% na identificação entre pacientes com TDAH e sujeitos do grupo controle.

Os estudos iniciais, como os de (COLBY *et al.*, 2012), (BROWN *et al.*, 2012), (DAI *et al.*, 2012), (DEY *et al.*, 2014) e (GHIASSIAN *et al.*, 2016), avaliaram diferentes características com SVMs para a classificação entre grupos TDAH e controle. Com esta abordagem estes autores obtiveram *accuracies* de 55,0%, 70,7%, 61,5%, 62,8% e 69,6%, respectivamente.

Em seu trabalho, (DEY *et al.*, 2014) experimentou treinar e aplicar a SVM separadamente para cada gênero, o que elevou em quase 11% a *accuracy* do classificador (de 62,8% para 73,6%). Na presente Tese, obteve-se maiores *accuracies*, mesmo sem explorar este recurso. Experimentos futuros podem investigar se a separação por gêneros pode produzir melhoras nas métricas ou auxiliar a encontrar biomarcadores específicos em cada gênero.

O trabalho de (ZOU *et al.*, 2017) utilizou uma fusão entre classificadores de sMRI e fMRI para o diagnóstico do TDAH. Estes autores utilizaram uma CNN com 4 camadas convolucionais e uma camada de neurônios para analisar mapas de características de fMRI e uma CNN com a mesma configuração para a análise dos dados de sMRI. Estas duas CNNs foram então unidas com uma última camada de neurônios. Os autores aplicaram esta metodologia em 77 pacientes com TDAH e 94 participantes do grupo controle. Contudo, esta estratégia resultou na obtenção de uma *accuracy* de apenas 69,2%.

Porém, (SHARMA *et al.*, 2023) obteve uma *accuracy* maior para a classificação de TDAH do que a apresentada nesta Tese, quando combinou as informações das características de fMRI com informações fenotípicas (dados pessoais) de cada sujeito. Um classificador foi designado para cada categoria de dados (fMRI ou dados fenotípicos). Estes dois classificadores foram unidos através de um meta-classificador de Voto Majoritário. No entanto, é importante ressaltar, que entre as características fenotípicas os autores incluíram o próprio diagnóstico de TDAH. Deste modo, no trabalho destes autores, o diagnóstico de TDAH através dos dados pessoais foi considerado, simultaneamente, como entrada e como saída de uma rede neural. Embora facilite o trabalho da rede neural, esta prática afasta muito a tarefa do classificador de um caso real, onde o diagnóstico de TDAH não seria conhecido antes de ser realizado com auxílio do próprio classificador. Por isso, estes autores constataram uma queda de *accuracy* ao unir o classificador fenotípico com o classificador de fMRI (de 99,2% para 95,2%). O classificador utilizando somente as características de fMRI, sem a informação prévia dos diagnósticos, obteve 86,7% de *accuracy*.

Assim como (SHARMA *et al.*, 2023) e (BROWN *et al.*, 2012), os trabalhos de (RIAZ

et al., 2017) e de (RIAZ *et al.*, 2018), também, utilizaram a estratégia de agregar características pessoais aos classificadores. Estes autores utilizaram esta técnica para elevar as *accuracies* de seus classificadores. Por exemplo, as *accuracies* da SVM de (RIAZ *et al.*, 2018) aumentaram de 67,4%, 25,4% e 85,3% (cada métrica corresponde a um local de coleta diferente) para 86,7%, 52,7% e 85,8%, respectivamente, ao usar esta abordagem. Nota-se que o aumento reportado variou, consideravelmente, em cada local de coleta. A possibilidade de utilizar os dados pessoais para elevar as métricas de classificação não foi explorada na presente Tese. Em trabalhos futuros, a combinação desta estratégia com os meta-classificadores apresentados nesta Tese pode ser útil para verificar se há melhora nas métricas de desempenho ou variação nos biomarcadores de acordo com diferenças nas características pessoais.

Assim como nesta Tese, (JHA *et al.*, 2020) trabalharam com rs-fMRI, separando o cérebro em ROIs e produzindo matrizes de correlação. Contudo, esses autores geraram grafos a partir das médias de cada sinal bruto de fMRI e da matriz de correlação. Então, aplicaram CNNs siamesas para detectar características de semelhanças e de diferenças entre os grafos de cada indivíduo. No entanto, o classificador produzido através dessa estratégia produziu apenas uma *accuracy* de 68,3% em uma amostra desbalanceada de 311 participantes (aproximadamente, 90 do grupo controle e 221 com TDAH).

No estudo de (MENG *et al.*, 2022) foram aplicados modelos clássicos de CNN, a ResNet e a VGG com aprendizagem por transferência para obter *accuracies* de 80,0% e de 82,0%, respectivamente, para a classificação entre indivíduos com TDAH e do grupo controle. Na presente Tese, experimentou-se o uso destas arquiteturas, conforme relatado no APÊNDICE E. Porém, nesta Tese as *accuracies* encontradas com a ResNet e com a VGG foram menores do que as relatadas por esses autores. Esta diferença pode estar relacionada a esses autores terem gerado e analisado mais uma matriz de conectividade funcional por participante através da correlação parcial. Não obstante, as CNNs projetadas e otimizadas nesta Tese obtiveram *accuracies* de até $96,3 \pm 0,6$.

Em fMRI, é comum os estudos considerarem uma maior quantidade de sujeitos em relação aos estudos com EEG. A presente pesquisa utilizou 632 indivíduos após as exclusões do controle de qualidade. A presença de mais de 100 indivíduos de cada classe também foi observada nos trabalhos de (MAO *et al.*, 2019), (LIU *et al.*, 2022) e (SHARMA *et al.*, 2023).

O trabalho de (MAO *et al.*, 2019) realizou seu treinamento e teste com dados de 788 sujeitos. Entretanto, o trabalho de (MAO *et al.*, 2019) obteve uma *accuracy* baixa (71,3%) a partir de características espaço temporais de rs-fMRI com uma CNN e não considerou a classificação de subtipos de TDAH.

Na literatura, há poucos trabalhos que também realizaram a classificação de subtipos de TDAH a partir de exames de fMRI. Este experimento foi realizado também por (DAI *et al.*, 2012), (BROWN *et al.*, 2012), (DESHPANDE *et al.*, 2015), (QURESHI *et al.*,

2017) e (SAHA & SARKAR, 2022).

Na pesquisa de (DAI *et al.*, 2012) foram empregados dados de 776 participantes, porém seu modelo de classificação baseado em SVM obteve *accuracies* de apenas 61,5% na identificação entre classes TDAH e Controle e de 54,1% na classificação de subtipos (TDAH-C e TDAH-I). Assim como (DAI *et al.*, 2012), o classificador de (BROWN *et al.*, 2012) empregou uma SVM e obteve *accuracies* mais baixas do que as da presente Tese para distinguir entre classes TDAH e Controle ($70,7 \pm 6,2\%$) e entre subtipos ($64,7 \pm 4,3\%$) em um banco de dados com grande quantidade de participantes (668).

O trabalho de (DESHPANDE *et al.*, 2015) foi o estudo que utilizou a maior base de dados para fMRI: 744 participantes do grupo controle e 433 com TDAH, dos quais 260 foram diagnosticados com TDAH-C e 173 com TDAH-I. A pesquisa de (DESHPANDE *et al.*, 2015) empregou uma RNA em cascata para obter *accuracies* médias em torno de 70% para a detecção entre classes TDAH e Controle e de 60% para a classificação de subtipos. Este trabalho mostrou que a análise das características de conectividade funcional conseguem discriminar os grupos de forma mais eficiente do que o uso do dado bruto. Na presente Tese, também trabalhamos com o uso das características de conectividade funcional, ao invés de usarmos o dado bruto. No entanto, as *accuracies* obtidas nesta Tese foram maiores para os dois cenários de classificação, entre TDAH e Controle e entre subtipos: $97,6 \pm 2,1\%$ e $94,5 \pm 1,3\%$, respectivamente.

A pesquisa de (QURESHI *et al.*, 2017) utilizou uma ELM para obter 76,2% de *accuracy* de classificação de subtipos e 92,9% na classificação entre TDAH e grupo controle. Esse estudo foi baseado em dados de 201 indivíduos, sendo 67 de cada uma das classes: controle, TDAH-C e TDAH-I. Estes indivíduos foram selecionados na base ADHD-200. Os autores selecionaram aproximadamente 79% dos indivíduos para treinamento e apenas cerca de 21% para teste, em vez de usar o padrão de aprendizado de máquina de 70% de amostras para treinamento. Este grupo usou as características de fMRI para auxiliar na seleção de características de sMRI. Contudo, para seu experimento de classificação entre grupos TDAH e controle analisaram apenas as características de sMRI. Já para a classificação de subtipos, a ANOVA apontou diferenças nas redes de atenção ventral, sensorio-motora, visual e cerebelar. A ELM indicou a conectividade global média, uma característica de sMRI, como um fator importante para distinguir os subtipos.

Em seu trabalho anterior, usando as mesmas amostras, (QURESHI *et al.*, 2016), alcançou 60,78% de *accuracy* na classificação de subtipos usando apenas características de sMRI. No entanto, o classificador do presente estudo superou a *accuracy* obtida por (QURESHI *et al.*, 2017) em cerca de 20% para classificação de subtipos com base em dados de mais que o dobro do número de sujeitos. Diferente do trabalho de (QURESHI *et al.*, 2017), na presente Tese a CNN projetada e os meta-classificadores baseados nesta arquitetura indicaram que a rede de atenção ventral tem menor relevância na distinção de subtipos. Os meta-classificadores desta Tese também não apontaram uma relevância par-

ticularmente alta para as redes de atenção dorsal e ventral, como indicado pelo trabalho de (QURESHI *et al.*, 2017).

A pesquisa de (PARK *et al.*, 2021) também utilizou a base de dados HBN. Nesta, estimou-se os estados de conectividade na rede *default* usando modelagem oculta de Markov (*hidden Markov modeling*). Através deste estudo, investigaram as diferenças entre 56 indivíduos com TDAH-C e 32 voluntários do grupo controle. Este estudo observou uma autoconexão inibitória mais forte no estado de conectividade funcional presente no TDAH-C.

O estudo de (ARIYARATHNE *et al.*, n.d.) encontrou diferenças na rede *default* de pacientes com TDAH usando uma CNN para avaliar recursos de conectividade funcional de fMRI em estado de repouso. Esta pesquisa foi baseada em dados de 40 indivíduos, metade destes com TDAH e metade saudável. Estes autores restringiram sua investigação a quatro zonas distintas da rede *default* e usaram correlação baseada em sementes (*seed-based correlation*) para investigar sua conectividade funcional. Através desta metodologia, este trabalho alcançou até 85,2% de *accuracy* na classificação entre grupos TDAH e controle quando sua CNN investigou o córtex pré-frontal medial, parte da rede *default*. Contudo, as *accuracies* encontradas tenham sido muito semelhantes usando as demais partes da rede *default*: córtex cingulado posterior, junção temporoparietal esquerda, junção temporoparietal direita.

No estudo seguinte do mesmo grupo, (DE SILVA *et al.*, 2021), os autores compararam a variação na *accuracy* obtida por uma CNN na classificação de TDAH *versus* grupo controle usando diferentes métodos de extração de características. Para tanto, empregaram o mesmo banco de dados de (ARIYARATHNE *et al.*, n.d.). Em (DE SILVA *et al.*, 2021), o uso da correlação baseada em sementes possibilitou que a CNN a obtivesse uma *accuracy* de 85,4%. Assim como em seu trabalho anterior, esta pesquisa investigou apenas o uso da rede de conectividade *default*. Sua precisão foi consideravelmente maior usando a CNN que projetaram do que quando aplicaram CNN baseadas nas arquiteturas VGG-16 e ResNet-50 (com as quais obtiveram *accuracies* de somente 68,1% e 67,1%, respectivamente). Estes autores empregaram técnicas de aumento de dados para ampliar seu conjunto de dados inicial em 120 vezes usando segmentação dos sinais de fMRI e transformações de imagens. Na presente Tese, porém, essas transformações de imagens não foram utilizadas, pois observou-se que estas poderiam provocar a perda das posições dos dados nas matrizes de correlação, as quais são informações relevantes para determinar a ocorrência de TDAH e para distinguir seus subtipos.

O trabalho apresentado em (LIU *et al.*, 2022) usou fMRIs adquiridas em estado de repouso de 289 pacientes com TDAH e de 389 voluntários saudáveis participantes da base ADHD-200. Este trabalho empregou um *autoencoder* convolucional com um bloco residual interno para extrair as características espaço-temporais e remover seu ruído. Em seguida, estes autores usaram uma unidade recorrente de porta convolucional seguida por

uma camada de *pooling* de média global e um classificador sigmoide para obter uma *accuracy* de 72,44% na distinção entre as classes TDAH e Controle.

Os autores do (SAHA & SARKAR, 2022) investigaram apenas a classificação de subtipos de TDAH (TDAH-C e TDAH-I) usando árvores de decisão. Esta técnica de aprendizado de máquina foi usada para procurar biomarcadores de conectividade funcional em fMRIs coletadas em estado de repouso. Este estudo indicou as regiões pré-frontal, visual, *default* e límbica, além das regiões do cerebelo, como marcadores biológicos. Na presente Tese, na classificação de subtipos identificou-se como biomarcador a rede límbica. Porém, as redes visual e *default* apenas foram consideradas entre as mais importantes para os meta-classificadores que apresentaram menores *accuracies*: Votação Majoritária, *Ridge* e, em menor medida, *AdaBoost*. A presente Tese apontou a relevância da rede Somatomotor para a distinção de subtipos, o que não foi identificado por (SAHA & SARKAR, 2022) e nem por (QURESHI *et al.*, 2017).

Em seu primeiro experimento, em (SAHA & SARKAR, 2022), os autores usaram uma amostra de 43 indivíduos do sexo masculino de cada classe: TDAH-C, TDAH-I e grupo controle. No segundo experimento, consideraram amostras de 22, 19 e 43 mulheres das classes TDAH-C, TDAH-I e grupo controle, respectivamente. Enquanto no primeiro experimento obtiveram uma *accuracy* de 67,5%, no segundo alcançaram uma *accuracy* de 77,1%. Entretanto, os autores não relataram o uso de uma técnica de balanceamento, a qual era necessária no segundo experimento para evitar que as amostras desequilibradas influenciassem as métricas do classificador. Por isso, o segundo experimento apresentou uma diferença de 15% entre as classes em especificidade (verdadeiros positivos/falsos negativos).

Em fMRI, a presente pesquisa mostrou que, usando os classificadores básicos, as redes de conectividade que melhor distinguem indivíduos com TDAH do grupo controle são as redes visual, somatomotor, padrão e límbica, seguidas das redes de atenção ventral e frontoparietal. No entanto, a Votação Majoritária alcançou a maior *accuracy* para classificar TDAH *versus* grupo controle (97,6%) combinando as redes límbica, visual, somatomotor e *default*. Já a região frontoparietal foi apontada como biomarcador para a distinção entre classes TDAH e Controle pelos classificadores *Ridge* e *AdaBoost*. Os subtipos de TDAH apresentam maiores diferenças nas redes frontoparietais e *default*. O meta-classificador *XGBoost* alcançou 94,5% de *accuracy* na classificação de subtipos, valorizando a região límbica e, em menor grau, a região somatomotor. Os meta-classificadores identificaram a rede límbica como um biomarcador comum para TDAH e seus subtipos.

Através de uma meta-análise de trabalhos de classificação de TDAH, a pesquisa de (RUBIA, 2018) observou alterações nas redes de atenção dorsal, de atenção ventral, frontoparietal, límbica e *default*. O estudo desta Tese não corroborou a importância atribuída por (RUBIA, 2018) da rede de atenção dorsal, a qual foi apontada pela CNN projetada como a rede de menor relevância para a distinção das classes TDAH e Controle (o

que também foi determinado pelos meta-classificadores). Para as redes de atenção ventral e frontoparietal as CNNs projetadas obtiveram *accuracies* de $90,7 \pm 1,7\%$ e de $90,0 \pm 1,5\%$, confirmando sua importância.

A pesquisa de (AARTS *et al.*, 2017) trabalhou com dados de 19 indivíduos com TDAH e 77 do grupo controle e identificou, através de fMRI, uma redução nas ativações cerebrais na rede de atenção ventral em experimentos de antecipação de recompensas em sujeitos com TDAH. A CNN da presente Tese apontou que há alterações na rede de atenção Ventral, também, em estado de repouso, embora o estudo com classificadores indique que estas diferenças sejam um pouco menores do que as encontradas nas redes visual, somatomotor, límbica e *default*. Por isso, o uso da rede de atenção ventral resulta em *accuracies* um pouco menores, como mostrou a Figura 44 no Capítulo 4 (Resultados). As diferenças entre o estudo da presente Tese e o de (AARTS *et al.*, 2017) podem estar relacionadas ao desbalanceamento de dados entre as classes do trabalho desses autores.

O estudo de (HART *et al.*, 2012) constatou diferenças na rede frontoparietal e no cerebelo através de uma meta-análise com dados de 150 sujeitos com TDAH e 145 do grupo controle coletados enquanto os participantes realizavam tarefas. Através de outra meta-análise, (HART *et al.*, 2013) observou também diferenças no tálamo (que é uma das regiões pertencentes à rede límbica) enquanto pacientes com TDAH e voluntários do grupo controle executavam tarefas. A presente Tese trabalhou com dados em estado de repouso, diferentemente de (HART *et al.*, 2012) e (HART *et al.*, 2013). Nestas condições, também, verificou-se com os meta-classificadores *Ridge* e *AdaBoost* diferenças entre TDAH e grupo controle na rede frontoparietal. A CNN, nesta pesquisa, apresentou uma alta *accuracy* na rede límbica, apontada também pelos métodos *Ridge*, *AdaBoost* e *XGBoost* como altamente influente para a classificação de TDAH.

A presente pesquisa obteve *accuracies* maiores nas classificações entre TDAH e grupo controle ($97,6 \pm 2,1\%$) e entre subtipos ($94,5 \pm 1,3\%$) do que os demais estudos baseados apenas nas características de fMRI. No entanto, características oriundas de outras fontes como a EEG ou a análise fenotípica, podem classificar os indivíduos com TDAH com maior *accuracy*. Porém, a investigação das características de fMRI possui a vantagem de proporcionar maior resolução espacial da funcionalidade do cérebro, localizando áreas específicas. Portanto, a análise da fMRI permite estabelecer vínculos entre o transtorno (ou seus subtipos) e áreas cerebrais que são diferentes nos pacientes de acordo com sua classe. A identificação da conexão entre estas áreas e o transtorno pode apontar um caminho para o desenvolvimento de tratamentos que visem modificar o comportamento dos cérebros com TDAH somente nestas determinadas áreas de modo que o seu funcionamento se aproxime do observado no grupo controle.

5.5 Limitações

O presente estudo, no entanto, possui algumas limitações. Os meta-classificadores, em sua maioria, obtiveram maior *accuracy* na classe TDAH-C para classificação de subtipos e maior *accuracy* no grupo controle para identificação entre classes TDAH e Controle. A razão para estas diferenças pode estar vinculada ao fato de que os voluntários do grupo controle e os pacientes com TDAH-C são classes minoritárias que foram aumentadas pela sobreamostragem que produziu amostras generativas, com a técnica SMOTE. O uso desta técnica pode ter provocado um pequeno viés (*bias*), uma vez que essas amostras artificiais podem ser ligeiramente mais semelhantes entre si do que as amostras naturais.

Existem poucas amostras de TDAH-H disponíveis na base de dados HBN. Isto não é exclusividade desta base e ocorre porque este subtipo de TDAH é mais raramente encontrado (GIBBINS *et al.*, 2010) (SMALLEY *et al.* (2007)). Diversas pesquisas de classificação de TDAH descartam o uso deste subtipo devido à baixa presença nos bancos de dados (QURESHI *et al.*, 2016), (QURESHI *et al.*, 2017), (VAHID *et al.*, 2019), (AHMADI *et al.*, 2021), (SAHA & SARKAR, 2022).

Nesta Tese, antes do controle de qualidade havia dados de EEG de 33 participantes e dados de fMRI de 43 indivíduos da HBN acometidos por TDAH-H. Isto corresponde a 3,6% e 5,0% da quantidade de dados disponível para cada um destes exames, respectivamente. Em outras palavras, há cerca de 10 vezes menos indivíduos pertencentes a esta classe. Após o controle de qualidade a quantidade de sujeitos com TDAH-H foi reduzida para 18 e 25 em EEG e fMRI. Desta forma, a amostra de TDAH-H é pouco representativa.

A classe TDAH-H sofre uma exclusão mais severa no número de sujeitos do que as demais nas etapas de controle de qualidade de dados tanto para EEG como para fMRI. Isto ocorre porque a hiperatividade agravada nestes indivíduos provoca uma vontade quase que irresistível de estes se movimentarem. E, este movimento produz artefatos de movimento nos exames, tornando-os muito ruidosos. Este ruído teria o potencial de influenciar nas características analisadas pelos classificadores e desta forma influenciar na classificação.

Não remover os pacientes, sinais e segmentos que possuem movimento excessivo poderia ser prejudicial para a busca por biomarcadores, embora pudesse aumentar as *accuracies* de classificação. Pois, os modelos em Tese poderiam decidir por classificar um determinado sujeito como TDAH-H em função de este apresentar sinais com maior quantidade de ruído, provocada por sua movimentação excessiva. Então, neste caso, os modelos poderiam apontar uma maior eficiência em biomarcadores mais predispostos a serem afetados por estes artefatos, ao invés dos biomarcadores correspondentes às alterações cerebrais.

Para EEG, a base HBN possui apenas 33 indivíduos com TDAH-H, dos quais 18 foram aprovados no controle de qualidade (Tabela 7, página 71) utilizado na presente pesquisa. Já em fMRI, a base HBN tem 43 participantes e apenas 25 destes possuíam qualidade

aceitável segundo os critérios adotados no presente estudo (Tabela 18, página 91).

A equipe da HBN prevê uma expansão de sua base de dados, através da fusão com outras bases coletadas em locais diferentes. Esta expansão, possivelmente, permitiria uma análise mais robusta dos biomarcadores e da classificação do subtipo TDAH-H. Contudo, este aumento de dados em termos de número de sujeitos está previsto para ocorrer após o término do tempo hábil para a conclusão do estudo apresentado nesta Tese de doutorado.

Na presente Tese, desconsiderou-se as variações nos níveis de gravidade do transtorno. Como apresentado no APÊNDICE F, os pacientes com TDAH podem ser classificados não apenas pela diversidade de sintomas, mas também pela intensidade destes sintomas. Uma análise futura considerando as divergências de intensidade sintomática pode auxiliar no entendimento do processo que provoca o transtorno e suas variações.

6 CONCLUSÃO

A presente pesquisa introduz uma abordagem original para o estudo da classificação de TDAH e de seus subtipos com a fusão de classificadores primários baseados em possíveis biomarcadores de EEG e de fMRI. Realizou-se, também, o objetivo de encontrar biomarcadores nestas duas modalidades de exames. O presente estudo utilizou um grande conjunto de dados e procedimentos cuidadosos de pré-processamento e de controle de qualidade, a fim de constituir uma fonte confiável de biomarcadores para futuros estudos de TDAH focados em tratamentos de subtipos. É também identificado o nível de conexão entre os biomarcadores vinculados à ocorrência de TDAH e à identificação de subtipos através da distribuição de pesos determinada pelos meta-classificadores.

Mesmo após as exclusões decorrentes dos procedimentos de controle de qualidade, o estudo apresentado nesta Tese ainda compreende o maior estudo (em quantidade de indivíduos) de classificação de subtipos de TDAH até o presente momento, dentre os estudos que se baseiam em dados coletados por EEG ou por fMRI. A combinação do grande número de indivíduos restantes (499 em EEG e 632 em fMRI após as exclusões) com um controle de qualidade rigoroso reduziu possíveis vieses que poderiam confundir a identificação do biomarcador. Portanto, a presente pesquisa apresenta resultados robustos e reproduzíveis que poderão servir de base para futuros estudos na área.

Este estudo empregou *ensembles* de classificadores CNN para melhorar a *accuracy* da classificação de grupos de TDAH/Controle e de classificações de TDAH-C/TDAH-I. Essa técnica não só aumentou o desempenho do classificador, mas também identificou as áreas e faixas de frequência do cérebro com maior conexão à ocorrência de TDAH, assim como as regiões e bandas de frequência com alterações que diferenciam os subtipos de TDAH.

As RFs obtiveram, para as análises de EEG, melhores *accuracies* que as SVMs em todos os 12 experimentos de classificação e, por isso, decidiu-se utilizar as RFs como classificadores primários nos *ensembles*. Os meta-classificadores experimentados foram o Voto Majoritário ponderado, o classificador linear *Ridge*, o *AdaBoost* e o *XGBoost*.

Além de aumentar as *accuracies*, os meta-classificadores reduziram consideravelmente o *overfitting* observado, e, conseqüentemente, aumentaram a capacidade de ge-

neralização dos modelos. No caso dos subtipos, a fusão através do *XGBoost* garantiu um equilíbrio quase perfeito, o que é desejado entre as *accuracies* das classes. O aumento nas *accuracies* totais com a fusão das informações das bandas indica que as informações das faixas de frequência são complementares. Ou seja, a presença do TDAH provoca alterações em mais de uma faixa de frequência e há diferenças entre os subtipos em mais de uma faixa de frequência de EEG.

A análise dos pesos utilizados por cada um dos métodos de fusão revelou que as alterações nas ondas Delta têm uma influência notável tanto para a classificação binária como para classificação de subtipos, o que significa que o TDAH modifica consideravelmente as regiões cerebrais ligadas à geração das ondas Delta, tipicamente, a região frontal. A combinação das ondas Alfa e Gama (ou Teta) pode ter influência considerável na classificação do TDAH/Controle. Para o diagnóstico de subtipos, a banda Teta, também, se mostrou um marcador biológico relevante.

Em fMRI, o uso combinado de forma ponderada das redes de conectividade elevou as *accuracies* de classificação em comparação ao uso separado dos biomarcadores, corroborando que o distúrbio afeta múltiplas áreas do cérebro. Esta metodologia obteve *accuracies* que superaram a utilização de toda a matriz de correlações como entrada de uma CNN (seja a CNN projetada ou os modelos clássicos, como mostra o APÊNDICE E). Isto ocorre em virtude da possibilidade da metodologia de fusão de atribuir pesos otimizados para cada região do cérebro. Isto foi análogo ao que foi observado com os dados de EEG.

Os resultados individuais dos classificadores primários, apontaram as redes que apresentam maiores *accuracies* quando analisadas isoladamente. No diagnóstico entre classes TDAH e Controle, o uso individual das regiões visual, somatomotor, *default* e límbica alcançou *accuracies* superiores a 93,0% com desvio padrão inferior a 2,6% com a CNN projetada. Porém, para os subtipos, os classificadores primários não indicaram biomarcadores de forma tão explícita, embora as redes de conectividade frontoparietal e *default* tenham sido mais relevantes.

Os *ensembles* indicaram que as redes límbica, visual, somatomotor, *default* e frontoparietal são relevantes para o diagnóstico de TDAH. O *XGBoost* identificou que a região límbica é a rede com maiores diferenças entre os subtipos.

A fusão das redes de conectividade aferidas através de fMRI possibilitou a obtenção de uma maior *accuracy* ($97,6 \pm 2,1\%$) do que a fusão das bandas de EEG ($92,4 \pm 0,5\%$) para a classificação entre TDAH e grupo controle, embora com maior desvio padrão. No entanto, na distinção dos subtipos, o *XGBoost* obteve maior sucesso analisando as faixas de frequência de EEG (*accuracy* de $97,8 \pm 0,3\%$).

O maior sucesso de fMRI na detecção de TDAH e de EEG na identificação de subtipos é uma informação relevante. Pois, pode indicar que as diferenças entre os subtipos tem maior relação com características espaço-temporais e que as diferenças entre TDAH e

grupo controle estão mais ligadas as relações entre as ROIs nas redes de conectividade.

A análise das distribuições de pesos obtidas pelas otimizações forneceram marcadores biológicos não redundantes tanto para EEG como para fMRI. O uso da metodologia apresentada nesta Tese levou à identificação de marcadores biológicos que podem orientar o desenvolvimento futuro de tratamentos especializados para o TDAH.

6.1 Trabalhos Futuros

A presente Tese abre novos caminhos para trabalhos futuros, além do uso dos biomarcadores identificados para monitorar e desenvolver novos tratamentos visando melhorar a qualidade de vida das pessoas com TDAH. A metodologia de classificação e identificação de biomarcadores através do uso dos *ensembles* pode ser aplicada para investigar separadamente partes menores de cada rede de conectividade de fMRI ou até mesmo de cada uma das 190 ROIs separadamente. Também, é possível usar os meta-classificadores para averiguar possíveis diferenças na forma como estas redes cerebrais se relacionam entre si analisando as regiões da matriz de conectividade que mostram as correlações entre as regiões cerebrais.

Em EEG, o método apresentado nesta Tese pode ser aprimorado separando os conjuntos de eletrodos conforme as áreas do escalpo cerebral onde se encontram. Isto permitirá avaliar de forma mais específica as conexões entre o transtorno e as áreas onde a modificação de cada banda de frequência ocorrem.

Outra possibilidade de trabalho futuro com os sinais de EEG está na investigação das diferenças entre os períodos em que os olhos dos participantes estão fechados ou abertos ao longo do exame. Além disso, é possível investigar as diferenças apresentadas pelos pacientes com TDAH ao abrir e fechar de olhos, como eventos associando-se às medições de EEG com o rastreamento ocular (*eyetracking*).

O uso dos *ensembles* pode também ser usado para a fusão dos classificadores de EEG e de fMRI, e, conseqüentemente, dos biomarcadores encontrados em EEG e em fMRI, para elevar as *accuracies* de classificação e para encontrar relações entre as alterações nas bandas de EEG e nas regiões das fMRI. Além disso, a metodologia desenvolvida neste trabalho também pode ser relevante para o estudo futuro de biomarcadores de outros transtornos cerebrais.

O uso de classificadores de primários mais complexos ou técnicas de diferentes *ensembles* pode contribuir para adquirir mais conhecimento sobre biomarcadores ou melhorar ainda mais a classificação do TDAH e de seus subtipos.

A técnica também poderia ser estendida para outras modalidades de exames ou outros paradigmas de coleta de dados, como, por exemplo, dados coletados em tarefas.

REFERÊNCIAS

AARTS, E. et al. **Gut microbiome in adhd and its relation to neural reward anticipation.** *PLOS one*, [S. l.], v. 12, n. 9, p. e0183509, abr. 2017.

ABBAS, A. K. et al. **Effective connectivity in brain networks estimated using eeg signals is altered in children with ADHD.** *Computers in Biology and Medicine*, [S. l.], n. 134, p. 104515, jun. 2021.

ADLER, L. A. et al. **Validity of pilot Adult ADHD Self- Report Scale (ASRS) to Rate Adult ADHD symptoms.** *Annals of Clinical Psychiatry*, [S. l.], v. 18, n. 3, p. 145–148, jul. 2006.

AHMADI, A. et al. **Computer aided diagnosis system using deep convolutional neural networks for ADHD subtypes.** *Biomedical Signal Processing and Control*, [S. l.], v. 63, p. 102227, jan. 2021.

ALCHALABI, A. E. et al. **Focus: Detecting adhd patients by an EEG-based serious game.** *IEEE Transactions on Instrumentation and Measurement*, [S. l.], v. 67, n. 7, p. 1512–1520, jun. 2018.

ALEXANDER, L. M. E. et al. **An open resource for transdiagnostic research in pediatric mental health and learning disorders.** *Scientific data*, [S. l.], v. 4, n. 1, p. 1–26, dez. 2017.

AMERICAN PSYCHIATRIC ASSOCIATION, A. **Diagnostic and statistical manual of mental disorders: DSM-IV.** Washington: American psychiatric association, v. 4, 1994. ISBN: 10:0890420610.

AMERICAN PSYCHIATRIC ASSOCIATION, A. **Diagnostic and statistical manual of mental disorders: DSM-5.** Washington: American psychiatric association, v. 5, n. 5, 2013. ISBN: 13: 9788582710883.

ANGRIMAN, M., BEGGIATO, A. , CORTESE, S. **Anatomical and functional brain imaging in childhood ADHD: Update 2013. Current Developmental Disorders Reports**, [S. l.], v. 1, p. 29–40, dez. 2014.

ANJUM, M. F. et al. **Linear predictive coding distinguishes spectral eeg features of parkinson’s disease. Parkinsonism & related disorders**, [S. l.], v. 79, p. 79–85, out. 2020.

ARBABSHIRANI, M. R. et al. **Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. NeuroImage**, [S. l.], v. 145, p. 137–165, jan. 2017.

ARIYARATHNE, G. et al. Adhd identification using convolutional neural network with seed-based approach for fmri data. In: 2020 9th International Conference on Software and Computer Applications, 24, 2020, Langkawi, Malaysia. **Proceedings...** Langkawi [Malaysia], 2020. p. 31–35. Disponível em: <<https://doi.org/10.1145/3384544.3384552>>. Acesso em: 20 abr. 2024.

AYDN, S. et al. **Comparison of domain specific connectivity metrics for estimation brain network indices in boys with ADHD-C. Biomedical Signal Processing and Control**, [S. l.], v. 76, p. 103626, jul. 2022.

BAI, Y. et al. **Resting-state functional connectivity of the sensory/somatomotor network associated with sleep quality: evidence from 202 young male samples. Brain Imaging and Behavior**, [S. l.], v. 16, n. 4, p. 1832–1841, abr. 2022.

BAKHTYARI, M. , MIRZAEI, S. **ADHD detection using dynamic connectivity patterns of eeg data and convlstm with attention framework. Biomedical Signal Processing and Control**, [S. l.], v. 76, p. 103708, jul. 2022.

BARKLEY, R. A. , PETERS, H. **The earliest reference to ADHD in the medical literature? melchior adam weikard’s description in 1775 of “attention deficit”(mangel der aufmerksamkeit, attentio volubilis). Journal of attention disorders**, [S. l.], v. 16, n. 8, p. 623–630, fev. 2012.

BARKLEY, R. A., MURPHY, K. R. , FISCHER, M. **ADHD in adults: What the science says**. New York: Guilford press, 2010. ISBN: 13:978-1593855864.

BELLEÇ, P. et al. **The neuro bureau adhd-200 preprocessed repository. Neuroimage**,

[S. l.], v. 144, p. 275–286, jan. 2017.

BIEDERMAN, J. Attention-deficit/hyperactivity disorder: a selective overview. *Biological psychiatry*, [S. l.], v. 57, n. 11, p. 1215–1220, jun. 2005.

BISWAS, S. D., CHAKRABORTY, R. , PRAMANIK, A. A brief survey on various prediction models for detection of adhd from brain-MRI images. In: 21st International Conference on Distributed Computing and Networking, 54, 2020, Kolkata, India. **Proceedings...** Kolkata [India], 2020. p. 1–5. Disponível em: <<https://doi.org/10.1145/3369740.3372775>>. Acesso em: 20 abr. 2024.

BOSER, B. E., GUYON, I. M. , VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: Fifth annual workshop on Computational learning theory, 57, 1992, Pittsburgh, Estados Unidos da America. **Proceedings...** Pittsburgh [Estados Unidos da America], 1992. p. 1215–1220. Disponível em: <<https://doi.org/10.1145/130385.130401>>. Acesso em: 20 abr. 2024.

BREIMAN, L. **Random forests. *Machine learning*, [S. l.], v. 45, p. 5–32, out. 2001.**

BROCA, P. Anatomie comparée des circonvolutions cérébrales. Le grand lobe limbique et la scissure limbique dans la série des mammifères. *Review of Anthropology*, [S. l.], v. 1, p. 385–498, 1978.

BROWN, M. R. et al. **ADHD-200 global competition: diagnosing adhd using personal characteristic data can outperform resting state fMRI measurements. *Frontiers in systems neuroscience*, [S. l.], v. 6, p. 69, set. 2012.**

BU, X. et al. **Exploring white matter functional networks in children with attention-deficit/hyperactivity disorder. *Brain Communications*, [S. l.], v. 2, n. 2, p. fcaa113, jul. 2020.**

CARROLL, J. **Imagination, the brain’s default mode network, and imaginative verbal artifacts. *Evolutionary perspectives on imaginative culture*, [S. l.], v. 2, n. 2, p. 31–52, ago. 2020.**

CATANI, M., DELL’ACQUA, F. , DE SCHOTTEN, M. T. **IA revised limbic system model for memory, emotion and behaviour. *Neuroscience & Biobehavioral Reviews*, [S. l.], v. 37, n. 8, p. 1724–1737, set. 2013.**

CETIN, F. H. et al. **A case study on eeg analysis: embedding entropy estimations indicate the decreased neuro-cortical complexity levels mediated by methylphenidate treatment in children with ADHD.** *Clinical EEG and Neuroscience*, [S. l.], v. 53, n. 5, p. 406–417, dez. 2022.

CHANG, Z. et al. **Serious transport accidents in adults with attention-deficit /hyperactivity disorder and the effect of medication: a population-based study.** *JAMA psychiatry*, [S. l.], v. 71, n. 3, p. 319–325, mar. 2014.

CHAWLA, N. V. et al. **Smote: synthetic minority over-sampling technique.** *Journal of artificial intelligence research*, [S. l.], v. 16, p. 321–357, jun. 2002.

CHEN, H., SONG, Y. , LI, X. **A deep learning framework for identifying children with ADHD using an EEG-based brain network.** *Neurocomputing*, [S. l.], v. 356, p. 83–96, set. 2019a.

CHEN, H., SONG, Y. , LI, X. **Use of deep learning to detect personalized spatial-frequency abnormalities in EEGs of children with ADHD.** *Journal of neural engineering*, [S. l.], v. 16, n. 6, p. 066046, nov. 2019b.

CHEN, T. , GUESTRIN, C. **Xgboost: A scalable tree boosting system.** In: 22nd acm sigkdd international conference on knowledge discovery and data mining, 16, 2016, São Francisco, Estados Unidos da America. **Proceedings...** São Francisco [Estados Unidos da America], 2016. p. 785–794. Disponível em: <<https://doi.org/10.1145/2939672.2939785>>. Acesso em: 20 abr. 2024.

CLARKE, A. R., BARRY, R. J. , JOHNSTONE, S. **Resting state eeg power research in attention-deficit/hyperactivity disorder: A review update.** *Clinical Neurophysiology*, [S. l.], v. 131, n. 7, p. 1463–1479, jul. 2020.

CLEVERT, D.-A., UNTERTHINER, T., HOCHREITER, S. **Xgboost: Fast and accurate deep network learning by exponential linear units (elus).** In: 4th International Conference on Learning Representations, 2016, San Juan, Puerto Rico. **arXiv:1511.07289** San Juan [Puerto Rico], 2016. Disponível em: <<https://doi.org/10.1145/2939672.2939785>>. Acesso em: 20 abr. 2024.

COLBY, J. B. et al. **Insights into multimodal imaging classification of adhd.** *Frontiers in systems neuroscience*, [S. l.], v. 6, p. 59, ago. 2012.

COLE, M. W. et al. **Multi-task connectivity reveals flexible hubs for adaptive task control.** *Nature neuroscience*, [S. l.], v. 16, n. 9, p. 1348–1355, jul. 2013.

CONSORTIUM, A. **The adhd-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience.** *Frontiers in systems neuroscience*, [S. l.], v. 6, p. 62, set. 2012.

CRADDOCK, R. C. et al. **A whole brain fmri atlas generated via spatially constrained spectral clustering.** *Human brain mapping*, [S. l.], v. 33, n. 8, p. 1914–1928, jul. 2012.

CRICHTON, A. **An inquiry into the nature and origin of mental derangement: comprehending a concise system of the physiology and pathology of the human mind and a history of the passions and their effects.** Londres: T. Cadell, junior, and W. Davies v. 2, 1798. ISBN-1: 978-1014837585.

DAI, D. et al. **Classification of adhd children through multimodal magnetic resonance imaging.** *Frontiers in systems neuroscience*, [S. l.], v. 6, p. 63, set. 2012.

DALSGAARD, S. et al. **Adhd, stimulant treatment in childhood and subsequent substance abuse in adulthood — a naturalistic long-term follow-up study.** *Addictive behavior*, [S. l.], v. 39, n. 1, p. 325–328, jan. 2014.

DALSGAARD, S. et al. **Mortality in children, adolescents, and adults with attention deficit hyperactivity disorder: a nationwide cohort study.** *The Lancet*, [S. l.], v. 385, n. 9983, p. 2190–2196, fev. 2015.

DE SILVA, S. et al. **fMRI feature extraction model for ADHD classification using convolutional neural network.** *International Journal of E-Health and Medical Communications*, [S. l.], v. 12, n. 1, p. 81–105, 2021.

DE SOUZA, I. G. et al. **Dificuldades no diagnóstico de tdah em crianças.** *Jornal brasileiro de psiquiatria*, [S. l.], v. 56, p. 14–18, 2007.

DELORME, A. , MAKEIG, S. **EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis.** *Journal of neuroscience methods*, [S. l.], v. 134, n. 1, p. 9–21, mar. 2004.

DEMONTIS, D. et al. **Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder.** *Nature genetics*, [S. l.], v. 51, n. 1, p. 63–75,

nov. 2019.

DESHPANDE, G. et al. **Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data.** *IEEE transactions on cybernetics*, [S. l.], v. 45, n. 12, p. 2668–2679, jan. 2015.

DEY, S., RAO, A. R. , SHAH, M. **Attributed graph distance measure for automatic detection of attention deficit hyperactive disordered subjects.** *Frontiers in neural circuits*, [S. l.], v. 8, p. 64, jun. 2014.

DIXON, M. L. et al. **Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks.** *Proceedings of the National Academy of Sciences*, [S. l.], v. 115, n. 7, p. 1598–1607, set. 2018.

DRECHSLER, R. et al. **ADHD: Current concepts and treatments in children and adolescents.** *Neuropediatrics*, [S. l.], v. 51, n. 5, p. 315–335, 2020.

DRON, N. et al. Preliminary fusion of eeg and mri with phenotypic scores in children with epilepsy based on the canonical polyadic decomposition. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, Berlim, Alemanha. **Proceedings...** Berlim [Alemanha], 2019. p. 3884–3887. Disponível em: <<https://doi.org/10.1109/EMBC.2019.88563285>>. Acesso em: 20 abr. 2024.

DU, Y., FU, Z. , CALHOUN, V. D. **Classification and prediction of brain disorders using functional connectivity: Promising but challenging.** *Frontiers in Neuroscience*, [S. l.], v. 12, ago. 2018.

DUBREUIL-VALL, L., RUFFINI, G. , CAMPRODON, J. A. **Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG.** *Frontiers in Neuroscience*, [S. l.], v. 14, p. 251, abr. 2020.

DUFFY, K. A. et al. **Increased integration between default mode and task-relevant networks in children with adhd is associated with impaired response control.** *Developmental Cognitive Neuroscience*, [S. l.], v. 50, p. 100980, ago. 2021.

DUPUY, F. E. et al. **Sex differences between the combined and inattentive types of**

attention-deficit/hyperactivity disorder: an eeg perspective. *International Journal of Psychophysiology*, [S. l.], v. 89, n. 3, p. 320–327, set. 2013.

FARAONE, S. V. et al. **Attention-deficit/hyperactivity disorder.** *Nature reviews Disease primers*, [S. l.], v. 1, p. 150202, ago. 2015.

FITZGERALD, M., BELLGROVE, M. , GILL, M. **Handbook of attention deficit hyperactivity disorder, Advances in neural information processing systems.** New York: John Wiley & Sons, Hoboken, 2007. ISBN-10: 1681081520.

FOX, M. D. , GREICIUS, M. **Clinical applications of resting state functional connectivity.** *Frontiers in systems neuroscience*, [S. l.], v. 4, p. 1443, jun. 2010.

FREUND, Y. , SCHAPIRE, R. E. Experiments with a new boosting algorithm. In: 13th International Conference on Machine Learning (ICML), 1996, Bari, Itália. **Proceedings...** Bari [Itália], 1996. p. 148–156. Disponível em: <<http://dl.acm.org/citation.cfm?id=3091696.3091715>>. Acesso em: 20 abr. 2024.

FREY, B. J. , DUECK, D. Mixture modeling by affinity propagation. In: Neural Information Processing Systems (NIPS), 2005, Vancouver, Canada. **Proceedings...** Vancouver [Canada], 2005. v. 18. Disponível em: <<https://dl.acm.org/doi/abs/10.5555/2976248.2976296>>. Acesso em: 20 abr. 2024.

FRISTON, K. J. et al. **Movement-related effects in fMRI time-series.** *Magnetic resonance in medicine*, [S. l.], v. 35, n. 3, p. 346–355, mar. 1996.

FUKUSHIMA, K. **Visual feature extraction by a multilayered network of analog threshold elements.** *IEEE Transactions on Systems Science and Cybernetics*, [S. l.], v. 5, n. 4, p. 322–333, out. 1969.

FURCZYK, K. , THOME, J. **Adult adhd and suicide.** *ADHD Attention Deficit and Hyperactivity Disorders*, [S. l.], v. 6, p. 153–158, jul. 2014.

GHIASSIAN, S. et al. **Using functional or structural magnetic resonance images and personal characteristic data to identify adhd and autism.** *PLOS one*, [S. l.], v. 11, n. 12, p. e0166934, dez. 2016.

GIBBINS, C. et al. **Adhd-hyperactive/impulsive subtype in adults.** *Mental Illness*, [S. l.], v. 2, n. 1, p. 41–45, jan. 2010.

GIZER, I. R., FICKS, C. , WALDMAN, I. D. **Candidate gene studies of ADHD: a meta-analytic review.** *Human genetics*, [S. l.], v. 126, p. 51–90, jun. 2009.

GREVE, D. N. , FISCHL, B. **Accurate and robust brain image alignment using boundary-based registration.** *Neuroimage*, [S. l.], v. 48, n. 1, p. 63–72, out. 2009.

HART, H. et al. **Meta-analysis of fMRI studies of timing in attention-deficit hyperactivity disorder (ADHD).** *Neuroscience & Biobehavioral Reviews*, [S. l.], v. 36, n. 10, p. 2248–2256, nov. 2012.

HART, H. et al. **Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects.** *JAMA psychiatry*, [S. l.], v. 70, n. 2, p. 185–198, fev. 2013.

HAYKIN, S. **Redes Neurais** Porto Alegre: Bookman, 2000. 2 ed. ISBN: 8573077182.

HE, K. et al. Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, 2016, Las Vegas, Estados Unidos da America. **Proceedings...** Las Vegas [Estados Unidos da America], 2016. p. 770–778. Disponível em: <<https://dl.acm.org/doi/abs/10.5555/2976248.2976296>>. Acesso em: 20 abr. 2024.

HERAULT, J., JUTTEN, C. Space or time adaptive signal processing by neural network models. In: American Institute of Physics (AIP) conference proceedings, 1986, Woodbury, Estados Unidos da America. **Proceedings...** Woodbury [Estados Unidos da America], 1986. v. 151, n. 1, , p. 206–211. Disponível em: <<https://pubs.aip.org/aip/acp/article-abstract/151/1/206/755690>>. Acesso em: 20 abr. 2024.

HERMAN, A. M., CRITCHLEY, H. D. , DUKA, T. **Trait impulsivity associated with altered resting-state functional connectivity within the somatomotor network.** *Frontiers in Behavioral Neuroscience*, [S. l.], v. 14, p. 111, jun. 2020.

HERRMANN, C. S., FRÜND, I. , LENZ, D. **Human gamma-band activity: a review on cognitive and behavioral correlates and network models.** *Neuroscience & Biobehavioral Reviews*, [S. l.], v. 34, n. 7, p. 981–992, jun. 2010.

HINSHAW, S. P. **Attention deficit hyperactivity disorder (ADHD): controversy, developmental mechanisms, and multiple levels of analysis.** *Annual review of clinical psychology*, [S. l.], v. 14, p. 291–316, dez. 2018.

HO, T. K. Random decision forests. In: 3rd international conference on document analysis and recognition, 1995, Montréal, Canada. **Proceedings...** Montréal [Canada], 1995. v. 1, p. 278–282. Disponível em: <<https://doi.org/10.1063/1.36258>>. Acesso em: 20 abr. 2024.

HOLKER, R. , SUSAN, S. **Computer-aided diagnosis framework for ADHD detection using quantitative EEG, Brain Informatics.** *Brain Informatics*, [S. l.], p. 229–240, ago. 2022.

HOOGMAN, M. et al. **Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis.** *The Lancet Psychiatry*, [S. l.], v. 4, n. 4, p. 310–319, fev. 2017.

HUGHES, C. et al. **Age differences in specific neural connections within the default mode network underlie theory of mind.** *NeuroImage*, [S. l.], v. 191, p. 269–277, maio 2019.

JATOI, M. A., KAMEL, N. **Brain source localization using EEG signal analysis.** Boca Raton, Estados Unidos da América: CRC Press, 2017. ISBN-13: 978-1498799348.

JENKINSON, M. et al. **Improved optimization for the robust and accurate linear registration and motion correction of brain images.** *NeuroImage*, [S. l.], v. 17, n. 2, p. 825–841, out. 2002.

JHA, R. R. et al. HLGSNET: Hierarchical and lightweight graph siamese network with triplet loss for fMRI-based classification of ADHD. In: 2020 International Joint Conference on Neural Networks (IJCNN), 2020, Glasgow, Escócia. **Proceedings...** Glasgow [Escócia], 2020. p. 1–7. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9207667>>. Acesso em: 20 abr. 2024.

JIANG, J.-A. et al. **An automatic analysis method for detecting and eliminating ECG artifacts in EEG.** *Computers in biology and medicine*, [S. l.], v. 37, n. 11, p. 1660–1671, nov. 2007.

JOY, C. R. et al. **Detection of ADHD from EEG signals using different entropy me-**

asures and ANN. **Clinical EEG and Neuroscience**, [S. l.], v. 53, n. 1, p. 12–23, ago. 2022.

KAMALI, A. et al. **The cortico-limbo-thalamo-cortical circuits: An update to the original papez circuit of the human limbic system**. **Brain Topography**, [S. l.], p. 1–19, abr. 2023.

KARCH, S. et al. **Neural correlates (ERP/fMRI) of voluntary selection in adult ADHD patients**. **European Archives of Psychiatry and Clinical Neuroscience**, [S. l.], v. 260, p. 427–440, nov. 2010.

KHARE, S. K., GAIKWAD, N. B. , BAJAJ, V. **VHERS: a novel variational mode decomposition and hilbert transform-based eeg rhythm separation for automatic ADHD detection**. **IEEE Transactions on Instrumentation and Measurement**, [S. l.], v. 71, p. 1–10, set. 2022.

KIM, H. E., AVRAHAM, G. , IVRY, R. B. **The psychology of reaching: action selection, movement implementation, and sensorimotor learning**. **Annual review of psychology**, [S. l.], v. 72, p. 61–95, jan. 2021.

KIM, S. B. et al. **The usefulness of quantitative electroencephalography in diagnosis and severity evaluation of delirium: A retrospective study**. **Psychiatry Investigation**, [S. l.], v. 20, n. 2, p. 144, fev. 2023.

KIRMIZI-ALSAN, E. et al. **Comparative analysis of event-related potentials during Go/NoGo and CPT: Decomposition of electrophysiological markers of response inhibition and sustained attention**. **Brain Research**, [S. l.], v. 1104, n. 1, p. 114–128, ago. 2006.

KLAMBAUER, G. et al. **Self-normalizing neural networks**. In: **Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)**, 2017, Long Beach, Estados Unidos da América. **Proceedings...** Long Beach [Estados Unidos da América], 2017. v. 30 p. 972–981. Disponível em: <<https://dl.acm.org/doi/10.5555/3294771.3294864>>. Acesso em: 20 abr. 2024.

KOFLER, M. J. et al. **Working memory deficits and social problems in children with ADHD**. **Journal of abnormal child psychology**, [S. l.], 39, p. 805–817, abr. 2011.

University of California (Estados Unidos da América). Christian Andreas Edgar Kothe,

Tzyy-Ping Jung. **Artifact removal techniques with signal reconstruction.** US n. PI US20160113587A1, 28 abr. 2016 (depósito),- .

LAHEY, B. B et al. **Instability of the dsm-iv subtypes of adhd from preschool through elementary school.** *Archives of general psychiatry*, [S. l.], v. 62, n. 8, p. 896–902, 2005.

LAI, T. K. et al. **Development of a peptide targeting dopamine transporter to improve ADHD-like deficits.** *Molecular brain*, [S. l.], v. 11, p. 1–14, nov. 2018.

LANGE, K. W. et al. **The history of attention deficit hyperactivity disorder, ADHD Attention Deficit and Hyperactivity Disorders.** *ADHD Attention Deficit and Hyperactivity Disorders*, [S. l.], v. 2, p. 241–255, nov. 2010.

LE CUN, Y. et al. Handwritten digit recognition with a back-propagation network. In: 2nd International Conference on Advances in Neural Information Processing Systems 2 (NIPS), 1989, Cambridge, Estados Unidos da América. **Proceedings...** Cambridge [Estados Unidos da América], 1989. v. 2, p. 396–404. Disponível em: <<https://dl.acm.org/doi/10.5555/109230.109279>>. Acesso em: 20 abr. 2024.

LENARTOWICZ, A. , LOO, S. K. **Use of EEG to diagnose ADHD.** *Current psychiatry reports*, [S. l.], v. 16, p. 1–11, set. 2014.

LENARTOWICZ, A. et al. **Alpha desynchronization and frontoparietal connectivity during spatial working memory encoding deficits in adhd: A simultaneous EEGfMRI study.** *NeuroImage: Clinical*, [S. l.], v. 11, p. 210–223, 2016.

LIN, M., CHEN, Q. , YAN, S. Network in network. In: First International Conference on Learning Representations, 2013, Scottsdale, Estados Unidos da América. **arXiv...** Ithaca [Estados Unidos da América], 2014. arXiv:1312.4400. Disponível em: <<https://api.semanticscholar.org/CorpusID:16636683>>. Acesso em: 20 abr. 2024.

LIU, S. et al. **Attention deficit/hyperactivity disorder classification based on deep spatio-temporal features of functional magnetic resonance imaging.** *Biomedical Signal Processing and Control*, [S. l.], v. 71, p. 103239, jan. 2022.

LOO, S. K. et al. **Characterization of the theta to beta ratio in adhd: identifying potential sources of heterogeneity.** *Journal of attention disorders*, [S. l.], v. 17, n. 5, p.

384–392, dez. 2013.

LUBAR, J. F. **Discourse on the development of eeg diagnostics and biofeedback for attention-deficit/hyperactivity disorders. Biofeedback and Self-regulation**, [S. l.], v. 16, p. 201–225, set. 1991.

MAO, Z. et al. **Spatio-temporal deep learning method for adhd fmri classification. Information Sciences**, [S. l.], v. 499, p. 1–11, out. 2019.

MAREK, S. , DOSENBACH, N. U. **The frontoparietal network: function, electrophysiology, and importance of individual precision mapping. Dialogues in clinical neuroscience**, [S. l.], v. 20, n. 2, p. 133–140, abr. 2018.

MARKOVSKA-SIMOSKA, S. , POP-JORDANOVA, N. **Quantitative eeg in children and adults with attention deficit hyperactivity disorder: comparison of absolute and relative power spectra and theta/beta ratio. Clinical EEG and neuroscience**, [S. l.], v. 48, n. 1, p. 20–32, jul. 2017.

MAZZIOTTA, J. C. et al. **A probabilistic atlas of the human brain: theory and rationale for its development. Neuroimage**, [S. l.], v. 2, n. 2, p. 89–101, abr. 1995.

MENG, X. et al. **Diagnostic model optimization method for adhd based on brain network analysis of resting-state fmri images and transfer learning neural network. Frontiers in Human Neuroscience**, [S. l.], v. 16, p. 1005425, out. 2022.

MENON, V. **Large-scale brain networks and psychopathology: a unifying triple network model. Trends in cognitive sciences**, [S. l.], v. 15, n. 10, p. 483–506, out. 2011.

MERCAN ISIK, C. et al. **The effect of methylphenidate treatment on olfactory function in children and adolescents with ADHD. Journal of Attention Disorders**, [S. l.], p. 10870547231171727, maio 2023.

MIRANDA, L. et al. **Functional MRI applications for psychiatric disease subtyping: a review. arXiv preprint arXiv:2007.00126**, [S. l.], p. 10870547231171727, maio 2020.

MOGHADDARI, M., LIGHVAN, M. Z. , DANISHVAR, S. **Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG. Computer Methods and Programs in Biomedicine**, [S. l.], v. 197, p. 105738,

dez. 2020.

MORGANE, P. J., GALLER, J. R. , MOKLER, D. J. **A review of systems and networks of the limbic forebrain/limbic midbrain.** *Progress in neurobiology*, [S. l.], v. 75, n. 2, p. 143–160, fev. 2005.

MULERT, C. , LEMIEUX, L. 2023. **EEG-fMRI: physiological basis, technique, and applications.** Berlin: Springer Nature, 2023. ISBN: 9783540879183.

MUSSO, M. W., GOUVIER, W. M. D. **“why is this so hard?” a review of detection of malingered adhd in college students.** *Journal of Attention Disorders*, [S. l.], v. 18, n. 3, p. 186–201, maio 2014.

MWAMBA, H. M., FOURIE, P. R., VAN DEN HEEVER, D. **Pandas: Paediatric attention-deficit/hyperactivity disorder application software.** *Applied Sciences*, [S. l.], v. 9, n. 8, p. 1645, abr. 2019.

NENTWICH, M. et al. **Functional connectivity of EEG is subject-specific, associated with phenotype, and different from fMRI.** *NeuroImage*, [S. l.], v. 218, p. 117001, maio 2020.

NIGG, J. T. et al. **Toward a revised nosology for attention-deficit/hyperactivity disorder heterogeneity.** *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, [S. l.], v. 5, n. 8, p. 726–737, ago. 2020.

NÚÑEZ-JARAMILLO, L., HERRERA-SOL S, A., HERRERA-MORALES, W. V. **ADHD: Reviewing the causes and evaluating solutions.** *Journal of personalized medicine*, [S. l.], v. 11, n. 3, p. 166, mar. 2021.

OGAWA, S. et al. **Brain magnetic resonance imaging with contrast dependent on blood oxygenation.** *Proceedings of the National Academy of Sciences*, [S. l.], v. 87, n. 24, p. 9868–9872, dez. 1990.

O’CONNOR, D. et al. **The healthy brain network serial scanning initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions.** *Gigascience*, [S. l.], v. 6, n. 2, p. giw011, fev. 2017.

PAARMANN, L. D. **Design and analysis of analog filters: a signal processing perspective.** Berlin: Springer Science Business Media, 2005. 617 p. ISBN-13: 978-0792373735.

PALLOYELIS, Y. et al. **Functional MRI in adhd: a systematic literature review. Expert review of neurotherapeutics**, [S. l.], v. 7, n. 10, p.1337–1356, jan. 2007.

PAN, Y. et al. **Operand-order-specific representation in simple addition and multiplication: Evidence from an ERP study. Social Science Research Network** , [S. l.], 4699552, jan. 2024. No prelo.

PARK, H.-J. et al. **State-dependent effective connectivity in resting-state fMRI. Frontiers in Neural Circuits**, [S. l.], p. 112, out. 2021.

PEDROLLO, G. R. et al. Spiking neural networks diagnosis of ADHD subtypes through EEG signals evaluation. In: 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, Glasgow, Escócia. **Proceedings...** Glasgow [Escócia], 2022. p. 3166–3169. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9871223>>. Acesso em: 20 abr. 2024.

PEREDA, E. et al. **The blessing of dimensionality: Feature selection outperforms functional connectivity-based feature transformation to classify adhd subjects from EEG patterns of phase synchronisation. PLOS one**, [S. l.], v. 13, n. 8, p. e0201660, ago. 2018.

PFEIFER, P. et al. **Prefrontal and striatal dopamine d2/d3 receptors correlate with fMRI bold activation during stopping. Brain imaging and behavior**, [S. l.], v. 16, n. 1, p. 186–198, ago. 2022.

PHAM, A. V., RIVIERE, A. **Specific learning disorders and adhd: current issues in diagnosis across clinical and educational settings. Current psychiatry reports**, [S. l.], v. 17, p. 1–7, abr. 2015.

PHILIPPI, C. L. et al. **Lesion network mapping demonstrates that mind-wandering is associated with the default mode network. Journal of neuroscience research**, [S. l.], v. 99, n. 1, p. 361–373, jun. 2021.

PINEAU, G. et al. **Dopamine transporter genotype modulates brain activity during a working memory task in children with ADHD. Research in developmental disabilities**, [S. l.], v. 92, p. 103430, set. 2019.

QUIST, J. et al. **The serotonin 5-ht1b receptor gene and attention deficit hyperactivity disorder.** *Molecular psychiatry*, [S. l.], v. 8, n. 1, p. 98–102, jan. 2003.

QURESHI, M. N. I. et al. **Multiclass classification for the differential diagnosis on the ADHD subtypes using recursive feature elimination and hierarchical extreme learning machine: structural MRI study.** *PLOS one*, [S. l.], v. 11, n. 8, p. e0160697, ago. 2016.

QURESHI, M. N. I. et al. **Multi-modal, multi-measure, and multi-class discrimination of ADHD with hierarchical feature extraction and extreme learning machine using structural and functional brain MRI.** *Frontiers in human neuroscience*, [S. l.], v. 11, p. 157, abr. 2017.

RAMALHO, R. et al. **Docosahexaenoic acid supplementation for children with attention deficit hyperactivity disorder: A comprehensive review of the evidence.** *Clinical nutrition ESPEN*, [S. l.], v. 25, p. 1–7, jun. 2018.

RIAZ, A. et al. **Fcnet: a convolutional neural network for calculating functional connectivity from functional MRI, Connectomics in NeuroImaging.** In: 20th International Conference on Medical Image Computing and Computer Assisted Intervention 2017, 2017, Quebec, Canada. **Proceedings...** Quebec [Canada], 2017. p. 70–78. Disponível em: <<https://doi.org/10.1109/EMBC48229.2022.9871223>>. Acesso em: 20 abr. 2024.

RIAZ, A. et al. **Fusion of fmri and non-imaging data for adhd classification.** *Computerized Medical Imaging and Graphics*, [S. l.], v. 65, p. 115–128, abr. 2018.

RIECANSK ‘Y, I. , LAMM, C. **The role of sensorimotor processes in pain empathy.** *Brain topography*, [S. l.], v. 32, n. 6, p. 965–976, nov. 2019.

RODRIGUEZ, A. , LAIO, A. **Clustering by fast search and find of density peaks.** *Science*, [S. l.], v. 344, n. 6191, p. 1492–1496, jun. 2014.

ROSTAMI, M. et al. **Classifying adhd subtypes/presentations considering the joint effect of three levels of investigation.** *Nordic Journal of Psychiatry*, [S. l.], v. 75, n. 1, p. 31–37, out. 2021.

RUBIA, K. **Cognitive neuroscience of attention deficit hyperactivity disorder (ADHD) and its clinical translation.** *Frontiers in Human Neuroscience*, [S. l.], v. 12, mar. 2018.

SAAD, J. F. et al. **A systematic review of imaging studies in the combined and inattentive subtypes of attention deficit hyperactivity disorder.** *Frontiers in Integrative Neuroscience*, [S. l.], v. 14, p. 31, jun. 2020.

SAAD, J. F. et al. **Is the theta/beta eeg marker for adhd inherently flawed?** *Journal of attention disorders*, [S. l.], v. 22, n. 9, p. 815–826, mar. 2018.

SAHA, P., SARKAR, D. **Characterization and classification of adhd subtypes: An approach based on the nodal distribution of eigenvector centrality and classification tree model.** *Child Psychiatry & Human Development*, [S. l.], p. 1–13, set. 2022.

SALVI, V. et al. **Adhd in adults: clinical subtypes and associated characteristics.** *Rivista di Psichiatria*, [S. l.], v. 54, n. 2, p. 84–89, abr. 2019.

SANDER, C. et al. **EEG-vigilance and response to stimulants in paediatric patients with attention deficit/hyperactivity disorder.** *Clinical Neurophysiology*, [S. l.], v. 121, n. 9, p. 1511–1518, set. 2010.

SANEI, S., CHAMBERS J. A. **EEG signal processing and machine learning.** Hoboken, EUA, EUA: John Wiley & Sons, 2021. ISBN-13: 978-1119386940.

SCHAPIRE, R. E. **Explaining adaboost, Empirical Inference: Festschrift in Honor of Vladimir , n. Vapnik** Hoboken, EUA, EUA: John Wiley & Sons, 2013. ISBN-13: 978-3642411373.

SCIBERRAS, E. et al. **Prenatal risk factors and the etiology of adhd—review of existing evidence.** *Current psychiatry reports*, [S. l.], v. 19, p. 1–8, 2017.

SHAH, P. , MIYAKE, A. **Models of Working Memory: An Introduction. Mechanisms of Active Maintenance and Executive Control**, [S. l.], p. 1–27, jun. 1999. Disponível em: <<https://doi.org/10.1017/CBO9781139174909.004>>. Acesso em: 23 abr. 2024.

SHAKUR, A. H. et al. **A rule-based exploratory analysis for discovery of multimodal biomarkers of adhd using eye movement and EEG data.** *IISE Transactions on Healthcare Systems Engineering*, [S. l.], v. 13, n. 1, p. 74–88, out. 2023.

SHARMA, A. et al. **iPal: A machine learning based smart healthcare framework for**

automatic diagnosis of attention deficit/hyperactivity disorder (ADHD). *SN Computer Science*, [S. l.], v. 5, n. 4, p. 1-19.

SIMONYAN, K. , ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: First International Conference on Learning Representations, 2013, Scottsdale, Estados Unidos da América. **arXiv...** Ithaca [Estados Unidos da América], 2014. arXiv:1409.1556. Disponível em: <<https://api.semanticscholar.org/CorpusID:16636683>>. Acesso em: 23 abr. 2024.

SIPPER, M. , MOORE, J. H. **Addgboost: A gradient boosting-style algorithm based on strong learners.** *Machine Learning with Applications*, [S. l.], v. 7, p. 100243, mar. 2022.

SLATER, J. et al. **Can electroencephalography (EEG) identify ADHD subtypes? A systematic review.** *Neuroscience & Biobehavioral Reviews*, [S. l.], p. 104752, jun. 2022.

SMALLEY, S. L. et al. **Prevalence and psychiatric comorbidity of attention-deficit /hyperactivity disorder in an adolescent finnish population.** *Journal of the American Academy of Child & Adolescent Psychiatry*, [S. l.], v. 46, n. 12, p. 1575–1583, dez. 2007.

SPRENG, R. N., MAR, R. A. , KIM, A. S. **The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis.** *Journal of cognitive neuroscience*, [S. l.], v. 21, n. 3, p. 489–510, mar. 2009.

SQUIRE, L. R. **Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans.** *Psychological review*, [S. l.], v. 99, n. 2, p. 195, abr. 1992.

STILL, G. F. **Some abnormal psychical conditions in children.** *Lancet*, [S. l.], v. 1, p. 1077–1082, 1902.

STRIK, W. et al. **Larger topographical variance and decreased duration of brain electric microstates in depression.** *Journal of Neural Transmission/General Section*, [S. l.], v. 99, p. 213–222, fev. 1995.

SU, S. et al. **Altered neurovascular coupling in the children with attention-deficit /hyperactivity disorder: a comprehensive fmri analysis.** *European Child & Adolescent*

Psychiatry, [S. l.], p. 1–11, maio 2023.

SWANSON, J. M. et al. **Categorical and dimensional definitions and evaluations of symptoms of adhd: history of the snap and the swan rating scales. The International journal of educational and psychological assessment**, [S. l.], v. 10, n. 1, p. 51, abr. 2012.

TAGHI BEYGLOU, B. et al. ADHD diagnosis in children using common spatial pattern and nonlinear analysis of filter banked EEG. In: 2020 28th Iranian Conference on Electrical Engineering (ICEE), 2020, Tabriz, Irã. **Proceedings...** Tabriz [Irã], 2020. p. 51. Disponível em: <<https://ieeexplore.ieee.org/document/9260711>>. Acesso em: 21 abr. 2024.

TOR, H. T. et al. **Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with eeg signals. Computer Methods and Programs in Biomedicine**, [S. l.], v. 200, p. 105941, mar. 2021.

TORRICO, T. J., ABDIJADID, S. **Neuroanatomy, limbic system. Treasure Island (FL): StatPearls Publishing**, [S. l.], PMID 30860726, mar. 2019.

TOSUN, M. **Effects of spectral features of EEG signals recorded with different channels and recording statuses on ADHD classification with deep learning. Physical and Engineering Sciences in Medicine**, [S. l.], v. 44, n. 3, p. 693–702, maio 2021.

URIGÜEN, J. A., GARCIA-ZAPIRAIN, B. **EEG artifact removal—state-of-the-art and guidelines. Journal of neural engineering**, [S. l.], v. 12, n. 3, p. 031001, abr. 2015.

VAHID, A. et al. **Deep learning based on event-related EEG differentiates children with adhd from healthy controls. Journal of clinical medicine**, [S. l.], v. 8, n. 7, p. 1055, jul. 2019.

VINAYAK, R. K., GILAD-BACHRACH, R. Dart: Dropouts meet multiple additive regression trees. In: Eighteenth International Conference on Artificial Intelligence and Statistics, 2015, San Diego, Estados Unidos da América. **Proceedings...** San Diego [Estados Unidos da América], 2015. p. 489–497. Disponível em: <<https://proceedings.mlr.press/v38/korlakaivinayak15.html>>. Acesso em: 21 abr. 2024.

VOLKOW, N. D. et al. **Evaluating dopamine reward pathway in ADHD: clinical implications.** *JAMA*, [S. l.], v. 302, n. 10, p. 1084–1091, set. 2009.

VOS, T. et al. **Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015.** *The Lancet*, [S. l.], v. 388, n. 10053, p. 1545–1602, out. 2016.

WANG, T., SUN, M. , HU, K. Dilated deep residual network for image denoising. In: *IEEE 29th international conference on tools with artificial intelligence (ICTAI)*, 2017, Boston, Estados Unidos da América. **Proceedings...** Boston [Estados Unidos da América], 2017. p. 1272–1279. Disponível em: <<https://ieeexplore.ieee.org/document/8372095>>. Acesso em: 21 abr. 2024.

WASSERSTEIN, J. **Diagnostic issues for adolescents and adults with adhd.** *Journal of clinical psychology*, [S. l.], v. 61, n. 5, p. 535–547, fev. 2005.

WEIKARD, M. A. **“der” philosophische arzt** Alemanha: J.T. von Trattner, Linz, 1787. v. 3. ISBN-13. 978-1162010533.

WEIS, R. et al. **Adhd assessment in college students: Psychologists’ adherence to dsm-5 criteria and multi-method/multi-informant assessment.** *Journal of Psychoeducational Assessment*, [S. l.], v. 37, n. 2, p. 209–225, out. 2019.

WICKRAMARATNA, J., HOLDEN, S., BUXTON, B. Performance degradation in boosting. In: *Multiple Classifier Systems: Second International Workshop, 2001*, Cambridge, Reino Unido. **Proceedings...** Cambridge [Reino Unido], 2001. p. 11–21. Disponível em: <https://link.springer.com/chapter/10.1007/3-540-48219-9_2>. Acesso em: 21 abr. 2024.

WIRSICH, J. et al. **Multi-timescale hybrid components of the functional brain connectome: A bimodal EEG-fMRI decomposition.** *Network Neuroscience*, [S. l.], v. 4, n. 3, p. 658–677, jul. 2020.

WOLRAICH, M. L. et al. **ADHD diagnosis and treatment guidelines: a historical perspective.** *Pediatrics*, [S. l.], v. 144, n. 4, out. 2019.

WOLTERING, S. et al. **Resting state EEG oscillatory power differences in adhd college students and their peers.** *Behavioral and Brain Functions*, [S. l.], v. 8, n. 1, p.

1–9, dez. 2012.

YANG, S., HOQUE, S., DERAVID, F. **Adaptive template enhancement for improved person recognition using small datasets.** *arXiv preprint*, [S. l.], arXiv:2201.01218, jan. 2022. No prelo.

YAZAN, E. , TALU, M. F. Comparison of the stochastic gradient descent based optimization techniques. In: International Artificial Intelligence and Data Processing Symposium, 2017, Malatya, Turquia. **Proceedings...** Malatya [Turquia], 2017. p. 1–5. Disponível em: <<https://doi.org/10.1109/IDAP.2017.8090299>>. Acesso em: 21 abr. 2024.

YEO, B. T. et al. **The organization of the human cerebral cortex estimated by intrinsic functional connectivity.** *Journal of neurophysiology*, [S. l.], v. 106, n. 3, p. 1125-1165, set. 2011.

ZHANG, Z. , LI, X. Use transfer learning to promote identification adhd children with eeg recordings. In: Chinese Automation Congress (CAC), 2019, Hangzhou, China. **Proceedings...** Hangzhou [China], 2019. p. 2809–2813. Disponível em: <<https://doi.org/10.1109/CAC48633.2019.8997426>>. Acesso em: 21 abr. 2024.

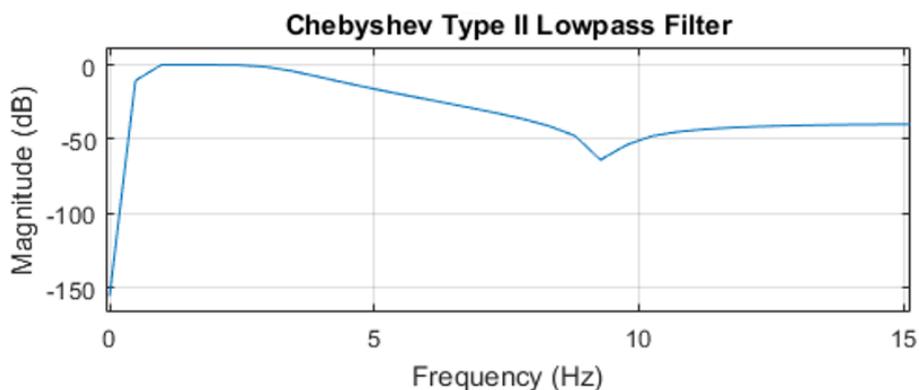
ZOU, L.. et al. **3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI.** *IEEE Access*, [S. l.], v. 5, p. 23626–23636, out. 2017.

APÊNDICE A GRÁFICOS DAS PCAS DAS BANDAS DE FREQUÊNCIA DE EEG

A.1 Gráficos das respostas em frequência dos filtros de EEG

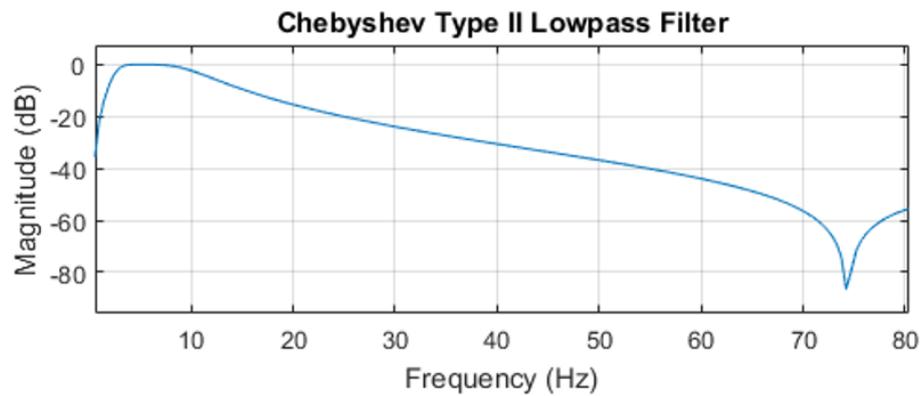
As figuras 49, 50, 51, 52 e 53 mostram as magnitudes das respostas em frequência dos filtros Chebyshev projetados, respectivamente, para permitir a passagem das bandas Delta, Teta, Alfa, Beta e Gama. Para eliminar as distorções causadas por deslocamento de fase, todos os filtros utilizados neste trabalho são filtros digitais de fase zero. Por isto, cada gráfico de fase da resposta em frequência provocada pela implementação destes filtros não é mostrado junto a magnitude correspondente.

Figura 49 – Filtro passa-banda na faixa de frequências Delta (0,53-3,9 Hz).



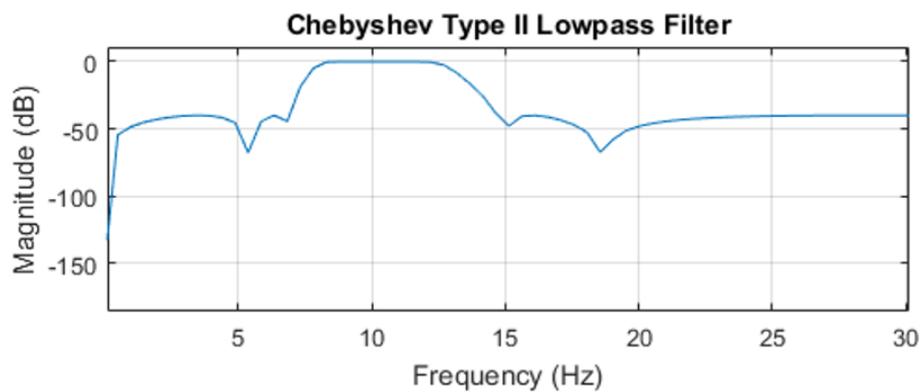
Fonte: O Autor

Figura 50 – Filtro passa-banda na faixa de frequências Teta (4-7,5 Hz).



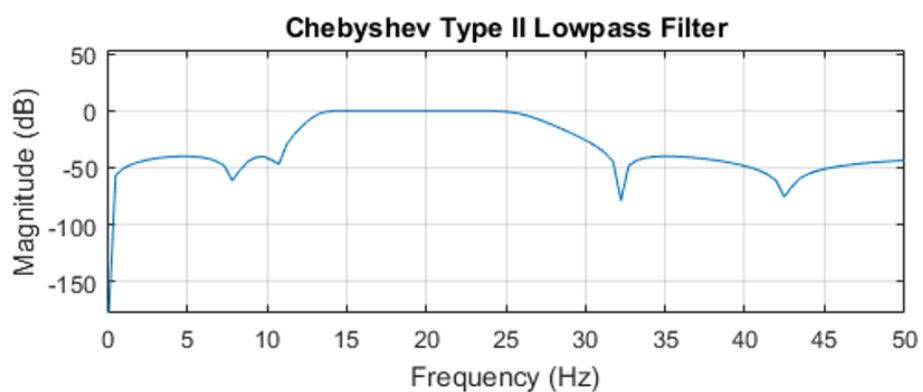
Fonte: O Autor

Figura 51 – Filtro passa-banda na faixa de frequências Alfa (8 - 13 Hz).



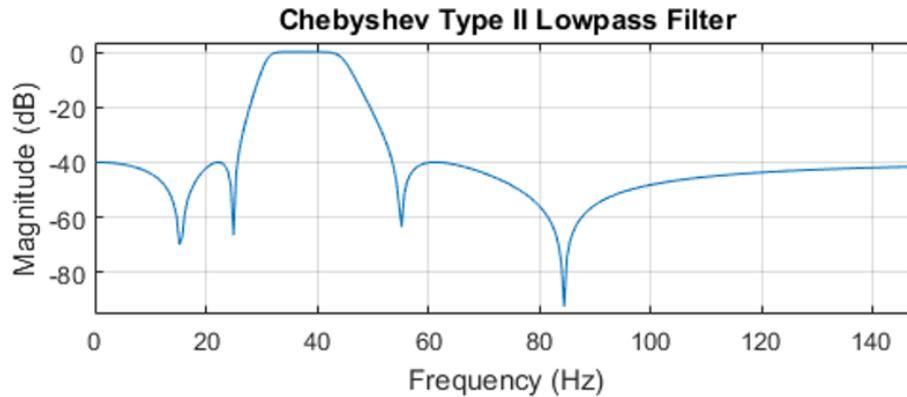
Fonte: O Autor

Figura 52 – Filtro passa-banda na faixa de frequências Beta (13,1 - 26 Hz).



Fonte: O Autor

Figura 53 – Filtro passa-banda na faixa de frequências Gama (32-45 Hz).

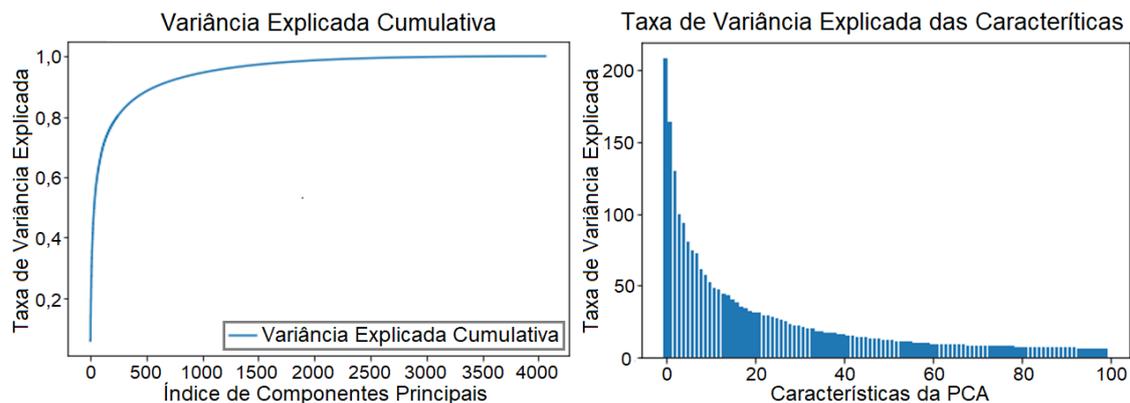


Fonte: O Autor

A.2 Gráficos das PCAs das Bandas de frequência de EEG para o diagnóstico entre grupos TDAH e controle

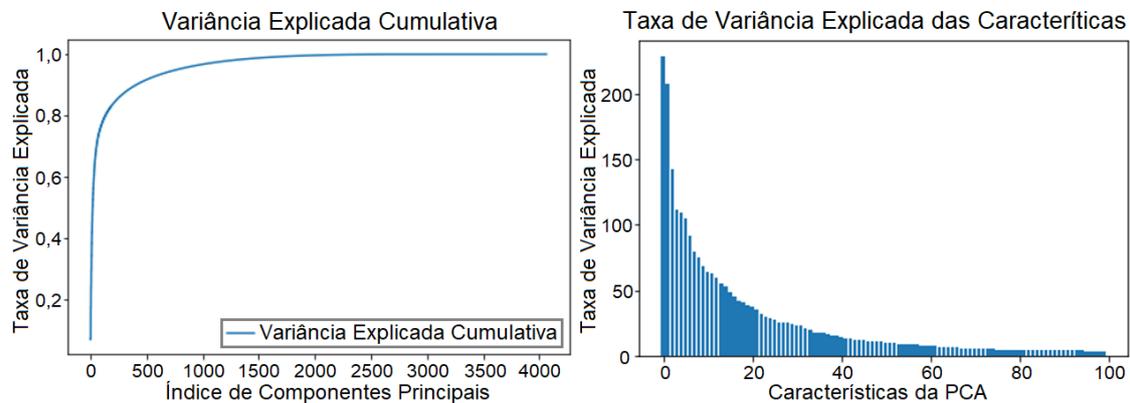
As figuras 54, 55, 56, 57, 58, 59 apresentam os gráficos das Variâncias Explicadas Cumulativas e das Taxa de Variância Explicada das Características das PCAs correspondentes às bandas larga, Delta, Teta, Alfa, Beta e Gama, respectivamente, considerando a classificação entre grupos TDAH e controle.

Figura 54 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda larga.



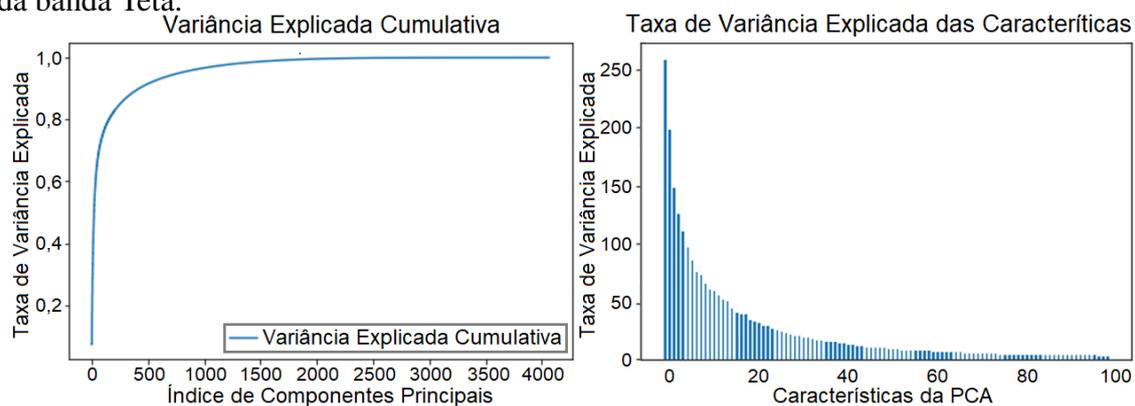
Fonte: O Autor

Figura 55 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Delta.



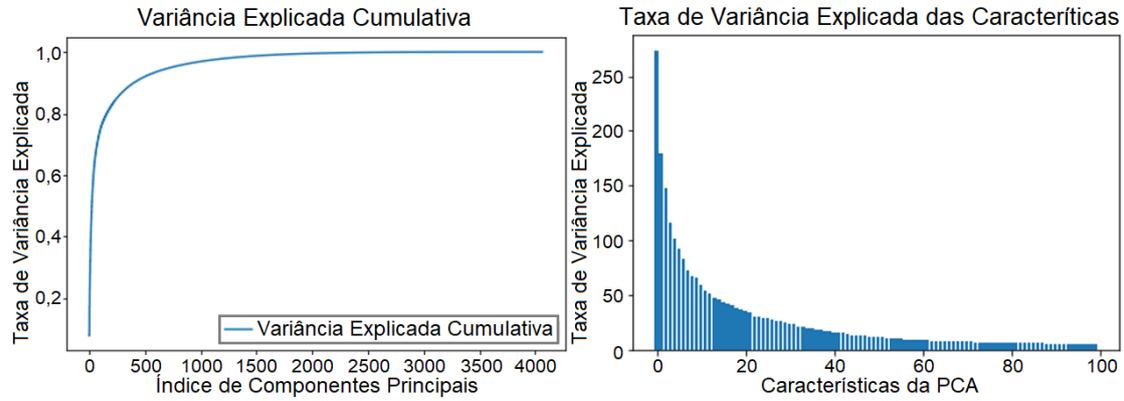
Fonte: O Autor

Figura 56 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Teta.



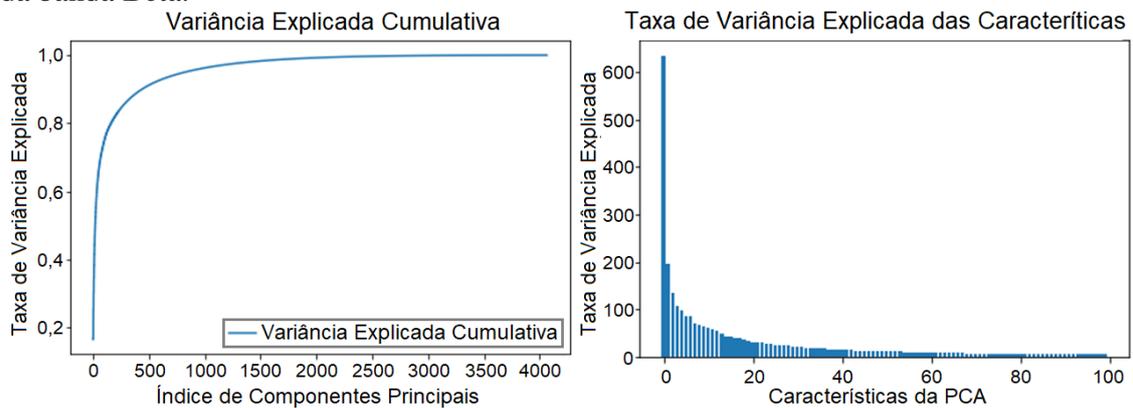
Fonte: O Autor

Figura 57 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Alfa.



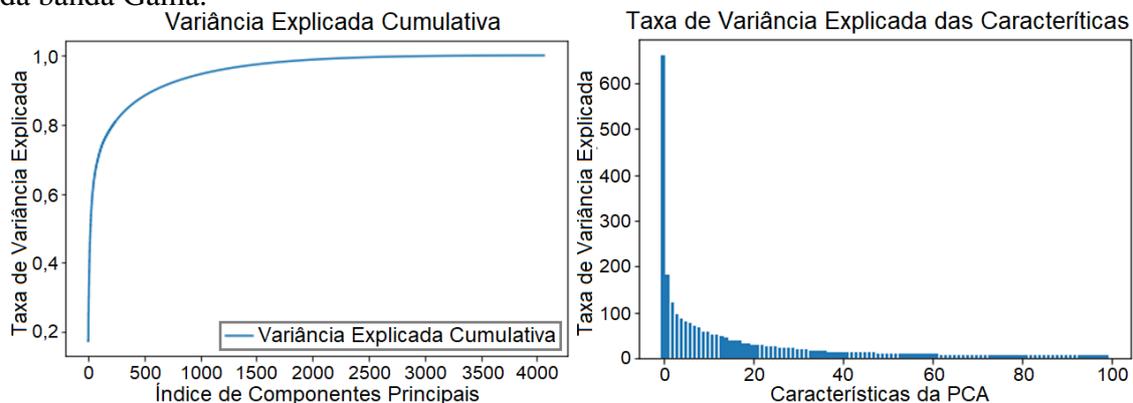
Fonte: O Autor

Figura 58 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Beta.



Fonte: O Autor

Figura 59 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH/Controle com características da banda Gama.

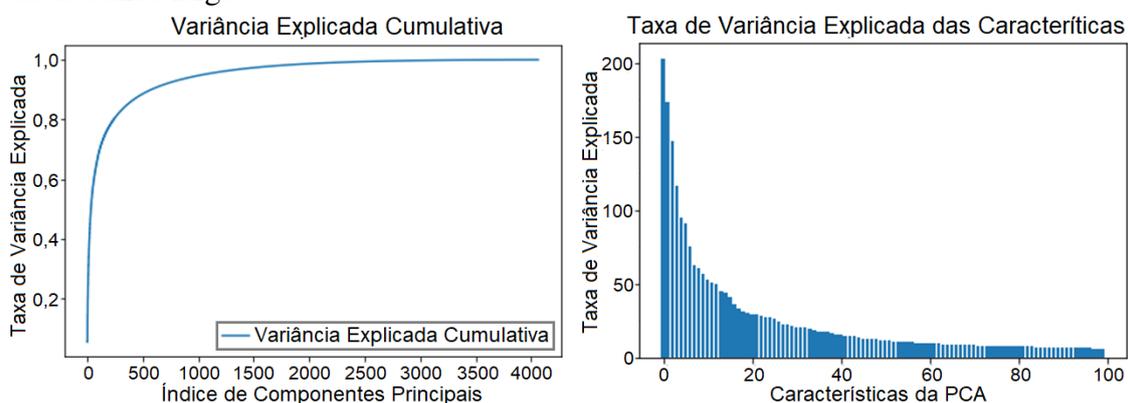


Fonte: O Autor

A.3 Gráficos das PCAs das Bandas de frequência de EEG diagnóstico TDAH-C/TDAH-I

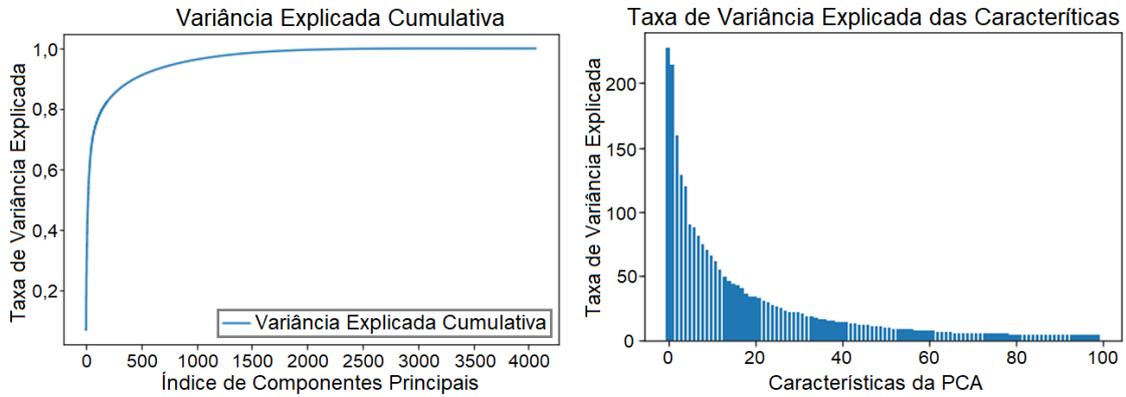
As figuras 60, 61, 62 63, 64, 65 apresentam os gráficos das Variâncias Explicadas Cumulativas e das Taxa de Variância Explicada das Características das PCAs correspondentes às bandas larga, Delta, Teta, Alfa, Beta e Gama, respectivamente considerando a classificação entre grupos TDAH-C e TDAH-I.

Figura 60 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH-C/TDAH-I com características da banda larga.



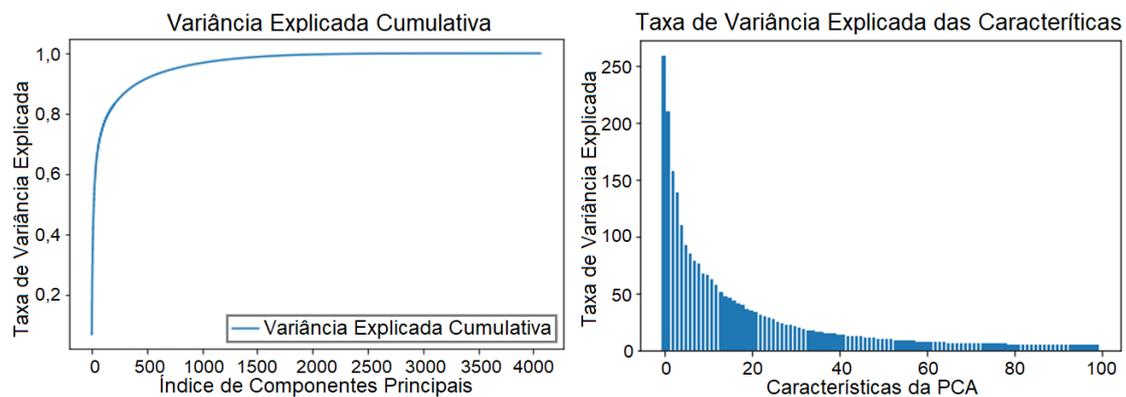
Fonte: O Autor

Figura 61 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH-C/TDAH-I com características da banda Delta.



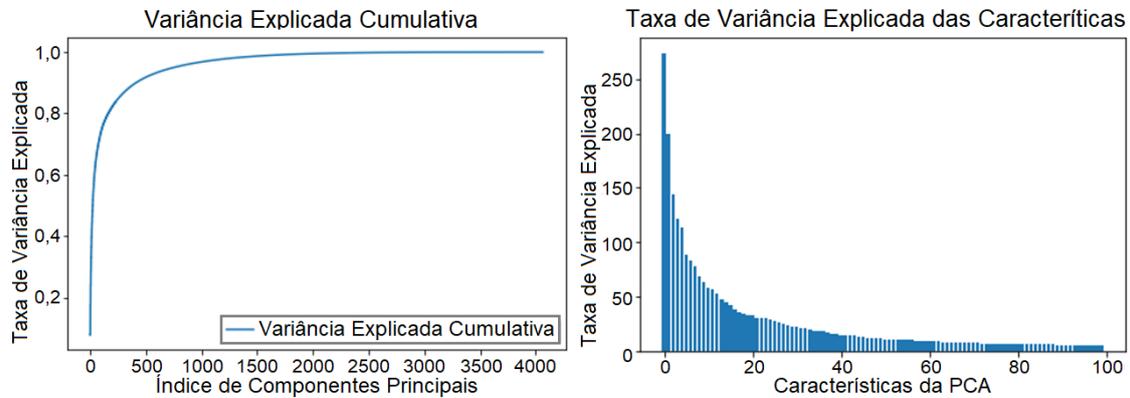
Fonte: O Autor

Figura 62 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH-C/TDAH-I com características da banda Teta.



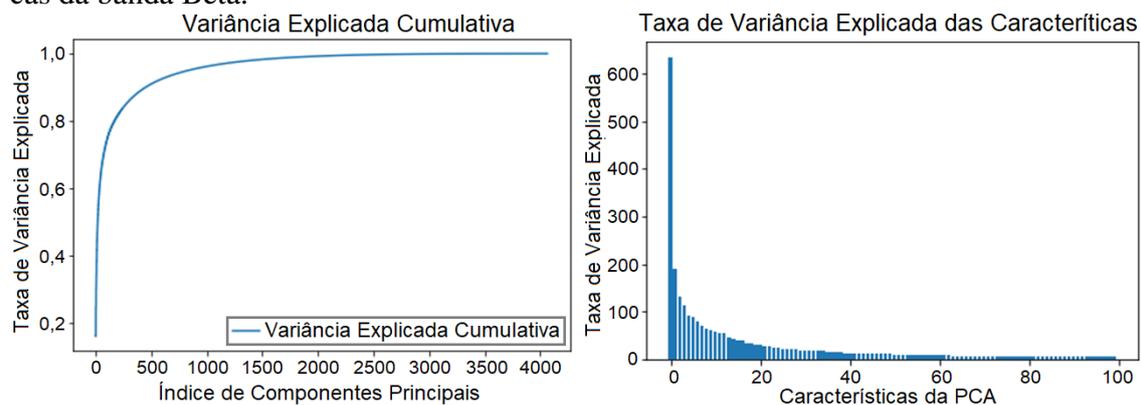
Fonte: O Autor

Figura 63 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH-C/TDAH-I com características da banda Alfa.



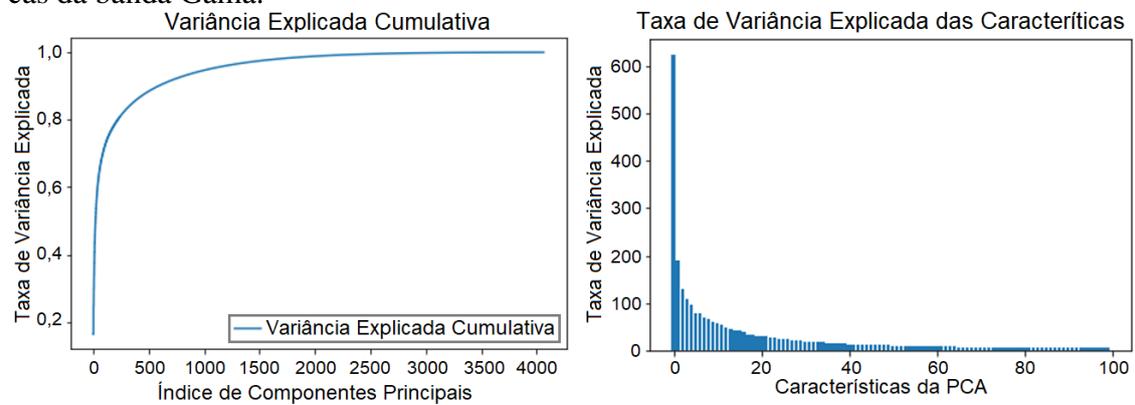
Fonte: O Autor

Figura 64 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH-C/TDAH-I com características da banda Beta.



Fonte: O Autor

Figura 65 – Variância Explicada Cumulativa e Taxa de Variância Explicada das Características apresentadas pela PCA para o diagnóstico de TDAH-C/TDAH-I com características da banda Gama.



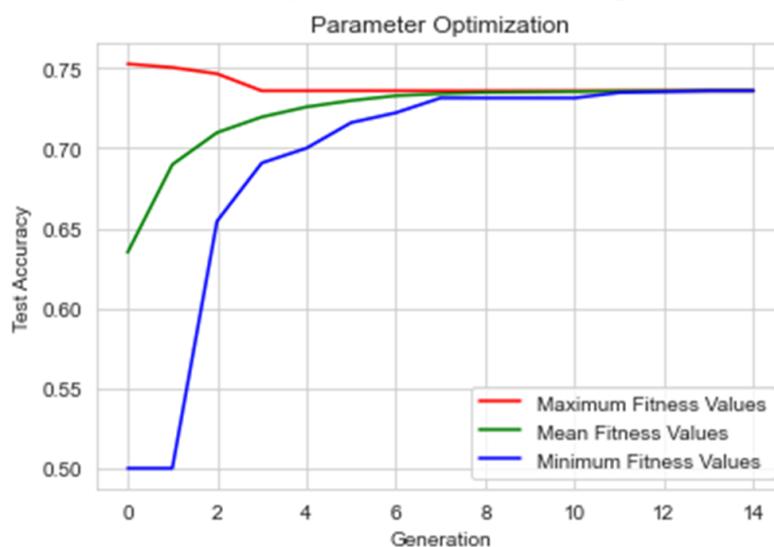
Fonte: O Autor

APÊNDICE B CURVAS DE OTIMIZAÇÃO DOS ALGORITMOS GENÉTICOS

B.1 Curvas de Otimização das RF

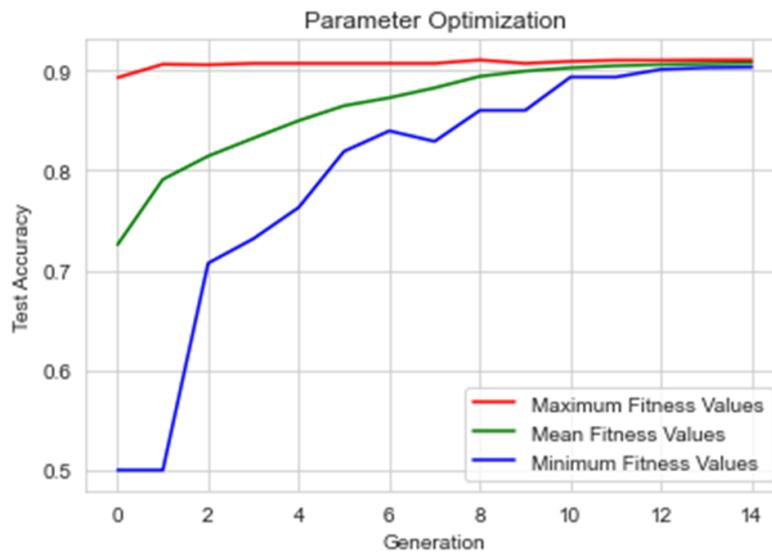
Nas figuras 66, 67, 68, 69, 70 e 71, mostra-se as curvas de otimização dos AGs das RFs para as bandas de frequência larga (0-100 Hz), Delta, Teta, Alfa, Beta, e Gama na classificação entre pacientes com TDAH e voluntários do grupo controle.

Figura 66 – Curva de Otimização da RF da banda larga para TDAH/Controle.



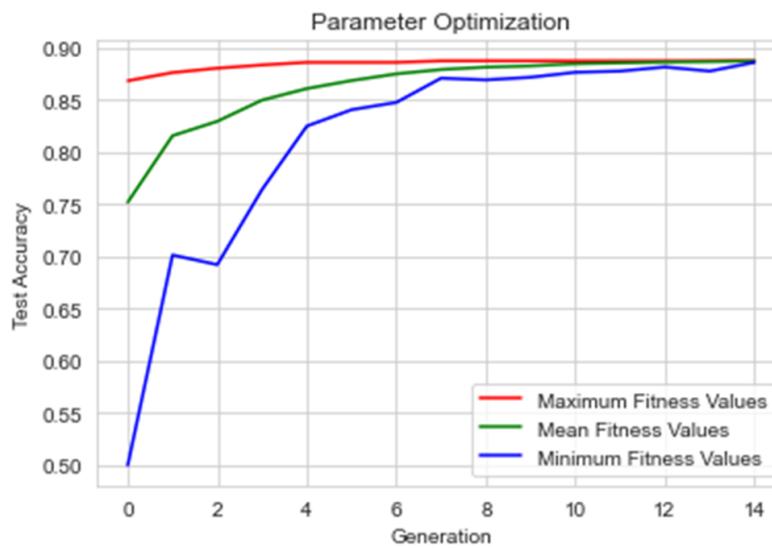
Fonte: O Autor

Figura 67 – Curva de Otimização da RF da banda Delta para TDAH/Controle.



Fonte: O Autor

Figura 68 – Curva de Otimização da RF da banda Teta para TDAH/Controle.



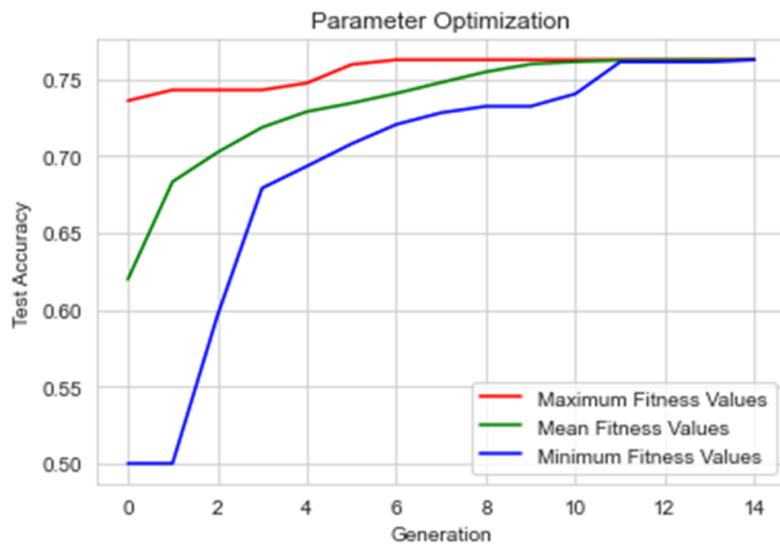
Fonte: O Autor

Figura 69 – Curva de Otimização da RF da banda Alfa para TDAH/Controle.



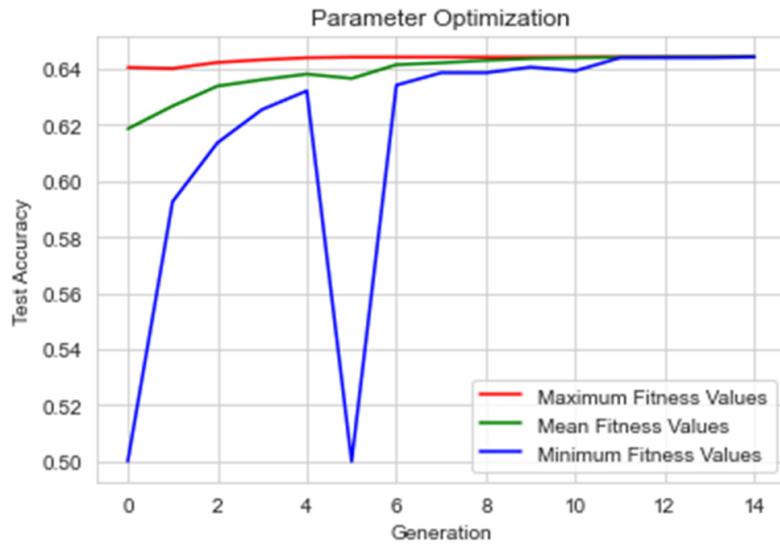
Fonte: O Autor

Figura 70 – Curva de Otimização da RF da banda Beta para TDAH/Controle.



Fonte: O Autor

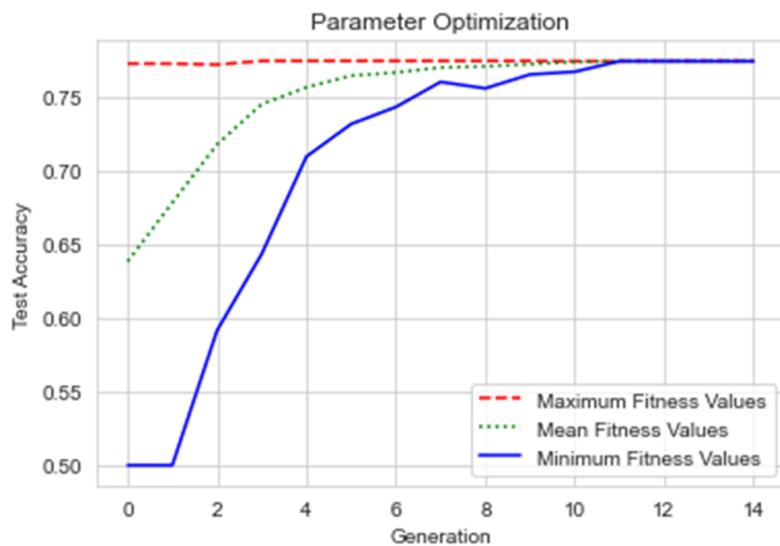
Figura 71 – Curva de Otimização da RF da banda Gama para TDAH/Controle.



Fonte: O Autor

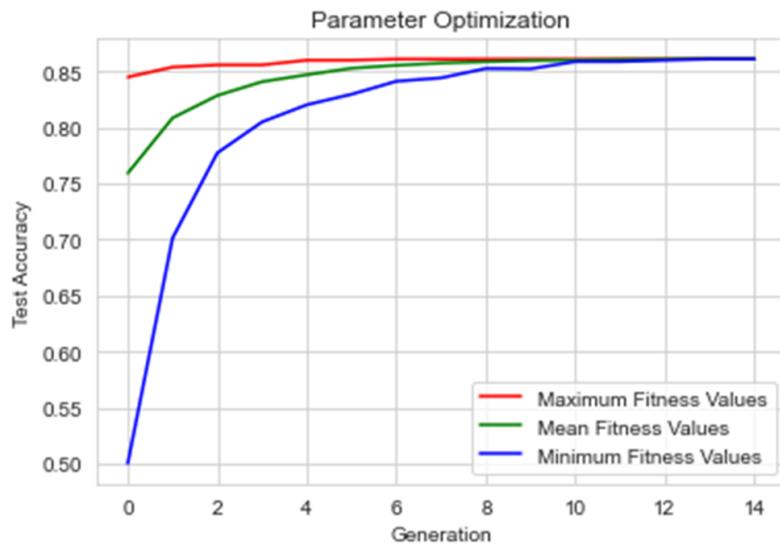
Nas figuras 72, 73, 74, 75, 76 e 77 apresenta-se as curvas de otimização dos AGs para as bandas de frequência larga (0-100 Hz), Delta, Teta, Alfa, Beta e Gama para a classificação TDAH-C/TDAH-I.

Figura 72 – Curva de Otimização da RF da banda larga para TDAH-C/TDAH-I.



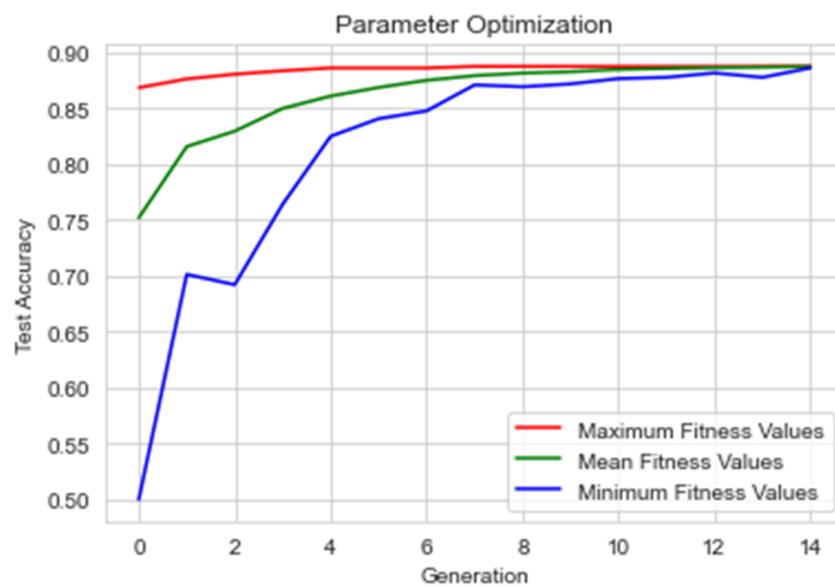
Fonte: O Autor

Figura 73 – Curva de Otimização da RF da banda Delta para TDAH-C/TDAH-I.



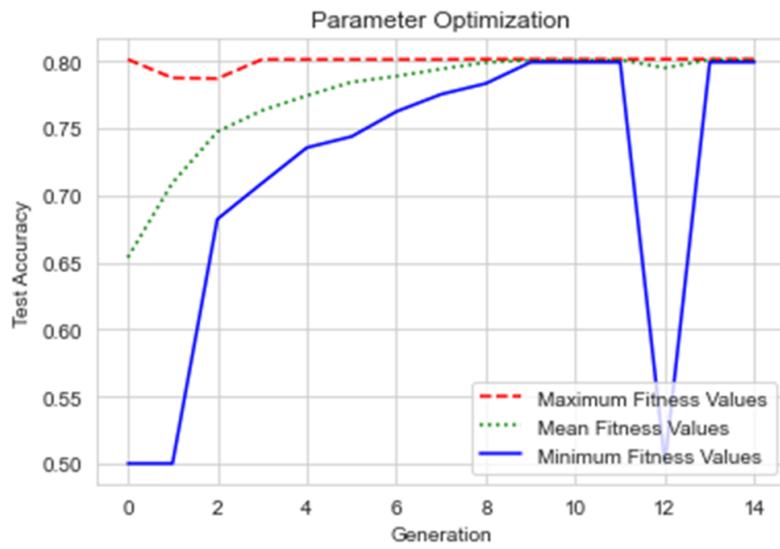
Fonte: O Autor

Figura 74 – Curva de Otimização da RF da banda Teta para TDAH-C/TDAH-I.



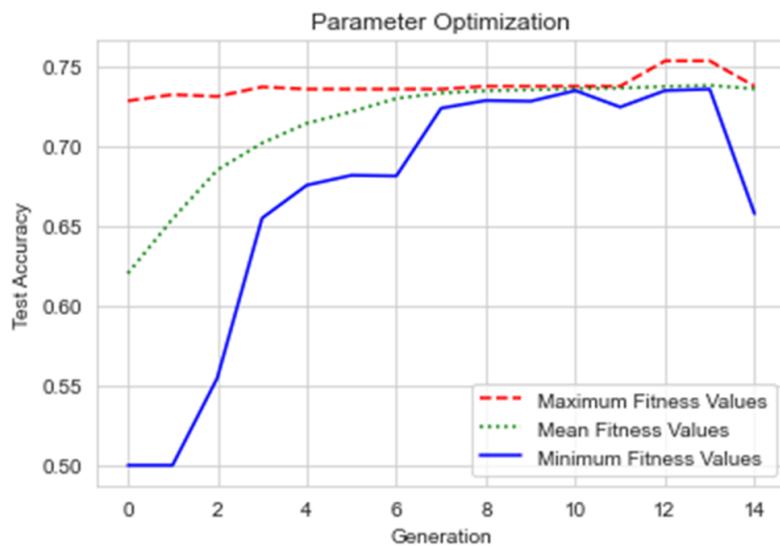
Fonte: O Autor

Figura 75 – Curva de Otimização da RF da banda Alfa para TDAH-C/TDAH-I.



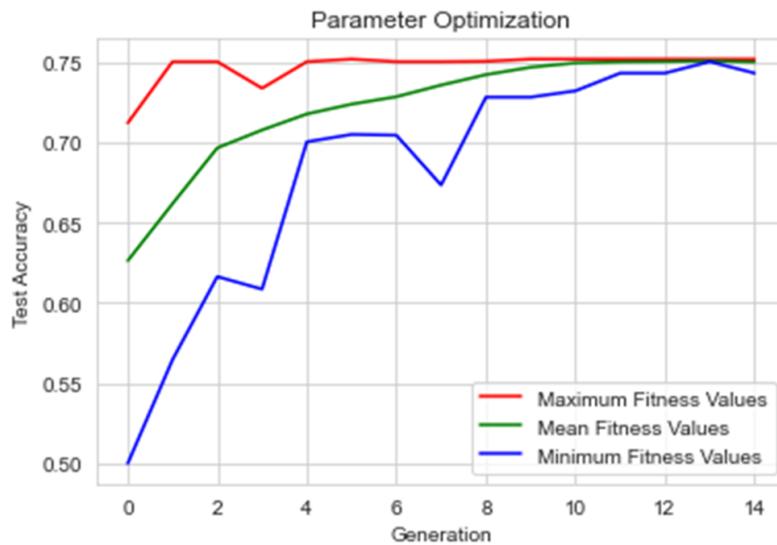
Fonte: O Autor

Figura 76 – Curva de Otimização da RF da banda Beta para TDAH-C/TDAH-I.



Fonte: O Autor

Figura 77 – Curva de Otimização da RF da banda Gama para TDAH-C/TDAH-I.

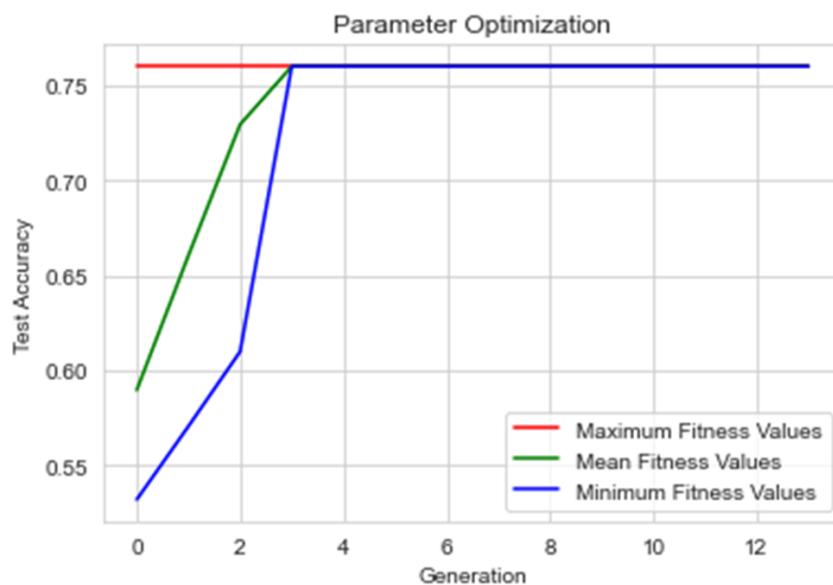


Fonte: O Autor

B.2 Curvas de Otimização das SVM

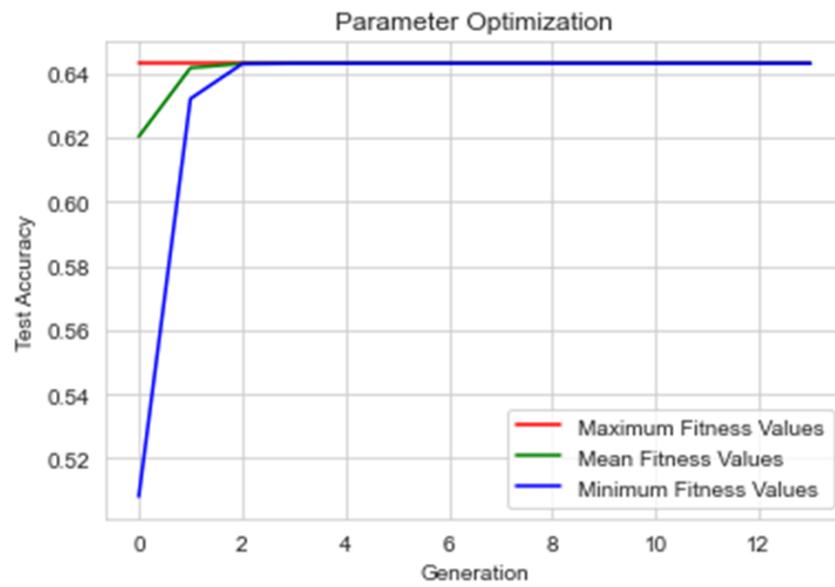
Nas figuras 78, 79, 81, 82, 80 e 83 mostra-se as curvas de otimização dos AGs das SVMs para as bandas de frequência larga (0-100 Hz), Delta, Teta, Alfa, Beta e Gama na classificação TDAH/Controle.

Figura 78 – Curva de Otimização da SVM da banda larga para TDAH/Controle.



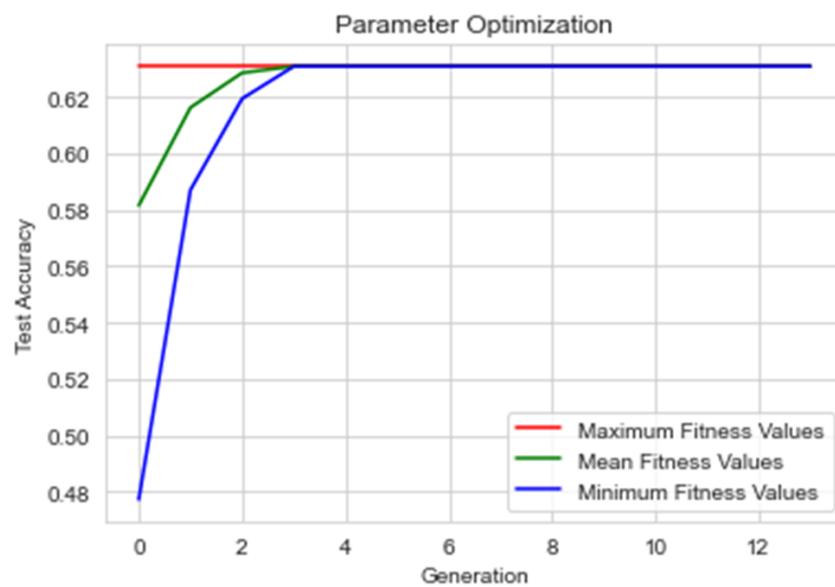
Fonte: O Autor

Figura 79 – Curva de Otimização da SVM da banda Delta para TDAH/Controle.



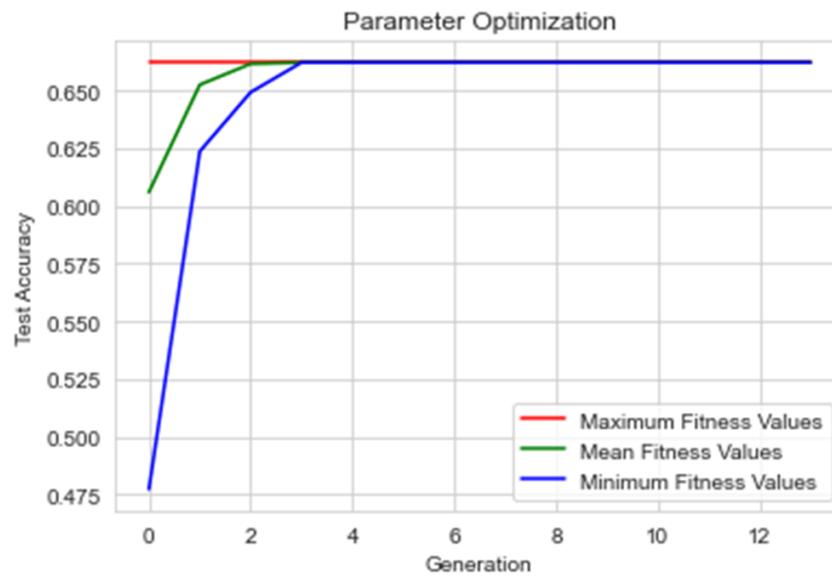
Fonte: O Autor

Figura 80 – Curva de Otimização da SVM da banda Teta para TDAH/Controle.



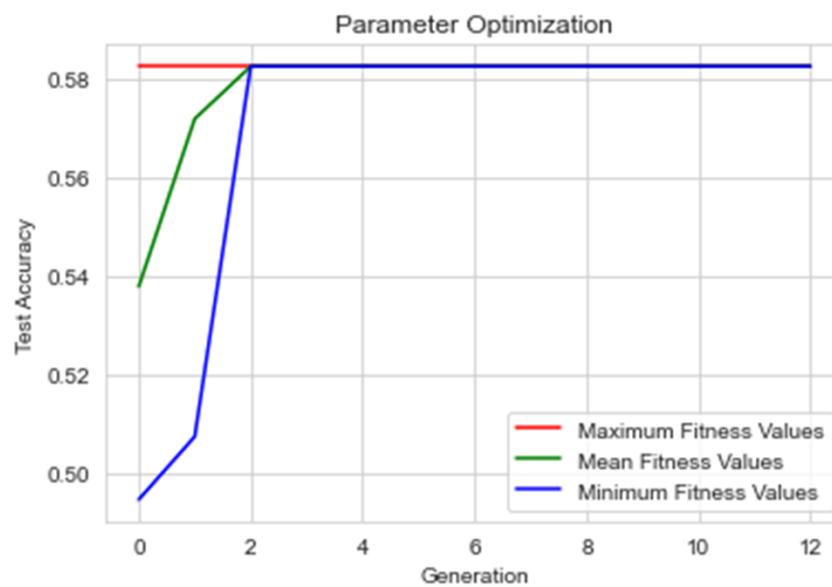
Fonte: O Autor

Figura 81 – Curva de Otimização da SVM da banda Alfa para TDAH/Controle.



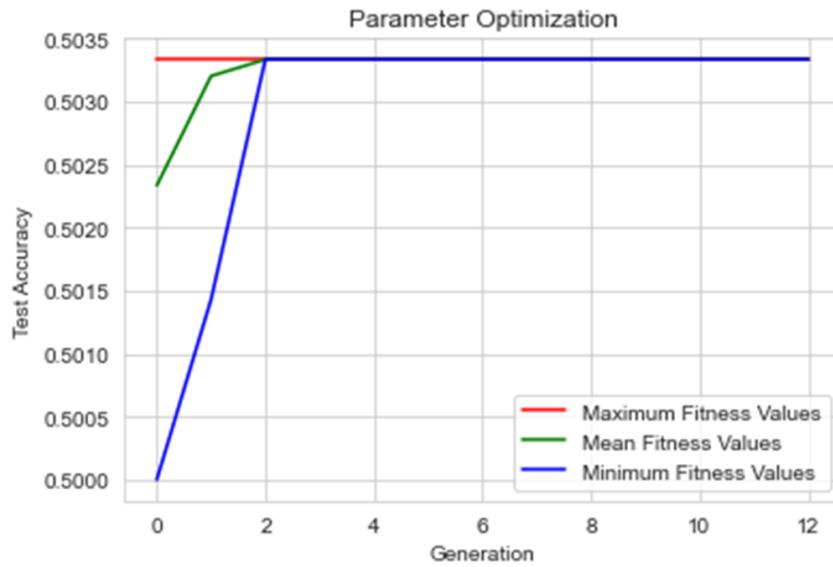
Fonte: O Autor

Figura 82 – Curva de Otimização da SVM da banda Beta para TDAH/Controle.



Fonte: O Autor

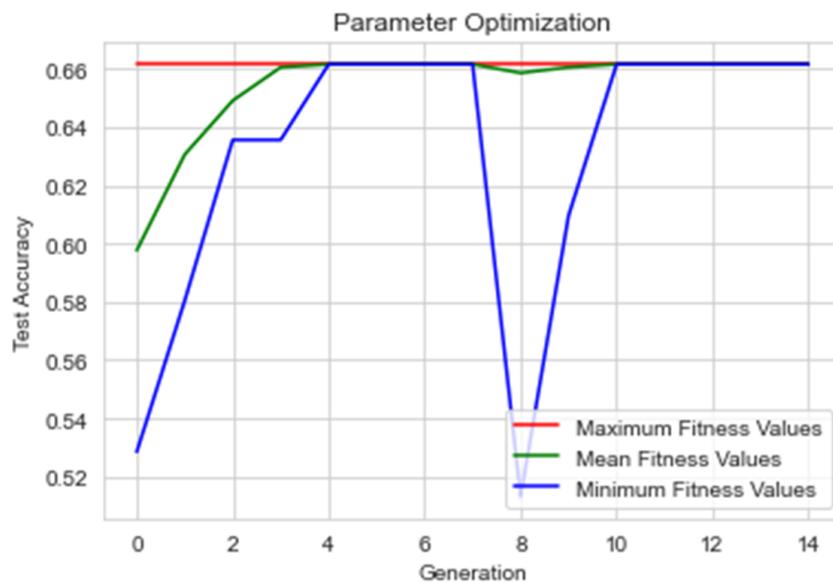
Figura 83 – Curva de Otimização da SVM da banda Gama para TDAH/Controle.



Fonte: O Autor

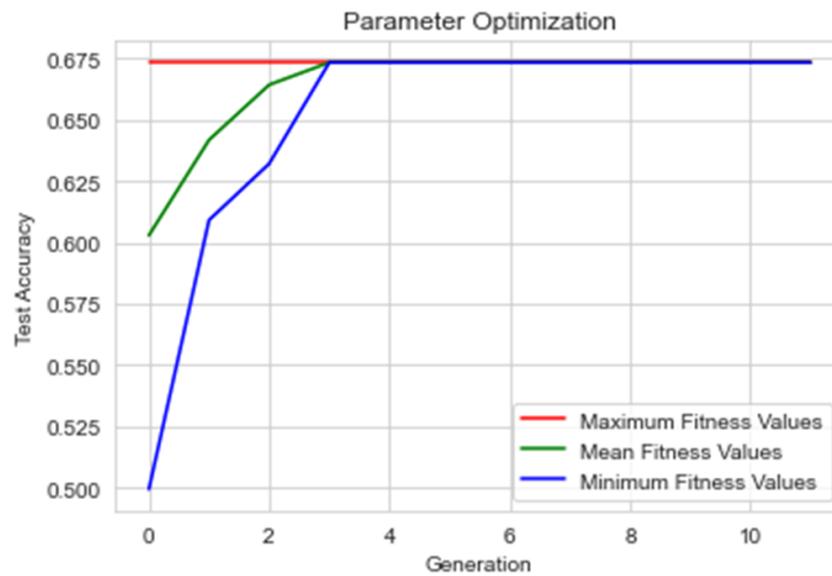
As figuras 84, 85, 86, 87, 88 e 89 apresentam as curvas de otimização dos AGs das SVMs para as bandas de frequência larga (0-100 Hz), Delta, Teta, Alfa, Beta e Gama para a classificação TDAH-C/TDAH-I.

Figura 84 – Curva de Otimização da SVM da banda larga para TDAH-C/TDAH-I.



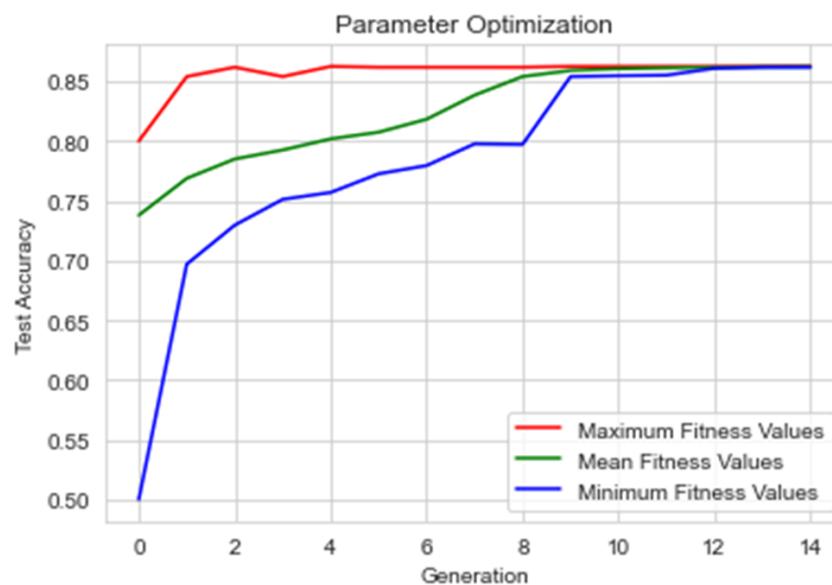
Fonte: O Autor

Figura 85 – Curva de Otimização da SVM da banda Delta para TDAH-C/TDAH-I.



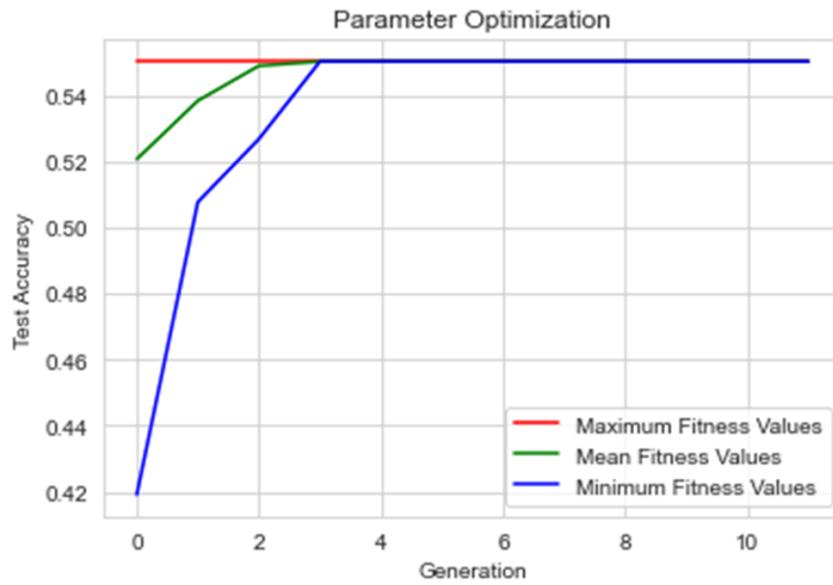
Fonte: O Autor

Figura 86 – Curva de Otimização da SVM da banda Teta para TDAH-C/TDAH-I.



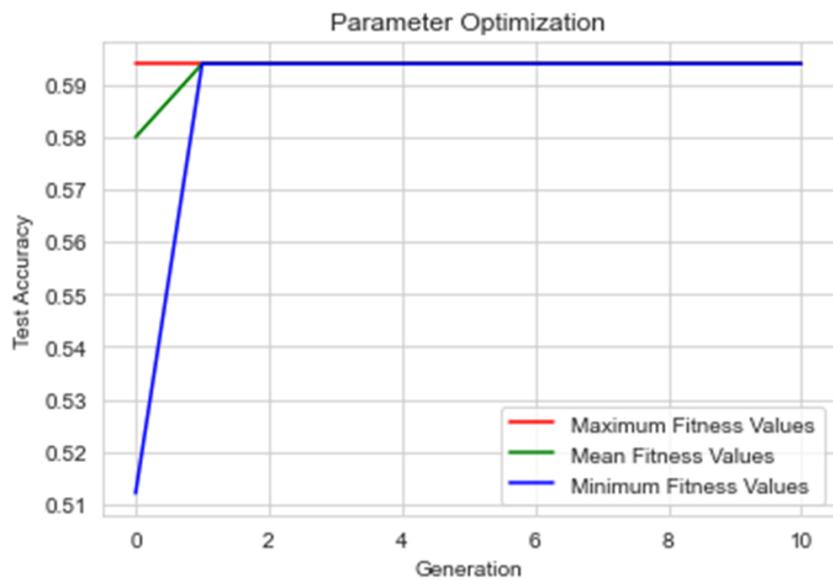
Fonte: O Autor

Figura 87 – Curva de Otimização da SVM da banda Alfa para TDAH-C/TDAH-I.



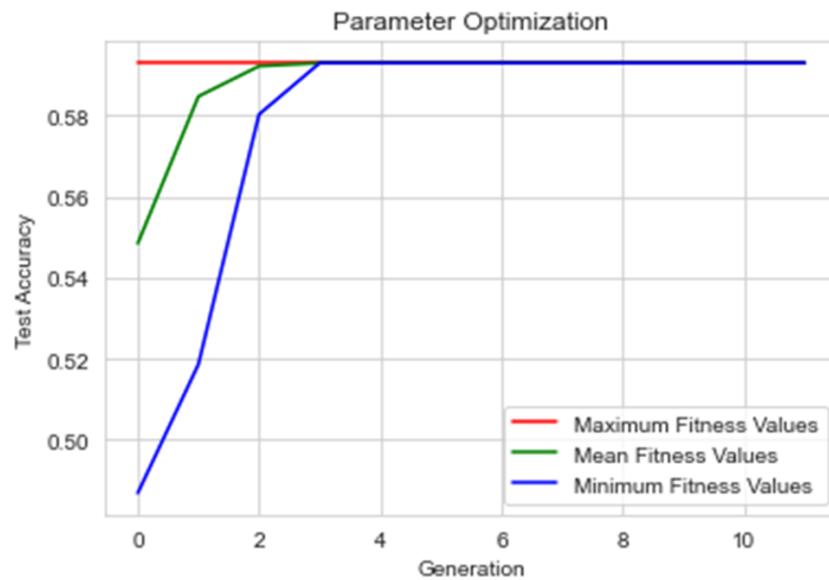
Fonte: O Autor

Figura 88 – Curva de Otimização da SVM da banda Beta para TDAH-C/TDAH-I.



Fonte: O Autor

Figura 89 – Curva de Otimização da SVM da banda Gama para TDAH-C/TDAH-I.

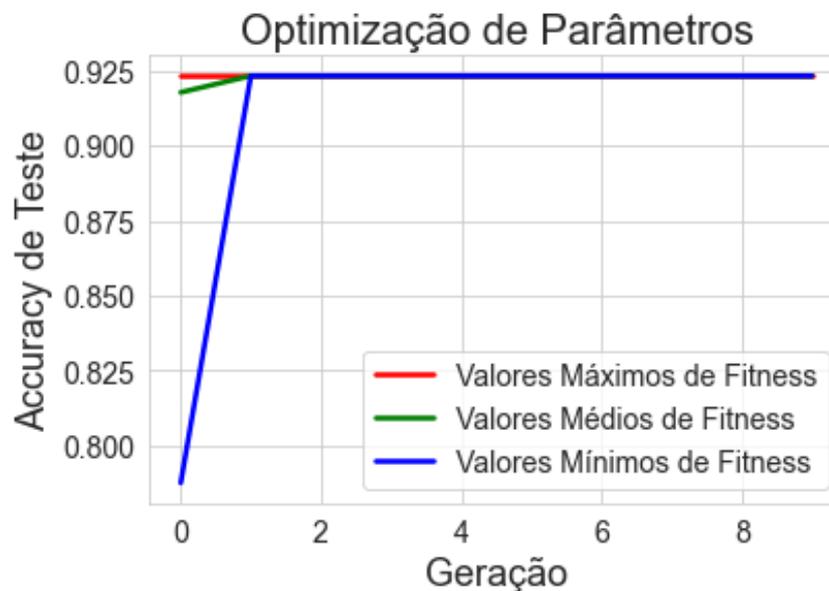


Fonte: O Autor

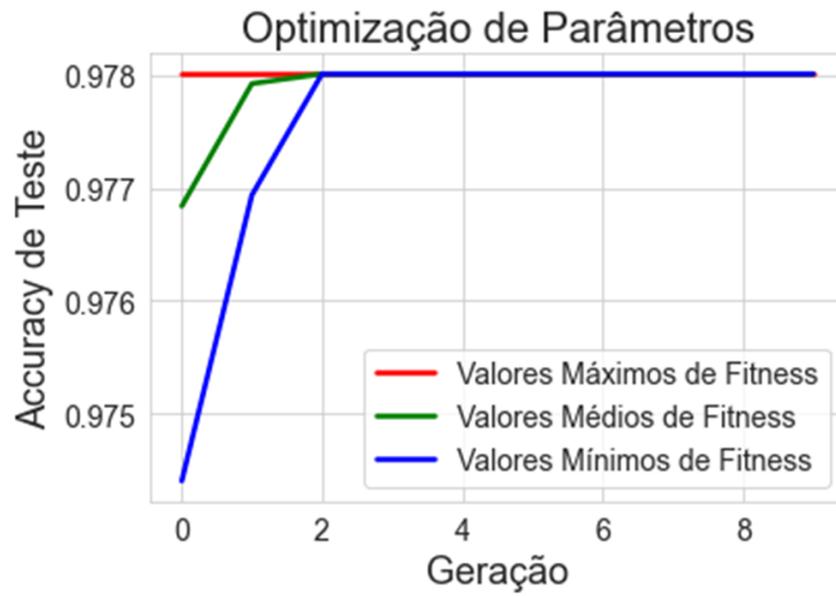
B.3 Curvas de Otimização dos XGBoost

As figuras 90 e 91 apresentam as curvas de otimização dos AGs para o método de fusão de bandas *XGBoost* para as classificações TDAH/Controle e TDAH-C/TDAH-I.

Figura 90 – Curva de Otimização da fusão *XGBoost* para TDAH/Controle.



Fonte: O Autor

Figura 91 – Curva de Otimização da fusão *XGBoost* para TDAH/Controle.

Fonte: O Autor

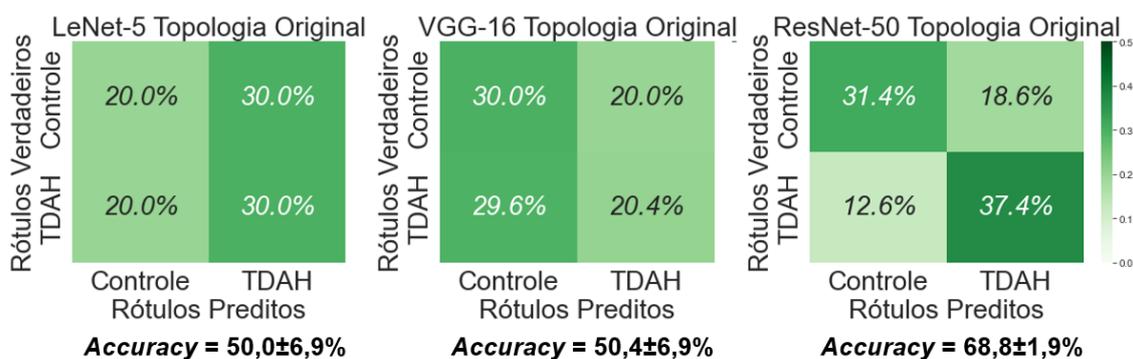
APÊNDICE C RESULTADOS DE CLASSIFICAÇÃO DE TDAH E DE SUBTIPOS COM OS MODELOS CLÁSSICOS DE CNN A PARTIR DE FMRI

C.1 Resultados da Classificação de TDAH baseada em fMRI: Modelos clássicos de CNN em suas arquiteturas originais

Para o caso da distinção entre pacientes com TDAH e indivíduos do grupo controle a partir da matriz de conectividade completa, utilizaram-se as topologia originais não otimizadas por GA dos modelos clássicos. Nestes testes, não utilizou-se aprendizado por transferência e não congelou-se camadas. Portanto, todas as camadas destas redes foram treinadas.

Nestas condições, o desempenho da LeNet-5 e da VGG-16 foi insatisfatório. Para a LeNet-5 obteve-se *accuracies* de treinamento e de teste de 49,6% e de 50,0%, respectivamente. Já a VGG-16 obteve *accuracies* de treinamento e de teste de 50,1% e de 50,4%, respectivamente. A Figura 92 apresenta as matrizes de confusão obtidas através do uso da redes LeNet-5, VGG-16 e ResNet-50.

Figura 92 – Matrizes de confusão com as *accuracies* de classificação obtidas pela LeNet-5, VGG-16 e ResNet-50 com suas topologias originais sem pesos pré-ajustados a partir da matriz de correlações funcionais completa de fMRI na classificação entre grupos TDAH e Controle.



Fonte: O Autor.

A ResNet-50 apresentou métricas melhores do que estas topologias com o treinamento completo de sua arquitetura original. Contudo, as *accuracies* obtidas com a ResNet-50 foram de 100% para o treinamento e de 68,8% para o teste, o que caracteriza a ocorrência de *overfitting*.

C.2 Resultados da Classificação de TDAH baseada em fMRI: Modelos Clássicos de CNN Otimizados

Nesta seção, descreve-se os resultados obtidos através da utilização dos modelos clássicos, LeNet-5, VGG-16 e ResNet-50 otimizados por AG para a classificação entre grupos TDAH e Controle usando aprendizado por transferência. Os resultados foram validados usando a validação cruzada do tipo *k-fold* com $k=10$. Assim como para os demais experimentos nesta Tese, os conjuntos de treinamento e de teste constituem 70% e 30% das amostras totais de cada classe após o balanceamento por sobreamostragem. Utilizou-se *k-fold* estratificado para garantir o balanço entre as classes.

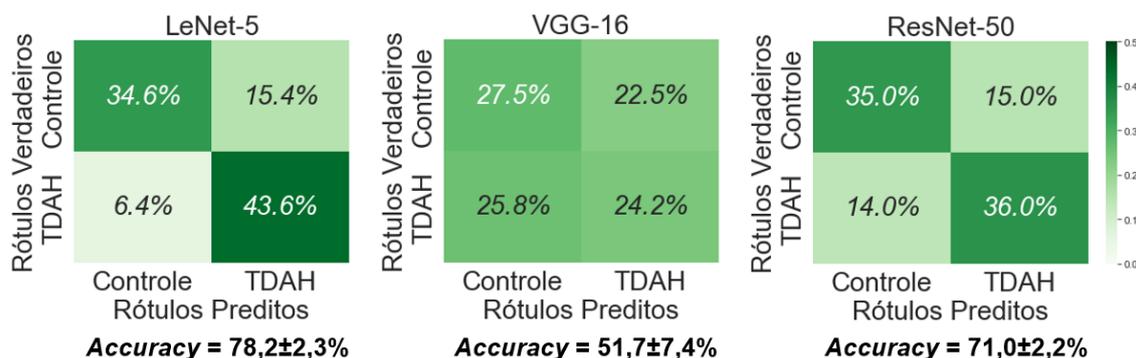
Os resultados da aplicação das arquiteturas originais destes modelos clássicos para a classificação entre grupos TDAH e controle a partir da matriz completa de conectividade encontram-se no APÊNDICE D desta Tese.

Para a matriz completa, que representa o conjunto de todas as redes de conectividade e as correlações entre estas redes, a arquitetura LeNet-5 obteve uma *accuracy* de 78,2±2,3% no teste. No entanto, com esta arquitetura obteve-se 69,1% de *accuracy* para o grupo controle e 87,3% para o grupo TDAH, evidenciando um desbalanço entre a quantidade de acertos em cada classe. No treinamento, o uso da LeNet-5 obteve uma *accuracy* de 80,0±2,5%. As precisões de treinamento e de teste foram de 75,0% e de 74,1%, respectivamente. No treinamento, a LeNet-5 atingiu um *recall* de 90%. O *recall* foi reduzido

para 87,3% no teste.

A Figura 93 apresenta as matrizes de confusão obtidas através da aplicação das arquiteturas LeNet-5, VGG-16 e ResNet-50 para a classificação entre grupos TDAH e Controle.

Figura 93 – Matrizes de confusão com as *accuracies* de teste obtidas pela LeNet-5, VGG-16 e ResNet-50 a partir da matriz de correlações funcionais completa de fMRI na classificação entre grupos TDAH e Controle.

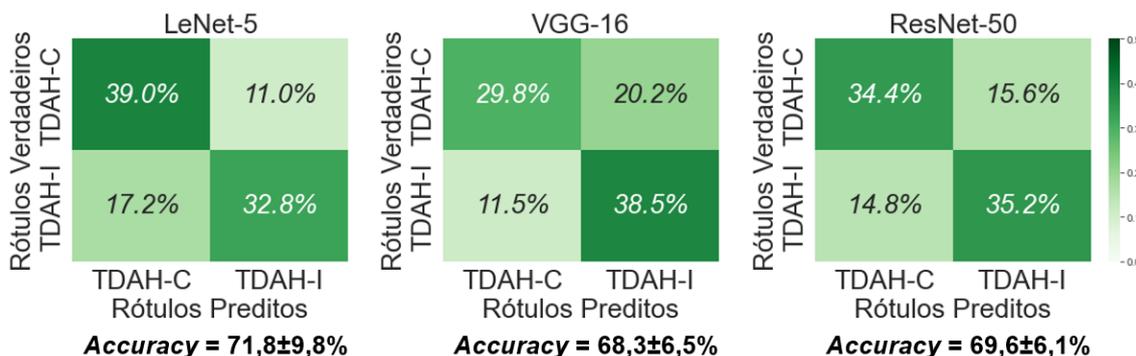


Fonte: O Autor.

O uso da arquitetura VGG-16 para a classificação TDAH/Controle rendeu resultados pouco satisfatórios. Obteve-se *accuracies* de treinamento e de teste de $56,0\pm 7,5\%$ e de $51,7\pm 7,4\%$, respectivamente. O classificador acertou somente 48,3% dos sujeitos com TDAH e 55,0% dos voluntários do grupo controle. A arquitetura ResNet-50 obteve uma *accuracy* de $71,1\pm 2,2\%$ no teste, sendo 70,1% de *accuracy* no grupo controle e 72,1% no grupo TDAH.

A Figura 94 apresenta as matrizes de confusão obtidas através da aplicação das arquiteturas LeNet-5, VGG-16 e ResNet-50 para a classificação entre subtipos de TDAH.

Figura 94 – Matrizes de confusão com as *accuracies* de teste obtidas pela LeNet-5, VGG-16 e ResNet-50 a partir da matriz de correlações funcionais completa de fMRI na classificação entre grupos TDAH-C e TDAH-I.



Fonte: O Autor.

Na classificação de subtipos, as 3 arquiteturas, LeNet-5, VGG-16 e ResNet-50, obtiveram resultados similares em termos de *accuracy* no grupo de teste: $71,8 \pm 9,8\%$, $68,3 \pm 6,5\%$ e $69,7 \pm 6,1\%$. A LeNet-5 obteve a *accuracy* mais alta, superando inclusive a *accuracy* da CNN projetada em pouco mais de 2%. Contudo, a LeNet-5 apresentou um desvio padrão de 9,8% na classificação de subtipos. O desvio padrão apresentado pela CNN desejada é de menos da metade deste valor: 4,8%. A VGG-16 e a ResNet-50 também apresentaram desvios-padrão menores (6,5% e 6,1%, respectivamente) do que os da LeNET-5.

O modelo baseado na LeNet-5 obteve 77,9% de *accuracy* na classificação de TDAH-C e 65,7% para TDAH-I. Em contrapartida, a VGG-16 apresentou 77,0% de *accuracy* para TDAH-I e somente 59,6% para TDAH-C. Já, com a arquitetura ResNet-50 obteve-se uma *accuracy* de 68,9% para TDAH-C e de 70,5% para TDAH-I. Portanto, usando a ResNet50 pode-se obter maior balanço de acertos entre as classes, o que é desejável.

A CNN baseada na arquitetura LeNet-5 obteve maiores *accuracies* classificando grupos TDAH e Controle quando recebeu informações das redes de conectividade separadamente. A maior *accuracy* no conjunto de teste, $94,9 \pm 2,5\%$, foi atingida a partir da rede *default*. Na rede *default*, a LeNet-5 obteve 97,7% e 92,1% para a classificação dos grupos Controle e TDAH, respectivamente. O *recall* e a precisão foram, respectivamente, iguais à 92,1% e 97,7%. No conjunto de treinamento, a *accuracy* foi de $100 \pm 0\%$ e, consequentemente, a precisão e o *recall* também foram de 100%.

Para a rede visual, o uso da LeNet-5 conduziu à obtenção de uma *accuracy* de $92,1 \pm 3,5\%$ no teste, com 96,6% de *accuracy* para o grupo controle e 87,6% para o grupo TDAH. Portanto, o *recall* foi de 87,0% e a precisão foi de 95,5% no conjunto de teste. No treinamento, a *accuracy*, o *recall* e a precisão foram $99,7 \pm 0,5\%$, 99,5% e 99,9%. Em comparação, a CNN projetada obteve métricas um pouco mais elevadas de *accuracy* e um desvio padrão consideravelmente menor ($94,5 \pm 0,1\%$).

Analisando a rede de conectividade somatomotor, a LeNet-5 obteve *accuracies* de $99,7 \pm 0,8\%$ e de $92,3 \pm 1,7\%$ no treinamento e no teste. No teste, a LeNet-5 acertou 94,8% dos indivíduos na classe Controle e 89,7% na classe TDAH. Por isso, o *recall* e a precisão no conjunto de teste foram de 89,7% e de 99,4%, respectivamente.

Para a rede de atenção ventral, o uso da arquitetura LeNet-5 permitiu a obtenção de *accuracies* de $96,3 \pm 1,0\%$ e de $88,4 \pm 2,5\%$ no treinamento e no teste. Na verificação, o classificador obteve um *recall* e uma precisão de 85,8% e de 90,9%. Observou-se um *overfitting* de 7,9%.

Trabalhando com a rede de atenção dorsal, a LeNet-5 alcançou *accuracies* no treinamento e no teste de $98,7 \pm 1,9\%$ e de $91,3 \pm 2,4\%$ (*recall* = 87,1% e precisão = 95,2%).

Assim como ocorreu para as demais redes de conectividade, na rede frontoparietal, a LeNet-5 também obteve uma maior *accuracy* para a classe Controle, 95,0% no teste (no caso da rede frontoparietal). Na classe TDAH, a LeNet-5 alcançou apenas 86,3% de

accuracy no teste. Por isto, a *accuracy* total obtida na rede frontoparietal foi de $90,7 \pm 2,5\%$ (*recall* = 86,3% e precisão = 94,8%). Uma vez que a *accuracy* no treinamento foi de $99,9 \pm 0,2\%$, há um *overfitting* de um pouco menos de 10%.

O classificador baseado na LeNet-5 que trabalhou com a rede de conectividade límbica, obteve *accuracies* para as classes Controle e TDAH de 90,8% e de 88,1%, cuja média resulta em uma *accuracy* de $89,5 \pm 2,5\%$, a qual foi obtida trabalhando com o conjunto de teste. O *recall* e a precisão no teste foram de 88,1% e 90,8%.

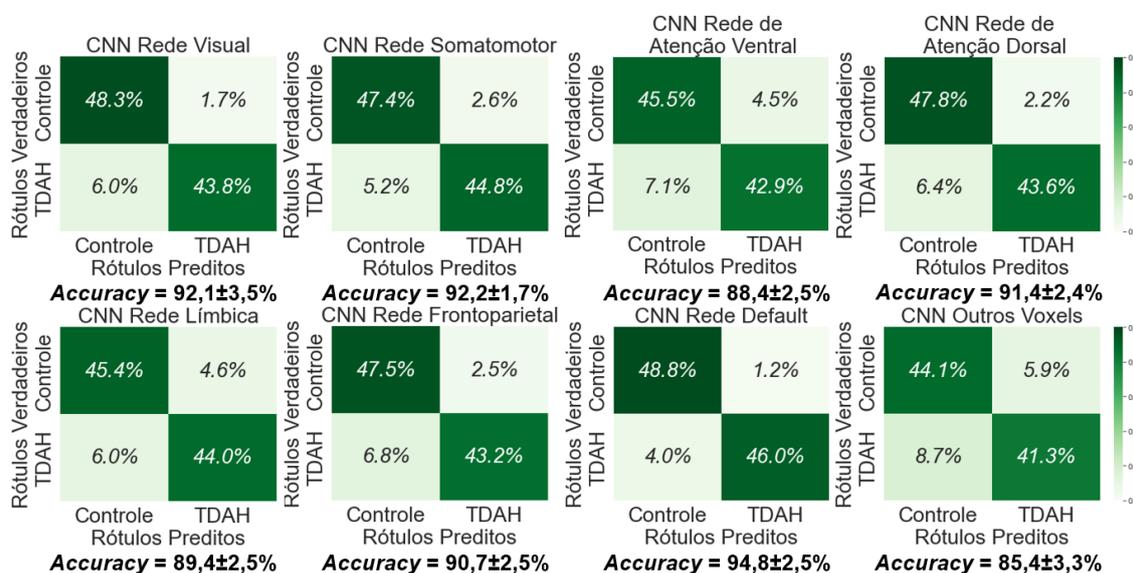
A LeNet-5 alcançou a menor *accuracy* na classificação entre grupos TDAH e Controle analisando a região dos 'outros voxels', onde obteve 88,2% na classe Controle e 82,6% na classe TDAH no conjunto de teste. Desta forma, totaliza-se uma *accuracy* média entre as classes de $85,4 \pm 3,3\%$. Já no treinamento, este classificador obteve uma *accuracy* de $91,7 \pm 1,8\%$.

A média das *accuracies* através da arquitetura LeNet-5 (90,6%) é similar à obtida com a CNN projetada (89,3%). No entanto, o desvio padrão médio (entre as redes de conectividade) resultante do uso da CNN projetada é de somente 1,7%, menor que o valor obtido pela LeNet-5 (2,6%). E, o *overfitting* médio na LeNet-5 foi de 7,1% contra os 4,5% observados na outra arquitetura. Isto indica que a CNN projetada possui maior capacidade de generalização na classificação entre indivíduos com TDAH ou do grupo controle.

As duas arquiteturas concordaram em apontar maiores relevâncias para as redes de conectividade visual, *default*, somatomotor, seguidas em importância pela rede frontoparietal. No entanto, houve discordância em relação as redes de atenção dorsal e límbica. A LeNet-5 apontou maior relevância para a rede dorsal, enquanto a CNN projetada obteve melhores resultados trabalhando com a rede límbica. Não obstante, a CNN projetada havia apontado baixa importância para a rede de atenção dorsal.

Para melhorar a visualização das métricas, a Figura 95 mostra as matrizes de confusão resultantes da aplicação da topologia inspirada na LeNet-5 na classificação entre TDAH e Controle a partir das redes de conectividade extraídas das imagens de fMRI.

Figura 95 – Matrizes de confusão com as *accuracies* de teste obtidas pela LeNet-5 a partir das redes de conectividade de fMRI na classificação de TDAH *versus* Controle.



Fonte: O Autor.

Na classificação de subtipos, a LeNet-5 obteve seus melhores resultados no conjunto de teste analisando as redes de atenção dorsal e visual, obtendo *accuracies* de $69,9\pm 4,9\%$ e de $69,5\pm 5,0\%$, respectivamente. Ou seja, com as redes de atenção dorsal, a LeNet-5 alcançou uma *accuracy* um pouco maior com uma incerteza, ligeiramente, menor. Para ambas as redes de conectividade, a LeNet-5 encontrou um forte balanço entre *accuracies* nas classes. No entanto, no conjunto de treinamento, as *accuracies* foram de $92,6\pm 6,5\%$ e de $99,9\pm 0,1\%$ para as redes de atenção dorsal e visual, respectivamente.

Na classificação de subtipos com a rede somatomotor, a LeNet-5 obteve *accuracies* de treinamento e de teste de $98,4\pm 2,3\%$ e de $65,4\pm 4,9\%$, respectivamente. Com a rede de atenção ventral, a LeNet-5 alcançou *accuracies* de treinamento e de teste de $85,8\pm 4,1\%$ e de $64,9\pm 4,0\%$, respectivamente.

Trabalhando com a rede límbica, a LeNet-5 obteve *accuracies* de treinamento e de teste, respectivamente, de $80,6\pm 6,6\%$ e de $64,3\pm 5,4\%$. As *accuracies* de treinamento e de teste, respectivamente, obtidas pela LeNet-5 foram de $99,1\pm 1,8\%$ e de $62,4\pm 0,7\%$ com a rede frontoparietal. Já para a rede *default*, as *accuracies* de $99,9\pm 0,4\%$ e de $67,5\pm 6,5\%$ foram alcançadas pela LeNet-5 para os conjuntos de treinamento e de teste, respectivamente.

O maior desequilíbrio entre as *accuracies* de classes ocorreu quando a LeNet-5 analisou a região dos 'outros voxels'. Nesta região, esta CNN acertou $67,5\%$ dos indivíduos com TDAH-I, mas somente $57,0\%$ daqueles com TDAH-C, o que corresponde à *accuracy* de teste de $62,3\pm 5,4\%$. Esta foi a menor *accuracy* no conjunto de teste encontrada para subtipos com classificadores baseados na arquitetura da LeNet-5. No conjunto de

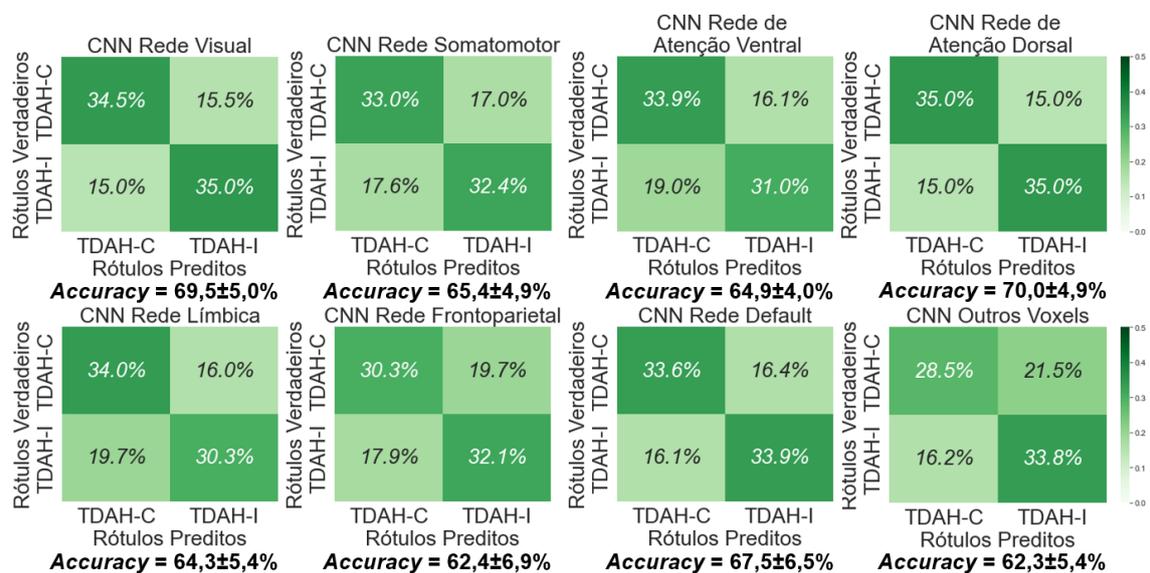
treinamento, a *accuracy* obtida foi de $76,8 \pm 5,1\%$.

A comparação entre as *accuracies* obtidas para a rede de atenção dorsal e para a região dos 'outros voxels' mostra que as diferenças entre estas métricas foram de apenas 7,7%. Estas ocorreram, principalmente, pela dificuldade de classificar-se sujeitos com TDAH-C através dos 'outros voxels'.

No entanto, na classificação de subtipos, alguns modelos tiveram maior facilidade para diagnosticar TDAH-C do que TDAH-I. Isto foi constatado com os classificadores que trabalharam com características das redes límbica, de atenção ventral e somatomotor. Já usando as características das redes visual, frontoparietal, *default* e 'outros voxels', as CNNs baseadas na LeNet-5 alcançaram maiores *accuracies* para TDAH-I, embora as diferenças entre as classes tenham sido de menos de 1% para as redes visual e *default*. Não obstante, o classificador que atuou junto à rede de atenção dorsal apresentou as maiores *accuracies* para ambas as classes (69,9%).

A Figura 96 apresenta as matrizes de confusão e as *accuracies* correspondentes ao uso da arquitetura LeNet-5 para classificar pacientes de TDAH entre os subtipos TDAH-C e TDAH-I, no conjunto teste, a partir das redes de conectividade extraídas das imagens de fMRI.

Figura 96 – Matrizes de confusão com as *accuracies* de teste obtidas pela LeNet-5 a partir das redes de conectividade de fMRI na classificação de TDAH *versus* Controle.



Fonte: O Autor.

A arquitetura LeNet-5 possibilitou a obtenção de uma média de *accuracies* entre todas as redes de conectividade de 65,8% para a classificação entre subtipos, o que supera a métrica correspondente obtida pela CNN projetada em 3%. As médias entre as redes de conectividade dos desvios-padrão são similares (5,3% para a CNN projetada e 5,4% para a LeNet-5).

No entanto, os menores *overfittings* obtidos com a LeNet-5 na classificação de subtipos ocorreram através do uso dos dados da rede límbica e da região dos 'outros voxels': 16,3% e 14,5%, respectivamente. E, o *overfitting* médio observado para a classificação de subtipos de TDAH com a LeNet-5 foi de 25,9%. Estes altos valores de *overfitting* demonstram que estes modelos possuem pouca capacidade de generalização, ao contrário do que foi constatado para a CNN projetada e otimizada por AG (Capítulo 4, seção 4.2.1).

O *overfitting* médio obtido pela CNN projetada na classificação de subtipos foi, somente, de 4,6%. Por isso, decidiu-se utilizar os modelos com a arquitetura da CNN projetada como classificadores primários dos *ensembles*, cujos resultados serão apresentados no Capítulo 4, de Resultados, na seção 4.2.2.

APÊNDICE D CONCEITOS BÁSICOS SOBRE TDAH

Em 1775, Melchior Weikard descreveu em seu livro (WEIKARD, 1787) pela primeira vez o transtorno que hoje se conhece como TDAH focando no aspecto da desatenção (BARKLEY & PETERS, 2012). O estudo do médico escocês Sir Alexander Crichton (CRICHTON, 1798) observou que algumas pessoas nasciam com esta condição e outras adquiririam esta condição ao longo da vida. Segundo Crichton, o transtorno poderia ser observado desde a infância e reduzia sua intensidade ao longo do tempo (CRICHTON, 1798); (LANGE *et al.*, 2010). Crichton observou que o transtorno dificultava muito a vida acadêmica, embora raramente a impedisse (CRICHTON, 1798).

Em 1902, o pediatra Sir George Frederic Still atribuiu os comportamentos típicos do TDAH a um defeito de controle moral que poderia ser nato ou em alguns casos estaria relacionado a doença tal como meningite, febre tifoide, epilepsia, ferimento na cabeça ou tumor cerebral (STILL, 1902). Deste modo, Still foi quem primeiro diferenciou o “defeito de controle moral” nato (como ele chamava o TDAH) do defeito relacionado a outras doenças e do defeito relacionado a capacidade intelectual reduzida (STILL, 1902), como eram chamados os transtornos na época.

Entre 1915 e 1926, ocorreu uma pandemia de encefalite letárgica. Esta doença infectou milhões de pessoas (LANGE *et al.*, 2010). Grande parte dos sobreviventes foram afetados por transtornos psiquiátricos ou doenças neurológicas, incluindo a hipercinesia e comportamento explosivo, disruptivo e distraído (FITZGERALD *et al.*, 2007) (LANGE *et al.*, 2010), o que na época foi associado ao comportamento do TDAH que havia sido, previamente, descrito por Still (STILL, 1902).

Uma vez que o transtorno foi se tornando cada vez menos associada ao dano cerebral, sendo chamada de “síndrome da criança hiperativa” ou “reação hipercinética da infância”, sendo adotado o nome atual a partir de 1994, no DSM-IV (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 1994), quando surgiu a diferenciação entre os três subtipos de acordo com a predominância de sintomas. E, esta definição foi mantida no DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013), assim como a subdivisão em três tipos: o desatento, no qual há predomínio de sintomas de desatenção, o hiperativo, no qual os sintomas predominantes são os de hiperatividade

e o combinado, no qual os sintomas de ambas as condições estão presentes.

D.1 Diagnóstico do TDAH

Na atualidade, o diagnóstico do TDAH é na maioria das vezes clínico e se dá de acordo com um critério predominantemente comportamental. Ademais, o diagnóstico do TDAH considera a história familiar, o rendimento escolar e fatores ambientais (WOLRAICH *et al.*, 2019)). O DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013), recomenda que o exame clínico seja realizado obtendo as respostas para os nove pontos dos Quadros 1 e 2. Se o paciente apresenta ao menos seis (cinco para maiores de 17 anos) dos comportamentos de desatenção da na Tabela 36, ele é diagnosticado como desatento.

Tabela 36 – Sintomas para Diagnóstico de TDAH relativos a desatenção.

Diagnóstico	Sintomas frequentes
Desatenção	Paciente não se atenta à detalhes ou comete erros por falta de cuidado em ambiente escolar ou no trabalho.
Desatenção	Paciente tem dificuldade de sustentar a atenção em tarefas.
Desatenção	Paciente não parece ouvir quando falam com ele diretamente.
Desatenção.	Paciente não segue instruções ou falha terminar em tarefas
Desatenção	Paciente tem dificuldades para organizar tarefas e atividades sequenciais; desorganizado com materiais e itens pessoais; gerencia mal o tempo.
Desatenção	Paciente evita, não gosta ou é relutante em executar tarefas que requeiram sustentar esforço mental.
Desatenção	Paciente perde objetos necessários para terminar tarefas.
Desatenção	Paciente é distraído por estímulos externos (ou pensamentos não relacionados).
Desatenção	Paciente esquece atividades diárias (manter compromissos, retornar chamadas, pagar contas).

Fonte: Adaptado de DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013).

Caso o paciente apresente no mínimo seis (cinco para maiores de 17 anos) dos comportamentos de hiperatividade da na Tabela 37, este é diagnosticado como hiperativo.

Se obtiver ambos os diagnósticos, então, o paciente possui TDAH do subtipo combinado. A análise destes fatores exige que eles estejam presentes na vida do paciente por

ao menos 6 meses, que eles estejam inconsistentes com o nível de desenvolvimento e que provoquem um impacto negativo na vida social, acadêmica ou ocupacional. É necessário que os sintomas estejam presentes antes dos 12 anos de idade e não estejam associados a outro transtorno.

Tabela 37 – Sintomas para Diagnóstico de TDAH relativa hiperatividade.

Diagnóstico	Sintomas frequentes
Hiperatividade	Paciente fica agitado ou fica se contorcendo enquanto está sentado.
Hiperatividade	Paciente se levanta em situações em que se espera que permaneça sentado
Hiperatividade	Paciente corre ou escala em situações inapropriadas ou se sente inquieto (adolescentes e adultos).
Hiperatividade	Paciente é incapaz de participar de atividades de lazer silenciosas.
Hiperatividade	Paciente se sente desconfortável em ficar parado por um longo tempo em restaurantes ou encontros.
Hiperatividade	Paciente fala excessivamente.
Hiperatividade	Em diálogos, o paciente responde antes da pergunta ser completamente formulada ou interrompe frases de outras pessoas.
Hiperatividade	Paciente tem dificuldades em esperar por sua vez.
Hiperatividade	Paciente interrompe ou se intromete em conversas e atividades, ou usa objetos de outras pessoas sem permissão.

Fonte: Adaptado de DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013).

A partir dos sintomas descritos no DSM-V, o trabalho de (SWANSON *et al.*, 2012) desenvolveu a escala SNAP (cujo nome provém das primeiras letras dos sobrenomes dos autores), na qual cada sintoma recebe uma pontuação segundo sua frequência (Nunca = 0, Um pouco = 1, Bastante = 2 e Muito = 3). Contudo, os autores notaram que os indivíduos analisados poderiam ter características nas quais tivessem desempenho superior ao considerado normal, o que poderia compensar os pontos nos quais apresentassem desvantagens. Por isso, desenvolveram uma nova escala, chamada de *Strengths and Weakness of ADHD-symptoms and Normal-behavior (SWAN)* com 7 possibilidades de classificação (Muito abaixo da média = 3, abaixo da média = 2, um pouco abaixo da média = 1, média = 0, um pouco acima da média = 1, acima da média = 2 e muito acima da média = 3) para as 18 características do DSM-V apresentadas anteriormente na Tabela 1.

D.2 Causas e consequências do TDAH

Atribuem-se múltiplas causas à ocorrência do TDAH, tais como fatores genéticos, e ambientais que contribuem e para o seu desenvolvimento. Os autores de (NÚÑEZ-JARAMILLO *et al.*, 2021) destacaram os fatores ambientais ocorridos antes do nascimento da criança, ou seja, nos períodos gestacional e pré-natal. Em sua revisão (NÚÑEZ-JARAMILLO *et al.*, 2021) ressaltaram que crianças nascidas prematuras possuem uma tendência 2,6-4 vezes maior do que as não-prematuras de possuírem o TDAH, o que está relacionado com a ocorrência de infecções no sistema nervoso central em bebês prematuros. O trabalho de (RAMALHO *et al.*, 2018) afirma que 25% dos diagnosticados com TDAH possuem um parente próximo também com TDAH e, quando o paciente tem um irmão ou irmã com este transtorno, a probabilidade este paciente ter TDAH é de 30-40% (90% para gêmeos monozigóticos). O estudo de (RAMALHO *et al.*, 2018) resalta os fatores ambientais que contribuem para o desenvolvimento da TDAH, tais como consumo de açúcar refinado, aditivos alimentares, deficiência de nutrientes e exposição à metais pesados. Complicações na gravidez e no parto, consumo de tabaco durante a gravidez e ambientes familiares adversos são fatores de risco importantes para o TDAH segundo (BIEDERMAN, 2005). A revisão de (SCIBERRAS *et al.*, 2017) apresenta outros fatores de risco pré-natal como o consumo de álcool, tabaco, drogas ilícitas, paracetamol, antidepressivos e até mesmo a baixa idade da mãe quando a criança nasce. Crianças entre 6 e 15 anos expostas a pesticidas tem duas vezes mais probabilidade de serem diagnosticadas com TDAH. A presença de maior concentração de metais pesados tais como manganês, mercúrio, chumbo, zinco, cromo, cobalto, níquel, cobre, molibdênio, estanho e bário está sendo investigada por diversos pesquisadores, por vezes resultando em conclusões contraditórias (NÚÑEZ-JARAMILLO *et al.*, 2021). De acordo com o estudo de revisão de (NÚÑEZ-JARAMILLO *et al.*, 2021), transtornos do sono costumam estar ligados ao TDAH sendo observados em 73-78% dos pacientes tanto em adultos como em crianças. Os transtornos do sono podem provocar sintomas similares ao TDAH, uma vez que para se manterem acordados, durante o dia, após uma noite mal dormida, os indivíduos podem apresentar desatenção, hiperatividade e impulsividade. A literatura sugere que os transtornos do sono podem ser causa, consequência ou doenças concomitantes provocadas pela mesma causa original. A pesquisa de (DEMONTIS *et al.*, 2019) demonstrou que o TDAH tem uma correlação muito forte com o genoma, o que mostra que está fortemente relacionada à traços herdados geneticamente. Conforme (WOLRAICH *et al.*, 2019), estima-se que em aproximadamente 76% dos casos o TDAH é hereditário. Dentre os genes cuja relação com o TDAH é mais investigada estão o gene de transporte de dopamina (DAT1), os genes de recepção de dopamina (DRD2, DRD3, DRD4 e DRD5), o gene de transporte de norepinefrina (SLC6A2), o receptor adrenérgico (ADRA2A), os genes de transporte (SLC6A4, 5HTT) e recepção (HTR1B, 5HT1B, HTR2A, 5HT2A) de serotonina,

os genes de Tryptophan hydroxylase (TPH1 e TPH2), os genes de Monoamine oxidase A (MAOA), os genes receptores da Acetilcolina nicotínica (CHRNA4), os genes associados com a proteína Synaptosomal (SNAP-25) e os genes do fator neurotrófico derivado do cérebro (BDNF), bem como as enzimas Catechol-O-methyl-transferase (COMT) e Dopamina beta hydroxylase (DH), responsável pela conversão de dopamina em norepinefrina. Alterações na enzima COMT, a qual é importante para regular o nível de dopamina no cérebro, também podem estar relacionadas o TDAH. Isto porque a síndrome velo-cardio-facial, onde ocorre a alteração da COMT, provoca sintomas similares aos do TDAH. A maioria das pesquisas genéticas busca uma relação entre o TDAH e genes relacionados ao processamento de dopamina, serotonina ou adrenalina, porém ainda não se identificou nenhum gene específico que provoque o TDAH e por isso o teste genético não é usado para fins de diagnóstico desta condição (GIZER *et al.*, 2009). Os genes DRD2, COMT, DRD4 e SLC6A2 são altamente representados nos lobos frontais do cérebro. Os demais genes se concentram no núcleo *accumbens*, substantia nigra, hipocampo, região pré-frontal (ADRA2A) ou se espalham equilibradamente entre as diferentes partes do cérebro. Os trabalhos de (PINEAU *et al.*, 2019); (PFEIFER *et al.*, 2022) correlacionaram o TDAH com genes ligados à transmissão e à recepção de dopamina, respectivamente. (QUIST *et al.*, 2003) verificou que desequilíbrios no gene 5-HT1B responsável por regular a dopamina através da serotonina estão relacionados à ocorrência do TDAH. O trabalho de (VOLKOW *et al.*, 2009) verificou que as pessoas com TDAH possuem menores níveis de genes de transporte e recepção de dopamina. Uma substância comumente utilizada para o tratamento do TDAH é o metilfenidato, o qual aumenta a dopamina nas sinapses cerebrais atuando sobre o gene de transporte de dopamina (LAI *et al.*, 2018). Conforme o DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013), o TDAH está associado ao mau desempenho na vida acadêmica e no trabalho, rejeição social, maior probabilidade de desemprego e muitos conflitos interpessoais. As dificuldades sociais estão associadas a interpretação dos demais sobre o comportamento do indivíduo com TDAH que é percebido como irresponsável, preguiçoso ou não-cooperativo. Assim, o indivíduo sofre com rejeições e provocações. Além de, possivelmente, prejudicar a vida acadêmica e profissional do indivíduo afetado, os sintomas provocados do TDAH elevam diversos riscos em sua vida, tais como o risco de suicídio (FURCZYK & THOME, 2014), o risco de abuso de substâncias (DALSGAARD *et al.*, 2014), o risco de envolvimento em acidentes de trânsito (CHANG *et al.*, 2014), o risco de sofrer ferimentos (DALSGAARD *et al.*, 2015), o risco de contrair doenças venéreas (BARKLEY *et al.*, 2010), o risco de se envolver em crimes (CHANG, *et al.*, 2014). Consequentemente, o indivíduo com TDAH possui um risco maior, que os indivíduos do grupo de controle, de morrer prematuramente (DALSGAARD *et al.*, 2015). Segundo o DSM-V (AMERICAN PSYCHIATRIC ASSOCIATION & Association, 2013), o mau desempenho acadêmico está mais relacionado a desatenção. Já a rejeição social e acidentes estão mais relacionados a hiperatividade.

Portanto, é fundamental para os indivíduos com TDAH que o diagnóstico da presença do transtorno seja obtido o quanto antes. E, assim, o indivíduo possa iniciar um tratamento adequado o mais rápido possível.

APÊNDICE E LISTA DE INDIVÍDUOS CUJOS DADOS FORAM UTILIZADOS NA CONFECÇÃO DESTA TESE

A base de dados HBN identifica os indivíduos por códigos. As identidades reais destes indivíduos são confidenciais e acessíveis apenas para poucos membros da equipe do *Child Minds Institute*.

E.1 Indivíduos que forneceram dados de EEG

Indivíduos da Classe Controle (Total - 193):

NDARAA075AMK, NDARAC904DMU, NDARAE012DGA, NDARAM277WZT
 NDARAM675UR8, NDARAP522AFK, NDARAT100AEQ, NDARAU447JZH
 NDARAV610EY3, NDARAX283MAK, NDARBA521RA8, NDARBN100LCD
 NDARBY518PRN, NDARCD401HGZ, NDARCH514JCT, NDARCJ363KLE
 NDARCX053GU5, NDARDJ092YKH, NDARDJ825GBP, NDARDL511UND
 NDARDR296XHN, NDARDR804MFE, NDARDX872VH6, NDARED047DTH
 NDAREF389RY2, NDAREG590BNY, NDAREH074NG8, NDAREK549XUQ
 NDAREL622FA6, NDAREM703YFD, NDARET653TAM, NDAREU551GPC
 NDAREW201WD9, NDAREW661NZJ, NDAREW976FNL, NDAREX065KJU
 NDARFG851ZNZ, NDARFK610GY5, NDARFL411AT1, NDARFW292PBD
 NDARFY075REK, NDARGB441VVD, NDARGD507TDZ, NDARGF543PM2
 NDARGG205WVN, NDARGH592NZ2, NDARGH790CEF, NDARGL800LDW
 NDARGR875AXY, NDARGT022BEW, NDARGT682ZWN, NDARGV956EGX
 NDARGX443CEU, NDARHB000YF8, NDARHB764VZ2, NDARHJ945PV0
 NDARHR443EHF, NDARHV527DNK, NDARJH441HJD, NDARJJ343TR0
 NDARJJ356DAL, NDARJM708VGE, NDARJV377HG4, NDARJW315BUA
 NDARKH837TB2, NDARKN509RP9, NDARKR010EN7, NDARKT811ATJ
 NDARKU278YRR, NDARLA395AG8, NDARLA516PH1, NDARLC358CYJ
 NDARLE091XAZ, NDARLJ886BFK, NDARLM196YRG, NDARLU111UYF
 NDARLU606ZDD, NDARLV387GP4, NDARLX816JUZ, NDARLX836EHJ

NDARLY114PDC, NDARLZ104NDT, NDARMC760NEC, NDARMF508PA2
 NDARMT882AWE, NDARMY301WFK, NDARNG689AAP, NDARNH147WGN
 NDARNK489GNR, NDARPE752VYE, NDARPL596YTD, NDARPW482TVE
 NDARPZ621ZLE, NDARTR840XP1, NDARUA421HT8, NDARUF069EHR
 NDARUG507AZX, NDARVY054WEA, NDARWW901MTC, NDARWX051KEV
 NDARXU253JHC, NDARXZ902NFM, NDARYH480GTD, NDARYL272HDW
 NDARYL758JGG, NDARZG690NH, NDARAP522AFK, NDARAV610EY3
 NDARAX283MAK, NDARBH024NH2, NDARBU928LV0, NDARBY518PRN
 NDARCD401HGZ, NDARCE721YB5, NDARCH514JCT, NDARCJ363KLE
 NDARCT472UJ7, NDARCW963FP9, NDARCX053GU5, NDARCY178KJP
 NDARDC814WW2, NDARDH086ZKK, NDARDJ092YKH, NDARDJ825GBP
 NDARDL511UND, NDARDN489EXJ, NDARDR296XHN, NDARDR804MFE
 NDARDX872VH6, NDAREF389RY2, NDAREG590BNY, NDAREH074NG8
 NDAREK549XUQ, NDAREM703YFD, NDARET653TAM, NDAREW661NZJ
 NDARGB441VVD, NDARGG205WVN, NDARGH592NZ2, NDARGH790CEF
 NDARGT634DUJ, NDARGT682ZWN, NDARET893AK8, NDAREY778CR9
 NDARFA860RPD, NDARFB322DRA, NDARFB908HVX, NDARFD822TY5
 NDARFL411AT1, NDARGV956EGX, NDARGX443CEU, NDARHP924ZHW
 NDARHR443EHF, NDARHV527DNK, NDARJ257ZU2, NDARJB355YHK
 NDARJH441HJD, NDARJJ343TR0, NDARJJ356DAL, NDARJM708VGE
 NDARJV377HG4, NDARJW315BUA, NDARJZ089HVP, NDARKH837TB2
 NDARKN509RP9, NDARKR010EN7, NDARKT811ATJ, NDARKU278YRR
 NDARLM196YRG, NDARLU606ZDD, NDARLZ104NDT, NDARPJ151ERK
 NDARPL596YTD, NDARZV421TCZ, NDARBL117AUV, NDARFK610GY5
 NDARFL411AT1, NDARFP813TL5, NDARGD507TDZ, NDARGF543PM2
 NDARGL800LDW, NDARGR875AXY, NDARGT022BEW, NDARGT634DUJ
 NDARHB000YF8, NDARHB764VZ2, NDARHJ945PV0, NDARKV800DTW
 NDARLA395AG8, NDARLA516PH1, NDARLC358CYJ, NDARLE091XAZ
 NDARLU111UYF

Indivíduos da Classe TDAH-C (Total - 329):

NDARAA112DMH, NDARAA117NEJ, NDARAB055BPR, NDARAB708LM5
 NDARAC462DZH, NDARAC853DTE, NDARAD224CRB, NDARAD615WLJ
 NDARAG340ERT, NDARAH793FBF, NDARAK738BGC, NDARAP176AD1
 NDARAP782TVC, NDARAT299YRR, NDARAU708TL8, NDARAV187GJ5
 NDARAV894XWD, NDARAY238DFL, NDARBC240LNC, NDARBH228RDW
 NDARBM433VER, NDARBU112XZE, NDARBU532YXZ, NDARBU607ZZ3
 NDARBV059CGE, NDARBV364MBC, NDARBW026UGE, NDARBX121UM9

NDARBX400RTC, NDARCA153NKE, NDARCE353XKH, NDARCF052AE0
NDARCG159AAP, NDARCH868WVX, NDARCJ475WJP, NDARCK647MU6
NDARCK661RZ6, NDARCL008HLA, NDARCT933HF2, NDARCU957PGQ
NDARCW071AU5, NDARCW497XW2, NDARCZ770BRG, NDARCZ915NC1
NDARCZ947WU5, NDARDC389TLD, NDARDC843HHM, NDARDE597GKQ
NDARDN702GXX, NDARDR458MR7, NDARDV245WJG, NDARDX770PJK
NDARDZ147ETZ, NDARDZ440NGK, NDAREF888EVV, NDAREK375DKR
NDAREK395BM3, NDAREK801BPB, NDAREK947FYP, NDAREM018TJQ
NDAREM141CKP, NDAREM731BYM, NDAREN151YXN, NDAREN417DJE
NDAREN667YTZ, NDARER115FTJ, NDARHT095YB4, NDARHU211PYD
NDARHU831NXU, NDARHX628PC6, NDARHX877BLQ, NDARHY099CUL
NDARHY255FVU, NDARHZ439EBY, NDAREC647MKW, NDARAA948VFH
NDARAD224CRB, NDARAG143ARJ, NDARAJ366ZFA, NDARAJ807UYR
NDARAN076TPG, NDARAV187GJ5, NDARAV189JGX, NDARAW320CGR
NDARBA839HLG, NDARBE091BGD, NDARBH228RDW, NDARBK638HLZ
NDARBM642JFT, NDARBU757VJK, NDARBV059CGE, NDARBZ475NKK
NDARCB370UK4, NDARCD661GL0, NDARCJ594BWQ, NDARCK162REX
NDARCK661RZ6, NDARDF373ZDW, NDARDF568GL5, NDARDH934MUV
NDAREN417DJE, NDAREV970HAN, NDAREW074ZM2, NDAREY512KVX
NDAREY741MG9, NDAREY897LB1, NDARFA402LMW, NDARFB757VY3
NDARFB969EMV, NDARFE163HVN, NDARFE299MWU, NDARFH674DWX
NDARKX701BJ4, NDARML024FF2, NDARML406ZB8, NDARMP977RHZ
NDARMT661WCR, NDARMU269TFY, NDARMZ200GVD, NDARNN164EA6
NDARNP370WGP, NDARNZ792HBN, NDARPB619PHV, NDARPL721WTD
NDARPV009TYX, NDARPV303LAX, NDARPW915RGD, NDARPY297MD8
NDARRA717GYV, NDARRA719CPH, NDARRG735TU5, NDARRK370YFX
NDARRL685WB7, NDARTC802PGL, NDARHX145ZKV, NDARJW373UE3
NDARKP815KPZ, NDARKT032RYR, NDARKU649DB7, NDARKW010CT2
NDARKW521EMY, NDARKZ479FC8, NDARLF142AF5, NDARLF283KJX
NDARLJ148JH2, NDARLJ311YNA, NDARLJ334YD0, NDARLJ549WTT
NDARLK551KXD, NDARLN778RYN, NDARLN961XGP, NDARLP624GFH
NDARLY125HMB, NDARLY585ZVN, NDARMF939FNX, NDARJE086YU0
NDARJF517HC8, NDARJG298YVA, NDARJG300HWH, NDARJH707GJM
NDARJJ767FBH, NDARJJ859RRW, NDARJM618NZH, NDARKC143BEE
NDARKD064EVC, NDARKD134TCX, NDARKF442GZ5, NDARKG460MWV
NDARKJ322ELL, NDARKL327YDQ, NDARXG670HFE, NDARFJ000DCY
NDARFL793LDE, NDARFP512MEK, NDARFR095UJK, NDARFR820KFF
NDARFT834NT1, NDARFU444HEF, NDARFW140WFX, NDARFX455JFY
NDARFX710UZA, NDARFZ296UNG, NDARGB000CW8, NDARGB040MGR

NDARGE536BGD, NDARGE994BMX, NDARGF192VD1, NDARGH407KVG
NDARGJ627BL2, NDARGK608LJ8, NDARGK786DWR, NDARGL085RTW
NDARGL107GFK, NDARGM540EV8, NDARGR125APK, NDARGR661YZG
NDARGU100JH4, NDARGU395RFP, NDARGU667MCK, NDARGV948BFZ
NDARGW404GJL, NDARGY542TJ5, NDARGZ282DLD, NDARHB772TE5
NDARHE199NGL, NDARHB993EV0, NDARHE236YLB, NDARHF351AA6
NDARHG321LLD, NDARHG594GKH, NDARKM497VKX, NDARKN175HWB
NDARHJ523XGB, NDARHM413HVA, NDARHN482HPM, NDARJF565ZRA
NDARJM618NZH, NDARJP489HCE, NDARJR579FW7, NDARJU787GF5
NDARJX939UCQ, NDARJY033DKZ, NDARJZ526HN3, NDARJZ549HU8
NDARKA958FAH, NDARKB290YNY, NDARXE684LUY, NDARRC190NKB
NDARRC355BM6, NDARRE445RHR, NDARRF897HB5, NDARRW100AU0
NDARRX443NVT, NDARRX819BZ0, NDARTB213ML4, NDARTB755MF5
NDARTC227AXA, NDARTD468TUH, NDARTE424BUF, NDARTE773XJ5
NDARTG679NKQ, NDARTK357VHL, NDARTK933KL9, NDARTN400TB1
NDARTN487KDK, NDARTV776ZJ8, NDARTW939GJF, NDARUB231LHN
NDARUC022JWT, NDARUD306BB0, NDARUE533XNH, NDARUF389TXE
NDARUJ779NM0, NDARUL454XRE, NDARUN221VCJ, NDARUV621PB0
NDARUZ206DRV, NDARXG670HFE, NDARXH017ZLE, NDARXN764AC7
NDARYB432MB5, NDARYJ536MCR, NDARUZ601TW4, NDARVC350WVG
NDARVD194JX2, NDARVL964HH1, NDARVM685MER, NDARVN469KP6
NDARVN883AU2, NDARVP109DF0, NDARVP135ZGE, NDARVP285RXV
NDARVX162AZU, NDARVZ525AA0, NDARWA351ZE2, NDARWB457JU6
NDARWC925WDX, NDARWF122UUJ, NDARWG911DZG, NDARWJ621JRZ
NDARWM319UU2, NDARWR881WMX, NDARWV677EFC, NDARYL771XDP
NDARYY218LU2, NDARYY694NE7, NDARZE542ZVH, NDARZL724HAX
NDARZN277NR6, NDARZT096AUP, NDARZT199MF6, NDARZU822WN3
NDARZX561DR9, NDARZY019GNZ, NDARWT694TXM, NDARXD388TTE
NDARXF623GEL, NDARXK164TNN, NDARXL023GFG, NDARXL023NG9
NDARXL523GHX, NDARXR123PFR, NDARYJ301DYN, NDARYM277DEA
NDARYM994YF0, NDARYN595JMA, NDARYT155NHX, NDARYW984FLT
NDARYZ909VND, NDARYZ986HEW, NDARZE850WXD, NDARZF949WTW
NDARZG081LYG, NDARZG861CXJ, NDARZH521MMD, NDARAM048AFP
NDARWV405ZW0, NDARWV938NKD, NDARWX338NDL, NDARXC534KT4
NDARNB674VMY, NDARN085WZK, NDARNW218ZBU, NDARPE551CK7
NDARPF319WV2, NDARPF640ZRG, NDARPL201YL4, NDARPP622WV4
NDARWP954GPJ

Indivíduos da Classe TDAH-I (Total - 288):

NDARAB458VK9, NDARAC349YUC, NDARAD774HAZ, NDARAG340ERT
NDARAJ977PRJ, NDARAK738BGC, NDARAP785CTE, NDARAR238RZ8
NDARAX272ZJL, NDARAX287BLD, NDARAY461TZZ, NDARBF176LPM
NDARBH512BHA, NDARBJ482HJL, NDARBJ674TVU, NDARBL214YLX
NDARBR128UFP, NDARBV577EE0, NDARBZ925PRF, NDARCA050RZL
NDARCG984YXZ, NDARCH001CN2, NDARCH889NUF, NDARCH881EFG
NDARCH889NUF, NDARCM135DVC, NDARCN669XPR, NDARCP292KPA
NDARCR499NE4, NDARCT974NAJ, NDARCU118LMX, NDARCU744XWL
NDARCU865PBV, NDARDD128NWL, NDARDG747ULB, NDARDL129FV3
NDARDU986RBM, NDARDV222JZ6, NDARDY714NV9, NDARDZ058NZN
NDARDZ627XFA, NDAREC078VFT, NDAREC542MH3, NDAREE196FEG
NDAREF893ZM8, NDAREK918EC2, NDAREP505XAD, NDARET550NXC
NDARFG118DW3, NDARFF644ZGD, NDAREV527ZRF, NDARFM619GTV
NDARHY177LY9, NDARJA157YB3, NDARLJ654TZ0, NDARLR030EW6
NDARNR459MUJ, NDARNT043XGH, NDARNT151UN1, NDARRC016UZE
NDARRH407MEY, NDARTB552HP0, NDARTB883GUN, NDARAN461TGQ
NDARAP359UM6, NDARAT244VGA, NDARAY283AU1, NDARBC296PLU
NDARBD879MBX, NDARBL444FBA, NDARBR128UFP, NDARBZ925PRF
NDARCL895UV8, NDARCU744XWL, NDARCX221CWA, NDARDG484ZJ1
NDARDM385EK2, NDARDU760YZC, NDAREF848YWD, NDAREG408DJ5
NDAREV601CE7, NDAREV848HWX, NDAREW201WD9, NDARFA780CUG
NDARFB506UJ3, NDARFD453NPR, NDARFE372VHL, NDARFE526NK3
NDARFF598HGT, NDARFG114VB1, NDARFG713PLR, NDARFK477WX9
NDARFL579BAP, NDARFM211KDJ, NDARFM229HHA, NDARFM619GTV
NDARFT581ZW5, NDARFT615JZ6, NDARFT964EU8, NDARFU395UBW
NDARFV780ABD, NDARFW444PN1, NDARFY009MNL, NDARGA507DCC
NDARGB324PM1, NDARGF367KVL, NDARGH775KF5, NDARGH842VFY
NDARGJ395FKP, NDARGM196JMM, NDARGN044FE4, NDARGR893PUA
NDARGT267UZR, NDARGU729WUR, NDARGV781AMW, NDARGX583YNZ
NDARGX760NYV, NDARGZ361JJ9, NDARHC462NGR, NDARHD059PEQ
NDARHF150FDY, NDARHF269ATU, NDARHF545HFW, NDARHF904CWB
NDARHG294AFG, NDARHL164JHH, NDARHL238VL2, NDARHN931CJV
NDARJA788CH7, NDARJB366HYJ, NDARJF579EH9, NDARJJ389CP2
NDARJL058RDV, NDARJN531EN3, NDARJT588ZAX, NDARJZ274PRQ
NDARJZ888UX6, NDARKA085YRG, NDARKE604RXR, NDARKG594WKT
NDARKM551DA4, NDARKN346XZH, NDARKU616UBL, NDARKV353MTD
NDARKX346VTV, NDARLB547HJD, NDARLB581AXF, NDARLD016VCA

NDARLE417RA4, NDARLE953JYU, NDARLF480DMW, NDARLG154ZA3
NDARLV086DEH, NDARLW312JT8, NDARLY687YDQ, NDARMH364EMW
NDARMH559LKK, NDARMJ495DE0, NDARMK416UUQ, NDARMK525TGP
NDARML148UCE, NDARMM005DLK, NDARMT271DT6, NDARMV410ZFY
NDARMV718DYL, NDARMV822KWR, NDARPP726XBF, NDARPV145KGL
NDARPW482TVE, NDARRA314KW4, NDARRA537FWW, NDARRK038CBD
NDARRL315KV3, NDARRN444MJG, NDARRN619WHY, NDARRU290EZV
NDARRU979UBW, NDARRW643DHJ, NDARRW680PHD, NDARRX800KW8
NDARRX897XV1, NDARRY112RAK, NDARRY268AF2, NDARRY715CY4
NDARRZ927VC3, NDARTJ418GP0, NDARTJ433TB5, NDARTX659HAF
NDARTY146DKE, NDARTY225EWV, NDARTY549YDG, NDARTZ901LDK
NDARUJ962NZZ, NDARUK329RUU, NDARUM791WPF, NDARUP067TM8
NDARUT470BM4, NDARUX086TVT, NDARUY529FNK, NDARVB151GXF
NDARVC711ZK0, NDARVG971CHH, NDARVH153RE7, NDARVM025NCF
NDARVN363NNQ, NDARWA622HHZ, NDARWB685NUG, NDARWJ414WB8
NDARWJ498CZY, NDARWK065NJ9, NDARWN691CG7, NDARWN709HA2
NDARWP815VUP, NDARWT086XMF, NDARWU018VDB, NDARWW480TU2
NDARWW820WZN, NDARWX711LBK, NDARWY198BVY, NDARXF441RHQ
NDARXP557DLJ, NDARXP717YMH, NDARXU715EGL, NDARXU902BRT
NDARXV294YRJ, NDARXY240WJC, NDARXY532ZTT, NDARYA030ZG7
NDARYF085LZQ, NDARYF135ZHU, NDARYF272EDC, NDARYG546ZFF
NDARYG665JUD, NDARYJ638RTN, NDARYN152YMW, NDARYN174NPH
NDARYN254VW6, NDARYR150UDP, NDARYX806FL1, NDARYY664KHF
NDARYZ871DXQ, NDARZA511FYF, NDARZA694WHF, NDARZE831DKC
NDARZN677EYE, NDARZN899JCM, NDARZV766YXP, NDARZW013RUH
NDARZW472CCF, NDAREX589EXG, NDARFA089ZZG, NDARMB216LA6
NDARME880JD7, NDARMM951YEH, NDARMR570GJ5, NDARMY533CYM
NDARNF766VW2, NDARNF873FCV, NDARNL587LVL, NDARNT245TP0
NDARNU770PM5, NDARNW182PJC, NDARNY023FWG, NDARNZ615UEU
NDARPB293EPC, NDARPF937BDQ, NDARPU329MDJ, NDARPX244EHJ
NDARPZ344ZGE, NDARTC710UM1, NDARTC865YNZ, NDARTN487KDK
NDARTT979FDD, NDARTV525XC0, NDARVP953BBF, NDARVT488RXJ
NDARVV199XK2, NDARWK983JFE, NDARWV938NKD, NDARWX338NDL
NDARWZ346RV4, NDARXD944ZXV, NDARXF497LYF, NDARXG646PM9
NDARXH393TMM, NDARXK009VGV, NDARXK986EK6, NDARYM994YF0

E.2 Indivíduos que forneceram dados de fMRI aprovados no Controle de Qualidade

Indivíduos da Classe Controle (Total - 57):

NDARAP522AFK, NDARAV610EY3, NDARAX283MAK, NDARBU928LV0
NDARCE721YB5, NDARCG808HDJ, NDARDC814WW2, NDARDN489EXJ
NDARDX872VH6, NDAREK549XUQ, NDARFG851ZNZ, NDARFK610GY5
NDARGA563XVW, NDARGB441VVD, NDARGD507TDZ, NDARGF543PM2
NDARGG205WVN, NDARGT682ZWN, NDARGX443CEU, NDARJJ343TR0
NDARJV377HG4, NDARJZ089HVP, NDARKH837TB2, NDARLF446MT5
NDARLM196YRG, NDARLV387GP4, NDARLY114PDC, NDARMT882AWE
NDARMZ518UH1, NDARND697FLK, NDARPE056ACA, NDARRB338YZ0
NDARRK882CLT, NDARRY280KNW, NDARTN173MXV, NDARTU777GVV
NDARUL456EER, NDARUP249AMD, NDARUW816MP3, NDARVC195NLH
NDARWM656UWL, NDARWR247CE1, NDARXF203DCD, NDARXF358XGE
NDARXM571XCC, NDARXZ685TU4, NDARXZ902NFM, NDARYA503XVQ
NDARYD546HCB, NDARYH697TPA, NDARYK164AEJ, NDARYL758JGG
NDARZG690NHH, NDARZM903TNL, NDARZV421TCZ, NDARPC817XZ5
NDARXV445NYZ

Indivíduos da Classe TDAH-C (Total - 137):

NDARAA948VFH, NDARAD224CRB, NDARAP176AD1, NDARAP782TVC
NDARAU708TL8, NDARBH228RDW, NDARBM433VER, NDARBU532YXZ
NDARBV059CGE, NDARBW026UGE, NDARBX400RTC, NDARCF052AE0
NDARCG159AAP, NDARCK647MU6, NDARCZ915NC1, NDARDC843HHM
NDARDE597GKQ, NDARDF568GL5, NDARDT528ZZJ, NDARDV245WJG
NDARDZ147ETZ, NDARED945TNQ, NDAREE015WGU, NDAREH852JE0
NDAREK395BM3, NDAREK801BPB, NDAREN417DJE, NDAREN667YTZ
NDAREY897LB1, NDARFA402LMW, NDARFB969EMV, NDARFE299MWU
NDARFL793LDE, NDARFR820KFF, NDARFT881VT6, NDARFU444HEF
NDARFU916UZY, NDARFX455JFY, NDARFX710UZA, NDARGA048BD8
NDARGA815EUZ, NDARGB000CW8, NDARGL085RTW, NDARGM439FZD
NDARGR125APK, NDARGV948BFZ, NDARGW404GJL, NDARHM413HVA
NDARHM611GDR, NDARHT095YB4, NDARHT095YB4, NDARHX628PC6
NDARHX934KJ5, NDARJF517HC8, NDARJG298YVA, NDARJG300HWH
NDARJH707GJM, NDARJM296MDW, NDARJT517YHZ, NDARKA958FAH
NDARKB290YNY, NDARKE456BVQ, NDARKF442GZ5, NDARKG460MWV

NDARKJ322ELL, NDARKK774VP5, NDARKM568HJK, NDARKP815KPZ
 NDARKU649DB7, NDARKW010CT2, NDARKW521EMY, NDARLF142AF5
 NDARLH881FBD, NDARLJ148JH2, NDARLJ334YD0, NDARLK551KXD
 NDARLK963HXP, NDARLL790WLF, NDARLN070VX7, NDARLN961XGP
 NDARLP624GFH, NDARMG196KV4, NDARML024FF2, NDARMN302YAV
 NDARMN695VFH, NDARMT661WCR, NDARNV694EMG, NDARNX125KJ1
 NDARNZ792HBN, NDARPL201YL4, NDARPV009TYX, NDARPV303LAX
 NDARPW915RGD, NDARRE445RHR, NDARRH285LR3, NDARRX819BZ0
 NDARRY006EMB, NDARTB203DU7, NDARTB755MF5, NDARTC802PGL
 NDARTG679NKQ, NDARTK357VHL, NDARTN487KDK, NDARTR464DCR
 NDARUJ779NM0, NDARUL454XRE, NDARUU615WPR, NDARUX534HZQ
 NDARUZ206DRV, NDARVC673HFA, NDARVD194JX2, NDARVD543CMZ
 NDARVL498YX5, NDARVL964HH1, NDARVN629CRD, NDARVN883AU2
 NDARWA351ZE2, NDARWT715BMK, NDARXD345GBN, NDARXL023NG9
 NDARXN764AC7, NDARXU376FDN, NDARYE379GGT, NDARYG391PMU
 NDARYM994YF0, NDARYT332LMM, NDARYZ770NA1, NDARZE850WXD
 NDARZG081LYG, NDARZG236AUT, NDARZH521MMD, NDARZJ460RWZ
 NDARZN277NR6, NDARZY668NMV, NDARGB528CP8, NDARHE199NGL
 NDARHY676RYH, NDARPP622WV4

Indivíduos da Classe TDAH-I (Total - 145):

NDARAB458VK9, NDARAC349YUC, NDARAG340ERT, NDARAK738BGC
 NDARAX272ZJL, NDARAY283AU1, NDARBC240LNC, NDARBF176LPM
 NDARBL444FBA, NDARBR128UFP, NDARBZ925PRF, NDARCH001CN2
 NDARCH881EFG, NDARCH889NUF, NDARCL590LPD, NDARCM811CV1
 NDARCU118LMX, NDARCU744XWL, NDARCU865PBV, NDARCZ947WU5
 NDARDV222JZ6, NDARDZ627XFA, NDAREF848YWD, NDAREF893ZM8
 NDAREP505XAD, NDARET550NXC, NDAREV527ZRF, NDARFA089ZZG
 NDARFB506UJ3, NDARFD453NPR, NDARFE526NK3, NDARFF644ZGD
 NDARFG118DW3, NDARFG713PLR, NDARFK477WX9, NDARFM229HHA
 NDARFM619GTV, NDARFT787HJW, NDARFU395UBW, NDARFV780ABD
 NDARFY009MNL, NDARGA507DCC, NDARGF367KVL, NDARGH775KF5
 NDARGK059JW7, NDARGM196JMM, NDARGR893PUA
 NDARGV781AMW, NDARGX760NYV, NDARHD931FNA, NDARHF150FDY
 NDARHJ034LDM, NDARHL164JHH, NDARHL822EJK, NDARHP841RMR
 NDARJA157YB3, NDARJA788CH7, NDARJC559WW5, NDARJF579EH9
 NDARJH492TVW, NDARJK910PDK, NDARJM036PVX, NDARJN531EN3
 NDARKG594WKT, NDARKH444AGK, NDARKK732JW3, NDARKN346XZH
 NDARKN477EBH, NDARKU616UBL, NDARKV353MTD, NDARKX528PH3

NDARLB547HJD, NDARLG154ZA3, NDARMB216LA6, NDARLR030EW6
NDARMB216LA6, NDARMM005DLK, NDARMM951YEH, NDARMY294VNE
NDARNND348HB3, NDARNF873FCV, NDARNK740ZVM, NDARNL587LVL
NDARNX223DY3, NDARNZ615UEU, NDARPD855ARC, NDARPF459CME
NDARPF937BDQ, NDARPG847LB8, NDARPL306LC6, NDARPP726XBF
NDARPV145KGL, NDARPZ344ZGE, NDARRK038CBD, NDARRU290EZV
NDARRU979UBW, NDARTC865YNZ, NDARTG035JK8, NDARTV525XC0
NDARUG278FNY, NDARUJ292JXV, NDARUK329RUU, NDARUM791WPF
NDARUT470BM4, NDARUX408KJ1, NDARVC711ZK0, NDARVK720WZ0
NDARVL697YGK, NDARVN363NNQ, NDARVP953BBF, NDARVT488RXJ
NDARVU320XJZ, NDARVV199XK2, NDARVX111CNA, NDARVY378DFK
NDARWA433WYZ, NDARWA622HHZ, NDARWJ414WB8, NDARWK983JFE
NDARWN691CG7, NDARWN709HA2, NDARWT086XMF, NDARWW480TU2
NDARWX711LBK, NDARXF497LYF, NDARXG646PM9, NDARXH393TMM
NDARXK986EK6, NDARXU018RGY, NDARXY532ZTT, NDARYA030ZG7
NDARYC466ER1, NDARYF699FW0, NDARYG665JUD, NDARYM334BZ5
NDARYN174NPH, NDARYN254VW6, NDARYY664KHF, NDARZL855WVA
NDARZN677EYE, NDARZW472CCF, NDARAY461TZZ, NDARCM135DVC
NDARPH923FN7