



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA BIOMÉDICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA

GABRIEL MIRANDA DE SOUZA

**Representação de Sinais de EEG em Imagens para o Diagnóstico de Alzheimer:
uma abordagem baseada em Deep Learning e Métodos de Interpretabilidade**

Recife

2025

GABRIEL MIRANDA DE SOUZA

**Representação de Sinais de EEG em Imagens para o Diagnóstico de Alzheimer:
uma abordagem baseada em Deep Learning e Métodos de Interpretabilidade**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Biomédica do Centro de Tecnologia e Geociências da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre em Engenharia Biomédica.

Área de Concentração: Computação Biomédica

Orientador (a): Prof. Dr. Wellington Pinheiro dos Santos

Coorientador (a): Prof. Dra. Juliana Carneiro Gomes

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Souza, Gabriel Miranda de.

Representação de sinais de EEG em imagens para o diagnóstico de Alzheimer: uma abordagem baseada em Deep Learning e Métodos de Interpretabilidade / Gabriel Miranda de Souza. - Recife, 2025.

72f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Engenharia Biomédica, 2025.

Orientação: Wellington Pinheiro dos Santos.

Coorientação: Juliana Carneiro Gomes Cassemiro.

Inclui referências e apêndices.

1. Deep Learning; 2. Doença de Alzheimer; 3. Eletroencefalografia; 4. ASTERI; 5. Inteligência Artificial Explicável. I. Santos, Wellington Pinheiro dos. II. Cassemiro, Juliana Carneiro Gomes. III. Título.

UFPE-Biblioteca Central



UNIVERSIDADE FEDERAL DE PERNAMBUCO

Ata da defesa/apresentação do Trabalho de Conclusão de Curso de Mestrado do Programa de Pós-graduação em Engenharia Biomédica - CTG da Universidade Federal de Pernambuco, no dia 25 de março de 2025.

ATA Nº 99

Aos 25 dias do mês de março de 2025, às 14:00, em sessão pública realizada de forma presencial, teve início a defesa/apresentação do Trabalho de Conclusão de Curso intitulada Representação de Sinais de EEG em Imagens para o Diagnóstico de Alzheimer: uma abordagem baseada em Deep Learning e Métodos de Interpretabilidade do(a) mestrando(a) GABRIEL MIRANDA DE SOUZA, na área de concentração Computação Biomédica, sob a orientação do(a) Prof.(a) WELLINGTON PINHEIRO DOS SANTOS e co-orientação do(a) Prof.(a) JULIANA CARNEIRO GOMES CASSEMIRO. A Comissão Examinadora foi aprovada pelo colegiado do programa de pós-graduação em 18/03/2025, sendo composta pelos examinadores: GISELLE MACHADO MAGALHAES MORENO, do(a) UFPE; WELLINGTON PINHEIRO DOS SANTOS, do(a) UFPE; JULIANA CARNEIRO GOMES CASSEMIRO, do(a) UFPE; MAIRA ARAUJO DE SANTANA, do(a) UFPE. Após cumpridas as formalidades conduzidas pelo(a) presidente(a) da comissão, professor(a) WELLINGTON PINHEIRO DOS SANTOS, o(a) candidato(a) ao grau de Mestre(a) foi convidado(a) a discorrer sobre o conteúdo do Trabalho de Conclusão de Curso. Concluída a explanação, o(a) candidato(a) foi arguido(a) pela Comissão Examinadora que, em seguida, reuniu-se para deliberar e conceder, ao mesmo, a menção APROVADO. Para a obtenção do grau de Mestre(a) em Engenharia Biomédica, o(a) concluinte deverá ter atendido todas às demais exigências estabelecidas no Regimento Interno e Normativas Internas do Programa, nas Resoluções e Portarias dos Órgãos Deliberativos Superiores, assim como no Estatuto e no Regimento Geral da Universidade, observando os prazos e procedimentos vigentes nas normas.

Dra. MAIRA ARAUJO DE SANTANA, UFPE

Examinadora Externa ao Programa

Dra. GISELLE MACHADO MAGALHAES MORENO, UFPE

Examinadora Interna

Dra. JULIANA CARNEIRO GOMES CASSEMIRO, UFPE

Examinadora Interna

Dr. WELLINGTON PINHEIRO DOS SANTOS, UFPE

Examinador Interno



UNIVERSIDADE FEDERAL DE PERNAMBUCO

Ata da defesa/apresentação do Trabalho de Conclusão de Curso de Mestrado do Programa de Pós-graduação em Engenharia Biomédica - CTG da Universidade Federal de Pernambuco, no dia 25 de março de 2025.

GABRIEL MIRANDA DE SOUZA

Mestrando(a)

À minha família e amigos, meu alicerce nessa caminhada.

AGRADECIMENTOS

Gostaria de agradecer à minha família, que sempre me apoia nas minhas decisões.

Gostaria de agradecer aos meus amigos, que sempre estão comigo até nos momentos mais difíceis.

Gostaria de agradecer ao meu orientador e co-orientadora, pelos muitos anos juntos nessa caminhada de trocas de conhecimento.

Gostaria de agradecer ao Departamento de Engenharia Biomédica, a todos os professores do Departamento e aos demais funcionários.

Agradeço também a todos com quem cruzei nessa jornada e me fizeram ser quem sou.

RESUMO

A Doença de Alzheimer (DA) é um distúrbio neurodegenerativo progressivo que afeta milhões de pessoas em todo o mundo. Os métodos tradicionais de diagnóstico dependem de técnicas de neuroimagem, como PET-CT, que são caras e de difícil acesso, especialmente em regiões com poucos recursos. O eletroencefalograma (EEG) surge como uma alternativa de baixo custo e portátil para auxiliar no diagnóstico da DA. Este estudo propõe uma abordagem que transforma sinais de EEG em imagens por meio do método ASTERI, permitindo a classificação de DA, demência frontotemporal (DFT) e indivíduos saudáveis utilizando modelos de deep learning. Foi utilizada uma base de dados composta por EEGs de 86 indivíduos (36 DA, 23 DFT e 29 saudáveis), gerando um total de 141,063 imagens. Diferentes modelos de *deep learning* foram avaliados, com o modelo ConvNeXt alcançando o melhor desempenho, atingindo acurácia e F1-score de até 0,9964 com intervalo de confiança de 95%. Os métodos de Mapas de Saliência e Gradientes Integrados foram aplicados para identificar as regiões das imagens mais relevantes no processo de classificação. Ao cruzar esses resultados com os Mapas de Influência propostos, foi possível determinar os canais de EEG que mais contribuíram para a decisão do modelo. Os resultados sugerem que os canais das regiões occipital e frontal desempenham um papel fundamental na diferenciação entre doenças neurodegenerativas. Além disso, foram observadas diferenças de contribuição entre os hemisférios cerebrais, corroborando estudos anteriores que indicam uma deterioração mais rápida do hemisfério esquerdo em pacientes com DA. Este estudo contribui para o desenvolvimento de ferramentas diagnósticas baseadas em inteligência artificial, reforçando o potencial do EEG como uma modalidade acessível e não invasiva para a avaliação de doenças neurológicas.

Palavras-chaves: *Deep Learning*. Doença de Alzheimer. Eletroencefalografia. ASTERI. Inteligência Artificial Explicável.

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that affects millions of people worldwide. Traditional diagnostic methods rely on neuroimaging techniques, such as PET-CT, which are expensive and not widely available, especially in resource-limited settings. Electroencephalography (EEG) presents a low-cost and portable alternative for supporting AD diagnosis. This study proposes an approach that transforms EEG signals into images using the ASTERI method, enabling deep learning models to classify AD, frontotemporal dementia (FTD), and healthy controls. A dataset comprising EEG recordings from 86 subjects (36 AD, 23 FTD, and 29 controls) was used, generating 141,063 images. Several deep learning models were evaluated, with ConvNeXt achieving the highest performance, reaching an accuracy and F1-score of up to 0.9964 with a 95% confidence interval. The methods of Saliency Maps and Integrated Gradients were applied to identify the most relevant image regions in the classification process. By cross-referencing these results with the proposed Influence Maps, it was possible to determine the EEG channels that contributed the most to the model's decision. The results suggest that EEG channels from occipital and frontal regions play a crucial role in distinguishing neurodegenerative diseases. Additionally, differences in hemispheric contributions were observed, aligning with existing evidence that the left hemisphere undergoes faster neurodegeneration in AD patients. This research contributes to the development of AI-driven diagnostic tools, reinforcing EEG as a promising modality for non-invasive and accessible neurological disease assessment.

Keywords: Deep Learning. Alzheimer's disease. Electroencephalography. ASTERI. Explainable AI.

LISTA DE FIGURAS

Figura 1 – Ilustração do sistema internacional 10-20.	21
Figura 2 – Comparação entre os neurônios biológicos e artificiais.	23
Figura 3 – Funções de ativação frequentemente utilizadas em redes neurais.	25
Figura 4 – Operação de convolução ilustrada.	28
Figura 5 – Operação de uma camada de <i>pooling</i> ilustrada.	29
Figura 6 – Diagrama de fluxo representando as etapas da metodologia do estudo. . . .	41
Figura 7 – Exemplos das imagens geradas a partir do método ASTERI.	42
Figura 8 – Exemplo das imagens geradas em diferentes estágios. A subfigura (a) ilustra a imagem original com todos os canais ativos; A subfigura (b) representa a imagem perturbada com um canal zerado; A subfigura (c) mostra o Mapa de influência gerado a partir da Imagem original (a) e da imagem perturbada (b).	45
Figura 9 – Gráficos Boxplots da entropia obtida na classificação do conjunto de teste dos modelos.	51
Figura 10 – Progressão da acurácia no conjunto de validação ao longo de cada época para os diferentes modelos treinados na base.	52
Figura 11 – Progressão do F1 score no conjunto de validação ao longo de cada época para os diferentes modelos treinados na base.	53
Figura 12 – Mapa de Influência gerado para cada canal por meio do método de perturbação descrito em Métodos. Nesta versão, a escala na legenda foi fixada para todos os canais, destacando os canais com maior influência com base em suas cores mais intensas.	54
Figura 13 – Mapas de Saliência gerados para três imagens testes de cada classe. . . .	55
Figura 14 – Gradientes Integrados gerados para três imagens testes de cada classe. . . .	55
Figura 15 – Distribuição da média de contribuições dos canais nas regiões dos crânio. .	57

LISTA DE TABELAS

Tabela 1 – Métricas obtidas para diferentes modelos após o treinamento da base. Os valores são apresentados com seus respectivos intervalos de confiança (CI) de 95%. Observa-se uma melhora significativa nas métricas após o descongelamento dos modelos.	50
Tabela 2 – Entropia obtida com os diferentes modelos testados.	50
Tabela 3 – Sensibilidade e especificidade divididos por classe obtidos para os diferentes modelos após o treinamento da base.	52
Tabela 4 – Contribuições de cada canal na classificação das imagens geradas, estimadas a partir dos mapas de saliência e gradientes integrados.	56
Tabela 5 – Contribuições de cada canal na classificação das imagens geradas, estimadas a partir dos mapas de saliência e gradientes integrados.	60

LISTA DE ABREVIATURAS E SIGLAS

3D	Tridimensional
BCI	Interface Cérebro-Computador
CCL	Comprometimento Cognitivo Leve
CNN	Convolutional Neural Network
CT	Tomografia Computadorizada
CV	Visão Computacional (<i>Computer Vision</i>)
DA	Doença de Alzheimer
DFT	Demência Frontotemporal
ECG	Eletrocardiograma
EEG	Eletroencefalograma
ETS	Ensemble Temperature Scaling
GAN	Generative Adversarial Network
GR	Taxa de Crescimento
IA	Inteligência Artificial
KNN	K-Nearest Neighbors
LCR	Líquido Cefalorraquidiano
MAE	Masked Autoencoder
MEEM	Mini-Exame do Estado Mental
NLP	Processamento de Linguagem Natural (<i>Natural Language Processing</i>)
PET-CT	Tomografia Computadorizada por Emissão de Pósitrons
RM	Ressonância Magnética
SVM	Support Vector Machine
ViT	Vision Transformer
XAI	Inteligência Artificial Explicável

LISTA DE SÍMBOLOS

x	Vetor de entrada de uma rede neural
$\hat{y}$	Saída prevista de uma rede neural
y	Saída que representa a classe verdadeira do vetor x
$\Phi(.)$	Função de Ativação
w	Pesos de uma rede neural
b	Bias
$\in$	Pertence
∞	Infinito
∂	Derivada parcial
σ	Função Sigmoid
$L(.)$	Função de Perda
$J(.)$	Função Custo
∇f	Gradiente de f
α	Taxa de Aprendizagem
I	Imagem
Σ	Somatório
$\int$	Integral

SUMÁRIO

1	INTRODUÇÃO	15
1.1	CONTEXTUALIZAÇÃO	15
1.2	OBJETIVOS	16
1.3	ORGANIZAÇÃO DO TRABALHO	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	DOENÇA DE ALZHEIMER	19
2.2	ELETOENCEFALOGRAMA	20
2.3	INTELIGÊNCIA ARTIFICIAL NA SAÚDE	21
2.4	REDES NEURAIS ARTIFICIAIS	22
2.4.1	Deep Learning	26
2.4.1.1	<i>Convolutional Neural Networks</i>	26
2.4.1.2	<i>Outras redes de Deep Learning</i>	29
2.4.1.2.1	Transformers	29
2.5	MÉTODOS DE CALIBRAÇÃO	30
2.5.1	Ensemble Temperature Scaling	31
2.6	ASTERI	32
2.7	INTELIGÊNCIA ARTIFICIAL EXPLICÁVEL	33
2.8	MAPAS DE SALIÊNCIA	33
2.9	GRADIENTES INTEGRADOS	35
3	TRABALHOS RELACIONADOS	36
3.1	REPRESENTAÇÃO DE SINAIS EM IMAGENS	36
3.2	UTILIZAÇÃO DE EEG E IA NO DIAGNÓSTICO DE ALZHEIMER	38
4	METODOLOGIA	41
4.1	BASE DE DADOS	42
4.2	MAPAS DE INFLUÊNCIA	44
4.3	TREINAMENTO	45
4.4	MÉTRICAS DE AVALIAÇÃO	46
4.5	CONTRIBUIÇÕES DOS CANAIS	47
5	RESULTADOS E DISCUSSÕES	49
6	CONCLUSÃO	58

6.1	DIFICULDADES ENCONTRADAS	59
6.2	CONTRIBUIÇÕES	59
6.3	TRABALHOS FUTUROS	61
	REFERÊNCIAS	62
	APÊNDICE A – ARTIGO DE REVISÃO DE ESCOPO	68

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

A Doença de Alzheimer (DA) e outras demências afetam cerca de 55 milhões de indivíduos globalmente, representando um desafio significativo e crescente para a saúde pública. Ainda é estimado que esse número triplique até o ano de 2050, devido ao envelhecimento da população (NICHOLS et al., 2022). Diagnosticar essas demências, especialmente a doença de Alzheimer, continua sendo um processo complexo.

As práticas de diagnóstico atuais para a Doença de Alzheimer baseiam-se principalmente na identificação do declínio progressivo da memória por meio de exames clínicos e avaliações neuropsicológicas (LANE; HARDY; SCHOTT, 2018; KNOPMAN et al., 2021). Antigamente, o único teste diagnóstico definitivo era a biópsia cerebral. No entanto, atualmente, outros tipos de testes, como os baseados em neuroimagem, são amplamente utilizados para auxiliar no diagnóstico da DA (PORSTEINSSON et al., 2021). Sabe-se que a DA se desenvolve de forma silenciosa ao longo de anos, antes de apresentar os primeiros sintomas clínicos de declínio cognitivo, até progredir para a demência (LISS et al., 2021). Dessa forma, aprimorar a precisão diagnóstica para a Doença de Alzheimer e outras demências, especialmente em estágios iniciais, continua sendo uma área crítica de pesquisa.

Os exames de imagem são os testes diagnósticos que mais acompanham o diagnóstico da DA. Porém esses testes não são plenamente disponíveis, principalmente em países de menor renda (LANE; HARDY; SCHOTT, 2018; PORSTEINSSON et al., 2021). Nesse contexto, exames de Eletroencefalograma (EEG) oferecem diversas vantagens em relação às técnicas tradicionais de imagem, como Ressonância Magnética (MRI), Tomografia Computadorizada (CT) e Tomografia Computadorizada por Emissão de Pósitrons (PET-CT). Equipamentos de EEG são significativamente mais baratos e portáteis, tornando viável levar essas ferramentas diagnósticas para áreas remotas onde instalações de MRI ou PET-CT não estão disponíveis (ORBAN et al., 2022). Essa portabilidade pode beneficiar populações em regiões desfavorecidas, garantindo que mais indivíduos que necessitam de um diagnóstico possam recebê-lo.

Nos últimos anos, os avanços tecnológicos possibilitaram o crescimento da inteligência artificial (IA) na área da saúde, promovendo transformações significativas no apoio a diagnósticos, decisões clínicas e descobertas terapêuticas (SECINARO et al., 2021). Embora esses modelos, especialmente os baseados em *deep learning*, tenham demonstrado alto desempenho

em tarefas como a classificação de imagens médicas, ainda enfrentam desafios importantes, como a necessidade de bases de dados robustas, a presença de vieses, a baixa interpretabilidade dos resultados e a dificuldade de integrar múltiplos contextos clínicos (YU; BEAM; KOHANE, 2018). Além disso, questões éticas e legais precisam ser cuidadosamente consideradas para garantir uma implementação segura e transparente (NAIK et al., 2022). Assim, a aplicação da IA na saúde apresenta grande potencial, mas requer cautela para alcançar um equilíbrio entre eficácia, segurança e responsabilidade.

Avanços recentes em *deep learning* têm impulsionado significativamente seu uso na área da saúde, especialmente no auxílio ao diagnóstico médico (MISHRA; DASH; JENA, 2021; AHMAD et al., 2023). Modelos de *deep learning* têm demonstrado desempenho excepcional em tarefas como detecção precoce de doenças, segmentação de imagens médicas, análise preditiva e triagem automatizada (CARDOSO et al., 2022; TOBON et al., 2022; KAUL; RAJU; TRIPATHY, 2022). Além disso, a capacidade desses modelos de processar grandes volumes de dados de forma rápida e eficiente permite a criação de ferramentas de suporte à decisão médica, que são fundamentais para otimizar os fluxos de trabalho em ambientes hospitalares e melhorar os desfechos clínicos.

Esses avanços também têm fomentado o desenvolvimento de métodos que utilizam sinais fisiológicos para a geração de representações visuais, ampliando o alcance da IA para além de imagens médicas convencionais. Por exemplo, a síntese de imagens a partir de sinais, como eletroencefalogramas, tem mostrado grande potencial tanto no diagnóstico quanto em aplicações em tecnologias de interface cérebro-computador (BCI) (KO; YANG, 2022; SINGH et al., 2023). A representação de sinais em imagens oferece uma nova dimensão de análise, permitindo a aplicação de modelos avançados de classificação e segmentação originalmente desenvolvidos para dados visuais. Essas abordagens abrem novas possibilidades para o uso de IA em contextos multidisciplinares, incluindo o desenvolvimento de ferramentas mais robustas e personalizadas para diagnósticos médicos e tratamentos direcionados (JOSE et al., 2021; TAWHID et al., 2021).

1.2 OBJETIVOS

Este trabalho tem como objetivo apresentar uma abordagem inovadora para melhorar a precisão diagnóstica e precoce da Doença de Alzheimer por meio da integração de sinais de EEG e métodos avançados de interpretabilidade de modelos de *deep learning*. A metodologia

proposta consiste em transformar sinais de EEG em representações de imagem, que serão analisadas por diferentes modelos de classificação baseados em *deep learning*. Após a avaliação dos modelos no conjunto de teste, aquele que apresentar o melhor desempenho nas métricas será selecionado para a segunda etapa da metodologia.

Na segunda etapa, serão identificadas as regiões das representações de imagem mais influentes para a classificação a partir de métodos de interpretabilidade. Essas regiões serão então relacionadas aos canais originais do EEG, a fim de localizar canais específicos que contribuem significativamente para o diagnóstico da Doença de Alzheimer. Essa abordagem não apenas aprimora a compreensão da distribuição espacial da atividade neural associada à DA, como também potencialmente identifica regiões ou redes cerebrais relevantes para fins diagnósticos.

Levanta-se a hipótese de que certos canais do EEG, potencialmente correspondentes a regiões influentes nas imagens, possam apresentar padrões diferenciados que estejam associados à doença de Alzheimer. Tais padrões têm o potencial de contribuir como possíveis biomarcadores na detecção precoce e no acompanhamento da progressão da patologia. A partir dessa abordagem integrada entre dados de EEG e análise baseada em imagens, busca-se explorar possíveis correlações com os padrões de ativação neural relacionados à DA, o que pode oferecer subsídios para intervenções mais direcionadas e estratégias terapêuticas mais individualizadas.

Este projeto tem os seguintes objetivos específicos ou metas:

- Implementar e aplicar o método ASTERI para transformar sinais de EEG em representações de imagem adequadas para análise por modelos de classificação baseados em *deep learning*;
- Selecionar e treinar diferentes arquiteturas de *deep learning* para classificação das imagens derivadas de sinais de EEG;
- Avaliar e comparar o desempenho dos modelos de classificação com base em métricas padronizadas, selecionando o modelo mais robusto e eficaz para diagnóstico da Doença de Alzheimer;
- Propor e implementar uma abordagem interpretável para identificar as regiões das imagens mais influentes para a classificação;
- Realizar uma análise de rastreamento dos dados de EEG originais para localizar canais específicos que contribuem significativamente para o diagnóstico da Doença de Alzheimer;

- Investigar padrões distintos em canais críticos do EEG que possam servir como biomarcadores para a detecção precoce e o monitoramento da progressão da Doença de Alzheimer.

1.3 ORGANIZAÇÃO DO TRABALHO

O presente trabalho está organizado na seguinte forma que se segue: o capítulo 2 trata da fundamentação teórica, apresentando conceitos e definições pertinentes para o pleno entendimento do trabalho. O capítulo 3 sumariza os trabalhos relacionados ao tema presentes no estado da arte no momento da escrita do presente trabalho. O capítulo 4 detalha a metodologia adotada para construção do trabalho. O capítulo 5 traz os resultados obtidos, juntamente com as discussões cabíveis e implicações dos resultados. Por fim, o capítulo 6 encerra o trabalho trazendo as considerações finais acerca dos achados, assim como as perspectivas para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 DOENÇA DE ALZHEIMER

A Doença de Alzheimer (DA) é um transtorno neurodegenerativo progressivo que resulta em perda de memória, dificuldade de raciocínio e mudanças no comportamento. É causada pelo acúmulo de proteínas específicas no cérebro, principalmente a beta-amiloide e tau (KNOPMAN et al., 2021; LANE; HARDY; SCHOTT, 2018). Os sintomas mais recorrentes incluem, além da perda de memória, dificuldade na comunicação verbal e deterioração cognitiva. A gravidade dos sintomas pode variar desde um comprometimento cognitivo leve (CCL) até um quadro de demência, impactando significativamente a independência do paciente (KNOPMAN et al., 2021). O declínio cognitivo nem sempre é causado pela doença de Alzheimer. Outros fatores, como depressão, uso inadequado de medicamentos, apneia do sono, deficiências vitamínicas ou lesões cerebrais, também podem levar a comprometimento cognitivo leve (CCL) (KNOPMAN et al., 2021; LANE; HARDY; SCHOTT, 2018). Embora a maioria dos casos não seja hereditária, fatores genéticos podem influenciar o desenvolvimento da doença (KNOPMAN et al., 2021).

O diagnóstico da DA baseia-se principalmente em avaliações clínicas, incluindo entrevistas com pacientes, familiares e cuidadores, avaliações cognitivas e exames físicos (LANE; HARDY; SCHOTT, 2018; KNOPMAN et al., 2021). Testes de sangue podem ser utilizados para excluir outras condições que também possam ter um comprometimento cognitivo, como deficiências vitamínicas, disfunções da tireoide ou infecções (LANE; HARDY; SCHOTT, 2018). O diagnóstico precoce tem avançado por meio de testes modernos, como a análise do líquido cefalorraquidiano (LCR) e técnicas de imagem que auxiliam na identificação de marcadores biológicos característicos da doença. Técnicas avançadas de imagem, como os exames de Tomografia Computadorizada por Emissão de Pósitrons (PET-CT) de amiloide e tau, permitem que profissionais visualizem os depósitos dessas proteínas relacionadas à DA no cérebro (PIETRZAK et al., 2018; LANE; HARDY; SCHOTT, 2018; KIM et al., 2022; KNOPMAN et al., 2021). Esses métodos oferecem potencial para detecção precoce da doença e seu monitoramento, porém ainda não estão amplamente disponíveis em muitos países devido aos altos custos (LANE; HARDY; SCHOTT, 2018). Inovações em biomarcadores baseados em sangue, embora ainda em desenvolvimento, podem tornar os testes diagnósticos mais acessíveis (KNOPMAN et al., 2021).

Em contraste com as técnicas complexas e de alto custo como PET-CT, há uma alternativa de baixo custo associado: Eletroencefalograma (EEG).

2.2 ELETROENCEFALOGRAMA

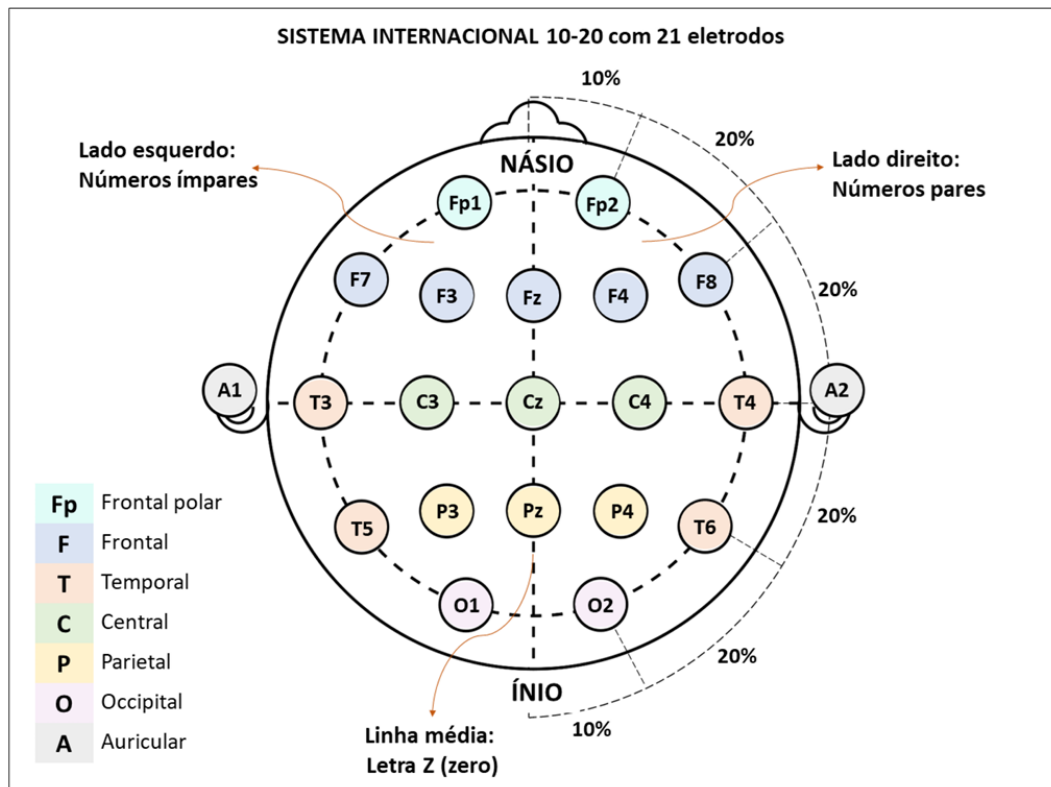
O eletroencefalograma (EEG) é uma técnica não invasiva que utiliza eletrodos colocados no couro cabeludo para detectar sinais elétricos gerados pelo cérebro. Esses sinais são produzidos pela atividade de grandes grupos de neurônios, especialmente os neurônios piramidais no córtex cerebral. Quando os neurônios são ativados, eles criam pequenas correntes elétricas devido ao movimento de íons carregados, como sódio (Na^+), potássio (K^+), cálcio (Ca^{++}) e cloreto (Cl^-) (TEPLAN, 2002). Os sinais captados são amplificados e exibidos em um eletroencefalograma, que mostra os padrões elétricos do cérebro.

Os eletrodos que são colocados no couro cabeludo durante o exame de EEG são comumente dispostos utilizando o sistema 10-20, que posiciona 21 eletrodos com base em porcentagens de separação em relação a quatro pontos de referência no crânio. As letras indicam a área cerebral, enquanto números pares representam o lado direito e ímpares, o esquerdo; eletrodos na linha média recebem "z" (GOMES, 2023; HOMAN; HERMAN; PURDY, 1987). A Figura 1 ilustra a disposição dos eletrodos e quais regiões do cérebro estão associadas a eles.

Na doença de Alzheimer, o EEG pode auxiliar a identificar mudanças na atividade cerebral causadas pela neurodegeneração, especialmente no córtex cerebral. O EEG em repouso, que mede a atividade cerebral sem a necessidade de tarefas cognitivas, é frequentemente utilizado em pacientes com DA, pois tem menos chances de causar desconforto ou dificuldade em realizar as tarefas (KIM et al., 2022; CASSANI et al., 2018). Pacientes com Alzheimer apresentam um aumento na atividade neuronal nas bandas de baixa frequência e uma diminuição na atividade nas bandas de alta frequência (PUTHANKATTIL et al., 2010; KIM et al., 2022; GOUW et al., 2017). Essas alterações podem ser um indício de que a disfunção sináptica ocorre desde os estágios iniciais da DA (KIM et al., 2022; GOUW et al., 2017).

O EEG tem como vantagem a alta resolução temporal, captando sinais em escalas de milissegundos, o que o diferencia de técnicas como a ressonância magnética (RM) e a tomografia por emissão de pósitrons (PET). No entanto, também possui limitações importantes: a atenuação dos sinais elétricos ao atravessarem camadas como o couro cabeludo e o crânio pode comprometer a precisão, e o processamento dos dados está sujeito a vieses e inconsistências. Além disso, os padrões cerebrais detectados pelo EEG podem se sobrepor aos de outras doenças neurodegenerativas, dificultando a detecção precoce da DA (KIM et al., 2022). Esses desafios reforçam a necessidade de aprimorar os métodos de análise e interpretação dos sinais de EEG, a fim de tornar essa ferramenta mais eficaz no diagnóstico e monitoramento

Figura 1 – Ilustração do sistema internacional 10-20.



Fonte: (GOMES, 2023)

da doença.

2.3 INTELIGÊNCIA ARTIFICIAL NA SAÚDE

Nos últimos anos o avanço tecnológico possibilitou um drástico aumento no poder computacional e um vasto poder de armazenamento e utilização de dados digitais de diferentes áreas, incluindo dados na saúde. Com isso, atualmente a inteligência artificial (IA) está revolucionando as práticas médicas, auxiliando diagnósticos, decisões clínicas, identificando possíveis novas drogas para tratamento de pacientes e revelando *insights* sobre dados médicos extensos (SECINARO et al., 2021).

O renascimento da prática médica assistida por IA é promissora, porém há diversos obstáculos que ainda precisam ser abordados com cautela. O alto desempenho, eficácia e confiabilidade de modelos baseados em *machine learning* dependem intrinsecamente da disponibilidade e uso de uma base de dados de treinamento robusta e de alta qualidade que represente a população fidedignamente (YU; BEAM; KOHANE, 2018). Vieses e ruídos inerentes a diferentes tipos de dados clínicos também podem comprometer a qualidade dos dados e prejudicar a capacidade

de generalização desses modelos. Embora os modelos de IA demonstrem desempenho superior em muitas das suas aplicações, comumente a interpretabilidade dos resultados é mínima e abstrata. Dissecar e entender a lógica por trás da tomada de decisão desses modelos ainda é um desafio, principalmente quando se trata de modelos de *deep learning* cuja aplicação não é direcionada a imagens (YU; BEAM; KOHANE, 2018).

Modelos de IA tornaram-se ferramentas robustas na resolução de problemas e tarefas específicas, como classificação de imagens médicas, porém o diagnóstico clínico e a prescrição de tratamento frequentemente possuem outros parâmetros a serem considerados, como as preferências do paciente, valores, histórico médico, condições contextuais que vão além de um simples problema de classificação no qual as redes de *deep learning* se destacam (YU; BEAM; KOHANE, 2018). Ainda é um desafio realizar essa integração de diferentes dados e contextos na alimentação de uma grande rede multimodal.

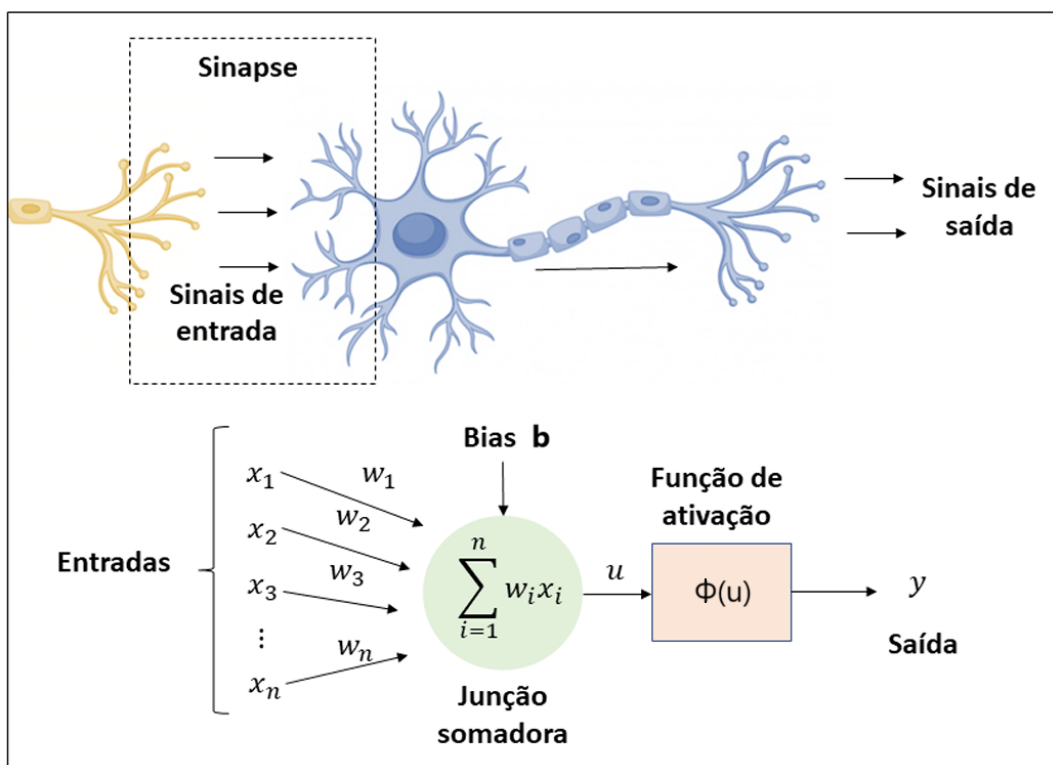
A utilização de IA na saúde também abre discussão para as questões éticas e legais quanto ao seu uso. Quatro principais problemas éticos podem ser mencionados: consentimento para utilização dos dados; segurança e transparência; a equidade dos algoritmos e seus vieses; privacidade dos dados (NAIK et al., 2022). Abordar e discutir esses desafios é essencial na construção de políticas públicas e internas, considerando os aspectos morais e legais do tema. A abstração dos modelos de IA também cria obstáculos acerca da responsabilidade, dificultando a capacidade de rastrear ações até seus responsáveis, quando necessário. Designar aos modelos de IA a responsabilidade pelas tomadas de decisão dificulta a prestação de contas. A integração da IA na tomada de decisões clínicas ao mesmo tempo que facilita, também aumenta a complexidade desse processo, exigindo um bom equilíbrio entre transparência e eficácia (NAIK et al., 2022).

2.4 REDES NEURAIIS ARTIFICIAIS

As redes neurais, inspiradas nos padrões básicos do funcionamento biológico dos neurônios, são modelos computacionais projetados para simular a maneira como os neurônios biológicos processam informações (RUSSELL; NORVIG, 2010). Assim como o cérebro é composto por neurônios interconectados que transmitem sinais eletroquímicos para processar informações sensoriais e possibilitar o pensamento, as redes neurais consistem em unidades ou "neurônios" conectados por links que propagam sinais por toda a rede. Cada unidade em uma rede neural recebe entradas, realiza uma soma ponderada e aplica uma função de ativação para

produzir uma saída, de forma semelhante ao modo como os neurônios biológicos disparam em resposta a estímulos (RUSSELL; NORVIG, 2010). Esse *design* permite que as redes neurais realizem tarefas como aprender com exemplos, classificação e reconhecimento de padrões. A ligação entre as redes neurais e o cérebro também é presente na capacidade do cérebro de se adaptar e modificar sua conectividade, uma característica que é refletida na capacidade de aprendizado das redes neurais. Essa semelhança entre os neurônios biológicos e artificiais está ilustrada na Figura 2.

Figura 2 – Comparação entre os neurônios biológicos e artificiais.



Fonte: Adaptado de Gomes (2023)

Considere um vetor de entrada x de atributos, onde $x = [x_1, x_2, \dots, x_n] \in R^n$ e um vetor de saída y , onde $y = [y_1, y_2, \dots, y_k] \in \{0, 1\}^k$. O vetor de saída y representa a classe associada ao vetor de atributos x . Por exemplo, se houver 3 classes e o vetor de atributos pertencer à classe do meio, o vetor y assume a forma $y = [0, 1, 0]$, ou seja, apenas um dos valores de y pode ser igual a 1, enquanto os demais são iguais a 0. Ao passar o vetor x por uma camada de uma rede neural cada valor de atributo x_i é multiplicado pelo seu respectivo peso associado w_i . Os resultados dessas multiplicações são somados, e adiciona-se um termo de *bias* b . Por fim, o resultado é passado por uma função de ativação $\Phi(\cdot)$ (RUSSELL; NORVIG, 2009). A equação que representa esse processo de passar um vetor de atributos pela camada da rede neural é

formalizado na Equação 2.1 abaixo:

$$\hat{y} = \Phi \left(\sum_{i=1}^n (w_i x_i) + b \right) \quad (2.1)$$

A função de ativação $\Phi(\cdot)$ pode assumir diversas formas, sendo que, comumente, utilizam-se funções de ativação não lineares devido aos seus benefícios. Funções de ativação não lineares permitem transformar os dados de entrada em representações onde pontos originalmente não linearmente separáveis tornam-se linearmente separáveis. Isso possibilita que as camadas da rede aprendam novas características, tornando classificadores lineares simples mais flexíveis. Sem funções de ativação não lineares, uma rede neural se comportaria como um modelo linear, independentemente de sua profundidade, e seria incapaz de lidar com padrões complexos. Ao introduzir a não linearidade, as funções de ativação aumentam o poder e a flexibilidade do modelo, permitindo que ele resolva problemas que modelos puramente lineares não conseguem (AGGARWAL et al., 2018). Uma das funções de ativação mais utilizadas é a função sigmoide, definida na Equação 2.2. Outros tipos comuns de funções de ativação estão representadas na Figura 3.

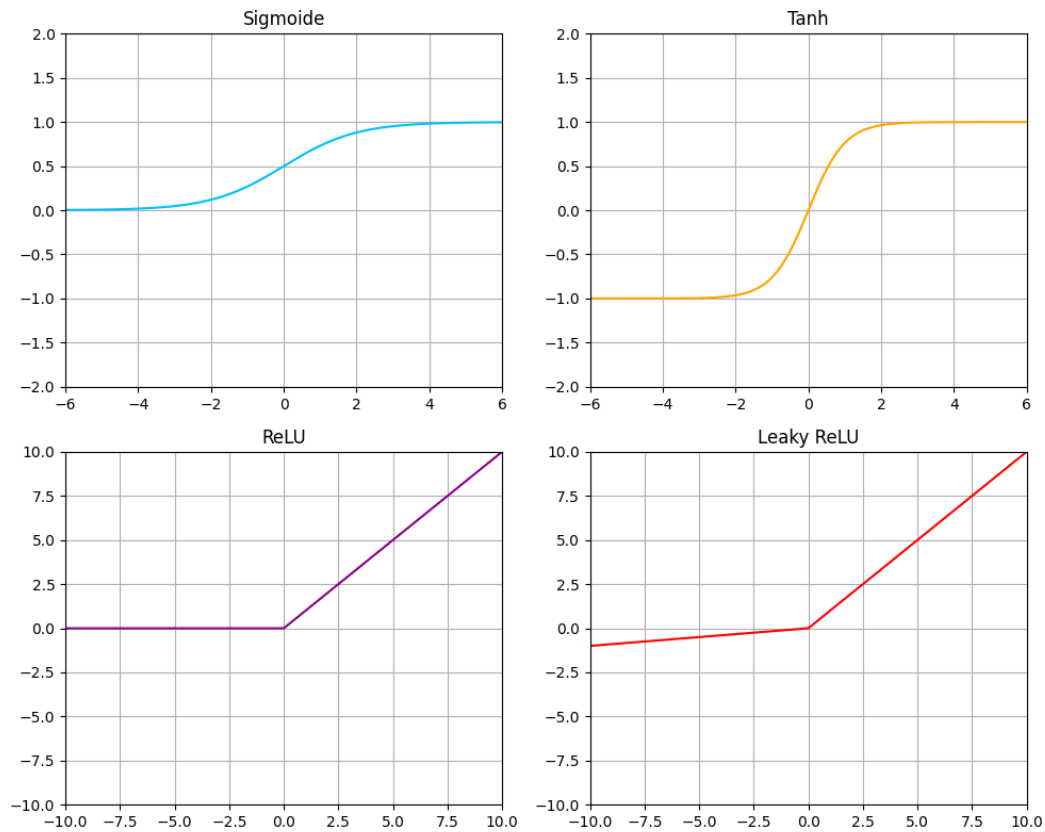
$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (2.2)$$

Na primeira iteração, os pesos w e o bias b da rede são inicializados aleatoriamente (NIELSEN, 2015). Para melhorar suas previsões, o modelo precisa aprender os valores de pesos e bias que maximizem os acertos da rede ao prever os valores de saída $\hat{y}$ em comparação com os valores reais de saída y . Para isso, os valores de saída preditos pela rede, $\hat{y}$, são comparados com os valores reais, y , por meio de uma função de perda. O objetivo da função de perda é calcular o quão distante os valores preditos estão dos valores reais (AGGARWAL et al., 2018; RUSSELL; NORVIG, 2009). Uma função de perda comum em problemas com mais de duas classes é a função *categorical cross-entropy*, definida a seguir:

$$L(\hat{y}, y) = - \sum_{k=1}^K y_k \log(\hat{y}_k) \quad (2.3)$$

onde K é a quantidade de classes. Nota-se que, quando $y_k = \hat{y}_k = 1$, a função de perda assume o valor $L = 0$, ou seja, quando o modelo acerta a classe do exemplo, a perda é zero. Quanto mais distante o valor de saída da rede $\hat{y}$ estiver do valor real y , maior será a perda. A função de perda calcula a distância entre y e $\hat{y}$ para apenas um exemplo de treinamento. Para calcular a distância para todos os m exemplos da base de dados, utiliza-se a função custo, definida como:

Figura 3 – Funções de ativação frequentemente utilizadas em redes neurais.



Fonte: o Autor (2025)

$$J(w, b) = \frac{1}{m} \sum_{i=1}^m L(\hat{y}^{(i)}, y^{(i)}) \quad (2.4)$$

onde $\hat{y}^{(i)}$ e $y^{(i)}$ são, respectivamente, a saída do modelo e a saída real do i -ésimo exemplo. Para melhorar as previsões da rede neural, é necessário otimizar os pesos w e o bias b a fim de minimizar a função custo J .

Para minimizar a função custo, utiliza-se o conceito de gradiente de uma função, ∇f . O gradiente de uma função é um vetor que fornece a direção de maior crescimento da função f , sendo calculado a partir das derivadas parciais da função em relação às suas variáveis (STEWART, 2012). Para minimizar a função custo J , devemos seguir a direção oposta ao gradiente de J , ou seja, $-\nabla f$. O processo que estima o gradiente da função custo em uma rede neural é conhecido como *backpropagation* (NIELSEN, 2015). Para atualizar os pesos w do modelo, utiliza-se a componente do gradiente na direção de w , que é dada pela derivada parcial de J em relação a w . De forma análoga, para atualizar o bias b , utiliza-se a componente do gradiente na direção de b (NIELSEN, 2015). As equações que regem a atualização dos pesos w e bias b são definidas a seguir:

$$w_i^{(t+1)} = w_i^{(t)} - \alpha \frac{\partial J(w, b)}{\partial w_i} \quad (2.5)$$

$$b^{(t+1)} = b^{(t)} - \alpha \frac{\partial J(w, b)}{\partial b} \quad (2.6)$$

onde $w_i^{(t+1)}$ é o valor atualizado do peso w_i , e $w_i^{(t)}$ é o valor do peso na iteração anterior. O parâmetro α , chamado de taxa de aprendizagem, define o tamanho dos passos dados na atualização dos pesos w e do bias b . Uma taxa de aprendizagem muito alta pode fazer com que J ultrapasse o mínimo, em vez de se aproximar dele, podendo até divergir. Por outro lado, uma taxa de aprendizagem muito baixa pode fazer com que o treinamento demore muito para convergir ao mínimo. Esse processo de atualização dos pesos e do bias da rede é conhecido como o algoritmo de gradiente descendente (AGGARWAL et al., 2018).

2.4.1 Deep Learning

Deep learning revolucionou a área de inteligência artificial ao permitir que computadores aprendam automaticamente representações a partir de dados brutos (LECUN; BENGIO; HINTON, 2015). Diferente das técnicas tradicionais de aprendizado de máquina, que exigem a engenharia manual de atributos, modelos de *deep learning* utilizam múltiplas camadas de transformações não lineares para extrair progressivamente atributos de entradas como imagens, textos e áudios (GOODFELLOW, 2016). Essa abordagem em camadas permite que as redes neurais capturem padrões e relações complexas em dados de alta dimensão, tornando *deep learning* altamente eficaz para tarefas como reconhecimento de imagens, transcrição de fala, processamento de linguagem natural e diagnósticos médicos (LECUN; BENGIO; HINTON, 2015; TOBON et al., 2022).

2.4.1.1 Convolutional Neural Networks

Uma das arquiteturas de *deep learning* mais utilizadas atualmente são as Redes Neurais Convolucionais (CNN). Essa arquitetura é composta, em geral, por três tipos principais de camadas: camadas convolucionais, camadas de *pooling* e camadas totalmente conectadas. Embora outros tipos de camadas possam estar presentes em uma CNN, essas três constituem a estrutura base desse tipo de arquitetura (CONG; ZHOU, 2022).

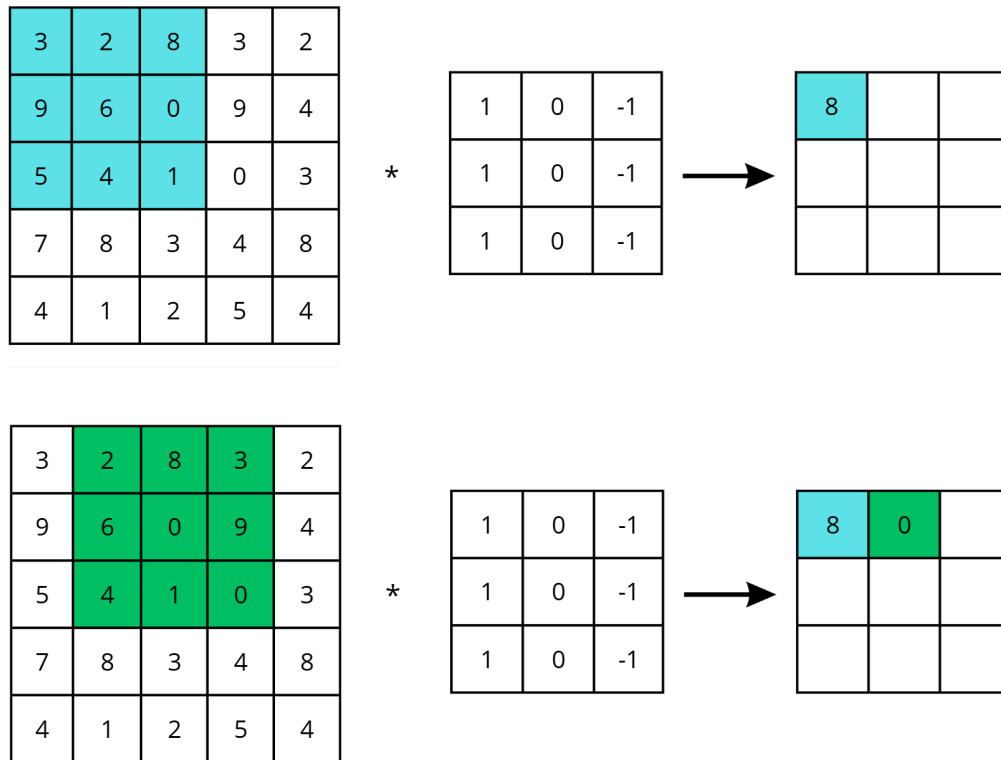
As camadas de convolução das CNN utilizam kernels aprendidos para extrair atributos locais dos dados de entrada, como imagens. Uma camada convolucional desliza uma janela de tamanho fixo, ou também chamado kernel, sobre a matriz de entrada, realizando uma multiplicação elemento a elemento e somando os resultados para produzir a matriz de saída (YAMASHITA et al., 2018). Esse processo permite que a rede identifique padrões espaciais, como bordas ou texturas (ZEILER; FERGUS, 2014). Cada kernel em uma camada de convolução se especializa na detecção de características específicas, contribuindo para uma compreensão mais profunda dos dados de entrada (YAMASHITA et al., 2018). A formulação matemática envolve a aplicação de uma função de ativação aos resultados da convolução, o que introduz não linearidade e ajuda a rede a modelar relações mais complexas (CONG; ZHOU, 2022). Ao agrupar várias camadas convolucionais, as redes neurais conseguem capturar características hierárquicas, desde bordas simples nas camadas iniciais até padrões complexos nas camadas mais profundas (ZEILER; FERGUS, 2014).

A Figura 4 ilustra as duas primeiras operações de convolução realizadas em uma imagem de entrada com dimensões (5,5) utilizando um kernel de dimensões (3,3). Na primeira operação, destacada em azul, os valores da coluna da esquerda da janela de entrada são multiplicados por 1, os valores da coluna do meio são multiplicados por 0, e os valores da coluna da direita são multiplicados por -1; os resultados são somados, e o valor final é atribuído ao primeiro pixel da matriz de saída. Na segunda operação, destacada em verde, o kernel desloca-se para a direita para calcular o valor do segundo pixel da matriz de saída. Esse processo é repetido sucessivamente até que o kernel percorra toda a matriz de entrada, realizando a operação de convolução para cada posição válida. Durante o treinamento de uma CNN, os valores no kernel, que correspondem aos pesos da rede neural, são aprendidos e ajustados automaticamente para otimizar o desempenho da rede.

As camadas de *pooling* em CNNs são essenciais para reduzir as dimensões espaciais dos dados de entrada, ao mesmo tempo que preserva informações cruciais. Elas realizam esse processo ao sintetizar regiões locais da entrada, geralmente por meio de operações como *max pooling*, que seleciona o maior valor de uma região para substituí-los, ou *average pooling*, que calcula a média dessa região. Esse processo contribui para a redução dos custos computacionais, melhora a eficiência estatística e aumenta a capacidade do modelo de generalizar, tornando-o mais invariante a pequenas transformações nos dados, como distorções, translações ou rotações (GOODFELLOW, 2016; LECUN; BENGIO; HINTON, 2015).

Além disso, a camada de *pooling* é útil para lidar com entradas de tamanhos variados,

Figura 4 – Operação de convolução ilustrada.

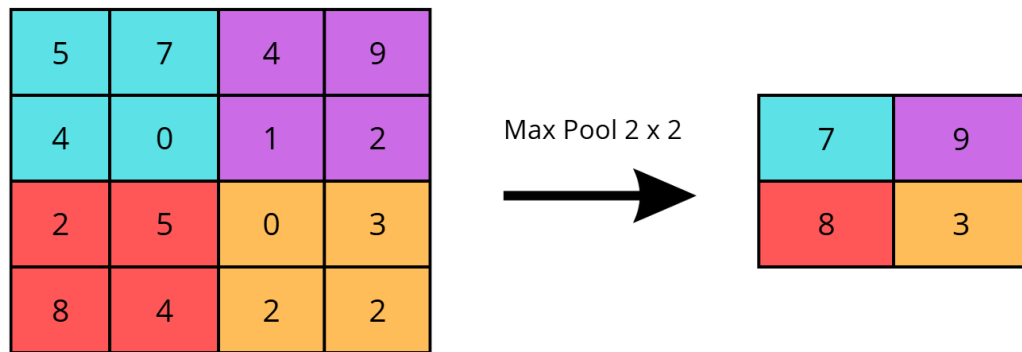


Fonte: o Autor (2025)

garantindo uma representação de tamanho fixo antes da etapa de classificação. No entanto, pode apresentar desafios para arquiteturas que dependem da reconstrução de detalhes finos, como *autoencoders* e modelos generativos, que exigem técnicas especializadas para reverter a operação de *pooling* (GOODFELLOW, 2016). A operação de *max pooling*, o tipo de *pooling* mais utilizado atualmente, está ilustrada na Figura 5, onde é possível observar que as diferentes regiões da matriz de entrada, representadas por cores distintas, resultam em um único valor na matriz de saída: o maior valor presente na região correspondente da matriz de entrada, justificando o nome *max pooling*.

Por fim, as camadas totalmente conectadas das CNNs constituem a porção final da rede, com o objetivo principal de transformar os dados multidimensionais das camadas anteriores em uma saída unidimensional, adequada para o estágio de classificação final (CONG; ZHOU, 2022).

Figura 5 – Operação de uma camada de *pooling* ilustrada.



Fonte: o Autor (2025)

2.4.1.2 Outras redes de Deep Learning

2.4.1.2.1 Transformers

Os *Transformers* são uma classe recente e robusta de modelos de *deep learning* que revolucionaram várias áreas, especialmente em processamento de linguagem natural (NLP), visão computacional (CV) e processamento de fala (LIN et al., 2022). Eles utilizam um mecanismo de autoatenção para processar e aprender eficientemente as relações dentro dos dados (KHAN et al., 2022). Ao contrário dos modelos recorrentes, que processam os dados de forma sequencial, os *Transformers* podem processar sequências inteiras simultaneamente, permitindo-lhes identificar dependências de longo alcance e melhorando sua escalabilidade (KHAN et al., 2022). O sucesso dos modelos pré-treinados baseados em *Transformers* em NLP abriu caminho para sua utilização em diversas aplicações, como reconhecimento de imagens, detecção de objetos e segmentação, consolidando os *Transformers* como uma arquitetura fundamental na pesquisa e nas aplicações modernas de IA (KHAN et al., 2022; LIN et al., 2022).

O modelo *Transformer* é composto por vários componentes essenciais que o tornam uma das arquiteturas mais robustas atualmente. Ele segue uma arquitetura de codificador-decodificador, onde tanto o codificador quanto o decodificador são formados por múltiplas camadas idênticas (KHAN et al., 2022). Cada bloco do codificador contém um mecanismo de autoatenção multi-cabeça, que permite ao modelo capturar relações entre todos os elementos de uma sequência, seguido por uma rede *feedforward* para transformação dos atributos. O decodificador possui camadas semelhantes, mas inclui um mecanismo adicional de atenção cruzada, permitindo que ele considere as saídas do codificador enquanto gera previsões (LIN et al., 2022). Conexões residuais e normalização de camadas são aplicadas após cada subcamada

para melhorar o fluxo do gradiente e a estabilidade (LIN et al., 2022). Como os *Transformers* não possuem recorrência nem convolução, eles dependem de codificações posicionais para preservar a ordem dos *tokens* de entrada (VASWANI et al., 2017). Essa arquitetura permite que o *Transformer* processe sequências de forma eficiente e em paralelo, tornando-o altamente eficaz para tarefas como tradução automática, geração de texto e aplicações em visão computacional (KHAN et al., 2022).

A autoatenção, introduzida pelo trabalho realizado por Vaswani et al. (2017), é um mecanismo onde cada elemento de uma sequência leva em conta todos os outros para gerar uma representação contextualizada (VASWANI et al., 2017). Esse processo ocorre projetando a entrada em vetores de *Query* (Q), *Key* (K) e *Value* (V), calculando pontuações de atenção usando o produto escalar de Q e K, normalizando-os com a função de ativação *softmax* e, em seguida, ponderando V (VASWANI et al., 2017). A autoatenção multi-cabeça expande esse conceito ao utilizar múltiplas cabeças de atenção em paralelo, cada uma com projeções aprendidas diferentes de Q, K e V, permitindo que o modelo capture diversas relações nos dados antes de concatenar os resultados em uma saída final (KHAN et al., 2022; LIN et al., 2022).

2.5 MÉTODOS DE CALIBRAÇÃO

Modelos de *deep learning* tornaram-se o padrão ouro em diversas aplicações. No entanto, alcançar alta acurácia nem sempre é suficiente em certas situações. Também é necessário garantir uma incerteza controlada para avaliar se as previsões podem ser confiadas (ZHANG; KAILKHURA; HAN, 2020). Em contextos como na área da saúde, previsões não confiáveis podem levar a resultados prejudiciais e perigosos.

A calibração é um processo que certifica-se que as previsões de um modelo estão de acordo com suas probabilidades verdadeiras, desempenhando um papel crucial na avaliação de confiabilidade. Modelos devidamente calibrados oferecem resultados mais confiáveis ao representar a incerteza de forma precisa. Sem calibração, as previsões podem se tornar excessivamente confiantes ou demasiadamente incertas, o que aumenta o risco de interpretações equivocadas ou decisões mal orientadas (MINDERER et al., 2021). Assim, a calibração é fundamental para refinar modelos de *deep learning*, melhorar sua usabilidade e fortalecer a confiança de suas previsões em aplicações do mundo real.

2.5.1 Ensemble Temperature Scaling

O *Ensemble Temperature Scaling* (ETS) é um método de calibração proposto por Zhang, Kailkhura e Han (2020) e projetado para aumentar a expressividade do *temperature scaling* (TS) padrão, preservando sua acurácia e eficiência (ZHANG; KAILKHURA; HAN, 2020). O *temperature scaling* tradicional é um método paramétrico simples que ajusta as probabilidades de previsão para melhorar a calibração sem alterar a acurácia da classificação. No entanto, sua expressividade é limitada, que é poder de representação suficiente para aproximar a função de calibração canônica, dado um número suficiente de dados de calibração (ZHANG; KAILKHURA; HAN, 2020). O ETS aborda essa limitação por meio de uma abordagem baseada em *ensemble*, onde múltiplos mapeamentos de calibração são combinados por meio de uma soma ponderada, aumentando assim a flexibilidade do modelo, garantindo que as probabilidades previstas permaneçam bem calibradas.

O método ETS é composto por três componentes. O primeiro componente é a função padrão de *temperature scaling*, que ajusta os escores de confiança elevando-os à potência de $1/t$, onde t é o parâmetro de temperatura. O segundo componente mantém as previsões não calibradas, garantindo estabilidade quando o classificador original já está bem calibrado. O terceiro componente consiste na forma extrema de *temperature scaling* com $t \mapsto \infty$, resultando em distribuições de probabilidade uniformes entre todas as classes. Ao combinar esses três componentes o método alcança um equilíbrio entre preservar a acurácia do modelo, suavizar previsões e ajustar dinamicamente os níveis de confiança com base nos dados de calibração.

Para determinar a temperatura t e os pesos w ideais no ETS, um problema de otimização convexa é resolvido usando um conjunto de dados de calibração. O objetivo é minimizar uma função de perda $R(\cdot)$ entre as previsões calibradas $T(z)$ e as saídas verdadeiras y . O problema de otimização convexa é definido na Equação 2.7 abaixo:

$$\min_{t, w} \sum_{i=1}^{n_c} R(w_1 T(z^{(i)}; t) + w_2 z^{(i)} + w_3 \frac{1}{L}, y^{(i)}) \quad (2.7)$$

sujeito às restrições $t > 0$; $1_{1 \times 3} w = 1$; $w > 0_{3 \times 1}$, onde L representa a quantidade de classes do problema.

Após a calibração das previsões, é possível estimar o intervalo de confiança das métricas de avaliação a partir do Erro Quadrático Médio (MSE) (ZHANG; KAILKHURA; HAN, 2020), definido

na Equação 2.8 abaixo:

$$MSE = \frac{1}{N} \sum_{i=1}^N (p(\hat{y}_i) - p(y_i))^2 \quad (2.8)$$

onde $p(\hat{y})$ representa a probabilidade estimada calibrada da predição do modelo para o i -ésimo exemplo. Com isso, o intervalo de confiança (CI) da métrica ME no conjunto de teste é dado por:

$$CI = ME \pm z \frac{\sqrt{MSE}}{\sqrt{N}} \quad (2.9)$$

onde N é a quantidade de exemplo no conjunto de teste, e z é o valor crítico correspondente ao nível de confiança desejado, por exemplo $z = 1,96$ para um intervalo de confiança de 95%.

2.6 ASTERI

ASTERI é um método de representações de sinais de EEG em imagens. O método utiliza a técnica de reconstrução de imagens por retroprojeção, originalmente desenvolvida para tomografia computadorizada (CT), para transformar janelas de sinais de EEG em imagens bidimensionais (GOMES; RODRIGUES; SANTOS, 2022). Sinais de EEG são dados estruturados. Essa transformação permite que modelos de *deep learning*, com arquiteturas especializadas para tarefas envolvendo dados não estruturados, processem e classifiquem sinais de EEG a partir de sua representação em formato de imagens (dados não estruturados).

O método ASTERI consiste em inicialmente realizar o janelamento dos sinais no tempo. Cada janela do sinal é então interpretada como um pseudo-sinograma, que é uma representação comumente usada na reconstrução de imagens de tomografia computadorizada. Em seguida esses pseudo-sinogramas são passados pelo algoritmo de retroprojeção, um método que reconstrói imagens ao somar as contribuições de projeções de diferentes ângulos para estimar a distribuição espacial original. Com isso, cada janela do sinal de EEG é reconstruído em uma imagem 2D ao somar essas projeções sobre o domínio espacial (GOMES; RODRIGUES; SANTOS, 2022).

2.7 INTELIGÊNCIA ARTIFICIAL EXPLICÁVEL

Redes de *deep learning* são comumente chamadas de *black-box* porque possuem muitas camadas ocultas, cujos pesos e operações não estão diretamente acessíveis aos usuários. O usuário final recebe apenas a previsão na saída da rede e não tem conhecimento sobre como a rede chegou àquela decisão. Essa falta de transparência pode gerar incerteza nas decisões (DAS; RAD, 2020).

Inteligência Artificial Explicável (XAI) refere-se a sistemas, métodos e técnicas de inteligência artificial que explicam de maneira clara e compreensível para humanos as suas previsões, decisões e ações. O objetivo da XAI é tornar o funcionamento da IA mais inteligível para os usuários, oferecendo *insights* sobre como esses modelos operam (GUNNING et al., 2019). Isso é especialmente importante em aplicações na área da saúde, onde os usuários precisam confiar e entender as previsões e os sistemas para tomar decisões orientadas por dados.

2.8 MAPAS DE SALIÊNCIA

O trabalho realizado por (SIMONYAN; VEDALDI; ZISSERMAN, 2014) introduziu dois métodos de visualização de classificação por modelos de *deep learning*. O primeiro, chamado de Visualização de Classe de Modelo (*Class Model Visualization*), tem como foco gerar uma imagem representativa de uma classe de interesse do modelo. Seja $S_c(I)$ a probabilidade (*score*) da classe c , calculada pela última camada de classificação do modelo para uma imagem I , o objetivo do método é encontrar uma imagem que maximize S_c através da equação definida abaixo:

$$\arg \max_I S_c(I) - \lambda \|I\|_2^2 \quad (2.10)$$

onde λ é um parâmetro de regularização. O processo de otimização é iniciado a partir de uma imagem zerada e refinado iterativamente utilizando *backpropagation* com relação à imagem de entrada (SIMONYAN; VEDALDI; ZISSERMAN, 2014). O resultado fornece uma compreensão visual das características que o modelo associa à classe, oferecendo *insights* sobre as representações aprendidas pelo modelo.

O segundo método apresentado por (SIMONYAN; VEDALDI; ZISSERMAN, 2014) é chamado de Visualização de Saliência de Classe de Imagens Específicas (*Image-Specific Class Saliency*

Visualisation), ou como é comumente chamado atualmente: mapas de saliência. Esse método identifica quais regiões de uma imagem de entrada são mais importantes para a previsão de uma classe específica. A probabilidade de classe $S_c(I)$ é uma função altamente não linear de I . Para lidar com essa não linearidade, é utilizada uma aproximação linear na vizinhança de I_0 , por meio de uma expansão de Taylor de primeira ordem, definida como:

$$S_c(I) \approx w^T I + b \quad (2.11)$$

onde w é a derivada de S_c em relação à imagem I no ponto I_0 :

$$w = \left. \frac{\partial S_c}{\partial I} \right|_{I_0} \quad (2.12)$$

A derivada w é calculada utilizando o método de *backpropagation*. O resultado dessa derivada é representado como um mapa de saliência, onde a magnitude da derivada de cada pixel indica sua influência na pontuação da classe. Esse mapa pode ser definido por:

$$M_{ij} = |w_{h(i,j)}| \quad (2.13)$$

ou, no caso de imagens com múltiplos canais de cor:

$$M_{ij} = \max_c |w_{h(i,j,c)}| \quad (2.14)$$

onde $h(i, j)$ é o índice de w correspondente ao pixel da imagem localizado na i —ésima linha e j —ésima coluna. A Equação 2.13 aplica-se a imagens em tons de cinza com apenas um canal de cor, enquanto a Equação 2.14 é utilizada para imagens com múltiplos canais de cor.

A utilização da derivada revela que os pixels mais destacados nos mapas de saliência são aqueles cuja mínima alteração resulta na maior variação na probabilidade de classe S_c . Devido a sua simplicidade, os mapas de saliência são um dos métodos de interpretabilidade de modelos mais utilizados na área de imagens médicas (HUFF; WEISMAN; JERAJ, 2021). Contudo, uma limitação considerável dos mapas de saliência é sua incapacidade de diferenciar se um pixel aumenta ou diminui a probabilidade de uma classe; eles apenas indicam quais pixels são mais sensíveis a essa probabilidade (HUFF; WEISMAN; JERAJ, 2021).

2.9 GRADIENTES INTEGRADOS

Gradientes integrados é um método para atribuir a previsão de modelos de *deep learning* às suas características de entrada. Proposto por (SUNDARARAJAN; TALY; YAN, 2017), o método calcula a importância de cada atributo de entrada ao agregar os gradientes da previsão da saída do modelo em relação aos atributos de entrada ao longo de um caminho linear entre uma entrada de referência (como, por exemplo, uma imagem zerada) e a entrada real. Para uma imagem de entrada, o método, então, representa a importância de cada pixel na previsão do modelo. O gradiente integrado ao longo da i -ésima dimensão para uma imagem de entrada x e uma imagem de referência x' é definido como:

$$IntegratedGrads_i(x) ::= (x_i - x'_i) \times \int_{\alpha=0}^1 \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x'} d\alpha \quad (2.15)$$

onde $\frac{\partial F(x)}{\partial x_i}$ é o gradiente de $F(x)$ ao longo da direção da i -ésima dimensão (SUNDARARAJAN; TALY; YAN, 2017). A Equação 2.15 pode ser aproximada pela soma:

$$IntegratedGrads_i^{approx}(x) ::= (x_i - x'_i) \times \sum_{k=1}^m \frac{\partial F(x' + \frac{k}{m} \times (x - x'))}{\partial x'} \times \frac{1}{m} \quad (2.16)$$

onde m é o número de passos na aproximação de Riemman da integral (SUNDARARAJAN; TALY; YAN, 2017). A implementação do método é computacionalmente eficiente, exigindo entre 20 e 300 cálculos de gradiente na prática para alcançar aproximações precisas.

O método de gradientes integrados satisfaz vários axiomas desejáveis para métodos de atribuição, aos quais outros métodos semelhantes falham em aderir. Por exemplo, os métodos de *Guided Backpropagation* e *Deconvolution Networks* violam o axioma de sensibilidade, enquanto *DeepLift* e *Layer-wise Relevance Propagation* violam o axioma de invariância de implementação (SUNDARARAJAN; TALY; YAN, 2017). Porém referências que desviam significativamente da distribuição de dados de entrada podem introduzir artefatos às atribuições.

3 TRABALHOS RELACIONADOS

Esse capítulo é dedicado à revisão de trabalhos relacionados ao presente estudo, publicados nos últimos cinco anos. A revisão está dividida em duas seções: trabalhos relacionados à representação de sinais em imagens e trabalhos relacionados à utilização de sinais de EEG no diagnóstico da Doença de Alzheimer.

3.1 REPRESENTAÇÃO DE SINAIS EM IMAGENS

O trabalho realizado por Naz et al. (NAZ et al., 2021) apresenta um método de representação de sinais de eletrocardiograma (ECG) em imagens para detecção de arritmias ventriculares. Inicialmente, os dados brutos do sinal de ECG, compostos por gravações de 24 horas com aproximadamente 127.356 pontos de dados por paciente, passam por um processo de normalização para garantir uniformidade. Essa normalização divide os dados do sinal em partes iguais e remove valores redundantes, resultando em 126.976 pontos de dados. Em seguida, o sinal é dividido em 124 segmentos, cada um contendo 1.024 pontos de dados. Cada segmento de 1024 pontos é transformado em uma imagem binária de 32×32 por meio de uma transformação matemática que pode ser considerada como o inverso da vetorização. Essas imagens binárias preservam os padrões críticos dos sinais de ECG enquanto eliminam a necessidade de etapas tradicionais de pré-processamento, como a detecção do complexo QRS, que costuma ser computacionalmente custosa e sujeita a erros. Essa abordagem gera 124 imagens binárias por paciente, criando um novo conjunto de dados com imagens classificadas como sinais normais e anormais. Em seguida, três modelos pré-treinados de CNN foram utilizados para extrair atributos das imagens. Os atributos extraídos por fim foram utilizados como entrada em redes de *Support Vector Machine* (SVM) e *K-Nearest Neighbors* (KNN). O método proposto obteve resultados expressivos a partir da concatenação de atributos extraídos dos modelos AlexNet, VGG19 e InceptionV3, e utilizando SVM Cúbico como classificador. A abordagem obteve uma acurácia de classificação de 96,6%, uma taxa de falsos negativos de 3,0%, sensibilidade de 97,12% e especificidade de 95,99%. Após isso, realizaram uma seleção de atributos baseada em entropia que refinou ainda mais os atributos concatenados, com resultados ótimos alcançados ao selecionar 5.120 atributos, que representava 50% do total original. Após essa seleção de atributos, o método obteve acurácia de 97,6%, taxa de falsos negativos de 2,3%, sensibilidade

de 98,2% e especificidade de 97,5%. O método atingiu maior acurácia que outros métodos existentes na época e também um menor tempo de processamento.

O trabalho de Singh et al. (SINGH et al., 2023) propõe um *framework* para a geração de imagens a partir de sinais de EEG. Os sinais são medidos a partir do couro cabeludo dos pacientes ao serem pedidos para visualizar certas classes de objetos e caracteres. O objetivo é reconstruir a imagem do que estava sendo visualizado a partir dos sinais de EEG gravados. O *framework* proposto tem uma abordagem em duas fases: extração de atributos e geração de imagens. Na primeira fase, é utilizado um método de aprendizagem contrastiva baseada em *triplet loss* para extrair atributos significativos dos sinais de EEG, garantindo que sinais de atividades cerebrais semelhantes sejam mapeados próximos no espaço de atributos, enquanto se mantém uma margem entre atividades diferentes. Na segunda fase, uma rede do tipo *Generative Adversarial Network* (GAN) condicional é utilizada para transformar esses atributos em imagens. Modificações na arquitetura da GAN, incluindo a função de perda *hinge loss* para treinamento estável, um bloco de *data augmentation* diferenciável para lidar com a escassez de dados, e termos de regularização de busca de modos, melhoram seu desempenho em pequenos conjuntos de dados e sintetizam imagens de alta qualidade. Essa abordagem contribui para o avanço da tecnologia de Interfaces Cérebro-Computador (BCI), com aplicações em neuro-reabilitação e comunicação para indivíduos com deficiências.

O trabalho realizado por Wang et al. (WANG et al., 2021) propõe uma abordagem para diagnóstico de falhas a partir do método EOSTI, que converte sinais de vibração no domínio do tempo em imagens RGB. Inicialmente, os dados brutos de sinal de vibração são processados utilizando o método EOSTI, que converte sinais unidimensionais no domínio do tempo em imagens 3D. Essas imagens são formadas pela extração de segmentos de sinal e mapeamento deles para os canais de vermelho, verde e azul de uma imagem RGB. Para aprimorar os atributos e reduzir o ruído, uma operação de erosão é aplicada às imagens. O conjunto de imagens gerado é então utilizado para treinar um modelo CNN baseado na rede AlexNet, que é otimizada para classificar as imagens em diagnóstico de falhas. Este método reduz a intervenção manual e melhora a acurácia no diagnóstico de falhas ao utilizar técnicas de *deep learning* para análise automatizada dos sinais de vibração. O método foi testado em duas bases de dados diferentes e alcança acurácias de até 99,67%, demonstrando resultados competitivos com o estado da arte, oferecendo uma solução inovadora para diagnóstico de falhas em aplicações industriais.

3.2 UTILIZAÇÃO DE EEG E IA NO DIAGNÓSTICO DE ALZHEIMER

O trabalho realizado por (ALSHARABI et al., 2022) estuda o desenvolvimento de um sistema de diagnóstico para doenças neurológicas utilizando sinais de EEG. A base de dados de EEG utilizada foi coletada de 86 participantes (grupos Controle, Alzheimer Leve e Alzheimer Moderado). Os dados são pré-processados a partir da remoção de ruídos e artefatos usando um filtro elíptico passa-banda (0,1–60 Hz). A extração de atributos utilizou a Transformada Wavelet Discreta (DWT) para decompor os sinais de EEG em sub-bandas de frequência (Delta, Teta, Alfa, Beta e Gama), também foi realizado o cálculo de atributos estatísticos, como potência logarítmica da banda, desvio padrão, variância, curtose, energia média, raiz quadrada média e norma. Os atributos extraídos foram classificados utilizando algoritmos de *machine learning*, incluindo SVM e *Random Forest*, com validação cruzada de 10 *folds*. Foram abordados cinco problemas de classificação, incluindo a distinção entre Alzheimer leve e moderado, controle versus Alzheimer leve e moderado, e controle versus DA leve versus DA moderado. O estudo demonstrou resultados promissores no diagnóstico da doença de Alzheimer, alcançando acurácias de classificação de até 99,98% em cenários binários e 98,1% no cenário de três classes. A análise comparativa revelou avanços significativos do sistema em relação a estudos anteriores na precisão diagnóstica. Apesar dos resultados promissores, o estudo destacou limitações, como o pequeno tamanho da base de dados e a dependência de um único método de extração de características, enfatizando a necessidade de trabalhos futuros que incluam conjuntos de dados maiores, inclusão de Alzheimer em estágio severo, abordagens baseadas em *deep learning* e aplicações em tempo real. A pesquisa ressalta o potencial da análise automatizada de EEG para o diagnóstico precoce e eficiente da doença de Alzheimer, reduzindo a dependência da avaliação manual.

O trabalho realizado por (MILTADOUS et al., 2023a) adotou uma metodologia para a classificação automática dos sinais de EEG que envolve quatro etapas: aquisição de dados, remoção de ruído do sinal, extração de características e classificação. A base de dados consiste em gravações de EEG de 88 participantes, incluindo indivíduos diagnosticados com DA, Demência Frontotemporal (FTD) e controles saudáveis (CN). Técnicas de pré-processamento, como o filtro Butterworth passa-banda (0,5-45 Hz), e métodos de rejeição de artefatos, como Artifact Subspace Reconstruction (ASR) e Análise de Componentes Independentes (ICA), foram aplicadas para limpar os dados. Os atributos extraídos incluem a Potência Relativa de Banda (RBP) e a Conectividade de Coerência Espectral (SCC), obtidos por meio da análise tempo-

frequência dos sinais de EEG utilizando métodos como o periodograma de Welch e a Transformada Wavelet de Morlet. Para a classificação, foi proposto um novo modelo DICE-net, que utiliza camadas de convolução em profundidade, camadas de codificador de *transformer* e redes feed-forward (FFN). O modelo processa dois blocos de entrada paralelos, um para RBP e um para SCC, para capturar relações espectro-temporais nos dados de EEG. O codificador *transformer*, com autoatenção multi-cabeça, modela as relações entre os canais, enquanto a FFN realiza a classificação final. O modelo é treinado utilizando *early stopping* e validação Leave-One-Subject-Out (LOSO). A densidade espectral de potência (PSD) mostrou uma redução na potência alfa (8-13 Hz) à medida que a gravidade da DA aumentava. A característica de Conectividade de Coerência Espectral (SCC) indicou uma conectividade delta mais baixa no grupo AD em comparação com CN, o que sugere sua relevância para a detecção de DA. O modelo DICE-net obteve o melhor desempenho, com uma acurácia de 83,28%, superando outros modelos. O desempenho do DICE-net foi estatisticamente significativo em comparação com outros métodos. Além disso, o modelo foi comparado com algoritmos de ponta, como EEGNet e DeepConvNet, que não tiveram um desempenho tão bom nas tarefas de classificação entre Alzheimer e controle e entre FTD e controle. A importância dos diferentes canais de EEG para a distinção entre DA e FTD foi avaliada examinando a magnitude dos pesos da camada de convolução. Os resultados indicaram que os atributos de RBP foram mais significativos que os de SCC, e o modelo focou nas regiões temporais e occipitais para a classificação de DA e nas regiões frontais e temporais para a classificação de FTD.

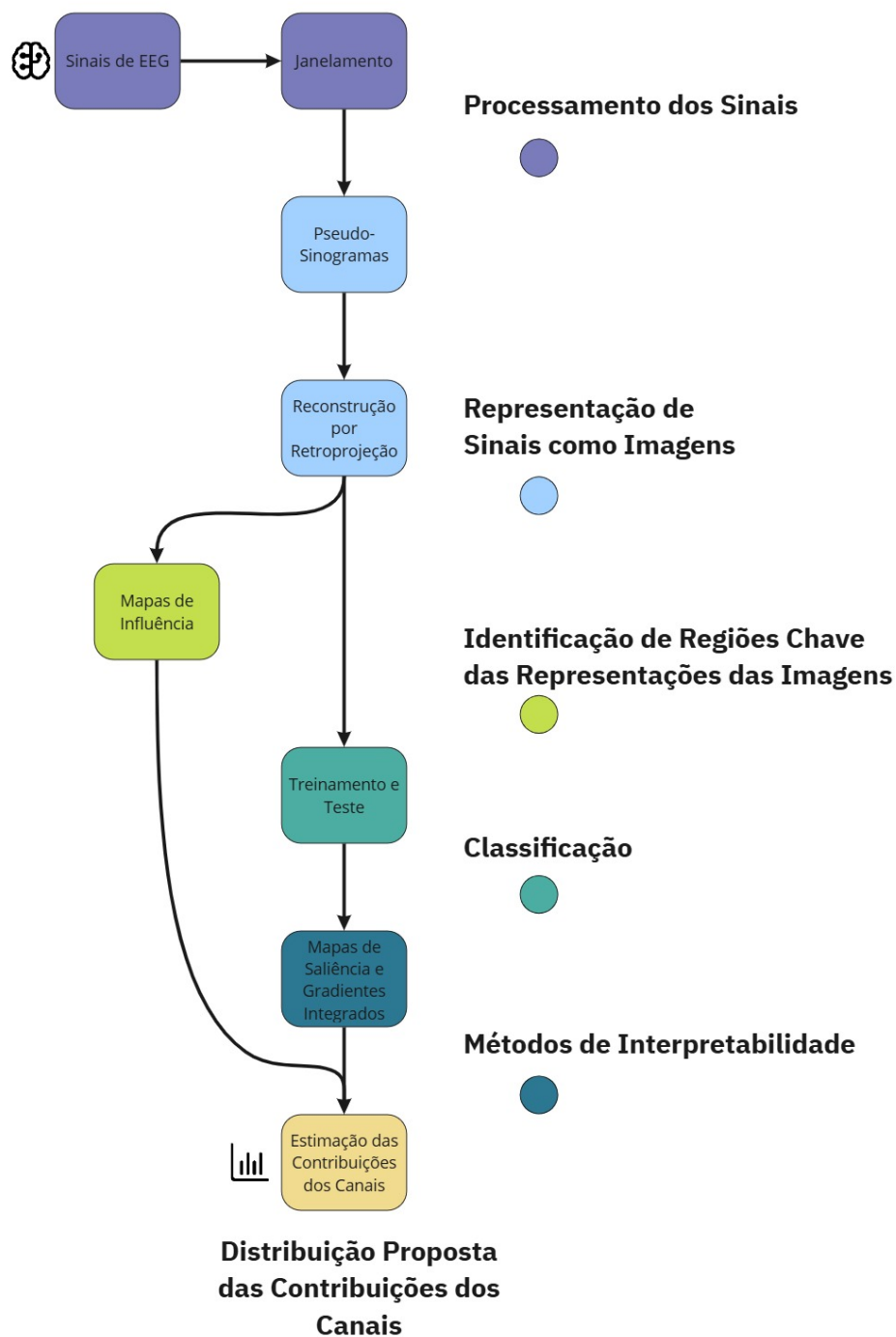
O trabalho realizado por (PURI et al., 2022) utilizou uma base de dados de EEG que consiste em dois grupos: pacientes com Doença de Alzheimer e controles saudáveis. Os dados foram processados para retirar artefatos através da análise de um clínico especialista. Para a extração de atributos, foi utilizada a Entropia Espectral (SE) para avaliar a distribuição de potência dos sinais de EEG. A SE é calculada derivando a distribuição de probabilidade do espectro de potência do sinal, aplicando a entropia de Shannon no domínio da frequência. A SE normalizada foi avaliada utilizando um espectrograma de potência tempo-frequência. Além disso, a Complexidade de Kolmogorov (KC) foi utilizada para medir a aleatoriedade dos dados de EEG, convertendo cada época em uma sequência binária. Essa medida de complexidade visa quantificar o conteúdo de informação dos sinais de EEG, podendo auxiliar no diagnóstico da DA. Os valores de SE foram significativamente menores nos pacientes com DA em comparação com os pacientes de controle em todos os 16 eletrodos do EEG, com diferenças notáveis nos eletrodos O1, O2, P3 e P4. A análise de KC indicou que a atividade de EEG dos pacientes com

DA é mais regular do que a dos do grupo de controle. O desempenho de seis classificadores, incluindo SVM e KNN, foi avaliado utilizando os atributos SE e KC, mostrando a maior acurácia de 90,8% para SE e 92,9% para KC. Ao combinar ambas atributos, os classificadores alcançaram uma acurácia de 95,6%, com o SVM apresentando o melhor desempenho em todos os experimentos. Os resultados sugerem que SE e KC são biomarcadores eficazes para identificar distinguir DA e controle. Apesar do tamanho pequeno do conjunto de dados, os achados indicam o potencial de usar esses atributos no diagnóstico de DA. Trabalhos futuros visam expandir o conjunto de dados e explorar a sincronia de EEG com outras entropias e complexidades para aprimorar a acurácia.

4 METODOLOGIA

A metodologia adotada neste estudo está representada por meio do diagrama de fluxo apresentado na Figura 6, o qual descreve as principais etapas conduzidas ao longo da pesquisa. Cada uma dessas etapas será detalhada na sequência desta seção.

Figura 6 – Diagrama de fluxo representando as etapas da metodologia do estudo.

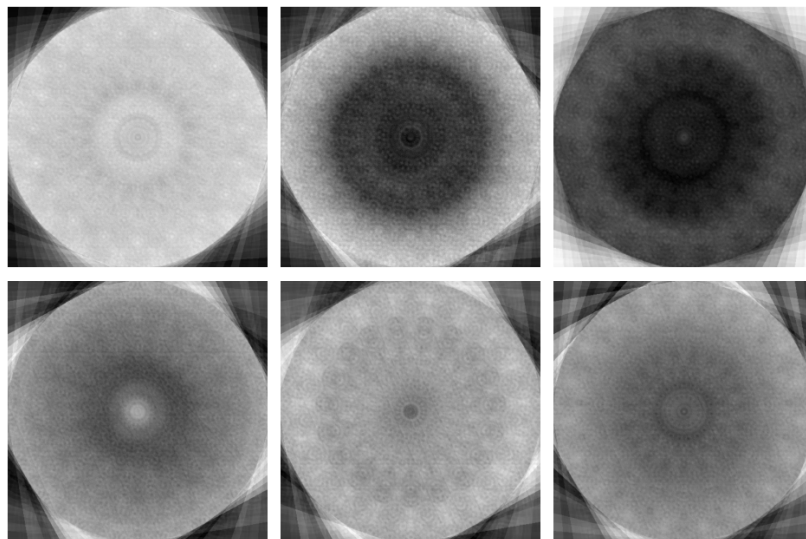


4.1 BASE DE DADOS

A base de dados utilizada é composta por sinais de EEG de 88 pacientes em estado de repouso, incluindo 36 pacientes com doença de Alzheimer (DA), 23 pacientes com demência frontotemporal (DFT) e 29 controles saudáveis. Os sinais foram coletados utilizando um sistema clínico de EEG com 19 eletrodos no couro cabeludo seguindo o sistema internacional 10–20. Durante a coleta os participantes estavam em estado de repouso com os olhos fechados. Cada gravação teve duração aproximada de 13,5 minutos para o grupo com DA, 12 minutos para o grupo com DFT e 13,8 minutos para o grupo controle. A base de dados também inclui pontuações do Mini-Exame do Estado Mental (MEEM) para todos os sujeitos, refletindo estados cognitivos e neuropsicológicos. As gravações foram coletadas a uma taxa de amostragem de 500Hz e resolução de $10\mu V/mm$ (MILTADOUS et al., 2023b).

Os sinais de EEG foram transformados em imagens utilizando o método ASTERI, descrito na Seção 2.6. Na geração das imagens, foram utilizados os sinais de EEG brutos, sem aplicação de filtragem. Cada imagem foi gerada a partir de um janelamento contendo 500 pontos de amostragem do sinal de EEG original, correspondendo a um segundo de gravação, considerando a taxa de amostragem de 500 Hz. Para evitar perda de informações contínuas do sinal, foi utilizada uma sobreposição de 250 pontos (0,5 segundos) entre as janelas. As imagens resultantes possuem dimensões de (500,500). A Figura 7 ilustra alguns exemplos das imagens geradas a partir do método ASTERI.

Figura 7 – Exemplos das imagens geradas a partir do método ASTERI.



Fonte: o Autor (2024)

Ao final do processo, foram geradas 58.758 imagens para a classe de Alzheimer, 33.477 para a classe DFT e 48.828 para a classe de controle, totalizando 141.063 imagens. Essas imagens foram divididas em conjuntos para treino, validação, calibração e teste. O conjunto de treino foi composto por 85% do total (119.903 imagens), enquanto os conjuntos de validação e calibração foram compostos por 3% cada (4.232 imagens). O conjunto de teste ficou com os 9% restantes (12.696 imagens). Foram selecionados seis diferentes modelos pré-treinados para comparação de desempenho. Os modelos selecionados e uma breve descrição de cada um são apresentados abaixo:

- Hiera é um *vision transformer* (ViT) hierárquico que simplifica arquiteturas modernas ao eliminar componentes desnecessários, como convoluções e janelas deslocadas ou cruzadas, mantendo apenas os elementos essenciais. Através do pré-treinamento com Masked Autoencoder (MAE), o modelo aprende viés espacial de forma eficiente, dispensando modificações complexas. Essa abordagem resulta em um modelo mais rápido tanto no treinamento quanto na inferência (RYALI et al., 2023);
- ResNet é uma arquitetura clássica tradicional de rede neural profunda que utiliza aprendizado residual para facilitar o treinamento de redes mais profundas. Em vez de aprender diretamente a função desejada, a rede aprende a função residual $F(x) := H(x) - x$, o que torna a otimização mais eficiente e permite que redes muito profundas (como 152 camadas) sejam treinadas com maior facilidade (HE et al., 2016);
- ConvNeXt é uma família de redes neurais puramente convolucionais que moderniza o design dos ConvNets tradicionais, alcançando desempenho competitivo com os Vision Transformers (ViTs) em tarefas de visão computacional. O modelo combina a eficiência e simplicidade dos ConvNets com melhorias inspiradas em Transformers, como técnicas de escalabilidade e treinamento avançado (LIU et al., 2022);
- EdgeNeXt é um modelo com uma arquitetura híbrida que combina as vantagens de ViTs e CNNs. A arquitetura utiliza convoluções adaptativas com *kernels* de tamanhos variáveis para capturar características em diferentes estágios da rede, otimizando o uso de recursos, alcançando alta acurácia com menos parâmetros e menor custo computacional (MAAZ et al., 2023);
- RDNet renova as arquiteturas *DenseNet* refinando os métodos de treinamento e elementos da arquitetura para melhorar a eficiência de memória, mantendo conexões densas.

Através de estratégias como aumento da taxa de crescimento (GR), camadas de transição refinadas e patchificação, RDNet demonstra ter resultados competitivos com o estado da arte, destacando o potencial das DenseNets para tarefas densas como segmentação semântica (KIM; HEO; HAN, 2024);

- MobileOne foca na otimização de redes neurais para dispositivos móveis, alcançando alta precisão ao mesmo tempo em que reduz a latência, tornando-o uma opção pertinente para aplicações em tempo real em dispositivos móveis (VASU et al., 2023).

A escolha desses seis modelos distintos foi motivada pela comparação do desempenho de diferentes tipos de arquiteturas. A seleção abrange desde uma arquitetura clássica, como a ResNet, até arquiteturas mais modernas, como o Hiera (um *vision transformer*), e também modelos otimizados para dispositivos móveis, como o MobileOne, que possui menos parâmetros. Além disso, os modelos variam no número de parâmetros. A versão do Hiera utilizada possui 27.143.523 parâmetros, enquanto a ResNet50 conta com 23.514.179 parâmetros. A versão da ConvNeXt utilizada tem 27.822.435 parâmetros, e a MobileOne, com foco em dispositivos móveis, possui 3.548.035 parâmetros. A RDNet tem 22.825.131 parâmetros, e, por fim, a EdgeNeXt apresenta 1.158.723 parâmetros.

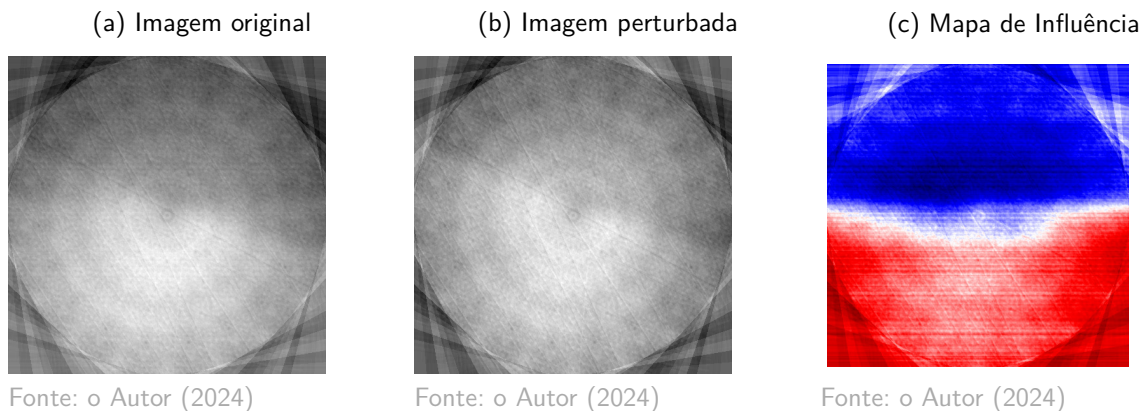
4.2 MAPAS DE INFLUÊNCIA

Foi realizada uma análise de perturbação modificando sequencialmente cada canal dos sinais de EEG para observar as mudanças resultantes nas imagens geradas. Essa abordagem tem intuito de identificar a influência individual que cada canal exerce na imagem final gerada. O processo de modificação envolveu zerar um canal por vez, gerar uma imagem com a configuração modificada do canal e, em seguida, compará-la com a imagem original com todos os canais ativos. A partir dessa comparação obteve-se mapas de calor que identificam as regiões mais afetadas da imagem por cada canal. Esses mapas de calor que indicam a influência de cada canal em cada região da imagem são denominados no presente trabalho como Mapas de Influência. A geração dos mapas de influência é realizada da seguinte maneira: seja I a matriz com dimensões (500,500) que representa a imagem original, e $\hat{I}_k$ a matriz com dimensões (500,500) que representa a imagem gerada com o canal k zerado. O mapa de influência M do canal k é definido como:

$$M^k = I - \hat{I}^k \quad (4.1)$$

A Figura 8 apresenta um exemplo de uma imagem original, de uma imagem com canal zerado e do mapa de influência respectivo. A subfigura 8a apresenta a imagem original construída a partir dos sinais de EEG com todos os canais ativos. Em seguida, a subfigura 8b mostra a mesma imagem, mas com um canal específico zerado, simulando uma perturbação no sinal. Por fim, a subfigura 8c exibe o mapa de influência resultante da comparação entre as imagens 8a e 8b, destacando visualmente as regiões mais afetadas pela ausência do canal em questão. Nesse mapa, as influências positivas são indicadas em vermelho, enquanto as influências negativas aparecem em azul.

Figura 8 – Exemplo das imagens geradas em diferentes estágios. A subfigura (a) ilustra a imagem original com todos os canais ativos; A subfigura (b) representa a imagem perturbada com um canal zerado; A subfigura (c) mostra o Mapa de influência gerado a partir da Imagem original (a) e da imagem perturbada (b).



4.3 TREINAMENTO

Todos os modelos foram pré-treinados na ImageNet-1K (RUSSAKOVSKY et al., 2015). A abordagem de treinamento foi realizada em duas etapas. Na primeira etapa, todos os modelos foram treinados com o modelo base congelado (todas as camadas, exceto a última, estavam congeladas) para preservar seus pesos pré-treinados originais. Os modelos foram treinados por 10 épocas com uma taxa de aprendizagem de 1×10^{-3} . O otimizador Adam foi utilizado em todas as etapas. Na segunda etapa, todas as camadas dos modelos foram descongeladas, e a taxa de aprendizagem foi reduzida para 1×10^{-5} , para ocorrer um processo de *fine-tuning* mais suave, permitindo ajustes de peso mais refinados.

4.4 MÉTRICAS DE AVALIAÇÃO

Para a avaliação dos modelos, foram utilizadas as seguintes métricas: Acurácia, F1-score, Sensibilidade, Especificidade e Entropia. Para uma melhor compreensão das métricas, consideram-se as seguintes definições:

- Verdadeiros Positivos (VP): exemplos corretamente classificados pelo modelo como positivos;
- Verdadeiros Negativos (VN): exemplos corretamente classificados pelo modelo como negativos;
- Falsos Positivos (FP): exemplos classificados pelo modelo como positivos, mas que, na realidade, são negativos.
- Falsos Negativos (FN): exemplos classificados pelo modelo como negativos, mas que, na realidade, são positivos.

A acurácia representa a taxa de predições corretas de um modelo em relação ao total de casos (GRANDINI; BAGLI; VISANI, 2020). A acurácia é calculada a partir da Equação 4.2 definida abaixo:

$$Acurácia = \frac{VP + VN}{VP + FP + VN + FN} \quad (4.2)$$

A sensibilidade representa a taxa de casos positivos que foram corretamente identificados pelo modelo em relação ao total de casos positivos (GRANDINI; BAGLI; VISANI, 2020). A sensibilidade é calculada a partir da Equação 4.3 definida abaixo:

$$Sensibilidade = \frac{VP}{VP + FN} \quad (4.3)$$

A especificidade representa o contrário da sensibilidade: é a taxa de casos negativos que foram corretamente identificados pelo modelo em relação ao total de casos negativos (GRANDINI; BAGLI; VISANI, 2020). A especificidade é calculada a partir da Equação 4.4 definida abaixo:

$$Especificidade = \frac{VN}{VN + FP} \quad (4.4)$$

O F1-score é a média harmônica entre a precisão, que mede a proporção de exemplos positivos previstos corretamente em relação a todos os exemplos positivos previstos, e a sensibilidade (também chamada de *recall*). Ele fornece uma avaliação equilibrada do desempenho

de um modelo, especialmente quando os valores de precisão e sensibilidade diferem (GRANDINI; BAGLI; VISANI, 2020). O F1-score é calculado a partir da Equação 4.5 definida abaixo:

$$F1 - score = \frac{2VP}{2VP + FP + FN} \quad (4.5)$$

A entropia de Shannon mede a incerteza ou imprevisibilidade de um sistema e é definida pela Equação 4.6 abaixo:

$$H(p(y)) = - \sum_{k=1}^K p(\hat{y}_k) \log p(\hat{y}_k) \quad (4.6)$$

onde $p(\hat{y}_k)$ é a probabilidade da saída $\hat{y}$ pertencer à classe k . A entropia aumenta à medida que as probabilidades p_i de diferentes classes se tornam mais iguais, atingindo seu valor máximo quando todas as classes são equiprováveis. Nesse caso, onde cada probabilidade $p_i = 1/K$, a entropia é igual a $\log K$, e ela aumenta com a quantidade de classes K . Quando há apenas um resultado possível (ou quando o modelo possui 100% de certeza da classe predita), o sistema é totalmente previsível e a entropia é zero. A entropia de Shannon é uma função contínua e simétrica, o que significa que pequenas mudanças nas probabilidades resultam em pequenas mudanças na entropia, e trocar os valores de duas probabilidades não afeta o resultado (OMAR; PLAPPER, 2020).

4.5 CONTRIBUIÇÕES DOS CANAIS

Para estimar a importância dos canais na classificação, os mapas de calor gerados pelas técnicas de mapas de saliência e gradientes integrados são multiplicados pelos mapas de influência de cada canal para cada imagem. O processo é realizado da seguinte forma: são selecionadas aleatoriamente 2.500 imagens de teste nas quais o modelo com melhor desempenho fez previsões corretas. Essas 2.500 imagens geram 47.500 mapas de influência (19 canais por imagem).

Mapas de Saliência e Gradientes Integrados: As 2.500 imagens são processadas pelo modelo, e, com base nas previsões do modelo, mapas de saliência e gradientes integrados são gerados para cada imagem. O modelo utilizado para realizar as previsões para gerar os mapas de saliência e gradientes integrados é o modelo com o melhor desempenho nas métricas após a etapa de treinamento. Os mapas de saliência e os gradientes integrados possuem dimensões de *array* (224, 224). Esses *arrays* são multiplicados pelos mapas de influência correspondentes para cada canal de suas respectivas imagens. Como os mapas de influência originalmente pos-

suas dimensões (500, 500), antes de realizar a multiplicação os mapas são pré-processados e têm suas dimensões alteradas para (224, 224) utilizando a mesma técnica usada para diminuir a imagem original antes de utilizá-la como entrada nos modelos. Os valores dos pixels são então somados para calcular uma pontuação geral para cada um dos 19 canais. O cálculo da contribuição percentual de cada canal é formalizado abaixo:

$$A^k = \sum_{j=1}^{D_2} \sum_{i=1}^{D_1} M_{ij}^k \times S_{ij} \quad (4.7)$$

$$T^k = \frac{1}{m} \sum_{i=1}^m A_i^k \quad (4.8)$$

$$C^k = 100 \frac{T^k}{\sum_{k=1}^N T^k} \quad (4.9)$$

Na Equação 4.7, M_{ij}^k representa o valor do pixel na linha i e coluna j do mapa de influência M do canal k . De forma análoga, S_{ij} corresponde ao valor do pixel na linha i e coluna j do mapa de saliência da imagem original. As dimensões das matrizes são representadas por D_1 e D_2 , sendo $D_1 = D_2 = 224$. A média da influência ponderada A do canal k sobre todas as imagens de teste é representada por T^k , considerando $m = 2500$. O número total de canais é dado por $N = 19$. Por fim, a contribuição percentual C do canal k é calculada com base na média das influências ponderadas normalizadas em relação à soma total.

5 RESULTADOS E DISCUSSÕES

Os valores das métricas de avaliação obtidos durante o treinamento estão apresentados na Tabela 1. A acurácia e o F1 Score são expressos com intervalos de confiança (CI) de 95%, enquanto a entropia é reportada em termos de média e desvio padrão.

Após o treinamento com os modelos congelados, os modelos ConvNeXt, ResNet, RDNet e MobileOne apresentaram desempenhos similares, com sobreposição de seus intervalos de confiança, indicando que não houve preferência estatística entre eles em relação à acurácia e ao F1 Score. As métricas de entropia desses modelos também apresentaram sobreposição. Os modelos Hiera e EdgeNext tiveram desempenhos ligeiramente inferiores, sendo que o Hiera superou o EdgeNext, mas sem sobreposição de CI entre eles.

Após o treinamento com os modelos descongelados, o ConvNeXt superou os demais, alcançando maior acurácia e F1 Score sem sobreposição de CI. Os modelos Hiera, EdgeNext, RDNet e MobileOne apresentaram desempenhos estatisticamente similares, com múltiplas sobreposições de CIs. O ResNet teve o pior desempenho, com seu CI não sobrepondo os dos outros modelos.

A entropia obtida na classificação dos diferentes modelos é apresentada na Tabela 2. Após o processo de descongelamento da base, houve uma redução significativa na entropia das previsões, um fenômeno também observado na acurácia e no *F1-score*, reforçando a importância do *fine-tuning* em modelos pré-treinados. Dentre os modelos avaliados, o ConvNeXt apresentou o menor valor médio de entropia (0.0195 ± 0.1237), indicando maior confiança nas classificações. No entanto, seu desvio padrão apresenta sobreposição com os intervalos dos demais modelos. É importante ressaltar que, embora o desvio padrão elevado possa dar a impressão de que a entropia poderia assumir valores negativos no limite inferior, isso não ocorre, pois a entropia é sempre um valor positivo. Esse desvio padrão elevado é causado por *outliers* positivos, que aumentam a dispersão dos dados. Para ilustrar esse comportamento, foram gerados *boxplots* da entropia dos modelos (Figura 9). Observa-se que o *boxplot* do ConvNeXt é extremamente estreito, quase indistinguível de uma linha próxima ao valor zero, com *outliers* dispersos acima do limite superior. Adicionalmente, o modelo ResNet foi o único cujo *boxplot* apresenta uma caixa visivelmente perceptível, indicando valores mais distantes de zero e, consequentemente, maior incerteza nas suas previsões em comparação com os demais modelos. Por outro lado, o ConvNeXt se destaca por possuir a menor quantidade de *outliers*, a julgar pela densidade

Tabela 1 – Métricas obtidas para diferentes modelos após o treinamento da base. Os valores são apresentados com seus respectivos intervalos de confiança (CI) de 95%. Observa-se uma melhora significativa nas métricas após o descongelamento dos modelos.

Modelo	Métricas			
	Modelo Congelado		Modelo Descongelado	
	Acurácia	F1 score	Acurácia	F1 score
Hiera	[0,5713; 0,5966]	[0,5643; 0,5896]	[0,9843; 0,9891]	[0,9843; 0,9891]
ConvNext	[0,6220; 0,6463]	[0,6196; 0,6438]	[0,9931; 0,9964]	[0,9931; 0,9964]
EdgeNext	[0,5377; 0,5638]	[0,5151; 0,5412]	[0,9792; 0,9849]	[0,9792; 0,9849]
ResNet	[0,6248; 0,6489]	[0,6229; 0,6470]	[0,9656; 0,9731]	[0,9656; 0,9732]
RDNet	[0,6315; 0,6554]	[0,6234; 0,6473]	[0,9886; 0,9928]	[0,9886; 0,9928]
MobileOne-S1	[0,6132; 0,6377]	[0,6027; 0,6272]	[0,9803; 0,9859]	[0,9804; 0,9859]

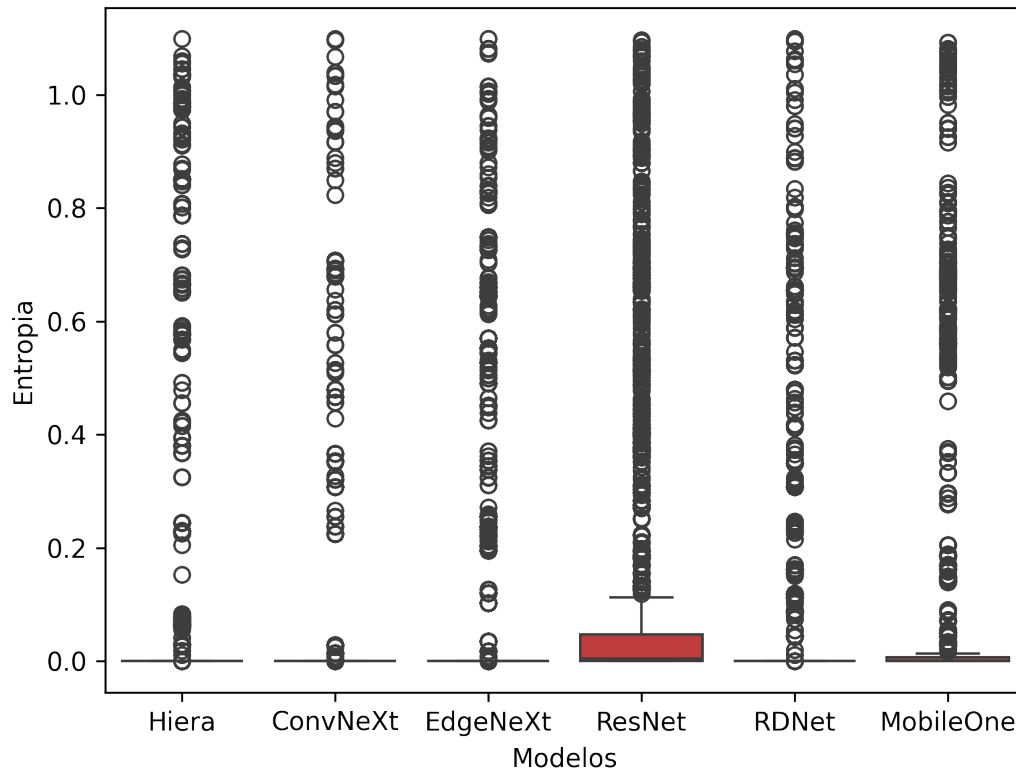
visual dos valores, corroborando os valores apresentados na Tabela 2, que indicam sua maior confiança e consistência nas classificações.

A Tabela 3 apresenta os valores de sensibilidade e especificidade para cada classe (DA, CN e DFT) obtidos pelos diferentes modelos após o treinamento da base. Com relação à sensibilidade, observa-se que o modelo ConvNeXt apresentou o melhor desempenho geral, com valores próximos a 1 em todas as classes (0,9936 para DA, 0,9966 para CN e 0,9894 para DFT). Esses resultados indicam que o ConvNeXt foi o mais eficaz em detectar os casos verdadeiramente positivos em cada grupo. O modelo RDNet também alcançou alta sensibilidade, destacando-se no grupo CN com um valor de 0,9986, superior a todos os demais na identificação do grupo de controle. Já o modelo ResNet teve desempenho ligeiramente inferior, apresentando valores de sensibilidade menores, particularmente no grupo DA (0,9675), indicando maior dificuldade em identificar positivamente os casos dessa classe. Quanto à especificidade, o ConvNeXt nova-

Tabela 2 – Entropia obtida com os diferentes modelos testados.

Modelo	Entropia	
	Modelo Congelado	Modelo Descongelado
Hiera	1,2827 $\pm$ 0,2572	0,0442 $\pm$ 0,1928
ConvNext	1,1429 $\pm$ 0,3250	0,0195 $\pm$ 0,1237
EdgeNext	1,3294 $\pm$ 0,2228	0,0596 $\pm$ 0,2069
ResNet	1,1753 $\pm$ 0,3037	0,1390 $\pm$ 0,2946
RDNet	1,1401 $\pm$ 0,3328	0,0350 $\pm$ 0,1575
MobileOne-S1	1,1772 $\pm$ 0,2943	0,0615 $\pm$ 0,1983

Figura 9 – Gráficos Boxplots da entropia obtida na classificação do conjunto de teste dos modelos.



Fonte: o Autor (2025)

mente se destacou como o modelo mais robusto, com valores de 0,9976, 0,9953 e 0,9975 para os grupos DA, CN e DFT, respectivamente. Da mesma forma, os modelos Hiera, RDNet e MobileOne apresentaram resultados robustos, com valores de especificidade bastante elevados e comparáveis aos do ConvNeXt, principalmente para os grupos de DA e DFT, indicando um desempenho consistente. O modelo ResNet, por outro lado, apresentou os menores valores de especificidade em todos os grupos, particularmente no grupo DA (0,9810).

O modelo MobileOne, apesar de sua estrutura otimizada para dispositivos móveis, apresentou valores elevados de sensibilidade e especificidade em todos os grupos, posicionando-se como uma alternativa eficiente, mesmo quando comparado a arquiteturas mais complexas e computacionalmente custosas. Por fim, destaca-se que o modelo ConvNeXt apresentou os resultados mais robustos em termos de sensibilidade e especificidade, enquanto o ResNet apresentou desempenho inferior comparativamente, especialmente na classe DA. Esses resultados reforçam que o ConvNeXt é a arquitetura mais promissora entre as avaliadas, com alta precisão tanto em identificar verdadeiros positivos quanto em evitar falsos positivos.

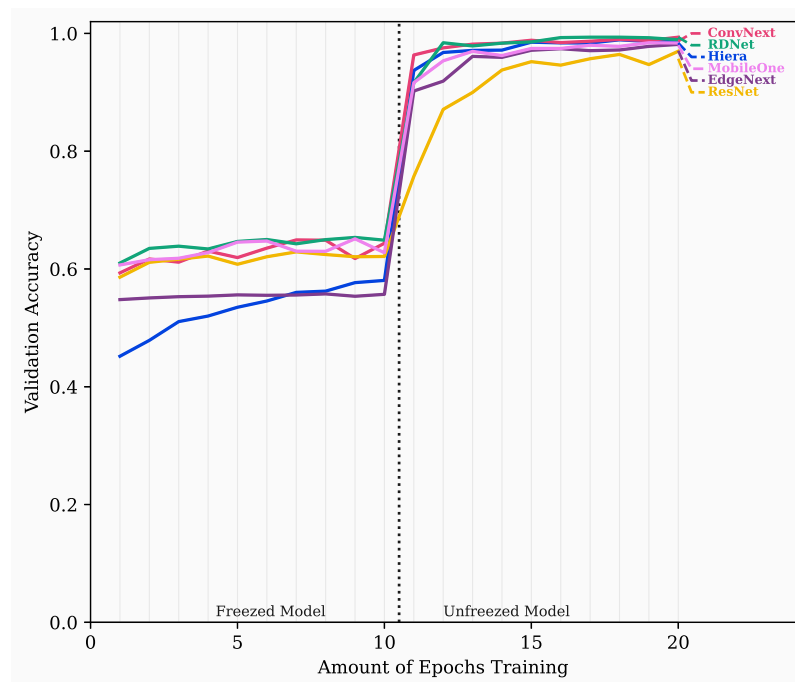
A progressão da acurácia e do F1-score no conjunto de validação ao longo de cada época

Tabela 3 – Sensibilidade e especificidade divididos por classe obtidos para os diferentes modelos após o treinamento da base.

Modelo	Métricas					
	Sensibilidade			Especificidade		
	Grupo DA	Grupo CN	Grupo DFT	Grupo DA	Grupo CN	Grupo DFT
Hiera	0,9925	0,9841	0,9856	0,9904	0,9932	0,9975
ConvNext	0,9936	0,9966	0,9894	0,9976	0,9953	0,9975
EdgeNext	0,9844	0,9856	0,9694	0,9869	0,9899	0,9942
ResNet	0,9675	0,9702	0,9732	0,9810	0,9860	0,9868
RDNet	0,9832	0,9986	0,9868	0,9975	0,9892	0,9972
MobileOne-S1	0,9899	0,9815	0,9756	0,9975	0,9892	0,9972

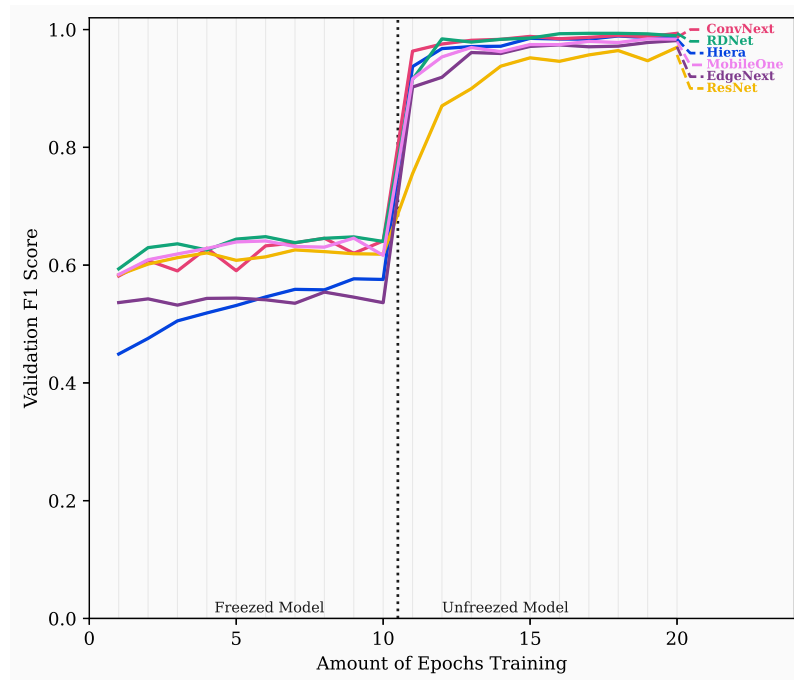
é apresentada nas Figuras 10 e 11, respectivamente. As métricas de validação melhoram significativamente durante as dez épocas de treinamento após o descongelamento do modelo, ressaltando sua importância. ConvNeXt, RDNet, ResNet e MobileOne alcançam acurácia e F1 Score significativamente maiores nas primeiras dez épocas de treinamento com a base do modelo congelada, demonstrando eficácia inicial em comparação com Hiera e EdgeNext. No treinamento após o descongelamento do modelo, as diferenças nas métricas tornam-se visualmente quase negligenciáveis.

Figura 10 – Progressão da acurácia no conjunto de validação ao longo de cada época para os diferentes modelos treinados na base.



Fonte: o Autor (2024)

Figura 11 – Progressão do F1 score no conjunto de validação ao longo de cada época para os diferentes modelos treinados na base.



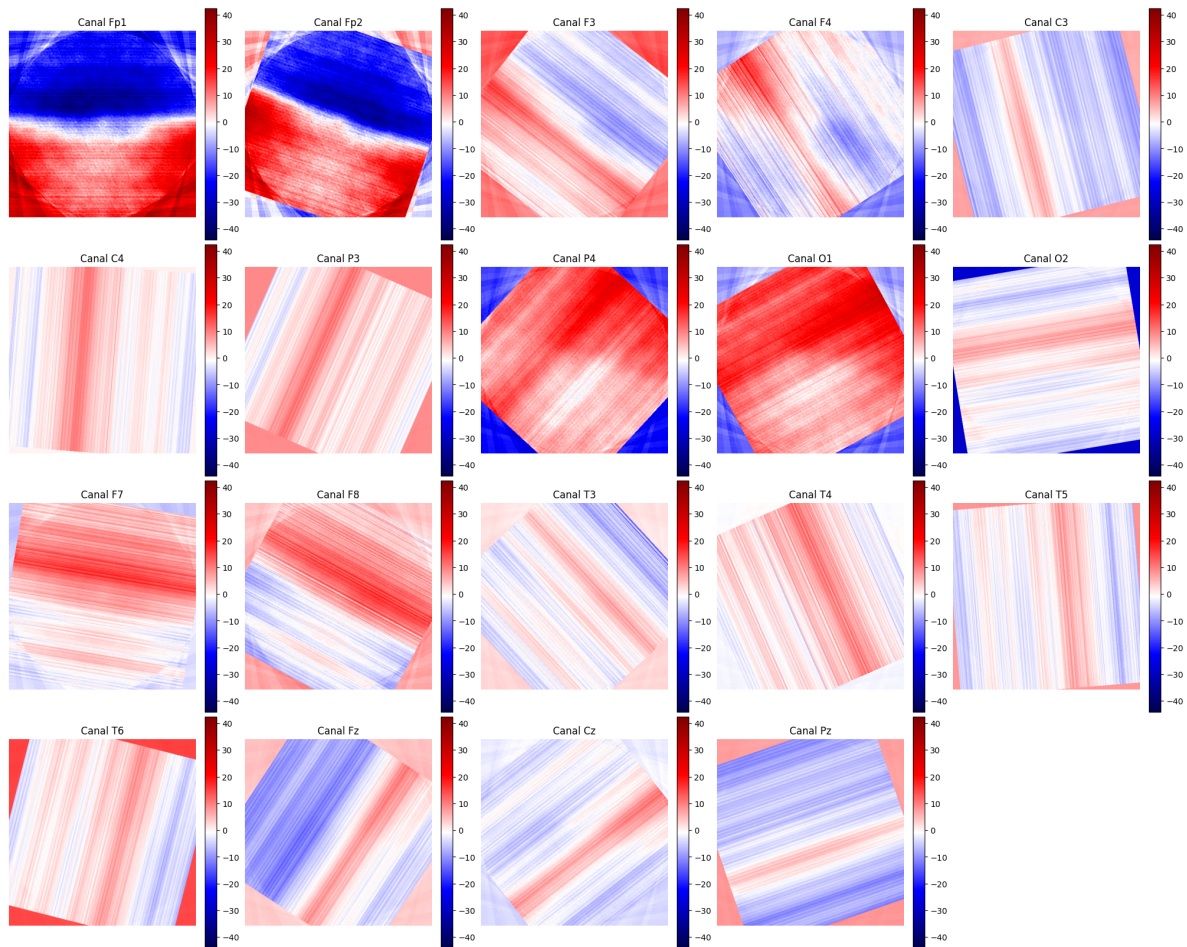
Fonte: o Autor (2024)

Observa-se também que, durante as dez primeiras épocas com os modelos congelados, os valores de acurácia e F1 Score atingem um platô rapidamente, logo nas primeiras épocas, mantendo-se estáveis ao longo das épocas subsequentes. Esse comportamento se repete após o descongelamento dos modelos: as métricas aumentam significativamente nas primeiras duas a três épocas e, em seguida, estabilizam-se em um novo platô até o final do treinamento. Esse padrão sugere que talvez não seja necessário realizar vinte épocas de treinamento — dez com o modelo congelado e dez com o modelo descongelado. Possivelmente, duas a três épocas em cada fase seriam suficientes para alcançar resultados similares. A redução no número de épocas poderia diminuir significativamente o tempo total de treinamento dos modelos, o que teria um impacto relevante, considerando que o treinamento de modelos modernos pode demandar períodos consideravelmente longos. A otimização da quantidade de épocas e outros hiperparâmetros durante o treinamento pode ser investigada mais a fundo em trabalhos futuros.

Como é possível observar a partir dos resultados, no geral o modelo ConvNeXt apresentou resultados superiores nas métricas em relação aos demais modelos. Portanto o modelo ConvNeXt foi utilizado como modelo base no processo de geração dos mapas de saliência e gradientes integrados para o cálculo das contribuições individuais de cada canal de EEG.

A Figura 12 apresenta os mapas de influência obtidos a partir de uma imagem teste. A

Figura 12 – Mapa de Influência gerado para cada canal por meio do método de perturbação descrito em Métodos. Nesta versão, a escala na legenda foi fixada para todos os canais, destacando os canais com maior influência com base em suas cores mais intensas.



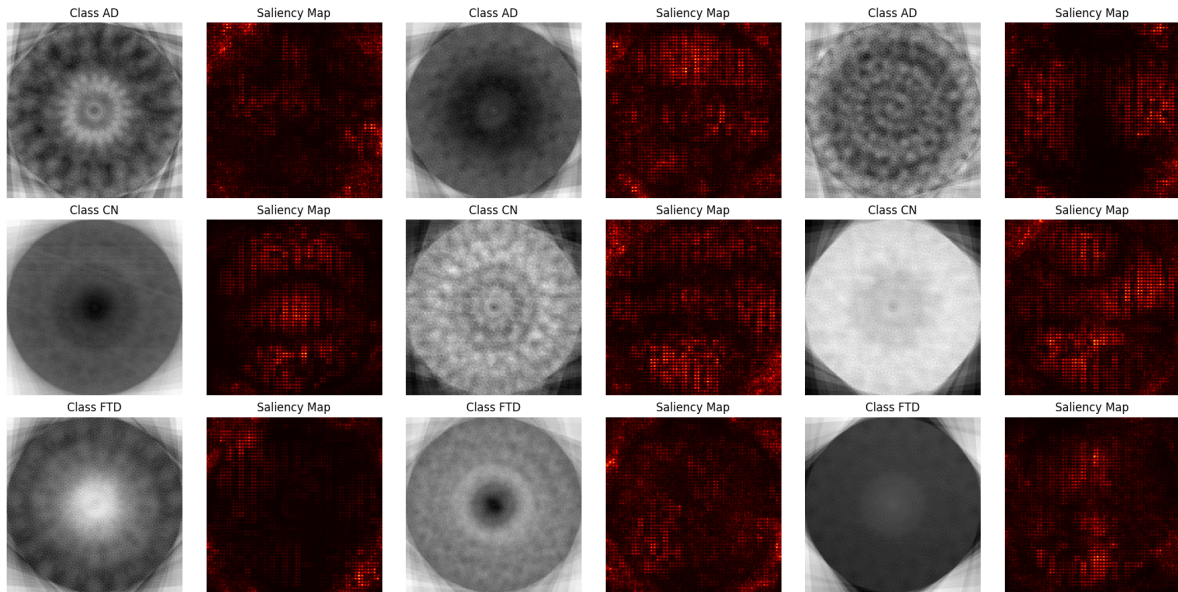
Fonte: o Autor (2024)

escala foi fixada para melhor visualização ao comparar diretamente a intensidade de cores à influência do canal na imagem original.

Mapas de saliência e gradientes integrados de três exemplos das imagens testes de cada classe são apresentados na Figura 13 e 14, respectivamente.

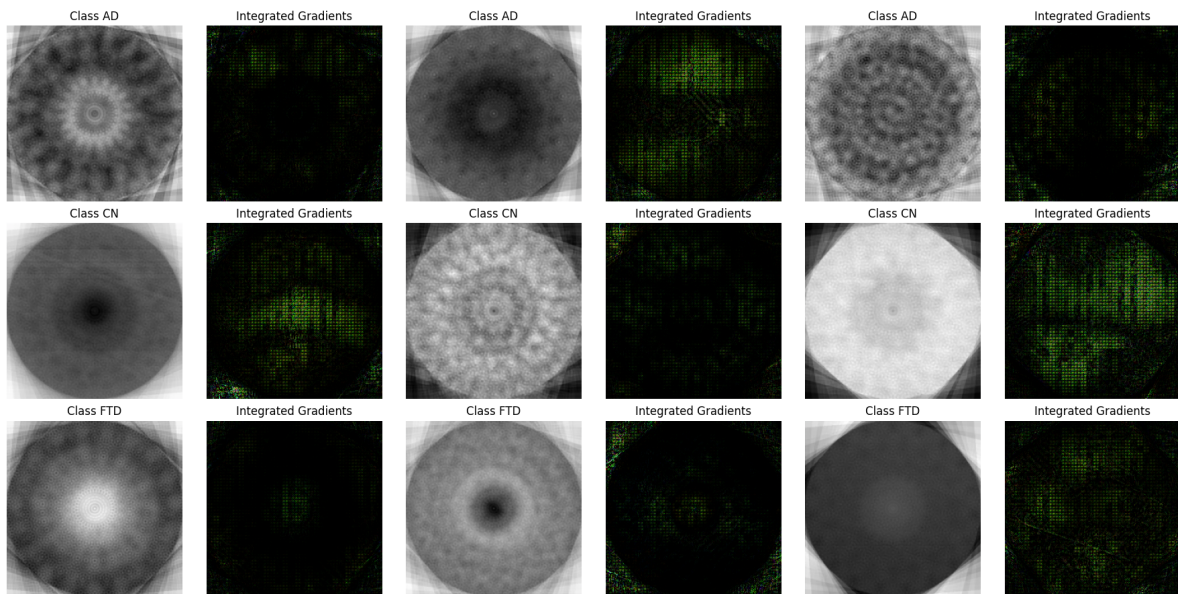
As contribuições percentuais de todos os canais para a tarefa de classificação estão apresentadas na Tabela 4. A tabela exhibe o ranqueamento dos canais com base nas contribuições obtidas pelos dois métodos testados, mapas de saliência e gradientes integrados, e também a partir da média de ambos. Observa-se que o canal O2 apresentou a maior contribuição média para a classificação, ficando em 1º lugar nos mapas de saliência e em 2º nos gradientes integrados, seguido do canal O1, que ocupou a 2ª posição na média. Esse resultado sugere um papel significativo da região occipital na diferenciação das classes, com uma margem considerável entre as contribuições dos dois canais occipitais e o terceiro colocado.

Figura 13 – Mapas de Saliência gerados para três imagens testes de cada classe.



Fonte: o Autor (2024)

Figura 14 – Gradientes Integrados gerados para três imagens testes de cada classe.



Fonte: o Autor (2024)

Outro ponto relevante é a alta contribuição dos canais da região frontal, com F7, F8, F4 e F3 entre os mais influentes. O único canal frontal abaixo da 10ª colocação foi Fz, que ocupou a 16ª posição. Além disso, houve uma discrepância entre os dois canais da região frontal polar: enquanto Fp1 (lado esquerdo) ficou em penúltimo na média das contribuições, Fp2 (lado direito) ocupou a 5ª posição. Essa diferença pode estar associada à assimetria hemisférica frequentemente observada em distúrbios neurodegenerativos, nos quais padrões de atividade podem se manifestar de forma diferenciada entre os hemisférios (MIZRAK et al., 2024;

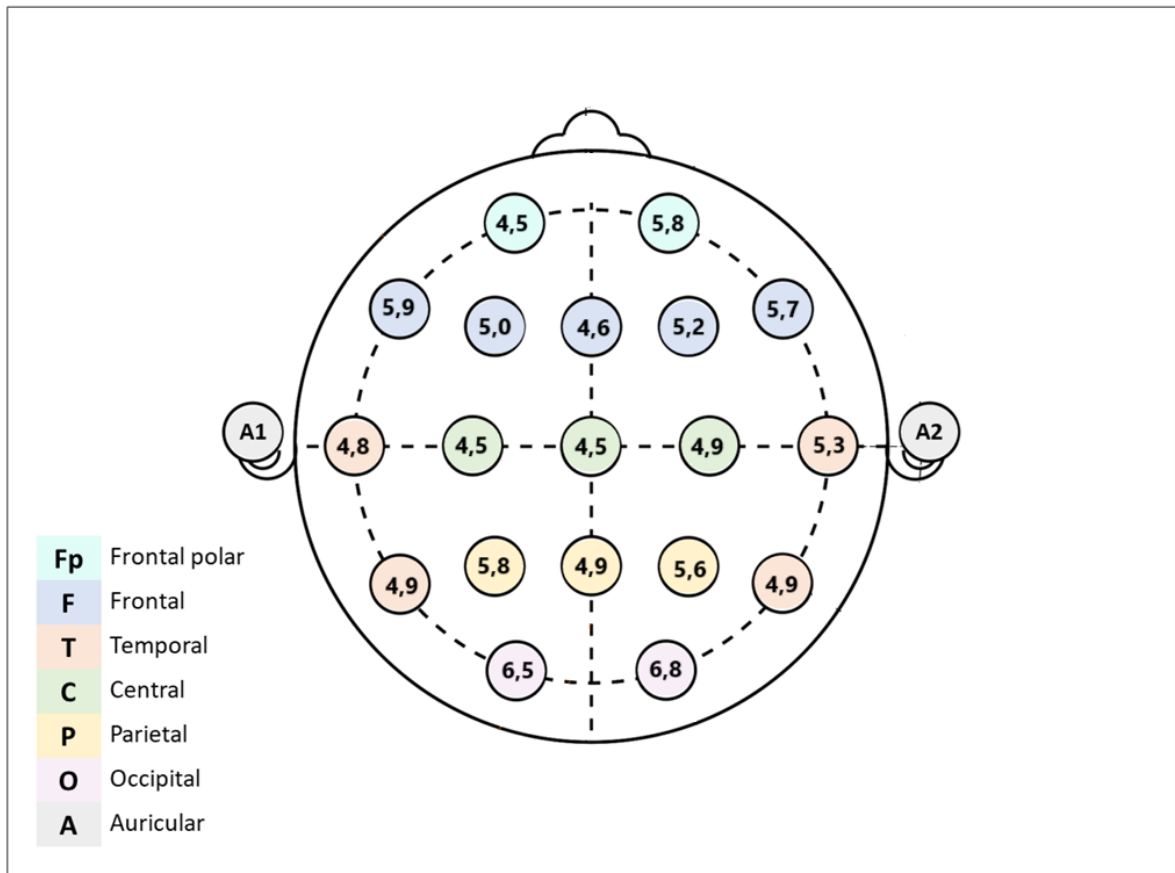
Tabela 4 – Contribuições de cada canal na classificação das imagens geradas, estimadas a partir dos mapas de saliência e gradientes integrados.

Ranque	Mapas de Saliência		Gradientes Integrados		Média	
	Canal	Contribuição(%)	Canal	Contribuição(%)	Canal	Contribuição(%)
1º	O2	7,2927	O1	7,2300	O2	6,8388
2º	P4	6,0796	O2	6,3849	O1	6,5187
3º	Fp2	5,9976	F7	6,0513	F7	5,8598
4º	F8	5,8280	P3	6,0447	P3	5,7720
5º	O1	5,8074	C4	5,8176	Fp2	5,7508
6º	F7	5,6683	F8	5,5680	F8	5,6660
7º	P3	5,4992	T4	5,5062	P4	5,6586
8º	T6	5,4924	Fp2	5,5040	T4	5,3476
9º	F4	5,4205	P4	5,2377	F4	5,1566
10º	T4	5,1891	T5	5,1244	F3	4,9746
11º	Fz	5,1301	F3	4,9608	C4	4,9375
12º	Pz	5,0352	C3	4,9025	T5	4,8852
13º	F3	4,9883	F4	4,8928	T6	4,8792
14º	Cz	4,8076	Fp1	4,7664	Pz	4,8776
15º	T3	4,7942	T3	4,7411	T3	4,7677
16º	T5	4,6460	Pz	4,7200	Fz	4,6410
17º	Fp1	4,1780	T6	4,2660	C3	4,4953
18º	C3	4,0881	Fz	4,1519	Fp1	4,4722
19º	C4	4,0574	Cz	4,1294	Cz	4,4685

LUBBEN et al., 2021). Há evidências que sugerem uma deterioração neuronal mais acelerada no hemisfério esquerdo em pacientes com Alzheimer (LUBBEN et al., 2021), o que poderia explicar por que o canal Fp2, no lado direito, apresentou uma contribuição mais significativa. Isso indica que o hemisfério direito pode ser mais representativo na identificação da doença de Alzheimer justamente por ser mais ativo que o hemisfério esquerdo, fazendo com que o padrão de ativação neuronal distorcida característica da DA seja mais facilmente identificável a partir desse lado.

Por fim, os resultados apresentam que os canais da linha média (Pz, Fz, Cz) tiveram menor contribuição para a classificação, ocupando as 14ª, 16ª e 19ª posições na média, respectivamente. Isso pode indicar que os sinais captados nesses pontos são menos discriminativos para diferenciar entre as classes avaliadas, possivelmente devido à distribuição da atividade neural associada a esses distúrbios. A Figura 15 ilustra a distribuição das médias das contribuições de cada canal de acordo com as regiões do crânio onde os eletrodos estão posicionados.

Figura 15 – Distribuição da média de contribuições dos canais nas regiões dos crânio.



Fonte: Adaptado de Gomes (2023)

Os canais com as maiores contribuições, O2 e O1, apresentam um padrão distinto nos mapas de influência da Figura 12. Enquanto O2 demonstra uma forte influência positiva (vermelho) no centro da imagem, O1 exibe uma forte influência negativa (azul) nos cantos.

Também é possível observar que, embora o canal Fp1 apresente um forte contraste no mapa de influência da imagem mostrada na Figura 12, ele não obteve um ranqueamento alto nas contribuições para a classificação, ficando em penúltimo lugar na média. Ambos os canais da região frontal polar exibem a maior intensidade nos mapas de influência. Isso sugere que a intensidade da influência de um canal na imagem gerada (intensidade do pixel, e consequentemente maior voltagem medida pelo EEG) nem sempre se traduz em uma maior contribuição no processo de classificação.

No contexto clínico, esses resultados podem ter implicações relevantes, sugerindo que a atividade occipital e frontal desempenha um papel fundamental na diferenciação entre DA, DFT e indivíduos de controle. Estudos futuros podem investigar se essas regiões também apresentam padrões específicos de conectividade funcional associados a essas condições.

6 CONCLUSÃO

A Doença de Alzheimer afeta milhões de pessoas em todo o mundo, e estima-se que o número de casos triplique até 2050. Os principais exames que acompanham o diagnóstico são exames de imagem, principalmente o PET-CT para identificação do acúmulo de proteínas b-amiloide e tau no cérebro. Porém esses exames são caros e não estão disponíveis em todos os lugares, principalmente em países com menos recursos. Nesse contexto, o EEG surge como uma alternativa viável de baixo custo e portátil para auxiliar no diagnóstico da doença.

O presente trabalho buscou contribuir para a construção de conhecimento e o desenvolvimento de tecnologias assistivas à saúde, visando auxiliar no diagnóstico da Doença de Alzheimer e outras doenças neurodegenerativas. Para isso, foi proposta uma abordagem baseada no uso de dados de EEG para a identificação da Doença de Alzheimer. A metodologia consiste na aplicação do método ASTERI para transformar sinais de EEG em imagens, possibilitando sua utilização por modelos de *deep learning* pré-treinados.

O conjunto de dados utilizado inclui registros de EEG de 86 pacientes, sendo 36 com Alzheimer, 23 com demência frontotemporal e 29 controles saudáveis. A transformação dos sinais resultou em 141.063 imagens, divididas em conjuntos de treino, validação, calibração e teste. Diferentes modelos de *deep learning* foram testados, e o modelo ConvNeXt apresentou o melhor desempenho nas métricas de avaliação, alcançando acurácia e F1-score de até 0,9964 com intervalo de confiança de 95%.

A partir desse modelo, foi realizada a segunda etapa da metodologia, que consiste na análise das contribuições dos canais de EEG para a classificação final. Foram construídos mapas de influência que representam o impacto de cada canal de EEG na formação das imagens geradas. Em seguida, os métodos de interpretabilidade Mapas de Saliência e Gradientes Integrados foram aplicados para identificar as regiões das imagens que mais contribuíram para o processo de classificação. Ao cruzar os resultados desses métodos com os mapas de influência, foi possível inferir quais canais de EEG tiveram maior relevância na tomada de decisão do modelo.

Os resultados sugerem que os canais de EEG das regiões occipital e frontal desempenham um papel crucial na identificação de doenças neurodegenerativas. Além disso, foram observadas diferenças de contribuição entre os hemisférios cerebrais, corroborando evidências prévias de que o hemisfério esquerdo sofre deterioração neuronal mais rapidamente do que o hemisfério direito em pacientes com Alzheimer.

Esta pesquisa contribui para o avanço de soluções baseadas em inteligência artificial aplicadas ao diagnóstico de doenças neurodegenerativas, ressaltando o potencial do EEG como uma ferramenta complementar aos métodos convencionais. Os achados reforçam a importância da análise de padrões específicos no EEG e indicam direções para estudos futuros voltados à otimização de modelos de classificação e interpretabilidade de redes neurais para aplicações clínicas.

6.1 DIFICULDADES ENCONTRADAS

Uma das principais dificuldades encontradas neste trabalho foi o tempo elevado necessário para o treinamento dos modelos de *deep learning*. Devido à complexidade desses modelos de *deep learning* e à grande quantidade de dados utilizados, cada modelo exigiu mais de uma semana de treinamento, o que impactou diretamente o ciclo de experimentação e ajustes. Esse longo tempo de processamento também limitou a possibilidade de testar um maior número de configurações e arquiteturas.

Além disso, a conversão dos sinais de EEG em imagens gerou um total de 141.063 imagens, o que resultou em uma alta demanda por armazenamento e recursos computacionais. A necessidade de gerenciar e processar esse grande volume de dados exigiu cuidados na alocação de memória e no uso do espaço de armazenamento. Essa limitação impôs desafios na manipulação eficiente dos dados e na execução dos experimentos dentro de um tempo viável.

6.2 CONTRIBUIÇÕES

Esta pesquisa gerou contribuições significativas para o avanço da aplicação de técnicas de inteligência artificial no apoio ao diagnóstico de doenças neurodegenerativas, especialmente a Doença de Alzheimer. As principais contribuições incluem:

- Proposição de uma metodologia robusta baseada na conversão de sinais eletroencefalográficos (EEG) em imagens utilizando o método ASTERI, viabilizando a aplicação de modelos de *deep learning* pré-treinados em grandes bases de dados;
- Avaliação sistemática de diferentes arquiteturas de *deep learning*, com destaque para o modelo ConvNeXt, que obteve desempenho superior nas métricas de acurácia e F1-score, alcançando valores de até 0,9964 com intervalo de confiança de 95%;

- Implementação de uma abordagem inovadora para análise da relevância espacial dos canais de EEG por meio da construção de mapas de influência, os quais possibilitam a identificação do impacto individual de cada canal na composição das imagens utilizadas pelo modelo;
- Aplicação de técnicas de interpretabilidade baseadas em Mapas de Saliência e Gradientes Integrados, permitindo a correlação entre regiões relevantes das imagens e os canais de EEG com maior influência na tomada de decisão do modelo;
- Identificação de padrões específicos de ativação cortical, com destaque para a contribuição das regiões occipital e frontal, bem como a assimetria entre hemisférios cerebrais, reforçando achados clínicos já descritos na literatura sobre deterioração neuronal em pacientes com Alzheimer;
- Validação do EEG como uma alternativa acessível, portátil e de baixo custo aos exames de imagem convencionais, com potencial para aplicações clínicas em larga escala, especialmente em contextos de recursos limitados;
- Estabelecimento de uma base metodológica que poderá ser expandida para o diagnóstico de outras doenças neurodegenerativas, contribuindo para o desenvolvimento de sistemas assistivos e modelos explicáveis em neurociência computacional.

Na tabela 5 estão as principais publicações alcançadas até a escrita do presente trabalho.

Tabela 5 – Contribuições de cada canal na classificação das imagens geradas, estimadas a partir dos mapas de saliência e gradientes integrados.

Título do Artigo	Evento/Periódico/Livro	Ano
Predição e identificação da diabetes a partir da análise de parâmetros clínicos e laboratoriais utilizando aprendizado de máquina (SOUZA et al., 2024a)	VII Simpósio de Inovação em Engenharia Biomédica - SABIO	2024
O Uso de Deep Learning aplicado à Imagem Motora: uma revisão (SOUZA et al., 2024b)	VII Simpósio de Inovação em Engenharia Biomédica - SABIO	2024

6.3 TRABALHOS FUTUROS

Futuros trabalhos poderão explorar a abordagem baseada no método ASTERI para representar dados de EEG em imagens, ampliando a análise com outros métodos de interpretabilidade. Isso permitirá validar a consistência dos resultados obtidos e aprofundar a compreensão dos padrões neurais associados às diferentes classes analisadas.

Além disso, a utilização de bases de dados mais amplas e diversificadas, contemplando um maior número de indivíduos e uma variedade maior de doenças neurológicas, poderá contribuir para uma análise mais detalhada das diferenças nos padrões de ativação neuronal. Isso ajudará a verificar se determinadas regiões cerebrais apresentam assinaturas específicas para diferentes condições, permitindo avanços na personalização do diagnóstico.

Outro aspecto a ser explorado é a realização de análises mais aprofundadas sobre a contribuição dos canais de EEG no processo de classificação, considerando não apenas sua contribuição na decisão final do modelo, mas também possíveis variações relacionadas a estágios da doença, lateralização hemisférica e respostas a estímulos específicos. Abordagens que integrem técnicas avançadas de processamento de sinais e aprendizado de máquina poderão aprimorar ainda mais a interpretação dos dados e a aplicabilidade clínica dos modelos desenvolvidos.

REFERÊNCIAS

- AGGARWAL, C. C. et al. *Neural networks and deep learning*. [S.l.]: Springer, 2018. v. 10.
- AHMAD, A.; TARIQ, A.; HUSSAIN, H. K.; GILL, A. Y. Revolutionizing healthcare: How deep learning is poised to change the landscape of medical diagnosis and treatment. *Journal of Computer Networks, Architecture and High Performance Computing*, v. 5, n. 2, p. 458–471, 2023.
- ALSHARABI, K.; SALAMAH, Y. B.; ABDURRAQUEEB, A. M.; ALJALAL, M.; ALTURKI, F. A. Eeg signal processing for alzheimer's disorders using discrete wavelet transform and machine learning approaches. *IEEE Access*, v. 10, p. 89781–89797, 2022.
- CARDOSO, M. J.; LI, W.; BROWN, R.; MA, N.; KERFOOT, E.; WANG, Y.; MURREY, B.; MYRONENKO, A.; ZHAO, C.; YANG, D.; NATH, V.; HE, Y.; XU, Z.; HATAMIZADEH, A.; MYRONENKO, A.; ZHU, W.; LIU, Y.; ZHENG, M.; TANG, Y.; YANG, I.; ZEPHYR, M.; HASHEMIAN, B.; ALLE, S.; DARESTANI, M. Z.; BUDD, C.; MODAT, M.; VERCAUTEREN, T.; WANG, G.; LI, Y.; HU, Y.; FU, Y.; GORMAN, B.; JOHNSON, H.; GENEREAUX, B.; ERDAL, B. S.; GUPTA, V.; DIAZ-PINTO, A.; DOURSON, A.; MAIER-HEIN, L.; JAEGER, P. F.; BAUMGARTNER, M.; KALPATHY-CRAMER, J.; FLORES, M.; KIRBY, J.; COOPER, L. A. D.; ROTH, H. R.; XU, D.; BERICAT, D.; FLOCA, R.; ZHOU, S. K.; SHUAIB, H.; FARAHANI, K.; MAIER-HEIN, K. H.; AYLWARD, S.; DOGRA, P.; OURSELIN, S.; FENG, A. *MONAI: An open-source framework for deep learning in healthcare*. 2022. Disponível em: <<https://arxiv.org/abs/2211.02701>>.
- CASSANI, R.; ESTARELLAS, M.; SAN-MARTIN, R.; FRAGA, F. J.; FALK, T. H. Systematic review on resting-state eeg for alzheimer's disease diagnosis and progression assessment. *Disease Markers*, v. 2018, n. 1, p. 5174815, 2018. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1155/2018/5174815>>.
- CONG, S.; ZHOU, Y. A review of convolutional neural network architectures and their optimizations. *Artificial Intelligence Review*, v. 56, p. 1–65, 06 2022.
- DAS, A.; RAD, P. *Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey*. 2020. Disponível em: <<https://arxiv.org/abs/2006.11371>>.
- GOMES, J.; RODRIGUES, M.; SANTOS, W. D. Asteri: image-based representation of eeg signals for motor imagery classification. *Research on Biomedical Engineering*, v. 38, 04 2022.
- GOMES, J. C. *ASTERI: representação de sinais de eletroencefalografia em imagens para utilização de arquiteturas profundas baseadas em CNN*. Tese (Doutorado) — Universidade Federal de Pernambuco, 2023.
- GOODFELLOW, I. *Deep learning*. [S.l.]: MIT press, 2016. v. 196.
- GOUW, A. A.; ALSEMA, A. M.; TIJMS, B. M.; BORTA, A.; SCHELTENS, P.; STAM, C. J.; van der Flier, W. M. Eeg spectral analysis as a putative early prognostic biomarker in nondemented, amyloid positive subjects. *Neurobiology of Aging*, v. 57, p. 133–142, 2017. ISSN 0197-4580. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0197458017301847>>.

GRANDINI, M.; BAGLI, E.; VISANI, G. Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*, 2020.

GUNNING, D.; STEFIK, M.; CHOI, J.; MILLER, T.; STUMPF, S.; YANG, G.-Z. Xai—explainable artificial intelligence. *Science Robotics*, v. 4, n. 37, p. eaay7120, 2019. Disponível em: <<https://www.science.org/doi/abs/10.1126/scirobotics.aay7120>>.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 770–778.

HOMAN, R. W.; HERMAN, J.; PURDY, P. Cerebral location of international 10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*, v. 66, n. 4, p. 376–382, 1987. ISSN 0013-4694. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0013469487902069>>.

HUFF, D. T.; WEISMAN, A. J.; JERAJ, R. Interpretation and visualization techniques for deep learning models in medical imaging. *Physics in Medicine Biology*, IOP Publishing, v. 66, n. 4, p. 04TR01, feb 2021. Disponível em: <<https://dx.doi.org/10.1088/1361-6560/abed17>>.

JOSE, J.; GAUTAM, N.; TIWARI, M.; TIWARI, T.; SURESH, A.; SUNDARARAJ, V.; REJEESH, M. An image quality enhancement scheme employing adolescent identity search algorithm in the nsst domain for multimodal medical image fusion. *Biomedical Signal Processing and Control*, Elsevier, v. 66, p. 102480, 2021.

KAUL, D.; RAJU, H.; TRIPATHY, B. Deep learning in healthcare. *Deep Learning in Data Analytics: Recent Techniques, Practices and Applications*, Springer, p. 97–115, 2022.

KHAN, S.; NASEER, M.; HAYAT, M.; ZAMIR, S. W.; KHAN, F. S.; SHAH, M. Transformers in vision: A survey. *ACM computing surveys (CSUR)*, ACM New York, NY, v. 54, n. 10s, p. 1–41, 2022.

KIM, D.; HEO, B.; HAN, D. Densenets reloaded: Paradigm shift beyond resnets and vits. In: *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part III*. Berlin, Heidelberg: Springer-Verlag, 2024. p. 395–415. ISBN 978-3-031-72645-3. Disponível em: <https://doi.org/10.1007/978-3-031-72646-0_23>.

KIM, J.; JEONG, M.; STILES, W. R.; CHOI, H. S. Neuroimaging modalities in alzheimer's disease: Diagnosis and clinical features. *International Journal of Molecular Sciences*, v. 23, n. 11, 2022. ISSN 1422-0067. Disponível em: <<https://www.mdpi.com/1422-0067/23/11/6079>>.

KNOPMAN, D.; AMIEVA, H.; PETERSEN, R.; CHÉTELAT, G.; HOLTZMAN, D.; HYMAN, B.; NIXON, R.; JONES, D. Alzheimer disease. *Nature Reviews Disease Primers*, Nature Publishing Group, v. 7, n. 1, dez. 2021. ISSN 2056-676X. Publisher Copyright: © 2021, Springer Nature Limited.

KO, D.-W.; YANG, J.-J. Eeg-based schizophrenia diagnosis through time series image conversion and deep learning. *Electronics*, MDPI, v. 11, n. 14, p. 2265, 2022.

LANE, C. A.; HARDY, J.; SCHOTT, J. M. Alzheimer's disease. *European Journal of Neurology*, v. 25, n. 1, p. 59–70, 2018. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/ene.13439>>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, maio 2015.

LIN, T.; WANG, Y.; LIU, X.; QIU, X. A survey of transformers. *AI open*, Elsevier, v. 3, p. 111–132, 2022.

LISS, J.; ASSUNÇÃO, S. S.; CUMMINGS, J.; ATRI, A.; GELDMACHER, D.; CANDELA, S.; DEVANAND, D.; FILLIT, H.; SUSMAN, J.; MINTZER, J. et al. Practical recommendations for timely, accurate diagnosis of symptomatic alzheimer's disease (mci and dementia) in primary care: a review and synthesis. *Journal of internal medicine*, Wiley Online Library, v. 290, n. 2, p. 310–334, 2021.

LIU, Z.; MAO, H.; WU, C.-Y.; FEICHTENHOFER, C.; DARRELL, T.; XIE, S. A ConvNet for the 2020s. In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 2022. p. 11966–11976. Disponível em: <<https://doi.ieeecomputersociety.org/10.1109/CVPR52688.2022.01167>>.

LUBBEN, N.; ENSINK, E.; COETZEE, G. A.; LABRIE, V. The enigma and implications of brain hemispheric asymmetry in neurodegenerative diseases. *Brain Communications*, v. 3, n. 3, p. fcab211, 09 2021. ISSN 2632-1297. Disponível em: <<https://doi.org/10.1093/braincomms/fcab211>>.

MAAZ, M.; SHAKER, A.; CHOLAKKAL, H.; KHAN, S.; ZAMIR, S. W.; ANWER, R. M.; KHAN, F. S. Edgenext: Efficiently amalgamated cnn-transformer architecture for mobile vision applications. In: KARLINSKY, L.; MICHAELI, T.; NISHINO, K. (Ed.). *Computer Vision – ECCV 2022 Workshops*. Cham: Springer Nature Switzerland, 2023. p. 3–20. ISBN 978-3-031-25082-8.

MILTADOUS, A.; GIONANIDIS, E.; TZIMOURTA, K. D.; GIANNAKEAS, N.; TZALLAS, A. T. Dice-net: A novel convolution-transformer architecture for alzheimer detection in eeg signals. *IEEE Access*, v. 11, p. 71840–71858, 2023.

MILTADOUS, A.; TZIMOURTA, K. D.; AFRANTOU, T.; IOANNIDIS, P.; GRIGORIADIS, N.; TSALIKAKIS, D. G.; ANGELIDIS, P.; TSIPOURAS, M. G.; GLAVAS, E.; GIANNAKEAS, N.; TZALLAS, A. T. A dataset of scalp eeg recordings of alzheimer's disease, frontotemporal dementia and healthy subjects from routine eeg. *Data*, v. 8, n. 6, 2023. ISSN 2306-5729. Disponível em: <<https://www.mdpi.com/2306-5729/8/6/95>>.

MINDERER, M.; DJOLONGA, J.; ROMIJNDERS, R.; HUBIS, F.; ZHAI, X.; HOULSBY, N.; TRAN, D.; LUCIC, M. *Revisiting the Calibration of Modern Neural Networks*. 2021. Disponível em: <<https://arxiv.org/abs/2106.07998>>.

MISHRA, S.; DASH, A.; JENA, L. Use of deep learning for disease detection and diagnosis. *Bio-inspired neurocomputing*, Springer, p. 181–201, 2021.

MIZRAK, H. G.; DIKMEN, M.; HANOĞLU, L.; ŞAKUL, B. U. Investigation of hemispheric asymmetry in alzheimer's disease patients during resting state revealed by fNIRS. *Scientific Reports*, v. 14, n. 1, p. 13454, jun. 2024.

NAIK, N.; HAMEED, B. M. Z.; SHETTY, D. K.; SWAIN, D.; SHAH, M.; PAUL, R.; AGGARWAL, K.; IBRAHIM, S.; PATIL, V.; SMRITI, K.; SHETTY, S.; RAI, B. P.; CHLOSTA, P.; SOMANI, B. K. Legal and ethical consideration in artificial intelligence in healthcare:

Who takes responsibility? *Frontiers in Surgery*, v. 9, 2022. ISSN 2296-875X. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fsurg.2022.862322>>.

NAZ, M.; SHAH, J.; KHAN, M.; SHARIF, M.; RAZA, M.; DAMAŠEVIČIUS, R. From ecg signals to images: a transformation based approach for deep learning. *PeerJ Computer Science*, v. 7, 02 2021.

NICHOLS, E.; STEINMETZ, J. D.; VOLLSET, S. E.; FUKUTAKI, K.; CHALEK, J.; ABD-ALLAH, F.; ABDOLI, A.; ABUALHASAN, A.; ABU-GHARBIEH, E.; AKRAM, T. T. et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. *The Lancet Public Health*, Elsevier, v. 7, n. 2, p. e105–e125, 2022.

NIELSEN, M. *Neural Networks and Deep Learning*. [S.l.]: Determination Press, 2015.

OMAR, Y. M.; PLAPPER, P. A survey of information entropy metrics for complex networks. *Entropy*, v. 22, n. 12, 2020. ISSN 1099-4300. Disponível em: <<https://www.mdpi.com/1099-4300/22/12/1417>>.

ORBAN, M.; ELSAMANTY, M.; GUO, K.; ZHANG, S.; YANG, H. A review of brain activity and eeg-based brain–computer interfaces for rehabilitation application. *Bioengineering*, v. 9, n. 12, 2022. ISSN 2306-5354. Disponível em: <<https://www.mdpi.com/2306-5354/9/12/768>>.

PIETRZAK, K.; CZARNECKA, K.; MIKICIUK-OLASIK, E.; SZYMAŃSKI, P. New perspectives of alzheimer disease diagnosis - the most popular and future methods. *Medicinal chemistry (Sharjah (United Arab Emirates))*, v. 14 1, p. 34–43, 2018. Disponível em: <<https://api.semanticscholar.org/CorpusID:207315130>>.

PORSTEINSSON, A. P.; ISAACSON, R.; KNOX, S.; SABBAGH, M. N.; RUBINO, I. Diagnosis of early alzheimer's disease: clinical practice in 2021. *The journal of prevention of Alzheimer's disease*, Springer, v. 8, p. 371–386, 2021.

PURI, D.; NALBALWAR, S.; NANDGAONKAR, A.; WAGH, A. Eeg-based diagnosis of alzheimer's disease using kolmogorov complexity. In: IYER, B.; GHOSH, D.; BALAS, V. E. (Ed.). *Applied Information Processing Systems*. Singapore: Springer Singapore, 2022. p. 157–165. ISBN 978-981-16-2008-9.

PUTHANKATTIL, S. D.; JOSEPH, P.; U, R.; LIM, C. Eeg signal analysis: a survey. *Journal of medical systems*, v. 34, p. 195–212, 04 2010.

RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M.; BERG, A. C.; FEI-FEI, L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211–252, 2015.

RUSSELL; NORVIG, S. Artificial intelligence: A modern approach. *Prentice Hall, Englewood Cliffs, NJ*, 01 2010.

RUSSELL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3rd. ed. USA: Prentice Hall Press, 2009. ISBN 0136042597.

- RYALI, C.; HU, Y.-T.; BOLYA, D.; WEI, C.; FAN, H.; HUANG, P.-Y.; AGGARWAL, V.; CHOWDHURY, A.; POURSAEED, O.; HOFFMAN, J.; MALIK, J.; LI, Y.; FEICHTENHOFER, C. Hiera: a hierarchical vision transformer without the bells-and-whistles. In: *Proceedings of the 40th International Conference on Machine Learning*. [S.l.]: JMLR.org, 2023. (ICML'23).
- SECINARO, S.; CALANDRA, D.; SECINARO, A.; MUTHURANGU, V.; BIANCONE, P. The role of artificial intelligence in healthcare: a structured literature review. *BMC Medical Informatics and Decision Making*, v. 21, 04 2021.
- SIMONYAN, K.; VEDALDI, A.; ZISSERMAN, A. Deep inside convolutional networks: Visualising image classification models and saliency maps. In: *Workshop at International Conference on Learning Representations*. [S.l.: s.n.], 2014.
- SINGH, P.; PANDEY, P.; MIYAPURAM, K.; RAMAN, S. Eeg2image: Image reconstruction from eeg brain signals. In: *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2023. p. 1–5.
- SOUZA, G. M.; ALBUQUERQUE, L. V. S.; BRITO, N. C.; ARAUJO, Y. C. S.; GOMES, J. C.; MORENO, G. M. M.; SANTOS, W. P. Predição e identificação da diabetes a partir da análise de parâmetros clínicos e laboratoriais utilizando aprendizado de máquina. In: *Anais do VII Simpósio de Inovação em Engenharia Biomédica - SABIO 2023*. Recife: [s.n.], 2024. p. 67–73. Trabalho apresentado no VII Simpósio de Inovação em Engenharia Biomédica - SABIO, 2023.
- SOUZA, G. M.; GUSMAO, C. M. G.; GOMES, J. C.; SANTOS, W. P. O uso de deep learning aplicado à imagética motora: uma revisão. In: *Anais do VII Simpósio de Inovação em Engenharia Biomédica - SABIO 2023*. Recife: [s.n.], 2024. p. 89–92. Trabalho apresentado no VII Simpósio de Inovação em Engenharia Biomédica - SABIO, 2023.
- STEWART, J. *Calculus: early transcendentals*. 7th. ed. [S.l.]: Cengage Learning, 2012.
- SUNDARARAJAN, M.; TALY, A.; YAN, Q. Axiomatic attribution for deep networks. In: *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. [S.l.]: JMLR.org, 2017. (ICML'17), p. 3319–3328.
- TAWHID, M. N. A.; SIJLY, S.; WANG, H.; WHITTAKER, F.; WANG, K.; ZHANG, Y. A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from eeg. *Plos one*, Public Library of Science San Francisco, CA USA, v. 16, n. 6, p. e0253094, 2021.
- TEPLAN, M. Fundamental of eeg measurement. *MEASUREMENT SCIENCE REVIEW*, v. 2, 01 2002.
- TOBON, D. P.; HOSSAIN, M. S.; MUHAMMAD, G.; BILBAO, J.; SADDIK, A. E. Deep learning in multimedia healthcare applications: a review. *Multimedia systems*, Springer, v. 28, n. 4, p. 1465–1479, 2022.
- VASU, P. K. A.; GABRIEL, J.; ZHU, J.; TUZEL, O.; RANJAN, A. MobileOne: An Improved One millisecond Mobile Backbone . In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 2023. p. 7907–7917. Disponível em: <<https://doi.ieeecomputersociety.org/10.1109/CVPR52729.2023.00764>>.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. Attention is all you need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates Inc., 2017. (NIPS'17), p. 6000–6010. ISBN 9781510860964.

WANG, Z.; ZHAO, W.; DU, W.; LI, N.; WANG, J. Data-driven fault diagnosis method based on the conversion of erosion operation signals into images and convolutional neural network. *Process Safety and Environmental Protection*, v. 149, p. 591–601, 2021. ISSN 0957-5820. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957582021001300>>.

YAMASHITA, R.; NISHIO, M.; DO, R.; TOGASHI, K. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, v. 9, 06 2018.

YU, K.-H.; BEAM, A.; KOHANE, I. Artificial intelligence in healthcare. *Nature Biomedical Engineering*, v. 2, 10 2018.

ZEILER, M. D.; FERGUS, R. Visualizing and understanding convolutional networks. In: FLEET, D.; PAJDLA, T.; SCHIELE, B.; TUYTELAARS, T. (Ed.). *Computer Vision – ECCV 2014*. Cham: Springer International Publishing, 2014. p. 818–833. ISBN 978-3-319-10590-1.

ZHANG, J.; KAILKHURA, B.; HAN, T. Y.-J. Mix-n-match : Ensemble and compositional methods for uncertainty calibration in deep learning. In: III, H. D.; SINGH, A. (Ed.). *Proceedings of the 37th International Conference on Machine Learning*. PMLR, 2020. (Proceedings of Machine Learning Research, v. 119), p. 11117–11128. Disponível em: <<https://proceedings.mlr.press/v119/zhang20k.html>>.

APÊNDICE A – ARTIGO DE REVISÃO DE ESCOPO

VII Simpósio de Inovação em
Engenharia Biomédica

O Uso de Deep Learning aplicado à Imagética Motora: uma revisão

Gabriel Miranda de Souza, Universidade Federal de Pernambuco (UFPE) (gabriel.msouza@ufpe.br)

Cristine Martins Gomes de Gusmão, Universidade Federal de Pernambuco (UFPE)

(cristine.gusmao@ufpe.br)

Wellington Pinheiro dos Santos, Universidade Federal de Pernambuco (UFPE) (wellington.santos@ufpe.br)

Abstract—Neurological trauma patients suffer from limitations in their daily lives due to motor impairment. However these patients still produce neural signals which can be acquired through the use of electroencephalogram (EEG). Motor Imagery (MI) is the act of imagining motor movement without performing the movement itself. Through the analysis of the EEG signals obtained during the MI exercises it is possible to, for instance, control a robot hand or a part of an exoskeleton, creating a brain-computer-interface (BCI) environment which could help trauma patients with daily activities. The usage of deep learning (DL) techniques has increased in the last years following its rapid growth and improvement. These algorithms can help the analysis of EEG signals with high accuracy. We performed a scoping review to access and evaluate the current state of deep learning applied to MI problems in the state of art. We found extensive usage of CNNs models as preferred architecture due to its high performance. However CNNs come with high computational costs which can cause it to be inadequate in real-time applications.

Keywords—Deep Learning, Motor Imagery, EEG, BCI

I. INTRODUÇÃO

Pacientes com algum tipo de trauma neurológico ou deficiência motora possuem um determinado grau de limitação em suas independências. Em muitos desses casos, apesar da perda da mobilidade motora, sinais cerebrais são gerados em regiões do córtex a partir da imaginação desses movimentos [1]. Sinais desse tipo podem ser monitorados pela técnica de eletroencefalografia, que pode ser não invasiva e portátil. Nesse sentido, a tradução desses sinais cerebrais é útil e promissora no controle de Interfaces Cérebro-Máquina (BCI) [2].

A Imagética Motora por Eletroencefalografia (IM-EEG) é uma das técnicas atualmente disponíveis na tradução de sinais cerebrais gerados a partir da imaginação do movimento. Nesse contexto, técnicas *deep learning* têm se tornado opções viáveis no problema de classificação de Imagética Motora. Porém, esses tipos de algoritmos estão associados a altos custos e complexidade computacional, além de exigir muita memória em aplicações em tempo real, limitando a sua usabilidade em determinados casos [3]-[5]. Há, portanto, uma necessidade no aprofundamento dos estudos em *deep learning* aplicados ao problema de Imagética Motora, fazendo-se necessário saber quais avanços têm sido feitos na área e quais limitações ainda estão sendo enfrentadas. Para isso, há 2 abordagens metodológicas principais que podem ser utilizadas: Revisão sistemática; Revisão de escopo.

A principal diferença entre as revisões sistemática e a de escopo consiste em que a primeira está mais interessada em responder uma questão específica, enquanto a segunda pretende dar uma visão mais geral sobre determinado tópico. O objetivo da revisão sistemática é buscar e organizar evidência acerca de uma determinada questão específica e sintetizar seus achados a fim de gerar e corroborar práticas e políticas públicas [6]. Já a revisão de escopo tem como objetivo identificar e mapear de maneira geral o que tem sido feito numa área, podendo identificar possíveis lacunas de conhecimento no tópico de interesse. O presente trabalho visa a partir de uma revisão de escopo identificar, entender e avaliar os últimos avanços feitos na utilização de técnicas de *deep learning* na área de Imagética Motora baseada em EEG, e quais limitações ainda são observadas nesse contexto.

O restante deste trabalho está organizado da seguinte maneira: Na seção 2 está descrita a metodologia do trabalho. Na seção 3 são apresentados os principais resultados encontrados, juntamente com uma discussão sobre suas respectivas implicações. Por fim, na seção 4 são feitas as considerações finais e perspectivas para trabalhos futuros.

II. MATERIAIS E MÉTODOS

A construção da metodologia deste trabalho consiste sumariamente na escolha das ferramentas de busca por artigos, na definição dos critérios de inclusão e exclusão, e por fim na leitura completa dos artigos selecionados para posterior discussão. O detalhamento dessas etapas será apresentado nas seções a seguir.

A. Ferramentas de Busca

De acordo com a análise de diversas ferramentas de busca realizada em [7], é possível dividir as ferramentas em Principais e Suplementares. O estudo discute um total de 27 critérios que consideram importantes que uma ferramenta de busca tenha implementado. Os autores separam os critérios como "Necessários" e "Desejáveis". Todos os critérios classificados como Necessários são considerados imprescindíveis e indispensáveis para que uma ferramenta de busca de artigos seja confiável e adequada numa revisão sistemática. No total há 12 critérios Necessários, e os demais desejáveis. Para que uma ferramenta seja considerada Principal, ela precisa atender aos 12 critérios Necessários propostos, caso contrário ela é considerada Suplementar. Das 28 ferramentas testadas, 14 foram

VII Simpósio de Inovação em Engenharia Biomédica



consideradas Principais, atendendo a todos os critérios Necessários discutidos no artigo.

Apesar do objetivo do presente trabalho não ser a realização de uma revisão sistemática, e sim uma revisão de escopo, foram considerados e ponderados os critérios discutidos em [7] para a escolha das ferramentas de busca utilizadas. Foram escolhidas 6 ferramentas Principais: ACM Digital Library; Cochrane Library; OVID; PubMed; ScienceDirect; Wiley Online Library. E 1 ferramenta Suplementar: IEEE Xplore. A escolha se deu priorizando as ferramentas Primárias de busca, buscando abranger o retorno de artigos em bases de dados na área da computação (ACM Digital Library, IEEE Xplore), da saúde (Cochrane, OVID, PubMed), e multidisciplinares (ScienceDirect, Wiley Online Library).

B. Critérios de Inclusão

A inclusão de artigos foi feita a partir da busca por termos específicos no título dos artigos. Todos os artigos, de 2019 em diante, retornados a partir da busca foram incluídos. Foram selecionados 3 termos julgados pertinentes para retornar artigos relevantes na área de interesse. Os termos "Deep Learning", "Motor Imagery" e um terceiro termo que consistia em "EEG", "Electroencephalogram" ou "Electroencephalography", pois, como a busca estava sendo realizada pelos termos exatos entre aspas, pretendia-se garantir que possíveis variações da escrita do termo EEG fossem encontradas. A sintaxe da *query* utilizada para realizar a busca pode ser utilizada como: "Title:('Deep learning' AND 'Motor Imagery' AND ('EEG' OR 'Electroencephalography' OR 'Electroencephalogram'))"

C. Critérios de Exclusão

Foram definidos os seguintes critérios de exclusão:

- Artigos escritos em idioma além do inglês;
- Artigos de Conferências, teses, dissertações e preprints;
- Análise subjetiva a partir da leitura do *abstract*, caso fosse deliberado como de pouca relevância para o presente trabalho por motivos variados, o artigo seria então excluído.

III. RESULTADOS E DISCUSSÃO

A quantidade total de artigos retornados por cada ferramenta de busca foi a seguinte: 3 artigos ACM Digital Library; 0 artigos da Cochrane Library; 14 artigos da IEEE Xplore; 12 artigos da PubMed; 10 artigos da OVID; 4 artigos da ScienceDirect; 1 artigo da Wiley Online Library. O fluxograma da Fig 1 sintetiza as etapas da metodologia até aqui e detalha quantos artigos foram excluídos e por quais motivos durante o processo de busca. Por fim, os artigos incluídos remanescentes foram lidos na íntegra.

As bases de dados utilizadas nos 12 artigos lidos são diversas. A mais utilizada é a base BCI Competition IV utilizada em metade (6) dos trabalhos. Essa base pode ser subdividida em que as mais utilizadas são as BCI Competition IV datasets 2a e 2b, que são utilizadas em três

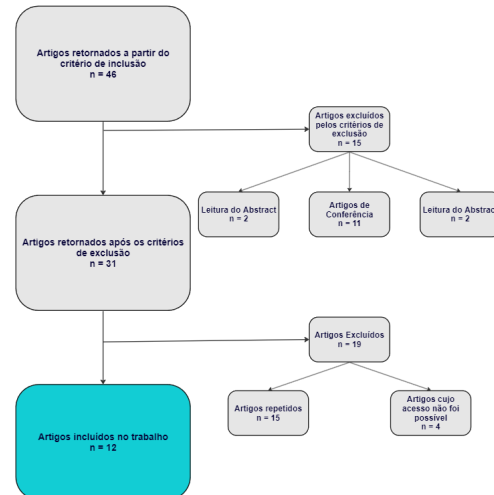


Fig. 1. Fluxograma da metodologia empregada na aplicação dos critérios de inclusão e exclusão do trabalho.

trabalhos cada. E a BCI Competition IV dataset 1 que é utilizada em 1 (um) trabalho. Sete artigos utilizam apenas uma (1) base de dados e cinco artigos utilizam duas (2) bases cada. Sete artigos (58,33%) utilizam sua própria base de dados unicamente ou em conjunto com uma segunda base de dados pública.

A maioria dos autores utilizam a função de ativação *softmax* ou a *Rectified Linear Unit Layer* (ReLU), e alguns utilizam a *Exponential Linear Units* (ELU). O método de otimização de gradiente mais utilizado é o Adam. A maioria dos autores utilizam como entrada para as redes os sinais de EEG dispostos em matrizes 2D. Porém há algumas aplicações, como mostrado em [8]-[10], em que os sinais de EEG são transformados em imagens antes de serem alimentados às redes. Os sinais podem ser transformados em imagens a partir de uma técnica como a *Quadratic Time-Frequency Distribution* (QTFD) utilizada em [8], ou o método de *Sparse Spectrotemporal Decomposition* (SSD) proposto em [9] para forçar a esparsidade de sinais de EEG no plano tempo-frequência, melhorando os *Event-Related Desynchronization/Synchronization* (ERD/ERS) observados.

A métrica mais utilizada como avaliação dos modelos é a acurácia, que encontra-se presente em 91,67% dos artigos avaliados, apenas 1 dos 12 totais não utilizou acurácia mas sim o índice kappa como métrica de avaliação de performance dos modelos. A segunda métrica mais utilizada foi o índice kappa, presente em 33,33% dos artigos (4/12). Algumas outras métricas também foram reportadas em alguns casos: *f1-score* em 16,67% (2/12); sensibilidade, especificidade e precisão apareceram em 1 artigo cada. Nem todos artigos reportaram algum tipo de medida ou gráfico de dispersão de valores. A mais utilizada foi o desvio padrão, que foi informado em metade dos artigos lidos (6/12).

VII Simpósio de Inovação em Engenharia Biomédica



Gráficos do tipo *boxplots* foram utilizados em 3 artigos (25%).

Todos os artigos realizam pelo menos um tipo de comparação dos seus resultados obtidos com algum parâmetro de referência: alguns comparam e testam vários modelos de diferentes arquiteturas entre si, outros comparam as métricas de avaliação escolhidas com os resultados obtidos em outros trabalhos no estado da arte. Porém nem todos utilizam algum teste estatístico para comprovar que os resultados são realmente superiores ou inferiores aos que estão sendo comparados. Alguns utilizam apenas a média das acurácias para constatar que um método teve desempenho melhor que outro, sem informar o desvio padrão para poder ser verificado que os limites superior e inferior dos modelos não se sobrepõem. Do total, 58,33% (7/12) dos artigos fazem uso de algum teste estatístico com esse propósito. Os testes mais utilizados são o Teste dos postos sinalizados de Wilcoxon, Teste t de Student e o Teste de Friedman, os três aparecem em 3 artigos cada (25%). O Teste de Shapiro-Wilk e Análise de Variância (ANOVA) ambos aparecem em 2 artigos cada.

Em 8 dos 12 artigos (66,67%) foram reportados os tempos de treinamento das arquiteturas avaliadas. Alguns entram em mais detalhes e discutem também um pouco do impacto do tempo de treinamento e do custo computacional que algumas redes complexas possuem, destacando que podem haver alternativas mais simples e rápidas que retornem uma performance semelhante ou comparáveis àquelas mais robustas.

Ao todo 91,67% dos artigos (11/12) utilizam Convolutional Neural Networks (CNN) em foco principal, ou em parte, dentro dos seus respectivos trabalhos. O único artigo que não utiliza CNN é o trabalho desenvolvido em [11], em que são testados dois modelos *end-to-end* (e2e) distintos de *Generative Deep Learning* (GDL): o primeiro a partir de *Deep Belief Network* (DBN) e o segundo a partir de *Stacked Sparse Autoencoder* (SSAE). O trabalho mostra uma performance otimizada e competitiva com os resultados do estado da arte, tendo a necessidade de pouco a nenhum pré-processamento da base de IM, visto que essa etapa é comumente uma das mais custosas nesse contexto.

Todos os demais 11 artigos utilizam CNN em alguma etapa dos seus desenvolvimentos. É visto que há atualmente um consenso geral entre a eficácia e robustez das CNNs comparadas às demais redes de *deep learning*. O trabalho realizado em [12] testa a performance das CNNs em relação às redes de DL Feed-Forward Neural Networks (FFNNs) e Recurrent Neural Networks (RNNs). O trabalho também utiliza algoritmos genéticos como etapa de entrada para as redes de DL. Pontuam que modelos de CNNs com apenas 1 camada escondida podem ter desempenho igual ou superar a DBNs com 6 camadas escondidas. O trabalho conclui que em média as CNNs produzem os melhores resultados com porém de possuir maior custo computacional, o que pode não ser adequado a todas as aplicações.

As CNNs são populares nesse campo de estudo também devido à sua versatilidade de aplicações, podendo ser utilizadas como algoritmo de classificação ou como

algoritmo de seleção de atributos, atuando como etapa de pré-processamento dos sinais de EEG. Tal característica faz com que a utilização de CNNs auxilie e reduza a complexidade de etapas como principalmente de *feature engineering*. Porém [13] pontua que apesar da utilização de CNNs, a etapa de pré-processamento é essencial e aperfeiçoa o desempenho de alto nível já apresentado por essas redes.

Outra vantagem das CNNs discutida em [14] é que a sua utilização melhora a eficácia de classificação de sujeitos que possuem baixa eficiência na realização de exercícios de IM. O trabalho também compara a aplicação de uma CNN a uma técnica mais tradicional de *machine learning* (ML) baseada em Common Spatial Patterns (CSP) e Linear Discriminant Analysis (LDA). Os resultados apresentados mostram que a rede CNN supera significativamente a rede de ML tradicional, aumentando a acurácia de classificação de todos os sujeitos, principalmente daqueles cujos desempenhos nos exercícios de IM foram baixos.

O baixo desempenho ou eficiência na realização dos exercícios de IM discutidos no parágrafo anterior pode ser devido a diversos fatores. Um desses fatores é a dificuldade de abstração do que é imaginar o movimento mas não o realizar. Nesse sentido, os trabalhos realizados em [15][16] trazem uma maneira alternativa de aquisição dos sinais de EEG durante a realização dos exercícios de IM, a partir da utilização de um ambiente de realidade aumentada (VR). Lee et al [15] argumenta que, sendo possível ver a ação ou movimento que pretende-se imaginar auxilia no processo da IM. Com a adoção dessa técnica os voluntários tiveram uma melhora significativa nos seus desempenhos durante os exercícios de IM.

Vários autores sinalizam que são necessárias poucas camadas de CNNs em relação a outras redes de DL. Em [17] é pontuado que CNNs rasas podem ser melhores em aprender as características espaço-temporais específicas do EEG como a técnica de *Filter Bank Common Spatial Pattern* (FBCSP). Já CNNs profundas podem ser boas em extrair atributos genéricos dos sinais de EEG.

IV. CONCLUSÃO

Nos últimos anos houve um aumento grande da utilização de *deep learning* nos problemas de imagética motora, campo em que até outrora era dominado pelas técnicas tradicionais de CSP, LDA e FBCSP. O avanço das CNNs nos recentemente catapultou os avanços, utilizações e aplicações dessas redes mais robustas na decodificação e classificação de sinais de EEG. As CNNs e redes de DL no geral fornecem um alto desempenho e versatilidade num problema difícil de classificação, garantindo muitas das melhores performances e acurácias no estado da arte.

Porém, ainda há limitações associadas a essas redes, principalmente em relação aos seus custos e complexidade computacionais. Muitas dessas redes robustas são lentas, o que inviabilizaria diversas aplicações práticas em tempo

VII Simpósio de Inovação em Engenharia Biomédica



real, que é um dos principais objetivos do campo da imagética motora. Limitações como o problema discutido em [14] da ineficiência de alguns usuários de realizar a imaginação corretamente ainda precisam ser contornados. A utilização de VR durante a imaginação e aquisição dos sinais de EEG, apesar de ser um recurso criativo auxiliar, é inviável numa aplicação prática do dia a dia de pessoas com deficiência motora ou traumas neurológicos, fazendo-se necessária a busca por uma alternativa prática.

As perspectivas de trabalhos futuros visam a longo ou a curto prazo avaliar e contornar as limitações previamente mencionadas, buscando o aperfeiçoamento das diferentes técnicas e aplicações para uma melhor acurácia e menor custo computacional.

REFERÊNCIAS

- [1] J. Munzert, B. Krüger, and K. Zentgraf, "Cognitive motor processes: The role of motor imagery in the study of motor representations," *Brain Research Reviews*, vol. 60, pp. 306–326, 05 2009.
- [2] N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "Eeg-based brain-computer interfaces using motor-imagery: Techniques and challenges," *Sensors (Switzerland)*, vol. 19, 03 2019.
- [3] V. Novak, R. Sigrist, N. Gerig, D. Wyss, R. Bauer, U. Götze, and R. Riener, "Benchmarking brain-computer interfaces outside the laboratory: The cybathlon 2016," *Frontiers in Neuroscience*, vol. 11, 01 2018.
- [4] R. Ortner, B. Allison, G. Korisek, H. Gaggli, and G. Pfurtscheller, "An ssvcp bci to control a hand orthosis for persons with tetraplegia," *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 19, pp. 1–5, 09 2010.
- [5] K. Statthaler, A. Schwarz, D. Steyerl, R. Kobler, M. Höller, J. Brandstetter, L. Hehenberger, M. Bigga, and G. Müller-Putz, "Cybathlon experiences of the graz bci racing team mirage91 in the brain-computer interface discipline," *Journal of NeuroEngineering and Rehabilitation*, vol. 14, 12 2017.
- [6] Z. Munn, M. Peters, C. Stern, C. Tufanaru, A. Mearthur, and E. Aromataris, "Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach," *BMC Medical Research Methodology*, vol. 18, 11 2018.
- [7] M. Gusenbauer and N. Haddaway, "Which academic search systems are suitable for systematic reviews or meta-analyses? evaluating retrieval qualities of google scholar, pubmed and 26 other resources," *Research Synthesis Methods*, vol. 11, pp. 181–217, 03 2020.
- [8] R. Alazrai, M. Abuhijleh, H. Alwanni, and M. Daoud, "A deep learning framework for decoding motor imagery tasks of the same hand using eeg signals," *IEEE Access*, vol. 7, pp. 109 612–109 627, 08 2019.
- [9] B. Sun, X. Zhao, H. Zhang, R. Bai, and T. Li, "Eeg motor imagery classification with sparse spectrotemporal decomposition and deep learning," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 541–551, 2021.
- [10] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "Eeg classification of motor imagery using a novel deep learning framework," *Sensors*, vol. 19, p. 551, 01 2019.
- [11] A. Hassanpour, M. Moradikia, H. Adeli, R. Khayami, and P. Shamsinejad Babaki, "A novel end-to-end deep learning scheme for classifying multi-class motor imagery electroencephalography signals," *Expert Systems*, vol. 36, 12 2019.
- [12] J. Leon, J. Escobar, A. Ortiz, J. Ortega, J. González, P. Smith, J. Gan, and M. Damas, "Deep learning for eeg-based motor imagery classification: Accuracy-cost trade-off," *PLOS ONE*, vol. 15, p. e0234178, 06 2020.
- [13] H. Zhu, D. Forenzo, and B. He, "On the deep learning models for eeg-based brain-computer interface using motor imagery," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2283–2291, 2022.
- [14] N. Tibrewal, N. Leeuwis, and M. Alimardani, "Classification of motor imagery eeg using deep learning increases performance in inefficient bci users," *PLOS ONE*, vol. 17, p. e0268880, 07 2022.
- [15] P.-L. Lee, S.-H. Chen, T.-C. Chang, W.-K. Lee, H.-T. Hsu, and H.-H. Chang, "Continual learning of a transformer-based deep learning classifier using an initial model from action observation eeg data to online motor imagery classification," *Bioengineering*, vol. 10, p. 186, 02 2023.
- [16] D. Achannarayana and M. Hayashibe, "Decoding hand motor imagery tasks within the same limb from eeg signals using deep learning," *IEEE Transactions on Medical Robotics and Bionics*, vol. 2, no. 4, pp. 692–699, 2020.
- [17] S. Amin, M. Alsulaiman, G. Muhammad, M. Amine, and M. S. Hossain, "Deep learning for eeg motor imagery classification based on multi-layer cnns feature fusion," *Future Generation Computer Systems*, vol. 101, 07 2019.