



UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MÉTODO DE CLASSIFICAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA
PARA AUXÍLIO AO DIAGNÓSTICO DO TRANSTORNO DO ESPECTRO
AUTISTA

Josias Lira de Sousa Neto

Manaus

Fevereiro de 2024



UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MÉTODO DE CLASSIFICAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA
PARA AUXÍLIO AO DIAGNÓSTICO DO TRANSTORNO DO ESPECTRO
AUTISTA

Josias Lira de Sousa Neto

Dissertação de Mestrado apresentada a
Programa de Pós-Graduação em Engenharia
Elétrica, da Universidade Federal do
Amazonas, como parte dos requisitos
necessários à obtenção do título de Mestre
em Engenharia Elétrica.

Orientadores: Waldir Sabino da Silva Júnior
Silvania da Conceição Furtado

Manaus

Fevereiro de 2024

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S725m Sousa Neto, Josias Lira de
Método de classificação de sinais de eletroencefalograma para auxílio ao diagnóstico do Transtorno do Espectro Autista / Josias Lira de Sousa Neto . 2024
102 f.: il. color; 31 cm.

Orientador: Waldir Sabino da Silva Júnior
Coorientadora: Silvania da Conceição Furtado
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. EEG - eletroencefalograma . 2. TEA - transtorno do espectro autista. 3. CNN - redes neurais convolucionais. 4. RNN - redes neurais recorrentes. 5. Ordenação por diferença percentual. I. Silva Júnior, Waldir Sabino da. II. Universidade Federal do Amazonas III. Título



Ministério da Educação
Universidade Federal do Amazonas
Coordenação do Programa de Pós-Graduação em Engenharia Elétrica

ATESTES

Poder Executivo Ministério da Educação
Universidade Federal do Amazonas
Faculdade de Tecnologia
Programa de Pós-graduação em Engenharia Elétrica

Pós-Graduação em Engenharia Elétrica. Av. General Rodrigo Octávio Jordão Ramos, nº 3.000 - Campus Universitário, Setor Norte - Coroado, Pavilhão do CETELI. Fone/Fax (92) 99271-8954 Ramal:2607. E-mail: ppgee@ufam.edu.br

JOSIAS LIRA DE SOUSA NETO

MÉTODO DE CLASSIFICAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA PARA AUXÍLIO AO DIAGNÓSTICO DO TRANSTORNO DO ESPECTRO AUTISTA

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Aprovada em 29 de fevereiro de 2024.

BANCA EXAMINADORA

Prof. Dr. Waldir Sabino da Silva Júnior - Presidente
Prof. Dr. Eddie Batista de Lima Filho - Membro Titular 2 - Interno
Prof. Dr. Robson Luis Oliveira de Amorim - Membro Titular 1 - Externo

Manaus, 14 de março de 2024.



Documento assinado eletronicamente por **Waldir Sabino da Silva Júnior, Professor do Magistério Superior**, em 14/03/2024, às 14:11, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Eddie Batista de Lima Filho, Usuário Externo**, em 14/03/2024, às 16:44, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Robson Luis Oliveira de Amorim, Professor do Magistério Superior**, em 15/03/2024, às 16:40, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1962286** e o código CRC **FDD79E67**.

Av. Octávio Hamilton Botelho Mourão - Bairro Coroado 1 Campus Universitário Senador Arthur Virgílio Filho, Setor Norte - Telefone: (92) 3305-1181
CEP 69080-900 Manaus/AM - mestrado_engeletrica@ufam.edu.br

Referência: Processo nº 23105.006475/2024-12

SEI nº 1962286

Agradecimentos

- Primeiramente à Deus por conceder saúde e intelecto para a produção deste trabalho.
- À minha mãe Rosângela e ao meu pai Pedro por serem meus alicerces e maiores orientadores desde a infância, não medindo esforços na busca pelo meu crescimento pessoal e profissional.
- Aos meus irmãos Jessé, Priscila e Esli pelo companheirismo, camaradagem e estima de sempre.
- Aos meus avós, Ruth e Josias, e tios, Pedro e Elisa, pelo apoio e incentivo durante minha estada em Manaus.
- À Raquel, minha ex-namorada, pela paciência, carinho e encorajamento durante os momentos decisivos.
- Ao meu orientador Waldir Sabino, pela excelente didática, paciência, palavras de incentivo e por tornar o caminho mais claro e simples.
- Aos demais professores do curso que ofereceram todo o embasamento teórico necessários à realização desta pesquisa.
- Aos colegas do PPGEE pela parceria e ajuda mútua durante essa jornada e a todos que contribuíram direto e indiretamente para a conclusão desta dissertação.

Resumo da Dissertação apresentada à UFAM como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica

MÉTODO DE CLASSIFICAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA
PARA AUXÍLIO AO DIAGNÓSTICO DO TRANSTORNO DO ESPECTRO
AUTISTA

Josias Lira de Sousa Neto

Orientadores: Waldir Sabino da Silva Júnior

Silvania da Conceição Furtado

Programa de Pós-Graduação em Engenharia Elétrica

Pesquisas recentes indicam um crescente número de crianças diagnosticadas com o transtorno do espectro autista (TEA), doença caracterizada por sintomas que impactam diretamente os campos de comportamento, comunicação e interação social. Segundo o centro de controle de prevenção de doenças (CDC), 1 em cada 36 crianças de 8 anos são autistas nos Estados Unidos, número 22% maior em relação à pesquisa anterior, de 2021. Estes fatos têm impulsionado o desenvolvimento de novas ferramentas para o diagnóstico da doença, o qual é essencialmente clínico (observação comportamental e entrevista), uma vez que não há exigência de exame específico. O presente trabalho explora uma abordagem com CNN e RNN, combinadas com técnicas de pré-processamento de sinais, tal como a ordenação por diferença percentual, com o objetivo de auxiliar no diagnóstico do TEA, a partir da classificação de sinais de eletroencefalograma (EEG). Utilizando-se a base de dados dos estudos de Milne *et al.*, da Universidade de *Sheffield*, mostra-se que o melhor modelo proposto alcançou uma acurácia de classificação de 99,63%, o que corrobora a capacidade do sistema em distinguir entre indivíduos com TEA e tipicamente desenvolvidos.

Palavras-chave: EEG; Transtorno do espectro autista; CNN; RNN; Ordenação por diferença percentual.

Abstract of Dissertation presented to UFAM as a partial fulfillment of the requirements for the degree of Master in Eletrical Engineering

METHOD FOR CLASSIFYING ELECTROENCEPHALOGRAM SIGNALS TO
AID THE DIAGNOSIS OF AUTISTIC SPECTRUM DISORDER

Josias Lira de Sousa Neto

Advisors: Waldir Sabino da Silva Júnior

Silvania da Conceição Furtado

Postgraduate Program in Electrical Engineering

Recent research indicates a growing number of children diagnosed with autism spectrum disorder (ASD), a disease characterized by symptoms that directly impact the fields of behavior, communication and social interaction. According to the Centers for Disease Control and Prevention (CDC), 1 in every 36 8-year-old children are autistic in the United States, a number 22% higher than in the previous survey in 2021. These facts have driven the development of new tools to the diagnosis of the disease, which is essentially clinical (behavioral observation and interview), since there is no requirement for a specific exam. The present work explores an approach with CNN and RNN, combined with signal pre-processing techniques, such as ordering by percentage difference, with the aim of assisting in the diagnosis of ASD, based on the classification of electroencephalogram (EEG) signals. Using the database of studies by Milne *et al.*, from the University of Sheffield, it is shown that the best proposed model achieved a classification accuracy of 99,63%, which corroborates the system's ability to distinguish between individuals with ASD and typically developed.

Keywords: EEG; Autism spectrum disorder; CNN; RNN; Ordering by percentage difference.

Sumário

1	Introdução	1
1.1	Objetivos da Dissertação	3
2	Fundamentos Teóricos	5
2.1	Sinais de Eletroencefalograma	5
2.2	Transtorno do Espectro Autista	8
2.3	Inteligência Artificial, <i>Machine Learning</i> e <i>Deep Learning</i>	11
2.4	Redes Neurais Convolucionais (CNN)	13
2.5	Redes Neurais Recorrentes (RNN)	15
2.6	A técnica de ordenação por diferença percentual	18
2.7	Trabalhos Relacionados	19
3	Metodologia Proposta	27
3.1	Introdução	27
3.2	Pré-processamento	28
3.3	Treinamento e Execução do Classificador	32
4	Base de Dados	34
4.1	Base de dados de Milne	34
4.2	Outras bases de dados	36
4.3	Síntese das bases de dados	38
5	Experimentos e Resultados	39
5.1	Introdução	39
5.2	Setup	40
5.3	Métricas de Desempenho	41

5.4	Projeto dos algoritmos de classificação	42
5.4.1	Configuração da arquitetura CNN	42
5.4.2	Configuração da arquitetura RNN (LSTM)	44
5.5	Análise de Desempenho	44
5.5.1	Sistema de classificação com CNN	46
5.5.2	Sistema de classificação com CNN e técnica de ODP	52
5.5.3	Sistema de classificação com CNN sem pré-processamento	56
5.5.4	Sistema de classificação com RNN	60
5.5.5	Sistema de classificação com RNN e técnica de ODP	65
5.5.6	Sistema de classificação com RNN sem pré-processamento	69
5.5.7	Comparação de resultados com estudos anteriores de diagnóstico de TEA	73
6	Conclusões	74
6.1	Proposta para Trabalhos Futuros	75
	Referências Bibliográficas	77
A	Camadas de CNN	87
B	Tabelas de arquiteturas dos sistemas de classificação	90
C	Artigos Publicados	102
C.1	Artigos Diretamente Relacionados ao Tema	102

Lista de Figuras

2.1	Tipos de ondas EEG mais comuns para um estado normal.	7
2.2	Disposição de eletrodos para EEG na configuração 10 – 20.	8
2.3	Correlação entre os conceitos de Inteligência Artificial, Aprendizado de Máquina e Aprendizado Profundo.	12
2.4	Arquitetura LeNet-5.	14
2.5	Processo de convolução.	15
2.6	Arquitetura RNN padrão.	16
2.7	Célula LSTM.	17
2.8	Amostra de resultado da ordenação por diferença percentual: (a) sinal EEG bidimensional original e (b) sinal EEG bidimensional reordenado. 19	
3.1	Diagrama em blocos geral do sistema proposto: (a) treinamento do classificador, (b) execução do modelo treinado.	28
3.2	Diagrama em blocos de pré-processamento para o sistema de classificação TEA.	31
3.3	Diagrama expandido do bloco <i>treinamento do classificador</i>	33
4.1	Amostra de sinal EEG de 64 canais, no domínio do tempo, do indivíduo tipicamente desenvolvido n° 54, no intervalo 0 – 5 s.	35
4.2	Amostra de sinal EEG de 64 canais, no domínio da frequência, do indivíduo tipicamente desenvolvido n° 54, na banda de frequência de 0 – 60 Hz.	36
4.3	Amostra de sinal EEG de 64 canais, no domínio do tempo, do indivíduo diagnosticado com TEA n° 11, no intervalo 0 – 5 s.	37

4.4	Amostra de sinal EEG de 64 canais, no domínio da frequência, do indivíduo diagnosticado com TEA n° 11, na banda de frequência de 0 – 60 Hz.	38
5.1	Modelo de Arquitetura CNN	43
5.2	Modelo de arquitetura RNN (LSTM).	44
5.3	Diagrama em blocos geral do sistema proposto: (a) treinamento do classificador, (b) execução do modelo treinado.	45
5.4	Curvas de acurácia e erro de treinamento e validação do experimento n° 8 com CNN.	48
5.5	Matriz de confusão do experimento n° 8 com CNN.	48
5.6	Curva ROC do experimento n° 8 CNN.	49
5.7	Curvas de acurácia e erro de treinamento e validação do experimento n° 9 com CNN e técnica de ODP.	52
5.8	Matriz de confusão do experimento n° 9 com CNN e técnica de ODP.	53
5.9	Curva ROC do experimento n° 9 com CNN e técnica de ODP.	54
5.10	Curvas de acurácia e erro de treinamento e validação do experimento n° 3 com CNN sem pré-processamento.	56
5.11	Matriz de confusão do experimento n° 3 com CNN sem pré-processamento.	57
5.12	Curva ROC do experimento n° 3 com CNN sem pré-processamento.	58
5.13	Curvas de acurácia e erro de treinamento e validação do experimento n° 9 com RNN.	61
5.14	Matriz de confusão do experimento n° 9 RNN	62
5.15	Curva ROC do experimento n° 9 RNN	62
5.16	Curvas de acurácia e erro de treinamento e validação do experimento n° 7 com RNN e técnica de ODP.	65
5.17	Matriz de confusão do experimento n° 7 com RNN e técnica de ODP.	66
5.18	Curva ROC do experimento n° 7 com RNN e técnica de ODP.	67
5.19	Curvas de acurácia e erro de treinamento e validação do experimento n° 2 com RNN sem pré-processamento.	70
5.20	Matriz de confusão do experimento n° 2 com RNN sem pré-processamento.	70

5.21 Curva ROC do experimento n° 2 com RNN sem pré-processamento. . 71

Lista de Tabelas

2.1	Faixas de ondas de EEG	6
2.2	Alterações mais relevantes observadas nos cérebros de autistas	9
2.3	Tabela comparativa entre trabalhos relacionados.	26
4.1	Síntese das bases de dados	38
5.1	Resultados do sistema de classificação TEA com CNN, no conjunto de teste.	50
5.2	Parâmetros do sistema de classificação TEA com CNN.	51
5.3	Resultados do sistema de classificação TEA com CNN e técnica de ODP, no conjunto de teste.	55
5.4	Parâmetros do sistema de classificação TEA com CNN e técnica de ODP.	55
5.5	Resultados do sistema de classificação TEA com CNN sem pré-processamento, no conjunto de teste.	59
5.6	Parâmetros do sistema de classificação TEA com CNN sem pré-processamento.	59
5.7	Resultados do sistema de classificação TEA com RNN, no conjunto de teste.	63
5.8	Parâmetros do sistema de classificação TEA com RNN.	64
5.9	Resultados do sistema de classificação TEA com RNN e técnica de ODP, no conjunto de teste.	68
5.10	Parâmetros do sistema de classificação TEA com RNN e técnica de ODP.	68
5.11	Resultados do sistema de classificação TEA com RNN sem pré-processamento, no conjunto de teste.	72

5.12	Parâmetros do sistema de classificação TEA com RNN sem pré-processamento.	72
5.13	Comparação com estudos anteriores de diagnóstico de TEA	73
B.1	Arquiteturas dos sistemas de classificação TEA com CNN.	91
B.2	Arquiteturas dos sistemas de classificação TEA com CNN e técnica de ODP.	95
B.3	Arquiteturas dos sistemas de classificação TEA com CNN sem Pré-processamento.	97
B.4	Arquiteturas dos sistemas de classificação TEA com RNN.	98
B.5	Arquiteturas dos sistemas de classificação TEA com RNN e técnica de ODP.	100
B.6	Arquiteturas dos sistemas de classificação TEA com RNN sem pré-processamento	101

Capítulo 1

Introdução

Pesquisas recentes da principal referência mundial a respeito da prevalência de autismo, o centro de controle de prevenção de doenças dos Estados Unidos (*centers for disease control and prevention*, CDC), apontam que 1 em cada 36 crianças de 8 anos são autistas nos Estados Unidos [1], número 22% maior em relação à pesquisa anterior, de 2021. O transtorno do espectro autista (TEA) é caracterizado por diversos sintomas, que impactam diretamente os campos de comportamento, comunicação e interação social [2,3]. O diagnóstico é essencialmente clínico, ou seja, observação comportamental e entrevista, uma vez que não há exigência de exame específico [4]. Dessa forma, alternativas para o diagnóstico precoce são fundamentais para a mitigação de prejuízos no campo social, profissional, e demais áreas de desenvolvimento destes indivíduos.

Sinais biológicos são registros das atividades elétricas dos diversos sistemas fisiológicos do corpo humano [5]. Os principais e mais utilizados são o eletrocardiograma (ECG), o eletromiograma (EMG) e o eletroencefalograma (EEG). Tais sinais são de extrema importância, tanto para a medicina quanto para a engenharia, pois permitem monitorar o funcionamento dos seus respectivos sistemas biológicos. Além disso, alterações nos seus parâmetros podem indicar diferentes estados clínicos. Normalmente, a informação de interesse, em sinais biológicos, está permeada por elementos irrelevantes (e.g., ruído) [6,7], o que pode dificultar seu uso. Para eliminar/mitigar esses problemas, diversos sistemas de processamento de sinais biológicos têm sido desenvolvidos. Atualmente, há diversas aplicações que utilizam, direta ou indiretamente, sinais biológicos, como identificação de condições patoló-

gicas [5, 8–10] e compressão de sinais [11–13]. O sinal biológico de EEG tem sido bastante explorado pela comunidade científica [5, 7, 11, 12, 14–16]. Isso é devido ao amplo número de aplicações médicas que o sinal possui. Como os sinais de EEG são registros da atividade elétrica do cérebro [10], pode-se empregá-los no diagnóstico e tratamento de doenças neurológicas, tais como epilepsia, Parkinson e Alzheimer [14], entre outras. Outra aplicação clínica do sinal é o auxílio na avaliação de transtornos mentais, e.g., depressão e autismo [15, 16].

Normalmente, os sinais de EEG são capturados em múltiplas derivações, a fim de permitir uma representação de qualidade da atividade de diversos grupos neurais, a uma taxa de amostragem de 256 Hz com resolução de 16 bits. Esse processo é uma das etapas dos sistemas de processamento de sinais de EEG, a qual é denominada aquisição do sinal. Três outras etapas integram tais sistemas: pré-processamento, extração de características, e classificação. O principal objetivo do pré-processamento é eliminar ou reduzir informações irrelevantes presentes no sinal. A etapa de extração de características consiste na detecção e captura de padrões e regularidades do sinal de entrada. Os métodos comumente usados nessa etapa são análise de componentes principais (PCA), análise de componentes independentes (ICA) [17–20], e algoritmos de aprendizado profundo [21]. Na etapa de classificação, as características extraídas são exploradas para a estimação de um estado clínico. Exemplos de algoritmos de classificação são máquina de vetores suporte (SVM), k-vizinhos mais próximos (kNN) e *random forest* [19, 22].

Atualmente, algoritmos de aprendizado profundo vêm sendo utilizados, com certa frequência, como ferramentas de auxílio para o diagnóstico de doenças neurológicas [3, 23]. Essas técnicas permitem a análise automática de imagens médicas ou sinais fisiológicos coletados de pacientes, como os sinais EEG, e possuem a capacidade de extrair e classificar recursos de forma seletiva e inteligente. As redes neurais convolucionais (CNN) e recorrentes (RNN) são exemplos típicos de técnicas de aprendizado profundo, as quais podem investigar a existência de padrões em sinais de EEG que possibilitem a distinção entre indivíduos diagnosticados com TEA e tipicamente desenvolvidos (TD). A rede CNN possui destaque em aplicações em que se deseja classificar imagens médicas, enquanto a rede RNN possui a capacidade de reconhecer padrões em sinais temporais.

Nesta pesquisa, tais redes são associadas a diversas técnicas de pré-processamento de sinais, utilizando como entrada, sinais de EEG bidimensionais gerados a partir da concatenação de canais. Dentre elas destaca-se a técnica de ordenação por diferença percentual, que reordena os segmentos de canais em busca do aumento da correlação intersegmentos. Esta abordagem inovadora se mostrou eficiente na tarefa de classificação de sinais EEG, para a base de dados analisada, produzindo aumentos expressivos na capacidade e velocidade de aprendizagem dos modelos com CNN.

Esta dissertação está organizada da seguinte forma: No Capítulo 2, apresentam-se os fundamentos teóricos da dissertação, onde são abordados os conceitos relativos ao sinal trabalhado, ao transtorno, a técnicas de aprendizagem profunda e ao pré-processamento, bem como explanação de trabalhos relacionados. No Capítulo 3, é descrita a metodologia proposta com diagramas em blocos. No Capítulo 4, a base de dados utilizada e, em seguida, no Capítulo 5, os experimentos e resultados para cada sistema proposto. Por fim, no Capítulo 6, apresentam-se as conclusões da pesquisa e propostas para trabalhos futuros.

1.1 Objetivos da Dissertação

O objetivo principal deste trabalho é explorar técnicas de pré-processamento, em conjunto com métodos de aprendizado de máquina, para realizar a classificação de sinais de EEG para o diagnóstico do TEA.

Objetivos Específicos

- Conceber um sistema completo, que utilize técnicas de pré-processamento e aprendizado de máquina, para auxílio ao diagnóstico de indivíduos com TEA.
- Explorar técnicas de pré-processamento (inseridas no sistema completo), tais como reamostragem, filtros passa-alta e passa-baixa, remoção de artefatos, podagem com ICA e ordenação por diferença percentual, para avaliação de desempenho.
- Utilizar diferentes métodos de aprendizado de máquina: 2D-CNN com e sem

ODP, RNN-LSTM com e sem ODP, 2D-CNN com *transfer learning* e 2D-CNN com *autoencoder*, inseridas no sistema completo, para avaliação de desempenho.

Capítulo 2

Fundamentos Teóricos

2.1 Sinais de Eletroencefalograma

Vários órgãos presentes no corpo humano tendem a manifestar suas funções por meio de atividades elétricas. No cérebro, por exemplo, tem-se a produção de um sinal denominado EEG, que representa informações diversas das atividades funcionais. O EEG registrado nos potenciais são devidamente gerados por meio da condução iônica, responsável pela circulação ao longo do processo de excitação sináptica de dendritos de uma ampla gama de neurônios piramidais presentes entre as camadas do córtex cerebral [24].

De acordo com Sanei *et al.* [25], tais elementos representam cerca de 80% de toda a massa encefálica e está presente na região acima da superfície do crânio. Tem-se ainda a soma dos chamados potenciais pós-sinápticos (PSP) com o fluido extracelular. A condução destes agentes ocorre por meio de inúmeras camadas dos tecidos não neurais, contemplando a pele, a gordura, os ossos do crânio, os fluidos, as meninges até chegar nos eletrodos devidamente ligados no amplificador do EEG. O potencial elétrico gerado em cada um dos neurônios piramidais existentes no córtex é extremamente pequeno, por essa razão, a aquisição e a amplificação desses sinais, através de um dispositivo responsável por registrar o EEG, apenas ocorre quando há a ativação de milhares de neurônios do tipo subjacente em conjunto. Desta forma, conclui-se que os potenciais do tipo pós-sinápticos, que integram as atividades do EEG, são síncronos, além de conter uma duração suficiente para a separação das cargas responsáveis pela aproximação de dipolos.

Tabela 2.1: Faixas de ondas de EEG

Tipos de onda	Faixa de frequência
Delta (sono profundo)	0,5 - 4 Hz
Teta (estados do sono)	4 - 8 Hz
Alfa (estados de vigília)	8 - 13 Hz
Beta (córtex ativado)	13 - 30 Hz

Além disso, o EEG consiste em um tipo de registro de potencial elétrico que, normalmente, conta com uma amplitude inferior à $300\ \mu V$. A aquisição de biopotenciais ocorre por meio de eletrodos especializados capazes de executar a interface do corpo ou do organismo, além de serem responsáveis pela transdução de sinais isentos de artefatos e com níveis de ruído reduzidos [25]. As ondas cerebrais oscilam entre 0,5 e 100 Hz. Em comparação com outros tipos de biopotencial, como é o caso do eletrocardiograma, o EEG tende a ser mais complexo e de difícil interpretação, especialmente quando o observador conta com uma experiência reduzida [26]. Há grandes oscilações nos ritmos dos EEG, sendo que estes se correlacionam intimamente com o estado atual do indivíduo, tais como em vigília, sono, atenção, entre outros. Esses ritmos são estratificados segundo uma faixa de frequência, cada uma delas representada por uma letra grega [23]. A Tabela 2.1 apresenta as faixas de ondas de EEG que podem ser obtidas através de eletrodos de superfície posicionados no escalpo e a Figura 2.1, as formas de onda para as faixas mais comuns, em um estado normal.

O processo de registro de sinal de um EEG adota um padrão universal que inexistia antes de 1958. Antes, cada laboratório utilizava um padrão para coleta de sinais, contudo, passou-se a empregar padrões internacionais de disposição de eletrodos no couro cabeludo, sendo o mais aceito o do tipo 10 – 20. Essa disposição tem como principal intuito garantir que grande parte das funções cerebrais, como, processamento de raciocínio, emoção, sistema motor, funções cerebrais, entre outras, sejam obtidas de maneira adequada a partir dos canais [26].

O posicionamento de eletrodos utiliza uma relação igual a 10 ou 20% da distância total que existe entre os chamados marcos anatômicos da cabeça, utilizando-os como referência, algo que propicia a cobertura uniforme das regiões superficiais do cérebro. Destaca-se que os marcos anatômicos da cabeça são os pré-auriculares (situados nas laterais direitas e esquerdas da cabeça), a protuberância occipital

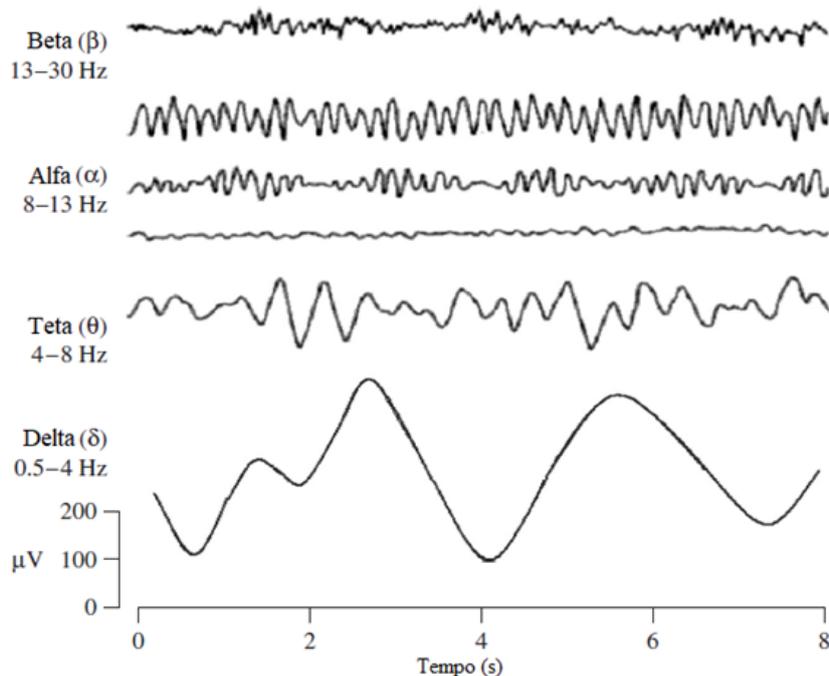


Figura 2.1: Tipos de ondas EEG mais comuns para um estado normal.
 Fonte: Adaptado de Sanei *et al.* [25]

(existente na região de trás da cabeça) e a nasal (localizada na parte frontal da cabeça) [24]. A Figura 2.2 permite a visualização das disposições dos eletrodos para EEG na configuração 10 – 20.

A nomenclatura dos eletrodos se dá segundo sua posição na cabeça, com isso, tem-se: *O* para o occipital, *T* para o temporal, *P* para o parietal, *C* para o central, *F* para o frontal e *Fp* para o frontal-polar. Os pontos presentes nas linhas médias do crânio podem ser indexados por meio da letra *z* que faz alusão ao zero, a exemplo dos canais *Pz* e *Fz*. No processo de determinação de posições faz-se o uso de dois pontos de referências: *inion*/occipital, presente na base do crânio (parte de trás da cabeça), e o *nasion*/frontal, situado na parte de topo de nariz, entre as sobrancelhas [25].

A distância entre os pontos descritos anteriormente é devidamente mensurada e, a partir disso, procede-se com a marcação de pontos na linha mediana do crânio, definidos como *Oz* e *Fpz*, os quais possuem uma distância para o *inion* e *nasion* de 10 %, respectivamente. Em relação aos pontos *Fz*, *Cz* e *Pz*, adota-se uma distância equivalente a 20 % [27]. Eletrodos pares como *T6*, *T4*, *P4*, entre outros, por exemplo, ficam situados no então hemisfério direito do crânio, enquanto os ímpares, como *T5*, *T3*, *P3*, no hemisfério esquerdo.

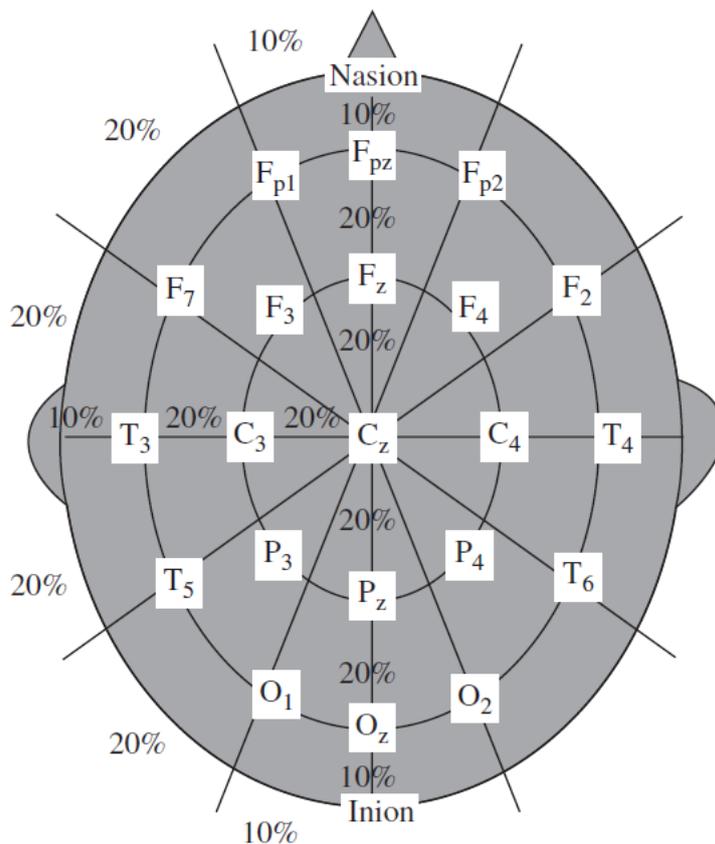


Figura 2.2: Disposição de eletrodos para EEG na configuração 10 – 20.
 Fonte: Adaptado de Sanei *et al.* [25].

2.2 Transtorno do Espectro Autista

O TEA corresponde a um transtorno de desenvolvimento global infantil que surge em crianças com menos de três anos de idade, contudo, se estende ao longo de toda a vida do indivíduo. É caracterizado por uma série de sintomas que impactam principalmente os campos de comportamento, de comunicação e de socialização, sendo a interação social a área mais afetada [28]. Existem diagnósticos médicos distintos que contemplam quadros de comportamento diversos, uma vez que há no autismo uma variação em grau de incidência de sintomas e de intensidade [29].

O conceito de autismo é relativamente novo e que tem exigido estudos contínuos. É comum utilizar o termo autismo para se referir a quaisquer variações do TEA que, em geral, circulam através da tríade de deficiências no campo de comportamento, de comunicação e social. De acordo com Whitman [30], as variações do TEA podem ser subdivididas em categorias que oscilam segundo o grau de comprometimento, indo do mais leve – com menor comprometimento – até o mais alto

– nível de comprometimento elevado. A Tabela 2.2 apresenta as alterações mais relevantes observadas nos cérebros de autistas.

Tabela 2.2: Alterações mais relevantes observadas nos cérebros de autistas

REGIÃO CEREBRAL	ALTERAÇÕES
HIPOCAMPO	Maior volume do hipocampo direito.
AMIGDALA	A amígdala inicialmente maior, onde existe correlação entre o volume da amígdala e a gravidade do quadro clínico.
CORPO MAMILAR	A hipoativação dessa região relacionada ao prejuízo da capacidade de aprendizagem espacial.
CÓRTEX ENTORRINAL	Alterações no volume e desenvolvimento dessa região.
SUBÍCULO	Como esta região atua na regulação do sistema de dopamina, anomalias na sua formação têm sido associadas ao autismo.
GIRO DO CÍNGULO	A hipoativação dessa região associada à necessidade de realizar julgamento social sobre os outros.
CÉLULAS DE PURKINJE	Redução na concentração de células de Purkinje.
CÓRTEX PRÉ-FRONTAL	Os padrões de maturação são mais lentos, o que é consistente com o padrão cognitivo dos mesmos.
LOBO TEMPORAL	Ativação anormal do hemisfério esquerdo pode estar envolvida nos prejuízos de linguagem e na resposta comportamental inadequada aos sons dos autistas.

Fonte: Adaptado de Siqueira *et al.* [31].

Portadores do TEA possuem dificuldades em entender as características faciais e comportamentais de terceiros, motivos pelos quais demoram a interagir com os demais indivíduos. Além disso, possuem dificuldades também em compreender atitudes, emoções, sensações e intenções. É comum ainda que apresentem dificuldades em realizar tarefas, se apegando, muitas vezes, às rotinas, pois isso lhes causa maior segurança e conforto [32]. No TEA, o comportamento pode estar associado a distúrbios no lobo temporal, o que esclarece a maioria dos sintomas clínicos percebidos, tais como, déficit perceptivo, cognitivo e emocional. Destaca-se ainda que as regiões do tipo associativas temporais se encontram interligadas de modo firme aos sistemas de sensores associativos frontal, parietais e límbicos [24]. Entre as principais alterações neuroquímicas presentes no cérebro de um autista, citam-se as modificações presentes no complexo dopaminérgico. Há também a liberação elevada de vasopressina, de oxitocina e serotonina, responsáveis pelo reconhecimento social. Com isso, passa-se a ter mudanças no seu modo de funcionamento [30].

Além das pesquisas realizadas pelo CDC, outros estudos [33] também apontam que 1% do total de crianças nos Estados Unidos são autistas. No Brasil, estima-se que existem aproximadamente 5,95 milhões de portadores do TEA [1], número

relativamente alto. Vale destacar que a quantidade de pacientes do sexo masculino é ainda quatro vezes superior à quantidade do sexo feminino. Graf *et al.* [33] evidenciam ainda que existem alguns fatores que influenciam no aumento do número de casos de TEA no mundo, como o aumento da presença desses sujeitos no ambiente social, que no passado se encontravam reclusos devido à discriminação existente e à ausência de políticas públicas de inclusão social; a elevação do número de diagnosticados, em decorrência do crescimento da busca por tratamento especializado e; o aumento real dos casos. A quantidade de autistas no Brasil e no mundo ainda é muito questionada, isso porque não há um número exato dessa população. Uma estratégia relativamente nova e que visa precisar esses números é a incorporação, pela primeira vez, do questionamento acerca do transtorno no Censo do Instituto Brasileiro de Geografia e Estatística (IBGE). Essa medida passou a vigorar no ano de 2020 após se sancionar a Lei 13.861 de 2019 [34], porém ainda não há resultados divulgados até o momento desta pesquisa.

Na contemporaneidade as informações acerca do TEA tendem a ser vagas. Os indivíduos que apresentam o transtorno se deparam com dificuldades no que tange à obtenção do diagnóstico e do tratamento. Outro ponto que merece atenção é que os diagnósticos, em muitos casos, não ocorrem ao longo dos primeiros anos de vida, culminando em várias dificuldades ao longo da infância, da adolescência e em uma parcela da vida adulta [35]. Existem diversos motivos que ensejam na dificuldade de se detectar o autismo, um deles é a falta de alterações visíveis em exames laboratoriais, bem como a impossibilidade de se afirmar que a ligação de sua incidência está relacionada a eventos genéticos específicos. Normalmente, o diagnóstico de pacientes ocorre por meio da observação, muito embora haja testes em desenvolvimento, seu nível de eficiência ainda é reduzido [4].

Aponta-se que o diagnóstico não se dá em curto prazo, este é advindo de um acompanhamento contínuo da vida do sujeito. Kerches [34] cita que o autismo é considerado um transtorno muito difícil de se diagnosticar e comumente confundido com um tipo de atraso comum ou introspecção no desenvolvimento, corroborando para o diagnóstico tardio. Mesmo com avanços significativos nos últimos anos em relação ao tempo de identificação de sinais de autismo em crianças, muitas ainda vivem por muito tempo sem um diagnóstico adequado ou fechado [36]. Diante disso, é im-

prescindível o desenvolvimento de novas metodologias para identificação automática dos transtornos ainda na fase inicial da criança, com o intuito de possibilitar uma intervenção precoce e, conseqüentemente, mitigar prejuízos no campo profissional, social e demais áreas de desenvolvimento humano.

2.3 Inteligência Artificial, *Machine Learning* e *Deep Learning*

A inteligência artificial (IA) é um campo da ciência da computação que se dedica ao desenvolvimento de sistemas e algoritmos capazes de realizar tarefas que normalmente requerem inteligência humana. Segundo Nilsson [37], a IA pode ser definida como o estudo de como fazer os computadores realizarem tarefas que, atualmente, os humanos desempenham melhor. Essas atividades abrangem desde o reconhecimento de padrões e tomada de decisões até a compreensão de linguagem natural e a capacidade de aprendizado. No entanto, estes algoritmos dependem de bases de dados produzidas por humanos, além de supervisão contínua, sobretudo em questões que envolvam tomada de decisões, uma vez que as escolhas da IA podem ser tendenciosas ou desconsiderar questões éticas e morais.

A IA engloba diversos subcampos e abordagens, incluindo aprendizado de máquina, redes neurais artificiais, lógica simbólica, processamento de linguagem natural e visão computacional. De acordo com Russell *et al.* [38], a IA busca criar agentes inteligentes que possam perceber o ambiente, raciocinar sobre ele e tomar ações adequadas para atingir objetivos específicos. Esses sistemas são alimentados com dados e utilizam algoritmos complexos para aprender e melhorar seu desempenho global. O aprendizado de máquina, por exemplo, permite que os sistemas de IA aprendam padrões e façam previsões, adaptando-se a novas situações e refinando suas respostas.

Quanto ao aprendizado de máquina (*machine learning*, ML), esta é uma subárea da inteligência artificial que se concentra no desenvolvimento de algoritmos e técnicas capazes de permitir que os sistemas aprendam e melhorem automaticamente a partir dos dados. Com o avanço da tecnologia e o crescimento exponencial da quantidade de dados disponíveis, o ML tornou-se uma ferramenta essencial em

diversas áreas, desde a análise de dados até a automação de processos. Uma das principais aplicações do ML é na classificação, em que um algoritmo é treinado para identificar padrões nos dados e atribuir rótulos ou categorias às instâncias não rotuladas [39]. Na grande maioria dos casos, o treinamento ocorre por aprendizagem supervisionada, em que os algoritmos são treinados em um conjunto de dados rotulados.

Deep learning (DL), ou aprendizado profundo, é uma abordagem da inteligência artificial que se baseia em redes neurais artificiais profundas para realizar tarefas complexas de aprendizado e reconhecimento de padrões. É um subcampo do ML que busca simular o funcionamento do cérebro humano, permitindo que um sistema computacional seja treinado para aprender e fazer previsões a partir de grandes conjuntos de dados. A principal característica do DL é o uso de redes neurais profundas, compostas por várias camadas de neurônios artificiais interconectados [40]. Essas camadas formam uma hierarquia de representações dos dados de entrada, permitindo que o sistema aprenda características cada vez mais complexas e abstratas à medida que avança pelas camadas. Essa capacidade de aprender representações hierárquicas é o que diferencia o DL de outros métodos de aprendizado de máquina. A Figura 2.3 ilustra a correlação existente entre os conceitos abordados.

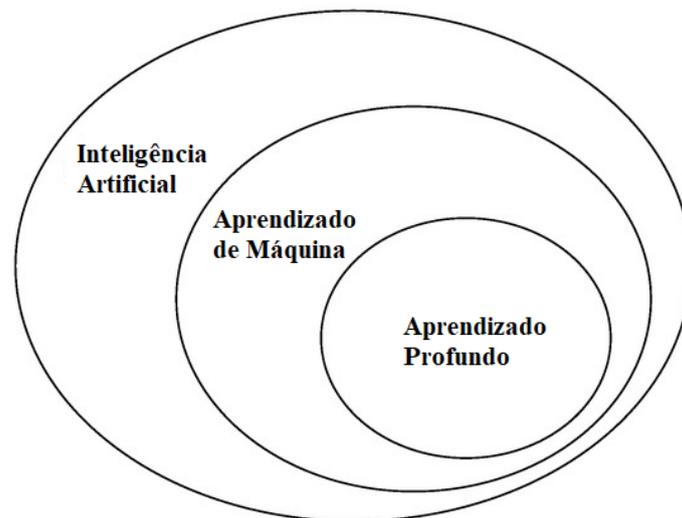


Figura 2.3: Correlação entre os conceitos de Inteligência Artificial, Aprendizado de Máquina e Aprendizado Profundo.

Ademais, a aplicação da IA na medicina tem sido uma área de grande interesse nas últimas décadas. Esta tem o potencial de transformar a prática médica,

auxiliando em diagnósticos mais precisos, tratamentos personalizados e descoberta de novos medicamentos. Cita-se que os modelos de aprendizado de máquina e profundo podem ajudar a melhorar a precisão do diagnóstico e aumentar a eficiência da triagem e rastreamento de doenças. Além disso, estes algoritmos podem analisar grandes quantidades de dados clínicos e de imagem, identificar padrões sutis e fornecer informações valiosas aos especialistas, ajudando-os a tomar decisões informadas [41]. Apesar desses avanços, é importante destacar que a implementação da IA na medicina também apresenta desafios. É fundamental garantir a privacidade e a segurança dos dados dos pacientes, além de lidar com questões éticas relacionadas à tomada de decisões baseadas em algoritmos. Como mencionado por Char *et al.* [42], a transparência e a responsabilidade da IA na medicina são questões críticas que devem ser abordadas. Portanto, é necessário um equilíbrio cuidadoso entre sua implementação e a consideração de questões éticas e regulatórias para garantir sua eficácia, segurança e aceitação generalizada na prática médica.

2.4 Redes Neurais Convolucionais (CNN)

Redes neurais convolucionais são redes biologicamente inspiradas e usadas na visão computacional. Sua forma de processamento é baseada no córtex visual dos mamíferos. Essa forma de processamento une em uma só abordagem a extração de características, podendo assim ser usada para tarefas de classificação, regressão e também para segmentação. Na arquitetura de uma rede neural convolucional, cada camada da rede é tridimensional. Na camada de entrada, a terceira dimensão diz respeito ao número de planos da entrada. Em relação às outras camadas, a terceira dimensão refere-se ao número de filtros convolutivos utilizados na camada [43].

Nas redes neurais convolucionais, os valores de pesos e polarizações para a extração e classificação são aprendidos no treinamento, ou seja, o algoritmo do gradiente descendente, ou alguma adaptação, é utilizado para obter filtros e classificar as características da imagem, sem a seleção de um filtro em específico. Diferentemente das redes totalmente conectadas, os pixels das redes convolutivas se conectam a uma pequena região da camada anterior.

As camadas mais importantes são a camada de convolução e a camada de

subamostragem. Na camada de convolução é definida uma operação na qual um filtro é usado para mapear as ativações de uma camada para a próxima, ou seja, a camada de convolução gera um mapa de características da camada. A camada de subamostragem é usada geralmente após uma camada de convolução, com a finalidade de reduzir o tamanho do mapa de características. Para isso, nessa camada é calculada a média ou encontrado o valor máximo entre os valores de uma região para compactar as dimensões espaciais das camadas. Outras camadas podem ser utilizadas em uma arquitetura de redes neurais convolucionais, como camada batch, relu, inteiramente conectada e de classificação. Na Figura 2.4 é ilustrada a arquitetura LeNet-5, umas das primeiras redes neurais convolucionais desenvolvidas, a qual utiliza imagens de entrada em escala de cinza. Essa arquitetura tem camadas convolutivas, de subamostragem, inteiramente conectada e classificação [44].

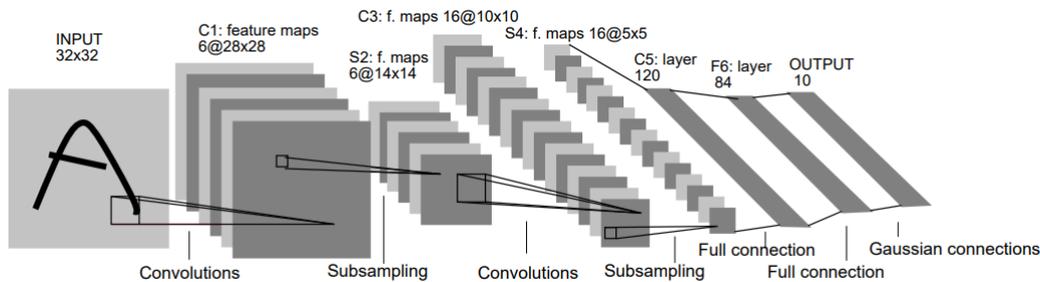


Figura 2.4: Arquitetura LeNet-5.
 Fonte: Adaptado de LeCun, Y. *et al.* [44].

Camada Convolutional

Esta camada é responsável por criar mapas de características a partir dos dados de entrada, a qual é utilizada uma operação matemática de convolução, através da imagem, para criar tais mapas. Os dados de entrada de uma camada são sempre os mapas de características gerado pela camada anterior. O processo de convolução é realizado ao se convoluir uma máscara através da imagem, tendo como resultado uma nova imagem com certas características, por exemplo, bordas reforçadas [43]. Isso faz com que, diferentemente das redes inteiramente conectadas, cada neurônio se conecte a apenas um subconjunto da camada que o antecede, como é mostrado na Figura 2.5. O tamanho espacial do volume de saída (TEVS) de cada imagem pode ser definido através da equação (2.1), em que W é o tamanho do volume de entrada,

F o campo receptivo, S o passo e P a quantidade de *padding*. O Apêndice A aborda as demais camadas de CNN utilizadas neste trabalho.

$$TEVS = \frac{W - F + 2P}{S} + 1, \quad (2.1)$$

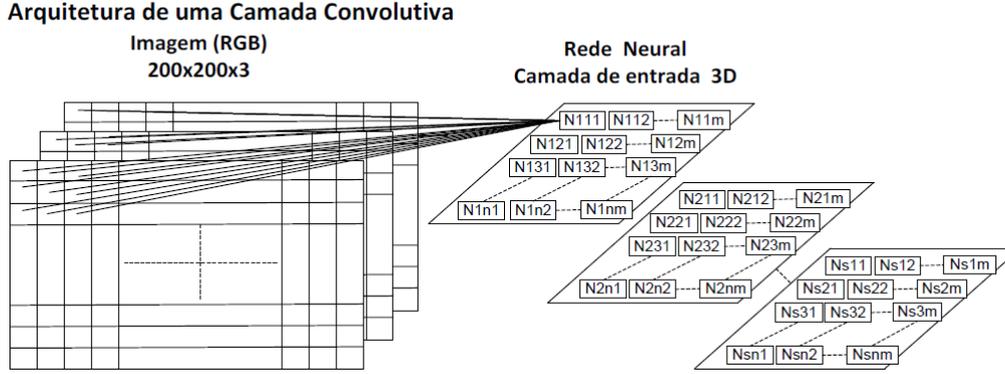


Figura 2.5: Processo de convolução.
Fonte: Adaptado de Aggarwal, C. *et al.* [43].

2.5 Redes Neurais Recorrentes (RNN)

A RNN é um tipo de rede neural artificial que usa dados sequenciais ou dados de séries temporais. Esses algoritmos de aprendizado profundo são comumente usados em problemas ordinais ou temporais, como tradução de linguagem, processamento de linguagem natural (NLP), reconhecimento de fala e legenda de imagem. Tais redes se distinguem por sua memória, pois recebem informações de entradas anteriores que influenciam nas entradas e saídas atuais [43].

As características essenciais das RNN são as conexões de *feedback* que acontecem internamente entre as unidades de processamento. Do ponto de vista do sistema, é uma alimentação de *feedback* que reflete nas características dos processos de cálculo. No modelo de rede neural *feedforward*, a conexão existe entre as camadas, mas não há conexão entre os nós de cada camada [45]. A rede neural recorrente pode processar qualquer comprimento de sequência usando o *self-feedback*. Esta é mais parecida com uma estrutura de rede biológica que uma rede neural *feedforward*. Um gráfico direcionado é formado ao longo da sequência de cada unidade na rede para representar a mudança dinâmica das séries temporais. A RNN usa sequências de

entrada de processamento de estado interno e, portanto, tem funções de memória, conforme ilustrado na Figura 2.6.

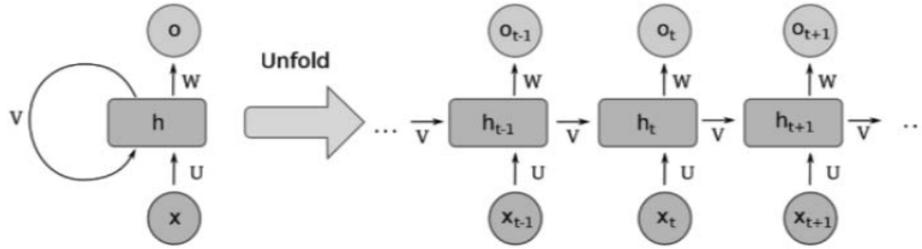


Figura 2.6: Arquitetura RNN padrão.
Fonte: Yang, S. [45].

A RNN tem uma conexão circular, em que a sequência de entrada, a sequência vetorial oculta e a sequência vetorial de saída são representadas por X , H e O , respectivamente. A sequência de entrada é dada por $X = (x_1, x_2, \dots, x_T)$. A RNN calcula a sequência vetorial oculta ($H = (h_1, h_2, \dots, h_T)$) e o vetor de sequência de saída ($O = (o_1, o_2, \dots, o_T)$) do tempo 1 a t , conforme equações (2.2) e (2.3).

$$h_t = \sigma(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (2.2)$$

$$o_t = W_{ho}h_t + b_o \quad (2.3)$$

onde a função σ é uma função não linear e W a matriz de pesos [45].

Rede LSTM

LSTM (*long short-term memory*) é uma arquitetura de RNN proposta para resolver o problema de explosão ou esvanecimento do gradiente; neste modelo, utiliza-se uma estrutura de portas e uma unidade de memória [46]. A Figura 2.7 mostra uma célula LSTM que descreve a operação de três estados de portas e de células, em que σ é a função sigmoide, I a porta de entrada, F a porta de esquecimento, O a porta de saída e C o status da célula. As três portas (I , F e O) controlam o fluxo da informação. Portas de entrada e esquecimento determinam o quanto da informação da célula anterior permanece na célula de memória atual, e a camada de saída determina o quanto de informação será passada para a próxima

célula [45].

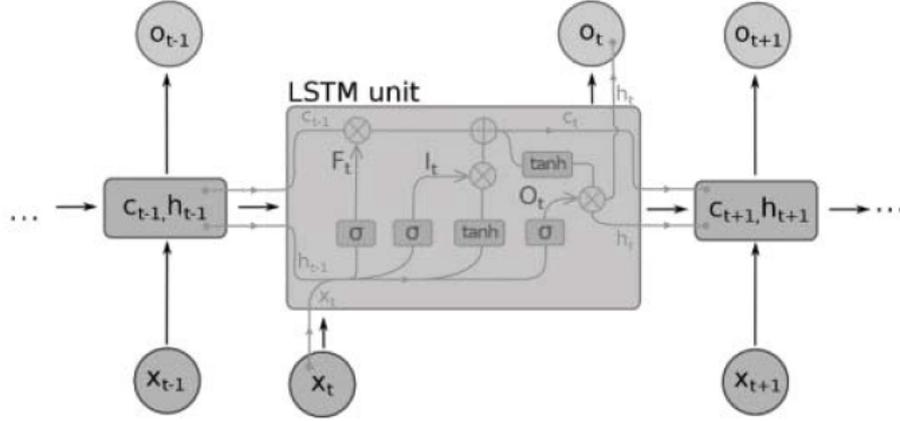


Figura 2.7: Célula LSTM.

Fonte: Yang, S. [45].

Sendo I , F , O e C as portas de entrada, esquecimento, saída e estado da célula, W_{ci} , W_{cf} e W_{co} são as matrizes de pesos. O processo específico é descrito a seguir. A informação de esquecimento do estado de célula, determinado pela função sigmoide da porta de esquecimento; a saída da entrada da camada atual; a saída da camada anterior como entrada; e o estado da célula no tempo $t - 1$, combinam-se por meio da equação (2.4).

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.4)$$

A informação armazenada no estado da célula é principalmente composta por duas partes: o resultado da sigmoide da porta de entrada, com a informação atualizada e; o novo vetor, criado por uma função \tanh e adicionado ao estado da célula [46]. Desta forma, multiplica-se o estado da célula antiga por f , com a nova informação candidata, e geram-se atualizações do estado da célula, conforme equações (2.5), (2.6) e (2.7) deduzidas a seguir.

$$i_t = \sigma(W_i[[h_{t-1}, x_t] + b_i]) \quad (2.5)$$

$$\hat{C}_t = \tanh(W_c[[h_{t-1}, x_t] + b_c]) \quad (2.6)$$

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (2.7)$$

A informação de saída é determinada pela porta de saída. Primeiro, a camada sigmoide é usada para determinar as informações do estado da célula a serem geradas, o estado da célula é tratado com \tanh e o produto de duas partes das informações é o valor da saída, de acordo com as equações (2.8) e (2.9).

$$o_t = \sigma(W_o[[h_{t-1}, x_t] + b_o]) \quad (2.8)$$

$$h_t = o_t * \tanh(C_t) \quad (2.9)$$

2.6 A técnica de ordenação por diferença percentual

No trabalho desenvolvido por Melo *et al.* [47], é apresentada uma técnica de pré-processamento denominada ordenação por diferença percentual (ODP), a qual reordena os segmentos de sinais utilizando a métrica de diferença percentual. Como é sabido, os sinais biológicos podem possuir um comportamento não estacionário, em que a média e a variância do sinal alteram-se no decorrer do tempo. Consequentemente, o processo de formatação do sinal em uma matriz bidimensional pode produzir um sinal com baixa correlação entre os segmentos, o que prejudica a exploração das dependências existentes em um sinal. Em contrapartida, a reordenação dos segmentos do sinal pode produzir áreas mais homogêneas dentro de uma imagem, contribuindo para o aumento da correlação intersegmento.

Em técnicas de ordenação baseadas em coeficientes de correlação, é possível haver dois segmentos com alta correlação, entretanto, suas amostras podem possuir amplitudes bastante distintas, o que pode comprometer a exploração das dependências intersegmentos nos sinais bidimensionais [47]. Diante disso, em vez de analisar a correlação linear dos segmentos, a técnica de ODP permite verificar a similaridade das amostras dos segmentos pela diferença percentual. O objetivo é ordenar os segmentos (canais do EEG) com base na semelhança, de forma que as redundâncias intersegmentos sejam aumentadas. Para tal, define-se a seguinte métrica:

$$DP(x, m) = \frac{\sum_{n=0}^{N-1} (x[n] - m[n])^2}{\sum_{n=0}^{N-1} x^2[n]}, \quad (2.10)$$

onde $DP(x, m)$ representa a ordenação por diferença percentual para os segmentos \mathbf{x} e \mathbf{m} , $x[n]$ é o n -ésimo elemento da última linha posicionada, $m[n]$ é o n -ésimo segmento em análise e N o número de amostras em cada linha.

Primeiramente, o segmento com menor variância é inserido na primeira linha da nova matriz a ser ordenada. Em seguida, a próxima linha recebe o segmento com menor valor de diferença percentual em relação à primeira linha e, por fim, os segmentos restantes são reordenados de acordo com o menor valor de diferença percentual em relação à última linha posicionada. A Figura 2.8 ilustra o resultado da aplicação da técnica em um sinal de EEG, onde é possível observar os efeitos da ordenação.

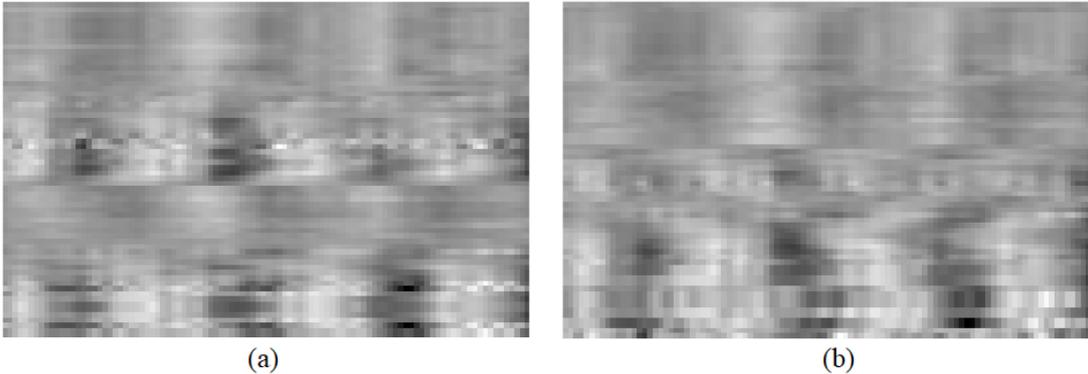


Figura 2.8: Amostra de resultado da ordenação por diferença percentual: (a) sinal EEG bidimensional original e (b) sinal EEG bidimensional reordenado.

2.7 Trabalhos Relacionados

O trabalho proposto por Bhaskarachary *et al.* [48] teve como objetivo explorar e analisar técnicas de pré-processamento de sinais de EEG, extração de características e aprendizado de máquina, visando identificar atividades anormais para o diagnóstico do TEA. Para isso, os autores utilizaram a base de dados de *child mind institute biobank* [49], a qual contém 88 sinais de EEG de crianças TD e 100 sinais

de EEG de crianças autistas. Esses sinais foram amostrados em 250 Hz e segmentados em trilhas de 20 s, extraindo-se apenas as partes relacionadas ao estado de olhos fechados. Depois, foram decompostos em quatro bandas de frequências: delta (0 – 4 Hz), teta (4 – 8 Hz), alfa (8 – 12 Hz) e beta (13 – 30 Hz), seguido do cálculo da densidade espectral de potência. A técnica de análise de componentes principais (PCA) [18, 19] foi implementada para reduzir a dimensionalidade das características, de 440 para 46 canais, os quais foram usados para alimentar dois classificadores, *extra trees* e *XGBoost* [50, 51]. Segundo os autores, o *XGBoost* foi usado em função da eficiência do framework e facilidade de implementação do algoritmo, ao passo que o classificador *extra trees* oferecia maior desempenho e evitava o overfitting dos dados, além de ser capaz de selecionar recursos que reduzem a redundância, levando a maior precisão e agilidade no treinamento. Os desempenhos dos modelos foram avaliados por meio de métricas, como, área sob a curva (AUC), acurácia, *recall*, precisão e especificidade, obtidas da matriz de confusão do conjunto de teste. Finalmente, observou-se que os resultados de desempenho do classificador *extra trees* foram superiores em relação ao *XGBoost*, com acurácia de 67,70 %, sensibilidade de 83,30 %, especificidade de 54,30 %, precisão de 61,00 % e AUC de 0,74.

Na pesquisa desenvolvida por Pistorius *et al.* [52], foi utilizada uma base de dados da Universidade de Califórnia, San Diego, que continha amostras de 5 crianças TD e 7 diagnosticadas com o TEA, ambas com idade entre 8 e 17 anos. Os dados foram obtidos usando o sistema Neuroscan Synamp 1, de 20 canais, com uma taxa de amostragem de 500 Hz e uma faixa de passagem de banda de 0,1 – 100 Hz, gravados por um intervalo de 10 minutos para cada indivíduo. Na etapa de pré-processamento, os sinais de EEG [25, 53] em estado de repouso – olhos fechados (EC) e olhos abertos (EO) – foram decompostos em cinco bandas de frequência: delta (0,5 – 4 Hz), teta (4 – 8 Hz), alfa (8 – 13 Hz), beta (13 – 30 Hz) e gama (30 – 40 Hz). Em seguida, os autores realizaram reamostragem para 250 Hz, filtros *notch* – para remoção de possíveis interferências da linha de energia –, remoção de canais inativos, correção de artefatos, seleção dos melhores segmentos e, extração de características com análise de quantificação de recorrência (RQA), incorporando séries temporais multivariadas e redução de dimensionalidade com PCA. Dois algoritmos de classificação foram empregados neste estudo: k-vizinhos mais próximos

(kNN) (com $k = 3$) e análise discriminante linear (LDA) [54,55]. O classificador kNN prevê o rótulo de classe de um ponto de dados de teste avaliando o rótulo de classe do número k de pontos de treinamento mais próximos a esse ponto e atribui a ele o rótulo mais comum, enquanto o classificador LDA primeiro determina uma linha de separação/hiperplano tentando maximizar a separabilidade linear entre pontos de dados pertencentes a diferentes classes. Os classificadores kNN e LDA obtiveram desempenhos idênticos para segmentos EC: uma precisão de 83,30 % (10/12 indivíduos classificados corretamente), com uma sensibilidade de 85,70 % (6/7 indivíduos TEA identificados corretamente) e especificidade de 80,0 % (4/5 TD corretamente identificados). Os resultados de classificação obtidos para os segmentos de EO revelam que o procedimento de correção de artefatos implementado neste estudo é insuficiente.

No trabalho publicado por Kang *et al.* [2], para a tarefa de identificação de crianças autistas, os autores utilizaram recursos extraídos de duas modalidades - sinais EEG e rastreamento ocular - como entrada para uma abordagem de aprendizado de máquina. Um total de 97 crianças de 3 a 6 anos foram incluídas no presente estudo, 49 com TEA e 48 TD. Após a gravação de dados de EEG em estado de repouso com um sistema de rede de sensores HydroCel de 128 canais, em uma taxa de amostragem de 1.000 Hz, durante 6 min, as crianças realizaram testes de rastreamento ocular com o sistema TX300, à medida que estímulos de rostos de estranhos de sua própria raça e de outras raças eram provocados por uma série de fotos em ordem aleatória. Por intermédio da *toolbox EEGLab* [56], do *Matlab* [57], os pesquisadores utilizaram as seguintes técnicas de pré-processamento: filtro passa-faixa entre 0,5 e 45 Hz, reamostragem para 250 Hz, filtro *notch* centrado em 50 Hz, remoção de artefatos com ICA e inspeção visual. Logo, sucederam com a análise da densidade de espectro de energia a partir da transformada rápida de Fourier para separação em bandas de frequências delta, teta, alfa, beta e gama. Para a análise dos dados de rastreamento ocular, 8 áreas de interesse (AOI) foram estabelecidas e, a fim de quantificar o engajamento da criança em cada AOI, definiram o índice de análise como a porcentagem do tempo de fixação na AOI em relação ao tempo total de fixação. Por fim, o método de seleção de recursos de relevância máxima e redundância mínima (MRMR), combinado com o classificador máquina de vetores

de suporte (SVM) - o qual tende a encontrar o hiperplano de separação com a maior distância de margem entre as categorias no espaço de características -, foi aplicado para promover a separação entre as duas classes. Os resultados mostraram que a acurácia de classificação usando a potência relativa das cinco bandas de frequência como entrada para o SVM foi de 68,0%. Usando individualmente as durações de AOI, ao olhar para estímulos faciais, a acurácia resultou em 75,89% e, combinando os dois tipos de dados como entrada, a acurácia atingiu o máximo de 85,44%.

Na pesquisa proposta por Jayawardana *et al.* [3], os sinais de EEG foram coletados de 8 crianças diagnosticadas com TEA e 9 TD, com idades entre 5 e 17 anos. Foi solicitado a essas crianças que completassem uma avaliação chamada *autism diagnostic observation schedule second edition* (ADOS-2) [58], enquanto simultaneamente gravavam os sinais de EEG. Esses sinais foram registrados usando o sistema EEG sem fio *LiveAmp* de 32 canais, com eletrodos ativos e uma taxa de amostragem digital de 250 Hz. Os autores derivaram os espectros de potência para cada eletrodo por meio da (i) decomposição da banda de frequência e (ii) da decomposição por transformadas Wavelet, as quais geram dois conjuntos de dados de treinamento que capturam tendências de longo e curto prazo, respectivamente. Por meio da ferramenta *EEGLab* [56], foram aplicadas as seguintes técnicas de pré-processamento: remoção do desvio da linha de base de baixa frequência, remoção de artefatos oculares, musculares e de movimento, referenciamento de canais, filtragem de ruídos AC de 50 a 60 Hz, rejeição de canais ruins, interpolação de canais removidos e análise de componentes independentes (ICA). Para classificação, os autores serviram-se de vários modelos de aprendizado de máquina para o diagnóstico do TEA e previsão do escore ADOS-2, tais como, *Random Forest*, *Logistic*, *Bagging*, *JRip*, *LMT*, *AdaBoostM1* e *CNN-1D*. Como resultado, em dependências de curto prazo, obteve-se um máximo de 56,0% de acurácia de classificação por meio de modelos lineares, ao passo que os modelos não lineares forneceram uma classificação acima de 92,0% de acurácia, com escore ADOS-2 previstas dentro de um RMSE de valor 4. Neste trabalho, destacam-se os classificadores *random forest*, para tendências de curto prazo, e *CNN-1D*, para tendências de longo prazo, ambos por fornecerem uma acurácia de classificação acima de 95,0%.

No trabalho desenvolvido por Ibrahim *et al.* [23], foram investigadas diferen-

tes técnicas de extração e classificação de características de EEG, para auxiliar no diagnóstico de epilepsia e TEA. Os pesquisadores experimentaram seus diferentes modelos em 4 bases de dados distintas, a saber: o conjunto de dados da Universidade de Bonn [59], o qual contém 5 pastas com 100 gravações de sinais epiléticos e normais de 23,60 s e taxa de amostragem de 173,61 Hz; o conjunto de dados do MIT [60], com 906 h de gravações de 23 pacientes epiléticos com taxa de amostragem de 256 Hz; o conjunto da Universidade King Abdulaziz [61], gravados de 10 jovens TD e 9 TEA, com taxa de amostragem de 256 Hz e; uma base de dados própria com gravações de EEG de 46 indivíduos [23]. Em todos os casos, os sinais de EEG foram pré-processados para remover os principais artefatos indesejados, antes de serem decompostos em várias sub-bandas de EEG por meio da transformada discreta de Wavelet (DWT). Em seguida, aplicaram-se as funções desvio padrão (SD), potência de banda (BP), entropia de *Shannon* (SE), e maior expoente de *Lyapunov* (LLE), sendo as duas últimas métodos não-lineares que medem complexidade e caoticidade no registro do EEG. Além disso, uma abordagem de correlação cruzada foi utilizada para medir a sincronização entre os canais de EEG agrupados nas regiões do cérebro anterior, central e posterior, que pôde revelar anormalidades na comunicação entre essas regiões. Diferentes métodos de classificação foram experimentados, tais como, redes neurais artificiais (RNAs), k-vizinho mais próximo (KNN), máquina de vetor de suporte (SVM) e análise discriminante linear (LDA). Por fim, a combinação das técnicas de DWT, entropia de *Shannon* e classificador KNN produziu o resultado de classificação mais promissor, com uma precisão geral de até 94,62% para o problema de classificação de três classes (multicanal).

Em [62], Milne *et al.* examinaram as características comportamentais do autismo juntamente com os métodos de coerência de fase entre as tentativas (ITC) e entropia multiescala (MSE), calculados a partir de sinais de EEG. Os dados para os experimentos foram obtidos de 99 adultos, sendo 38 diagnósticos com TEA (19 mulheres), adquiridos via *BioSemi ActiveTwo* em uma câmara blindada eletricamente durante estimulação visual e repouso com os olhos fechados. Após, foram filtrados em tempo real com uma passagem de banda de 0,01 – 140 Hz e digitalizados a uma taxa de amostragem de 2.048 Hz. Os sinais contínuos foram reamostrados para 512 Hz, referenciados ao canal *Cz* e refiltrados com cortes entre 1,5 Hz e 50 Hz

usando a função *eegfiltnew* dentro da *toolbox EEGLab*, do *Matlab*, a largura de banda de transição e as bordas da banda passante foram ambas de 1 Hz. Canais e segmentos de dados contaminados por artefatos grosseiros foram identificados e removidos por meio de inspeção visual. Além disso, os dados foram decompostos em componentes independentes (IC) usando o algoritmo *runica* do *EEGLab*. Para inclusão no grupo de diagnosticados, os participantes deveriam ter um diagnóstico clínico e obter uma pontuação do *ritvo autism asperger's diagnostic schedule* (RAADS-R) acima do ponto de corte para TEA. O RAADS-R é um questionário de 80 questões que foi desenvolvido para auxiliar no diagnóstico de TEA em adultos. Outros questionários também foram utilizados, a exemplo do *social responsiveness questionnaire* (SRS-2). Diante disso, os autores descobriram que ITC e MSE têm maior probabilidade de serem reduzidos em pessoas com TEA. A análise da função discriminante (DFA) inserindo MSE e ITC como preditores de associação de grupo (diagnosticados e não-diagnosticados) foi realizada. A associação entre grupo e preditores foi significativa, no entanto, enquanto 18 dos 22 (81,80%) participantes não-diagnosticados foram classificados corretamente, apenas 12 dos 22 (54,50%) participantes diagnosticados foram classificados corretamente. Na associação entre grupo e ITC, 81,80% dos participantes não-diagnosticados foram classificados corretamente e 50,00% no grupo dos diagnosticados. Por fim, a associação entre grupo e MSE também foi significativa, porém apenas 63,60% dos participantes não-diagnosticados e 59,10% dos participantes diagnosticados foram classificados corretamente. A DFA mostrou que as variáveis do EEG não distinguem claramente os participantes diagnosticados dos não-diagnosticados.

Na Tabela 2.3, apresenta-se uma comparação resumida das principais informações dos trabalhos relacionados, levando-se em consideração a metodologia proposta. Nesta, destaca-se que os resultados das métricas utilizadas pelos diversos autores não consideram a base de dados de testes usada nesta dissertação, com exceção do trabalho de Milne *et al.* [62]. Em decorrência da indisponibilidade e escassez de bases públicas, os resultados não podem ser avaliados de maneira igualitária, pois cada base tem sua peculiaridade e nível de dificuldade. Não obstante, dentre os trabalhos analisados, as métricas para avaliação dos diferentes modelos utilizados foram similares, o que possibilita a comparação aproximada dos trabalhos escolhi-

dos. Os melhores resultados foram alcançados pelos autores Jayawardana *et al.* [3] e Ibrahim *et al.* [23], com valores de acurácia acima de 94,0% e precisão, *recall*, *F1 Score* acima de 95,0%, testificando que diferentes técnicas de aprendizado de máquina podem ser aplicadas para o diagnóstico do TEA.

Tabela 2.3: Tabela comparativa entre trabalhos relacionados.

AUTORIA	BASE DE DADOS	PRÉ-PROCESSAMENTO	MÉTODO	RESULTADOS	
Bhaskarachary <i>et al.</i> [48]	Child Mind Institute Biobank: 88 sinais de EEG de indivíduos TD e 100 sinais de indivíduos com TEA	Filtro passa-faixa	Análise de espectro de potência + Análise de componentes principais + Classificadores Extra Trees e XGboost	EXTRA TREES: Acurácia = 67.70%, Precisão = 61%, Recall = 83.30%, Especificidade = 54.30%, AUC = 0.74	XGBoost: Acurácia = 60.00%, Precisão = 54.80%, Recall = 76.70%, Especificidade = 45.70%, AUC = 0.72
Pistorius <i>et al.</i> [52]	Obtido da Universidade de Califórnia: sinais de 12 indivíduos entre 8 e 17 anos - 7 TEA e 5 TD	Reamostragem para 250Hz + Filtros + Remoção de canais inativos + Correção de artefatos + Seleção do melhor segmento	PCA + RQA + Classificadores LDA e KNN	Acurácia = 83.30%, Sensibilidade = 85.70%, Especificidade = 80%	
Kang <i>et al.</i> [2]	Base própria: 97 crianças entre 3 e 6 anos - 49 com TEA e 48 TD	Filtro passa-faixa + Reamostragem para 250 Hz + Filtro Notch centralizado em 50 Hz + Remoção de artefatos com ICA + Inspeção visual	Análise de espectro de potência + Análise de dados de rastreamento ocular + Máquina de vetores de suporte (SVM)	Acurácia = 85.44%, AUC= 0.93	
Jayawardana <i>et al.</i> [3]	Base própria: 17 crianças entre 5 e 17 anos - 8 diagnosticadas com TEA e 9 TD	Remoção do desvio da linha de base de baixa frequência + Remoção de artefatos oculares, musculares e de movimento + ASR + Channel Re-reference + Filtragem de ruídos AC de 50-60 Hz + Rejeição de canais ruins + Interpolação de canal removida + ICA	Filtro passa-faixa + Decomposição de Banda de Frequência com Transformada Wavelet + Classificadores Random Forest, Logistic, Bagging, JRip, LMT, AdaBoostM1, CNN 1D	SHORT-TERM (Random Forest): Acurácia = 97%, Precisão = 97%, Recall = 97%, F1 Score = 97%	LONG-TERM (CNN 1D): Acurácia = 95%, Precisão = 95%, Recall = 95%, F1 Score = 95%
Ibrahim <i>et al.</i> [23]	The University of Bonn dataset, the MIT dataset, the King Abdulaziz University dataset, e base próprio (46 indivíduos)	Filtro passa-banda Elíptico (0.5 ~60 Hz)	Transformada wavelet discreta (DWT) + Desvio padrão (SD), Potência de banda (BP), Shannon entropia (SE) e Maior expoente de Lyapunov (LLE) + Correlação cruzada + Classificadores Redes neurais artificiais (ANNs), k-vizinho mais próximo (KNN), Máquina de vetores de suporte (SVM) e Análise discriminante linear (LDA)	Melhor Resultado de Acurácia (DWT+SE+KNN): 94.62%	
Milne <i>et al.</i> [62]	Base própria: 99 adultos - 38 diagnosticadas com TEA e 61 TD	Reamostragem para 512Hz + Filtros passa-banda (1.5 ~50 Hz) + Remoção de artefatos grosseiros + Remoção de ruídos com ICA	Análise da função discriminante (DFA) inserindo Coerência de fase entre as tentativas (ITC) e Entropia multiescala (MSE)	Classificações corretas da 1) associação entre grupo e preditores: 81,8 % (não-diagnosticados) e 54,5% (diagnosticados); 2) Associação entre grupo e ITC: 81,8 % (não-diagnosticados) e 50% (diagnosticados); 3) Associação entre grupo e MSE: 63,6 % (não-diagnosticados) e 59,1 % (diagnosticados).	

Capítulo 3

Metodologia Proposta

3.1 Introdução

Neste trabalho, proponho uma metodologia de classificação de sinais de EEG para o auxílio ao diagnóstico do TEA utilizando técnicas de pré-processamento e aprendizagem profunda. A metodologia proposta busca investigar a existência de padrões em sinais de EEG bidimensionais (imagens), gerados a partir de sinais unidimensionais, que distingam indivíduos diagnosticados com TEA dos TD.

O sistema proposto nesta dissertação se desenvolve em duas etapas: uma para o treinamento do classificador e outra para execução do modelo treinado. A Figura 3.1 a seguir apresenta o diagrama em blocos geral do sistema proposto. Na primeira etapa, representada na Figura 3.1(a), realizo, no bloco *pré-processamento*, as tarefas de reamostragem, filtragem e atenuação de ruídos nos sinais originais de entrada, as quais, em conjunto, possibilitam a diminuição da quantidade de dados a serem analisados, redução de inconsistências e redundâncias, identificação dos subconjuntos mais significativos, o aumento de qualidade dos sinais, bem como a redução do tempo de treinamento dos algoritmos. No mesmo bloco, aplico também a técnica de ordenação por diferença percentual para reordenar, por similaridade, os segmentos dos sinais. Em seguida, no bloco denominado *concatenação*, realizo a concatenação dos sinais EEG unidimensionais para formação de K imagens de dimensão $M \times N$, a serem utilizadas como sinais bidimensionais de entrada no bloco *treinamento do classificador*, onde os dados são reparticionados em conjunto de treinamento e validação. No mesmo bloco, defino os parâmetros e executo os treinamentos dos classificadores

(que nesta dissertação são as redes 2D-CNN e RNN-LSTM, detalhadas nas próximas seções) resultando em um modelo treinado. Na segunda etapa, representada na Figura 3.1(b), utilizo como entrada um conjunto de dados destinado ao teste. Os blocos *pré-processamento* e *concatenação* desempenham as mesmas funções descritas na primeira etapa. Em seguida, executo o bloco *modelo treinado* e avalio o desempenho do modelo a partir do bloco *métricas para avaliação*, que utiliza as métricas acurácia, AUC, erro, precisão, *recall* e *f1-score*.

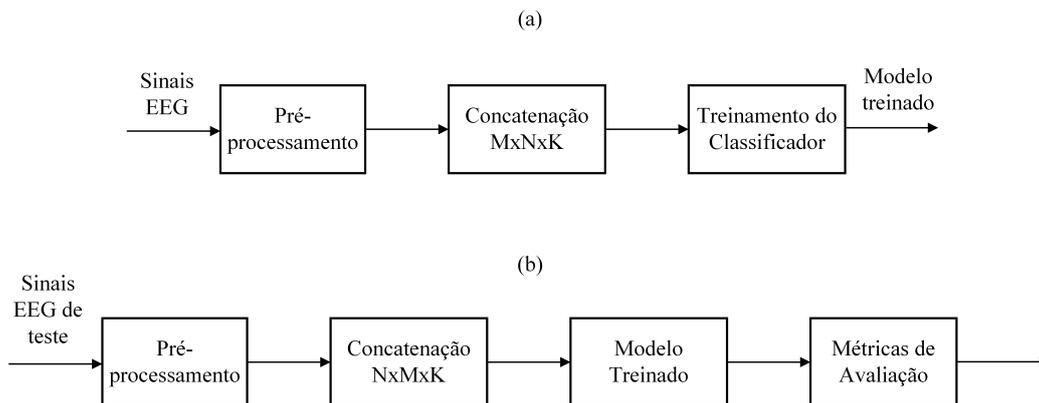


Figura 3.1: Diagrama em blocos geral do sistema proposto: (a) treinamento do classificador, (b) execução do modelo treinado.

3.2 Pré-processamento

Nesta dissertação, utilizo como entrada os sinais EEG da base de dados disponibilizada por Milne *et al.*, da Universidade de *Sheffield* [63], a qual contém assinaturas eletrofisiológicas de 56 indivíduos (28 TEA e 28 TD). Especificamente, na etapa de pré-processamento, utilizo seis blocos, conforme apresentado na Figura 3.2. O primeiro, denominado por *reamostragem*, efetua a reamostragem dos sinais de EEG, onde realizo a redução da taxa de amostragem de 512 Hz em um fator de 0,5 e 0,25 a fim de investigar o possível aumento na velocidade de processamento e diminuição da memória de armazenamento, sem comprometer a extração de características e o reconhecimento de padrões nos sinais de EEG. É importante destacar que, antes da efetiva reamostragem por decimação, um filtro passa-baixa na metade da frequência de reamostragem é aplicado automaticamente (pela ferramenta *EEGLab*) aos dados para evitar efeitos de *aliasing*. A seguir, nos blocos *filtro passa-baixa* e *filtro passa-alta*, busco delimitar entre 1 e 60 Hz a frequência dos sinais,

quando estes são visualizados pela perspectiva do domínio da frequência. Vale destacar que se um filtro passa-faixa fosse aplicado, as partes passa-baixa e passa-alta teriam as mesmas inclinações, como configuração padrão no *EEGLab*. Frequentemente, a inclinação da passagem baixa é mais íngreme do que o necessário. Para evitar este problema, optei por aplicar primeiro o filtro passa-baixa e depois, em um segundo bloco, o filtro passa-alta. Por convenção, adoto a seguinte decomposição em cinco bandas de frequência: delta (1 – 4 Hz), teta (4 – 8 Hz), alfa (8 – 13 Hz), beta (13 – 30 Hz) e gama (30 – 60 Hz), onde estão concentradas as principais atividades neurais em um cérebro humano, de acordo com Sanei *et al.* [25]. No bloco seguinte, denominado por *remoção de artefatos*, utilizo uma ferramenta de rejeição automática de artefatos [64] para remover canais planos maiores que 5 s, canais com ruídos de alta frequência com desvio padrão maiores que 4 em relação aos demais e canais com correlação inferior a 80% em relação aos demais. No bloco *podagem com ICA*, aplico a técnica de podagem com análise de componentes independentes (ICA) para remoção de artefatos com 80% de chances ou mais de estarem relacionados a movimentos de músculos ou olhos, quando da coleta de dados.

No bloco *padronização das amostras e ODP*, em razão dos procedimentos aplicados nos blocos precedentes e, conseqüentemente, das variações nas dimensões de linha e coluna de cada sinal pré-processado individualmente, realizo a padronização dos 56 sinais em número de linhas e colunas iguais para posterior concatenação e entrada no bloco de treinamento do classificador. Para tanto, adiciono *padding* simétricos aos sinais EEG para padronização das amostras em 64 linhas, visto ser este o número de canais dos sinais originais antes da etapa de pré-processamento. O tempo de gravação dos sinais EEG, a taxa de amostragem original, o fator de reamostragem utilizado e a remoção de fragmentos nos sinais, em conjunto, influenciam na quantidade total de colunas de cada sinal. Para fins de padronização, os 56 sinais são reparticionados em segmentos de 1000 colunas, podendo resultar na formação de mais de 4000 amostras de imagens de dimensão 64×1000 para os experimentos, a depender da taxa de amostragem utilizada. O número de colunas foi definido com base na avaliação de desempenho e custo de processamento nos procedimentos experimentais, satisfazendo a melhor relação tamanho de imagem/quantidade de amostras para as redes de treinamento. Finalmente, é aplicada

a técnica de ODP, com o objetivo de reordenar os segmentos de canais das imagens formadas e, conseqüentemente, fortalecer as redundâncias intersegmentos.

Ainda na Figura 3.2, apesar dos sinais possuírem 64 canais cada, exibe-se nos quadros apenas 5 canais de um sinal aleatório para melhor visualização dos efeitos de cada bloco. Inicialmente, comparando-se os quadros dos sinais de entrada e saída do bloco de reamostragem, é notório os efeitos da redução da frequência de amostragem do sinal, tornando-o mais suavizado, sem perda das características essenciais para uma redução de fator de 0,5. Em particular, nos quadros de entrada e saída do bloco de filtro passa-alta, não é possível observar alterações relevantes na amostra do sinal escolhido, uma vez que este filtro permite a passagem de todas as frequências acima de 1 Hz e, no quadro apresentado, não há quantidade significativa de frequências abaixo da frequência de corte; ao contrário do quadro de saída do bloco de filtro passa-baixa, onde é possível notar distintamente os efeitos de atenuação das frequências acima de 60 Hz, tornando o sinal ainda mais suavizado e sem transições abruptas. Para as amostras apresentadas nos quadros de saída dos blocos de remoção de artefatos e podagem com ICA, não foram encontrados canais planos maiores que 5 s, canais com ruídos de alta frequência com desvio padrão maiores que 4 em relação aos demais, canais com correlação inferior a 80 % ou artefatos associados a movimentos de músculos ou olhos que carecessem de remoção, isto porque tais problemas são pontuais e não estão presentes ao longo de todo o sinal gravado, ao que não foi possível, portanto, perceber na pequena amostra selecionada.

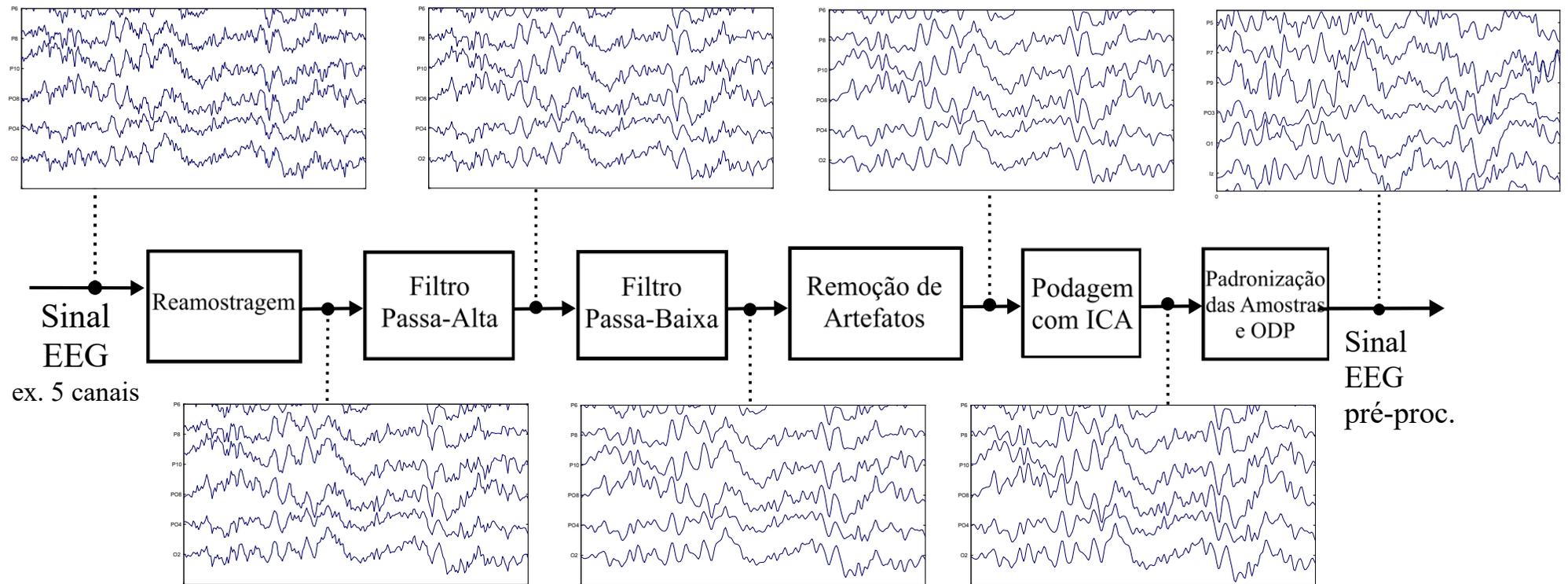


Figura 3.2: Diagrama em blocos de pré-processamento para o sistema de classificação TEA.

3.3 Treinamento e Execução do Classificador

Para treinamento do classificador, conforme mencionado, utilizei o diagrama apresentado na Figura 3.1(a). Para efetivamente realizar o treinamento, é necessário aplicar o bloco *concatenação*, onde os sinais de EEG unidimensionais pré-processados no bloco anterior são concatenados em segmentos de linhas, representadas por diferentes canais, para formação de K imagens de dimensões $M \times N$, utilizadas posteriormente como sinais bidimensionais de entrada no bloco de treinamento dos classificadores 2D-CNN e RNN-LSTM. Esta conversão transforma os 56 sinais, originalmente gravados ao longo de 150s, em múltiplas amostras de imagens, alcançando valores acima de 4000 amostras, a depender da taxa de amostragem utilizada. Em complemento, no mesmo bloco, produzo um vetor de rótulos de dimensão $n \times 1$ para treinamento e classificação, onde a primeira metade das amostras assume valor binário 0 e a segunda metade, 1, referindo aos sinais TEA e TD, respectivamente. Saliento que a quantidade de amostras geradas, K , para o bloco de treinamento das redes, varia de acordo com o fator de reamostragem e os demais filtros aplicados no bloco de pré-processamento, que podem eliminar canais inteiros ou partes destes.

A Figura 3.3 a seguir apresenta um diagrama expandido do bloco *treinamento do classificador* (Figura 3.1(a)), com vistas a minuciar o processo de treinamento das redes neurais. O primeiro bloco, denominado *divisão de dados*, recebe os sinais de EEG bidimensionais concatenados, bem como o vetor de rótulos, e os divide percentualmente em conjuntos distintos para treinamento e validação. No bloco *definição de arquitetura e parâmetros*, constituo a arquitetura da rede neural 2D-CNN (ou RNN-LSTM), composta por diversas camadas, e defino os parâmetros de treinamento, além dos otimizadores e funções de custo. Em suma, as principais camadas utilizadas na rede 2D-CNN aumentam ou diminuem a quantidade de canais das imagens (camada convolucional) [65], executam operações de limite não-linear, em que qualquer valor de entrada menor que zero é multiplicado por um fator de escala fixo (camada de ativação) [66] e, simplificam as informações da camada anterior reduzindo sua dimensionalidade a partir dos valores máximos em uma determinada região (camada de *pooling*) [67]; enquanto na rede RNN-LSTM, as principais camadas desempenham as funções de aprendizagem de dependências

de ordem em problemas de predição de sequência (camada LSTM) [68] e redução de *overfitting* da rede (camada de *dropout*) [69]. Finalmente, no bloco *treinamento da rede 2D-CNN ou RNN-LSTM*, executo o treinamento das redes definidas no bloco antecedente, a partir dos conjuntos de treinamento e validação separados para este fim. Neste bloco a rede é treinada sucessivas vezes de maneira a reduzir o erro e elevar a acurácia de validação, até uma determinada quantidade de épocas. O melhor modelo alcançado durante às épocas de treinamento é utilizado para avaliação de desempenho nas etapas posteriores do sistema proposto.

Por fim, para execução do classificador, utilizo como entrada um conjunto de dados destinado ao teste, que passará de igual forma pelos blocos de pré-processamento e concatenação, os quais desempenham as mesmas funções da etapa de treinamento descrita. A partir dos dados de teste pré-processados, o melhor modelo salvo no treinamento do classificador é executado para análise de desempenho do modelo através das métricas acurácia, AUC, erro, precisão, *recall* e *f1-score*.



Figura 3.3: Diagrama expandido do bloco *treinamento do classificador*.

Capítulo 4

Base de Dados

4.1 Base de dados de Milne

A base de dados, contendo sinais de EEG, utilizada nesta dissertação, foi concebida por Milne *et al.*, da Universidade de *Sheffield* [63]. Para criação da base, originalmente, um total de 102 pessoas foram recrutadas. No entanto, devido a problemas técnicos, os sinais foram coletados para 99 participantes: 38 diagnosticados previamente com autismo (sendo 19 mulheres) e 61 não diagnosticados (sendo 32 mulheres), com idades entre 18 e 68 anos [62]. Os dados dos sinais de EEG foram coletados pelo sistema *biosemi active two EEG* [70] em uma câmara blindada eletricamente durante uma estimulação visual e, depois, em repouso de olhos fechados. Após isso, estes foram filtrados online com uma passagem de banda de 0,01 – 140 Hz e digitalizados a uma taxa de amostragem de 2.048 Hz. Todos os desvios de canal foram mantidos abaixo de 25Ω .

Ainda segundo Milne *et al.* [62], a estimulação visual envolvia a apresentação de um tabuleiro de xadrez preto e branco, gerado no *Psychtoolbox* [71] em um monitor de 20 pol dentro do *Matlab*. Após 200 iterações, os dados de estado de repouso foram adquiridos: os participantes permaneceram sentados e foram solicitados a fechar os olhos enquanto o EEG era registrado por 150 s. A mesma base de dados foi utilizada posteriormente por Dickinson *et al.* [72], em que se examinam os padrões de lentidão oscilatória relacionada à idade em adultos com TEA, capturados por reduções no pico de frequência alfa (PAF) do cérebro.

Em função da limitação de participantes que forneceram consentimento as-

sinado para compartilhamento de dados, com aprovação ética pela *health research authority* [73], apenas os sinais de 28 indivíduos diagnosticados previamente com TEA e outros 28 TD tornaram-se públicos, com reamostragem para 512 Hz. As gravações originais foram convertidas em arquivos *.set* e *.fdt* via *EEGLab*. Há um arquivo *.fdt* e um arquivo *.set* para cada gravação, o primeiro contém os dados e o segundo informações sobre os parâmetros da gravação. Por fim, vale ressaltar que atualmente existem poucos conjuntos de dados públicos disponíveis de sinais EEG obtidos de indivíduos com TEA, sendo este, aqui apresentado, o único encontrado disponível para experimentação dos modelos desenvolvidos neste trabalho. As Figuras 4.1 e 4.2 a seguir expõem amostras de sinais no domínio do tempo e da frequência, respectivamente, de um indivíduo TD. No domínio do tempo (Figura 4.1), os valores da diferença de potencial são apresentados em função do tempo ($\mu V/s$), enquanto na transformação para o domínio da frequência (Figura 4.2), o sinal é plotado em termos de densidade espectral de potência logarítmica (eixo y) e frequência (eixo x), com destaque de um mapa de intensidade de atividade cerebral por região para as frequências de 6, 10 e 22 Hz.

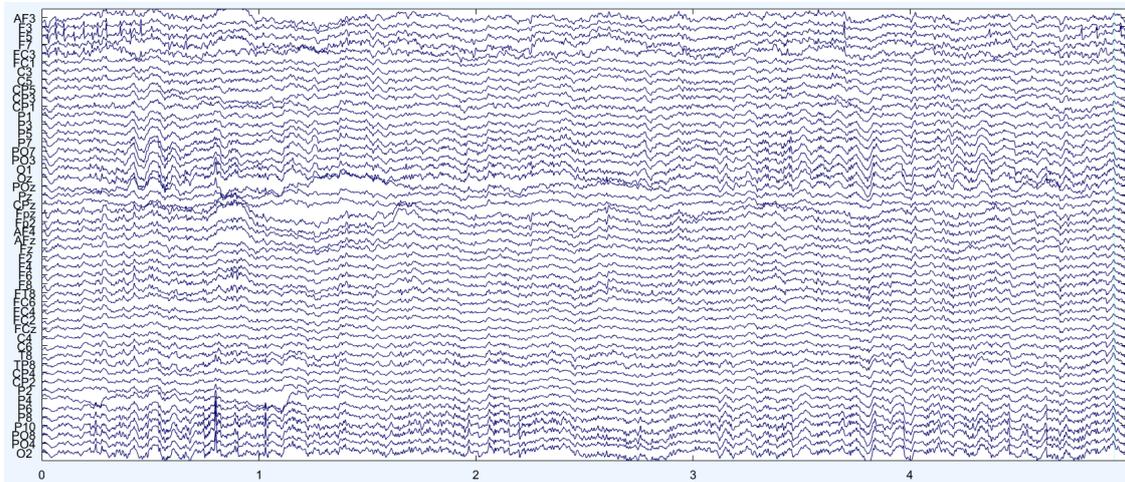


Figura 4.1: Amostra de sinal EEG de 64 canais, no domínio do tempo, do indivíduo tipicamente desenvolvido n° 54, no intervalo 0 – 5 s.

De forma equivalente, as Figuras 4.3 e 4.4 expõem amostras de sinais no domínio do tempo e da frequência, respectivamente, de um indivíduo diagnosticado com TEA.

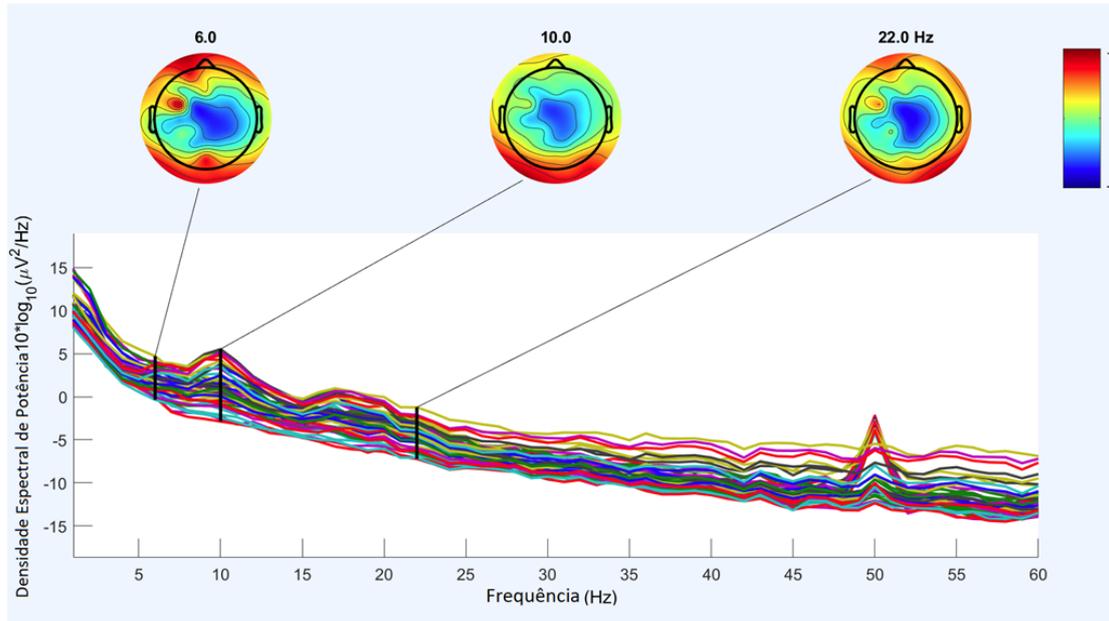


Figura 4.2: Amostra de sinal EEG de 64 canais, no domínio da frequência, do indivíduo tipicamente desenvolvido n° 54, na banda de frequência de 0 – 60 Hz.

4.2 Outras bases de dados

Embora a base apresentada na seção anterior seja a única lograda para experimentação dos modelos desenvolvidos neste trabalho, outras bases de dados de EEG foram prospectadas em conjunto. No trabalho proposto por Bhaskarachary *et al.* [48], os autores utilizaram a base de dados de acesso limitado de *child mind institute biobank* [49], a qual contém 88 sinais de EEG de crianças TD e 100 com TEA. A fim de empregar esta base nos experimentos dos modelos propostos, houve inúmeras tentativas de contato com o Instituto para concessão da referida base, porém, sem sucesso. De igual modo, não foi possível obter a base dos estudos de Kang *et al.* [2] (realizado com 49 crianças TEA e 48 TD) e Jayawardana *et al.* [3] (realizado com 8 crianças TEA e 9 TD) por se tratarem de bases de domínio privado.

Além disso, consultei o repositório de bases públicas *autism brain imaging data exchange* (ABIDE) [74], a qual agrega dados funcionais e estruturais de imagens cerebrais coletados de diversos laboratórios, no entanto, as coleções disponíveis possuem apenas imagens de ressonância magnética funcional em estado de repouso (R-fMRI). Outro repositório de dados pesquisado, o *national database for autism research* (NDAR) [75], combina a função de repositório de dados, que contém dados genéticos, fenotípicos, clínicos e de imagens médicas, e a função de plataforma de

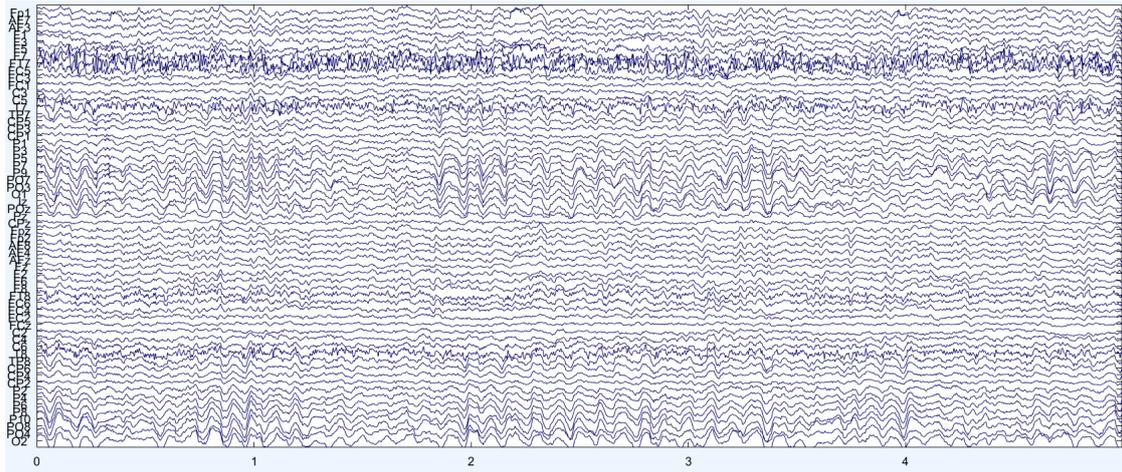


Figura 4.3: Amostra de sinal EEG de 64 canais, no domínio do tempo, do indivíduo diagnosticado com TEA n° 11, no intervalo 0 – 5 s.

comunidade científica, que define as ferramentas e políticas padrão para integrar os recursos computacionais desenvolvidos por instituições de pesquisa científica, privadas fundações e outras agências federais e estaduais que apoiam a pesquisa sobre TEA, porém não foram encontrados registros de bases de EEG até o momento desta pesquisa.

Dentre as dificuldades apontadas, inclui-se o sigilo médico, as questões éticas e de consentimento ao trabalhar com dados de EEG, especialmente quando se trata de dados de indivíduos com TEA. Neste sentido, é fundamental obter o consentimento informado dos participantes ou de seus responsáveis legais; os dados de EEG devem ser coletados e armazenados de forma a proteger a privacidade dos participantes; os participantes devem ser informados especificamente sobre o uso do EEG, incluindo a colocação de eletrodos no couro cabeludo e a natureza não invasiva da técnica; os pesquisadores devem implementar medidas rigorosas de segurança de dados para proteger os dados de EEG coletados e explicar claramente os possíveis benefícios da pesquisa aos participantes, bem como os benefícios potenciais para a sociedade em geral; por fim, pesquisas envolvendo seres humanos devem ser submetidas a uma revisão ética por um comitê de ética de pesquisa. Tais barreiras dificultam a aquisição e o compartilhamento de dados, motivo pelo qual, neste trabalho, apenas uma base pública disponível foi encontrada.

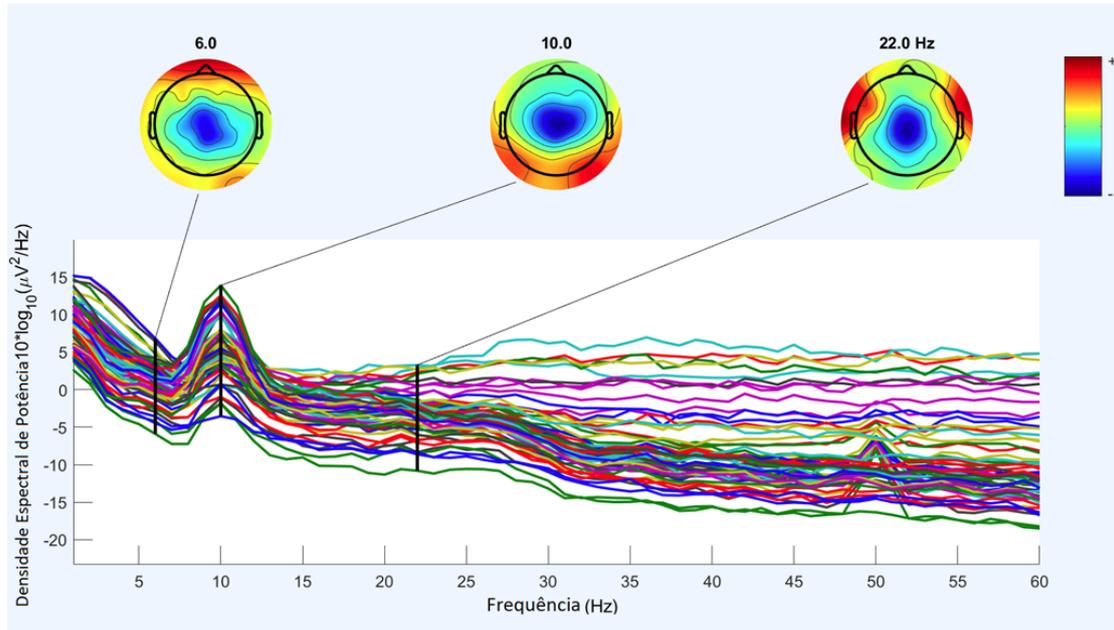


Figura 4.4: Amostra de sinal EEG de 64 canais, no domínio da frequência, do indivíduo diagnosticado com TEA n° 11, na banda de frequência de 0 – 60 Hz.

4.3 Síntese das bases de dados

As principais bases de dados expostas nas seções anteriores deste capítulo encontram-se sintetizadas na Tabela 4.1 a seguir, a qual reúne informações do quantitativo de indivíduos TD e TEA usados nos experimentos, faixa etária, taxa de amostragem de coleta, tempo de gravação em cada indivíduo e o domínio da base de dados, se público ou privado. Saliento que não foram encontrados dados de EEG nas bases ABIDE e NDAR pesquisadas.

Tabela 4.1: Síntese das bases de dados

Base de dados	Quant. indivíduos TD	Quant. indivíduos TEA	Faixa etária (anos)	Taxa de amostragem (Hz)	Tempo de gravação (s)	Domínio
Child Mind Institute Biobank [49]	88	100	5 - 19	250	20	Privado
Base de Kang <i>et al.</i> [2]	48	49	3 - 6	1000	360	Privado
Autism Brain Imaging Data Exchange (ABIDE) [74]	-	-	-	-	-	Público
Jayawardana <i>et al.</i> [3]	9	8	5 - 17	250	180	Privado
National Database for Autism Research (NDAR) [75]	-	-	-	-	-	Público
Base de Milne <i>et al.</i> [63]	28	28	18 - 68	512	150	Público

Capítulo 5

Experimentos e Resultados

5.1 Introdução

Nesta seção, apresentam-se as etapas do procedimento experimental, como também a análise dos resultados gerados por tais experimentos de modo a validar a metodologia proposta para classificação de sinais EEG para o diagnóstico do TEA, a partir da observação dos resultados das principais métricas para avaliação de desempenho das redes CNN e RNN. Para fins de validação do modelo proposto, utilizo a base de dados disponibilizada por Milne *et al.*, da Universidade de *Sheffield* [63], a qual contém assinaturas eletrofisiológicas de 28 indivíduos diagnosticados previamente com TEA e outros 28 TD, com idades entre 18 e 68 anos e duração de 150s para cada sinal (*vide* Capítulo 4). Pretendo, com os experimentos desenvolvidos, responder aos seguintes questionamentos da pesquisa:

- Transformando sinais EEG unidimensionais em imagens bidimensionais a partir da concatenação de canais, é possível extrair características destas imagens que distingam indivíduos previamente diagnosticados com TEA daqueles não-diagnosticados, com base em um modelo de CNN?
- Utilizando as mesmas etapas iniciais do modelo CNN, como se comportam as métricas para um modelo RNN com arquitetura LSTM?
- Frente à maioria dos estudos citados na Seção 2.7, os quais utilizam em seus experimentos sinais de EEG coletados de crianças (< 17 anos), os resultados

alcançados para os sinais de indivíduos adultos (18 – 68 anos) são igualmente relevantes para o auxílio ao diagnóstico do TEA?

- A utilização das técnicas de *transfer learning* e *autoencoder* refletem em ganhos consideráveis nas principais métricas quando comparadas aos do modelo base?
- Como se comportam as métricas de desempenho das redes quando aplicada a técnica de ODP no pré-processamento? O que ocorre com o desempenho das redes se retirado totalmente o bloco de pré-processamento?
- A partir da observação dos resultados das principais métricas de desempenho das redes propostas, é concebível a generalização do modelo para outras bases de dados de sinais EEG coletadas de indivíduos diagnosticados com o mesmo transtorno?

5.2 Setup

O *setup* dos experimentos deste trabalho é realizado conforme a seguir:

1. Ambiente de desenvolvimento: Os algoritmos são implementados utilizando o sistema operacional Microsoft Windows 10, na versão 21H2 de 64 bits, em um processador Intel(R) Core(TM) i7-7500U CPU @ 2.70 GHz 2.90 GHz. Utilizo o serviço em nuvem Google Colaboratory [76] como ambiente de desenvolvimento, com habilitação de GPU na seguinte configuração: NVIDIA Tesla K80, 12GB, CUDA versão 11.1, cuDNN versão 8.0.5.
2. Para concepção dos algoritmos, emprego a linguagem de programação Python [77], na versão 3.11.0, juntamente com o software Matlab [57] na sua versão 3.2.
3. Bibliotecas: Em conjunto, utilizo as seguintes bibliotecas: toolbox *EEGLab* [56] na versão 4.3 para pré-processamento dos dados, Keras [78] e Tensorflow [79] na versão 2.12.0 para implementação das redes neurais profundas, dentre outras bibliotecas auxiliares como Numpy [80], Matplotlib [81] e Scikit-learn [82].

5.3 Métricas de Desempenho

As métricas empregadas neste trabalho são similares àquelas apresentadas em [2, 3, 23, 48, 52]. Consideram-se, para avaliação do desempenho dos sistemas propostos, as métricas: acurácia, área sob a curva ROC (AUC, do inglês *area under the ROC curve*), erro, precisão, *recall* e *f1-score*, todas comumente adotadas na literatura [83]. A métrica acurácia, a qual mede a proporção de previsões corretas sobre o número total de instâncias avaliadas, é definida conforme equação (5.1).

$$\text{Acurácia} = \frac{TP + TN}{TP + FP + TN + FN}, \quad (5.1)$$

onde TP é o *true-positive*, TN o *true-negative*, FP o *false-positive* e FN o *false-negative*. Estas variáveis, quando recombinaadas, formam outras métricas de desempenho, que podem ser utilizadas em conjunto para avaliação de um modelo.

A métrica erro [84] mede a proporção de previsões incorretas em relação ao número total de instâncias avaliadas, conforme define a equação (5.2).

$$\text{Erro} = \frac{FP + FN}{TP + FP + TN + FN}, \quad (5.2)$$

A precisão [85], métrica usada para medir os padrões positivos que são previstos corretamente a partir do total de padrões previstos em uma classe positiva, pode ser definida matematicamente através da equação (5.3).

$$\text{Precisão} = \frac{TP}{TP + FP}, \quad (5.3)$$

Na equação (5.4), define-se a métrica *recall* [86], entendida como a medida da fração de padrões positivos que são classificados corretamente.

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (5.4)$$

Em sequência, a métrica *f1-score* [87] representa a média harmônica entre os valores de *recall* e precisão, conforme equação (5.5).

$$f1-score = \frac{2 * p * r}{p + r}, \quad (5.5)$$

Finalmente, a métrica AUC [88], usada para medir a área abaixo da curva ROC (do inglês, *receiver operating characteristic*), pode ser calculada a partir da equação (5.6), para um problema de classificação binária.

$$AUC = \frac{Sp - Np(Nn + 1)/2}{Np \times Nn}, \quad (5.6)$$

onde Sp é a soma de todos os exemplos positivos classificados, enquanto Np e Nn denotam o número de exemplos positivos e negativos, respectivamente. De acordo com Han *et al.* [89], a métrica AUC foi comprovada teórica e empiricamente melhor que a métrica de acurácia para avaliar o desempenho do classificador e discriminar uma solução ótima durante o treinamento.

5.4 Projeto dos algoritmos de classificação

5.4.1 Configuração da arquitetura CNN

Para o projeto dos algoritmos de classificação, utilizo a biblioteca de rede neural de código aberto *Keras* [78], em *Python*. O diagrama da Figura 5.1 a seguir apresenta o modelo da arquitetura base da rede neural convolucional proposta, formada por diversas camadas. A *camada de entrada de dados* recebe os dados de entrada padronizados em $64 \times 1000 \times 1$ para n amostras disponíveis. A camada *conv2D* 3×3 é composta por três partes: a convolução *2D*, a qual foi projetada neste trabalho para manter a dimensão de entrada e aumentar ou diminuir a quantidade de canais [65]; a função de ativação *leakyRelu*, que executa uma operação de limite não linear em que qualquer valor de entrada menor que zero é multiplicado por um fator de escala fixo igual a 0,1 [66] e; o *maxpooling*, o qual simplifica as informações da camada anterior reduzindo sua dimensionalidade a partir dos valores máximos em uma determinada região [67]. Essa camada composta pode ser repetida por diversas vezes, alterando-se apenas a quantidade de canais de convolução. Em seguida, utilizo a camada *GlobalAveragePooling2D*, para geração de um mapa de

recursos para cada categoria correspondente da tarefa de classificação, e as camadas *dense leaky relu* e *dense sigmoid*, com suas respectivas funções de ativação, para o propósito de classificação binária [90].

Em geral, para o treinamento da CNN proposta, utilizo o otimizador *Adam*, método de descida de gradiente estocástico baseado na estimativa adaptativa de momentos de primeira e segunda ordem, usado em substituição ao procedimento estocástico clássico de gradiente descendente para atualizar pesos de rede de forma iterativa [91], com base em dados de treinamento. Em conjunto, utilizo a função de custo *categorical_crossentropy* para o modelo de classificação multiclases, onde há dois ou mais rótulos de saída. Alternativamente, neste trabalho foram utilizados nos experimentos outros modelos de otimizadores, como, *stochastic gradient descent* (SGD), *adaptive gradient algorithm* (AdaGrad) e *root mean square propagation* (RMSProp).

Além da arquitetura transcrita nos parágrafos anteriores, algumas modificações na rede convolucional foram empreendidas. Cito o uso da técnica de transferência de aprendizagem (do inglês, *transfer learning*), a qual reutiliza o conhecimento aprendido de uma tarefa com muitos dados de treinamento rotulados disponíveis em uma nova tarefa que não possui tantos dados [92]. Neste trabalho, experimento os modelos pré-treinados *ResNet-50*, *MobileNet* e *MobileNetV2*. Por se tratar de imagens com 3 canais de entrada (RGB), adiciono à frente uma camada de convolução *2D* com 3 mapas de características para conversão dos dados de entrada padronizados de 1 canal, sem alteração da dimensão linha/coluna. Em outro experimento, a técnica de *autoencoder* foi explorada para investigação de possíveis ganhos com a compactação/extração de recursos de entrada e redução de ruídos a partir da adição de camadas densas com diferentes neurônios decrescentes.

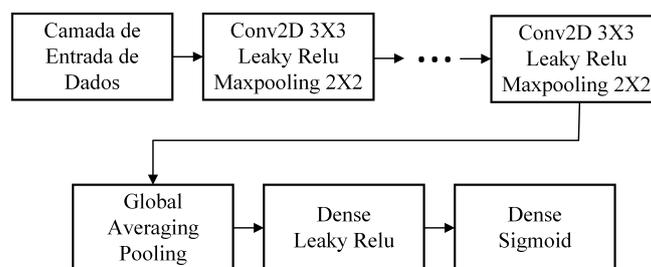


Figura 5.1: Modelo de Arquitetura CNN

5.4.2 Configuração da arquitetura RNN (LSTM)

A RNN utilizada para contrastar com o método anteriormente explanado foi a LSTM [68], conhecida por sua capacidade de aprender dependências de ordem em problemas de predição de sequência. A Figura 5.2 apresenta o modelo de arquitetura utilizado. Tal como na CNN, a *camada de entrada de dados* recebe os dados padronizados em 64×1.000 para n amostras disponíveis. As camadas *LSTM dropout* seguintes realizam o aprendizado dos dados sequenciais e são experimentadas com diversos parâmetros configuráveis, na intenção de se encontrar o melhor modelo para a base de dados disponível. A camada *dense sigmoid*, de 2 neurônios, recebe os dados de saída da última camada LSTM de forma inteiramente conectada para fins de classificação binária [90]. De maneira semelhante, para o treinamento da rede RNN, diferentes otimizadores foram utilizados, tais como, *Adam*, *SGD*, *RMSProp* e função de custo *binary_crossentropy*, usada recorrentemente em modelos de classificação binária.



Figura 5.2: Modelo de arquitetura RNN (LSTM).

5.5 Análise de Desempenho

O procedimento experimental para as Subseções 5.5.1 a 5.5.6 adiante segue etapas comuns e bem definidas, com modificações apenas no que se refere às particularidades de cada sistema, baseando-se na metodologia apresentada no Capítulo 3, sobretudo no diagrama em blocos geral do sistema proposto da Figura 3.1, repetido a seguir para maior comodidade ao leitor. As etapas são:

1. Execução do bloco de pré-processamento no *Matlab*, com carregamento e leitura da base de dados por meio da *toolbox EEGLab*, conforme metodologia descrita na Seção 3.2. Aplicam-se diferentes reamostragens para verificação de desempenho dos modelos, o que está ligado diretamente à quantidade de amostras geradas para a fase de treinamento da rede neural. Os sinais bidimensionais são padronizados em 64×1000 em todos os experimentos, por conter

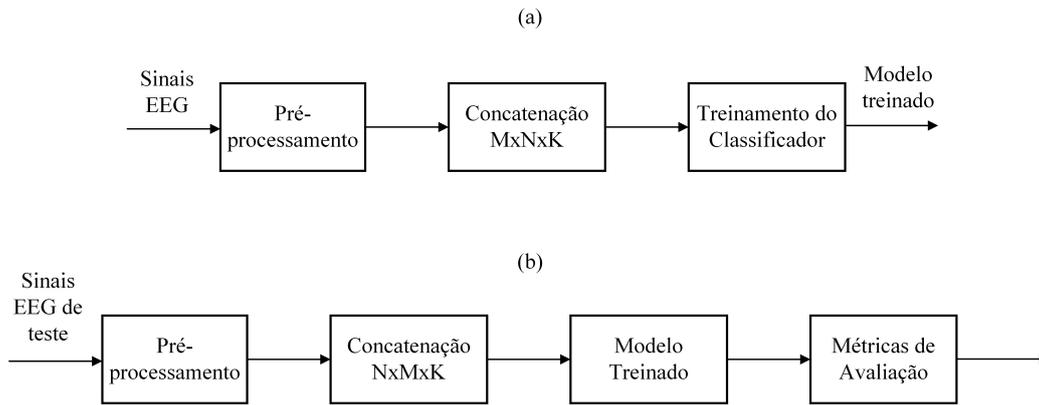


Figura 5.3: Diagrama em blocos geral do sistema proposto: (a) treinamento do classificador, (b) execução do modelo treinado.

uma quantidade suficiente de informações para treinamento, equiparando-se a uma imagem de dimensão 256 x 256 pixels. Para fins de experimentação, nos sistemas abordados nas Subseções 5.5.1 e 5.5.4, o bloco de pré-processamento é executado sem a técnica de ODP, enquanto nas Subseções 5.5.2 e 5.5.5, são incluídas as técnicas de ODP para a reordenação dos segmentos de canais. Em complemento, nas Subseções 5.5.3 e 5.5.6, o bloco de pré-processamento é totalmente retirado para avaliação dos efeitos da ausência deste.

2. Carregamento e leitura, no bloco de classificação executado em *Python*, das variáveis pré-processadas na etapa anterior, conforme metodologia da Seção 3.3. Os dados são particionados em treinamento, validação e teste, com diferentes atribuições percentuais. Os parâmetros de treinamento, como, função de custo, otimizador, taxa de aprendizagem, tamanho do *batch*, épocas, etc. são modificados para os diversos treinamentos da rede, a fim de se encontrar o melhor modelo classificador. Para fins de experimentação e avaliação, adicionam-se ao sistema da Subseção 5.5.1, separadamente, as técnicas de *autoencoder* e *transfer learning*, utilizando-se para esta última os modelos pré-treinados, ResNet-50, MobileNet e MobileNetV2.
3. Por fim, plotagens gráficas para visualização e análise dos resultados de todos os sistemas, tais como, curvas de acurácia de treinamento e validação, curvas de erro de treinamento e validação, matrizes de confusão e curvas ROC, além dos resultados numéricos das principais métricas utilizadas.

Salienta-se que, em virtude do grande número de parâmetros alteráveis e

das inúmeras possibilidades de combinação entre esses, inicialmente, os parâmetros configurados que ofereciam os melhores resultados na base de testes do primeiro sistema (Subseção 5.5.1), tornavam-se fixos para as experimentações dos demais sistemas, e assim sucessivamente. Ao final, a Subseção 5.5.7 compara os resultados dos sistemas propostos com estudos anteriores de diagnóstico de TEA.

5.5.1 Sistema de classificação com CNN

As Tabelas 5.1, 5.2 e Apêndice B.1, apresentam, respectivamente, os resultados, parâmetros e arquiteturas para 30 experimentações distintas, utilizando a abordagem de redes neurais convolucionais para a tarefa de classificação do transtorno do espectro autista. É importante observar que as referidas tabelas se complementam e suas linhas são resultantes de um único modelo. Dentre as 30 experimentações, destaca-se, em negrito, o melhor resultado alcançado. Os resultados dos experimentos de 1 a 12 referem-se à arquitetura CNN base, de 6 camadas; 13 a 24 à CNN com os modelos pré-treinados ResNet-50, MobileNet e MobileNetV2 e; 25 a 30 à CNN precedida de *autoencoders*.

Nota-se, a partir das principais métricas de desempenho para avaliação, que o experimento de n° 8 alcançou os melhores resultados em geral, com acurácia de classificação de 99,51 %, AUC de 99,95 %, erro de 0,0266 e aproximadamente 100 % para precisão, *recall* e *f1-score* (Tabela 5.1), seguido pelos experimentos n° 7 e 9, com desempenhos semelhantes. Os três melhores experimentos citados possuem a mesma arquitetura base, com 4 conjuntos de camadas convolucionais e 2 densas, conforme detalhado no Apêndice B.1. Também possuem configurações de parâmetros semelhantes, utilizando os dados com frequência de amostragem de 512 Hz, função de custo *categorical crossentropy* e otimizador Adam com taxa de aprendizagem de 0,001, diferenciando-se, portanto, apenas nas divisões percentuais de treinamento, validação e teste. O melhor modelo (n° 8) possui a divisão 60|20|20.

De modo geral, observa-se uma depreciação nos desempenhos quando os dados de entrada são reamostrados para 128 Hz, indicando uma maior dificuldade da rede em reconhecer padrões na etapa de treinamento, devido à redução de informações disponíveis na imagem. Além disso, nota-se uma tendência de geração de melhores resultados quando aplicados os otimizadores Adam e RMSprop, em de-

trimento do método mais simples de descida do gradiente estocástico, SGD. Não houve alterações significativas nos resultados para as diferentes funções de custo, *categorical crossentropy* e *binary crossentropy*.

Os experimentos de 13 a 24, com os modelos pré-treinados ResNet-50, MobileNet e MobileNetV2, não ensejaram em ganhos expressivos observados nas principais métricas, quando comparados ao modelo base inicialmente proposto de 6 camadas. Mostra-se que, a despeito das diferentes combinações de parâmetros (Tabela 5.2) e arquiteturas (Apêndice B.1), as redes formadas por esta abordagem possuem certa limitação para a tarefa de classificação. O melhor modelo encontrado (experimento nº 19), com a ResNet-50, resultou em uma acurácia de 84,48%, AUC de 90,25% e erro consideravelmente alto de 13,84, resultados em harmonia com o estudo de Zhang *et al.* [93], indicando que os modelos de aprendizagem por transferência, em geral treinados a partir da base de dados *imagenet*, podem não conferir benefícios significativos aos modelos treinados com imagens formadas a partir de sinais EEG.

Nos experimentos de 25 a 30 são apresentados os principais ensaios com a técnica de *autoencoders*, com o objetivo de se criar uma representação dos dados originais de entrada para verificação de possíveis ganhos de desempenho, dada a sua capacidade de extrair os principais recursos dos sinais [94]. No presente caso, para as diferentes combinações de frequências de amostragem, divisão de dados, otimizadores, arquiteturas etc., o melhor modelo encontrado (experimento nº 30) alcançou a máxima de 82,27% de acurácia na base de testes, utilizando duas camadas densas sequenciais de 6400 e 640 neurônios no treinamento do codificador (Apêndice B.1), o que sugere a perda de características importantes para o treinamento no processo da compactação.

A Figura 5.4 apresenta as curvas de acurácia e erro no processo de treinamento e validação do melhor modelo obtido, durante as 200 épocas. Verifica-se, a partir dos gráficos, o aumento gradual da acurácia de treinamento da CNN proposta, chegando-se ao valor máximo de 100% aproximadamente na 80ª época de treinamento, seguida logo abaixo pela curva de acurácia na base de validação, que oscila vagamente entre 98 e 99% de forma estável. À direita da mesma figura, percebe-se a diminuição constante do erro de treinamento e validação, aproximando-se de zero na 80ª época de aprendizagem, evidenciando a capacidade do modelo em reconhecer

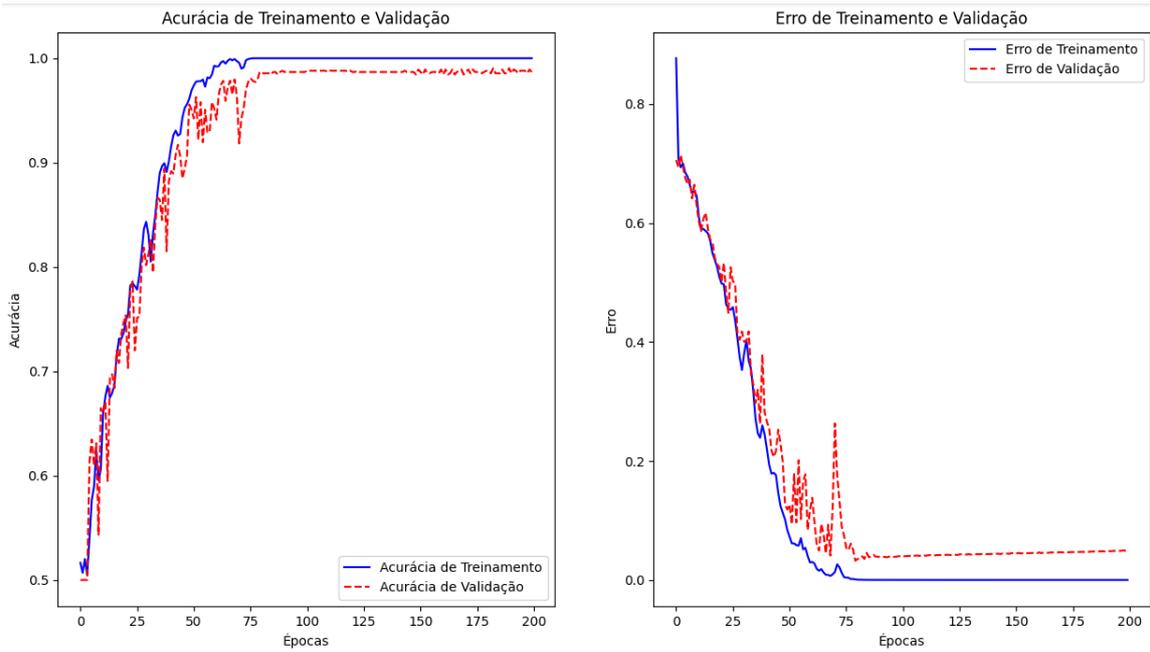


Figura 5.4: Curvas de acurácia e erro de treinamento e validação do experimento n° 8 com CNN.

padrões e classificar corretamente os sinais EEG.



Figura 5.5: Matriz de confusão do experimento n° 8 com CNN.

A Figura 5.5 permite a visualização do desempenho de classificação do melhor modelo proposto com CNN. Esta matriz relaciona as classificações reais ou verdadeiras (eixo y) com as preditas pela rede (eixo x). As classes 0 e 1 referem-se às

classes TEA e TD, respectivamente. Dentre as 832 amostras (416, TEA e 416, TD) da base de dados separadas para testes (20% do total), constata-se que o algoritmo classificou corretamente, como TEA, 414 amostras das 416 disponíveis, taxando erroneamente como negativo (ou TD) apenas 2 amostras. Observa-se ainda que o mesmo algoritmo classificou 415 amostras corretamente como TD, falhando apenas para 1 amostra. Estes resultados revelam a habilidade do modelo em classificar, com baixa margem de erro, as duas classes de indivíduos.

Finalmente, a curva ROC da Figura 5.6 para o melhor modelo proposto traça as taxas de verdadeiros positivos (eixo y) versus falsos positivos (eixo x), a qual estabelece o quanto o modelo é capaz de distinguir entre duas classes. No caso em questão, é possível notar uma curva muito próxima da ideal, confirmado pelo valor da área da Curva ROC igual a 0,9996.

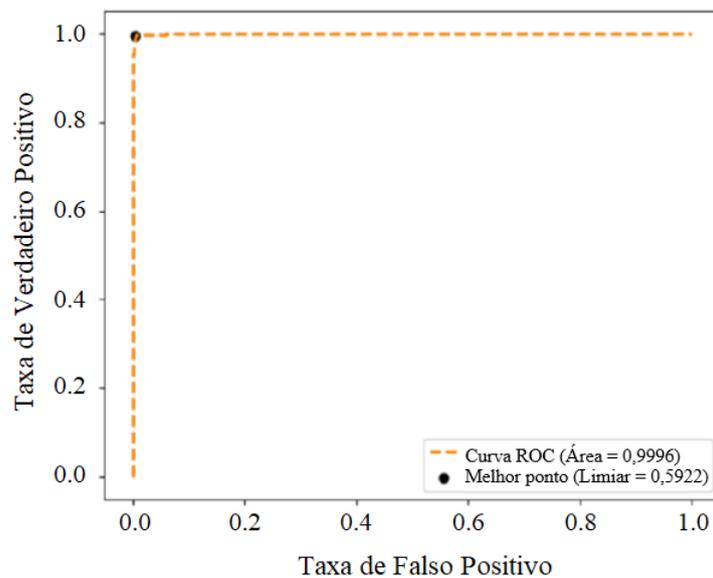


Figura 5.6: Curva ROC do experimento n° 8 CNN.

Tabela 5.1: Resultados do sistema de classificação TEA com CNN, no conjunto de teste.

Experimento N°	RESULTADOS						Tempo de Treino (s)
	Acurácia	AUC	Erro	Precisão	Recall	F1-score	
1	0,9702	0,9977	0,0549	0,95	0,98	0,96	150,9
2	0,9653	0,9885	0,1711	0,91	0,99	0,95	130,81
3	0,9703	0,995	0,0843	0,91	1	0,95	812,12
4	0,9802	0,9925	0,0739	0,99	0,97	0,98	803,76
5	0,8143	0,8719	0,4494	0,86	0,74	0,8	277,58
6	0,9752	0,9969	0,0726	0,98	0,97	0,98	325,28
7	0,9919	0,9997	0,0143	0,93	1	0,96	447,63
8	0,9951	0,9995	0,0266	1	1	1	250,71
9	0,9817	0,9979	0,0578	0,94	1	0,97	226,97
10	0,9783	0,997	0,0726	0,97	0,99	0,98	257,91
11	0,9803	0,9983	0,0354	1	0,98	0,99	246,23
12	0,9599	0,9882	0,1092	0,74	1	0,85	384,43
13	0,8399	0,6383	33,011	0,5	1	0,67	318,44
14	0,805	0,8733	0,4536	0,75	0,94	0,83	260,59
15	0,54	0,4966	12,307	0,5	1	0,67	811,034
16	0,6	0,6422	51,529	0,57	0,67	0,61	137,7
17	0,6133	0,5	0,641	0,5	1	0,67	203,49
18	0,675	0,7205	30,602	0,62	0,73	0,67	165,95
19	0,8448	0,9025	13,847	0,84	0,88	0,86	592,46
20	0,5321	0,5	0,8664	0,5	1	0,67	185,74
21	0,5321	0,5	0,8664	0,5	1	0,67	185,74
22	0,8397	0,834	0,5715	0,53	1	0,69	714,72
23	0,5528	0,5	0,7023	0,5	1	0,67	360,51
24	0,5961	0,5168	0,6985	0,5	1	0,67	440,43
25	0,6732	0,6327	0,6873	0,5	1	0,67	158,96
26	0,7237	0,7312	0,7243	0,65	1	0,69	148,96
27	0,5531	0,5392	0,5223	0,5	1	0,65	132,96
28	0,5231	0,5372	0,5142	0,63	1	0,67	128,96
29	0,8136	0,8317	0,7819	0,86	1	0,87	455,72
30	0,8227	0,8377	0,7915	0,78	1	0,86	498,16

Tabela 5.2: Parâmetros do sistema de classificação TEA com CNN.

Experi- mento N°	PARÂMETROS										
	Dimensão de entrada X	Frequência de Amostra gem	Labels Y	Treina- mento (%)	Teste (%)	Vali- dação (%)	Função de custo	Otimi- zador	Taxa de apren- dizagem	Batch Size	Épo- cas
1	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	64	200
2	64x1000x2020	256 Hz	2020x1	60	20	20	CC	Adam	0,001	64	200
3	64x1000x2020	256 Hz	2020x1	60	20	20	CC	Adam	0,001	64	200
4	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	64	200
5	64x1000x2020	256 Hz	2020x1	60	20	20	BC	SGD	0,01	64	200
6	64x1000x2020	256 Hz	2020x1	60	20	20	BC	RMSprop	0,01	64	200
7	64x1000x4158	512 Hz	4158x1	70	15	15	CC	Adam	0,001	64	200
8	64x1000x4158	512 Hz	4158x1	60	20	20	CC	Adam	0,001	64	200
9	64x1000x4158	512 Hz	4158x1	50	25	25	CC	Adam	0,001	64	200
10	64x1000x4158	512 Hz	4158x1	60	40	40	BC	SGD	0,01	64	200
11	64x1000x4158	512 Hz	4158x1	60	40	40	BC	RMSprop	0,01	64	200
12	64x1000x1002	128 Hz	1002x1	70	15	15	CC	Adam	0,001	64	200
13	64x1000x1002	128 Hz	1002x1	70	15	15	CC	Adam	0,001	64	200
14	64x1000x1002	128 Hz	1002x1	60	20	20	CC	Adam	0,001	64	200
15	64x1000x1002	128 Hz	1002x1	70	15	15	CC	SGD	0,01	64	200
16	64x1000x1002	128 Hz	1002x1	60	20	20	BC	Adam	0,001	64	200
17	64x1000x1002	128 Hz	1002x1	70	15	15	CC	Adam	0,001	64	200
18	64x1000x1002	128 Hz	1002x1	60	20	20	BC	Adam	0,001	64	200
19	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	64	200
20	64x1000x2020	256 Hz	2020x1	60	20	20	CC	Adam	0,001	64	200
21	64x1000x2020	256 Hz	2020x1	60	20	20	CC	RMSprop	0,001	64	200
22	64x1000x4158	512 Hz	4158x1	70	15	15	CC	Adam	0,001	64	200
23	64x1000x4158	512 Hz	4158x1	70	15	15	CC	Adam	0,001	64	200
24	64x1000x4158	512 Hz	4158x1	70	15	15	CC	SGD	0,01	64	200
25	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	64	200
26	64x1000x2020	256 Hz	2020x1	60	20	20	BC	Adam	0,001	64	200
27	64x1000x1002	128 Hz	1002x1	70	15	15	BC	RMSprop	0,001	64	200
28	64x1000x1002	128 Hz	1002x1	60	20	20	CC	Adam	0,001	64	200
29	64x1000x4158	512 Hz	4158x1	70	15	15	CC	SGD	0,01	64	200
30	64x1000x4158	512 Hz	4158x1	60	20	20	BC	RMSprop	0,001	64	200

CC: Categorical crossentropy / BC: Binary crossentropy

5.5.2 Sistema de classificação com CNN e técnica de ODP

Nesta nova abordagem, a técnica de ODP é adicionada ao bloco de pré-processamento descrito na Seção 3.2, utilizando o mesmo sistema de classificação com CNN. As Tabelas 5.3, 5.4 e Apêndice B.2, apresentam, respectivamente, os resultados, parâmetros e arquiteturas para 10 experimentações distintas, desta vez escolhidas com base no desempenho dos melhores modelos da subseção anterior. Novamente, é importante observar que as referidas tabelas se complementam e suas linhas são resultantes de um único modelo.

Observa-se, a partir das principais métricas de desempenho para avaliação, que o experimento de n° 9, em destaque, alcançou os melhores resultados dentre as configurações experimentadas, com acurácia de classificação de 99,63 %, AUC de 99,99 %, erro de 0,0091 e aproximadamente 100 % para precisão, *recall* e *f1-score* (Tabela 5.3), resultados ligeiramente superiores aos do sistema CNN sem a técnica de ODP. Além disso, nota-se uma considerável redução no tempo de treinamento dos modelos em geral (para uma mesma quantidade de épocas), alcançando uma redução de aproximadamente 30s na comparação direta com o melhor modelo treinado sem ODP (Tabela 5.1), que possui os mesmos parâmetros e arquitetura.

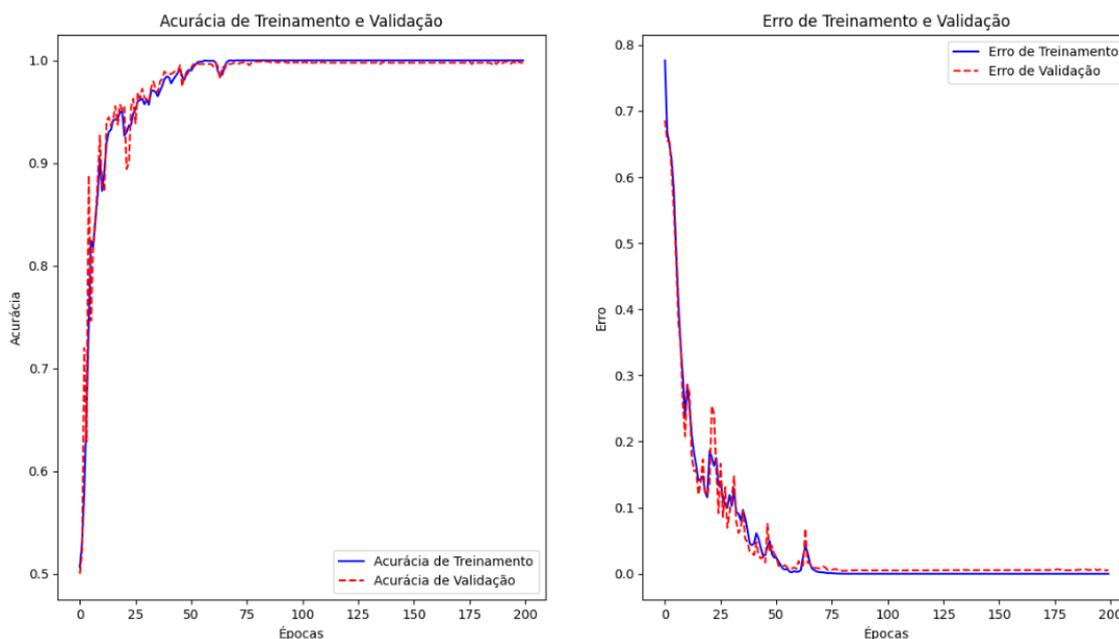


Figura 5.7: Curvas de acurácia e erro de treinamento e validação do experimento n° 9 com CNN e técnica de ODP.

A Figura 5.7 apresenta as curvas de acurácia e erro no processo de treina-

mento e validação do melhor modelo obtido para o sistema CNN com ODP (experimento de n° 9), durante as 200 épocas. Verifica-se, a partir dos gráficos, um aumento abrupto da acurácia de treinamento do modelo CNN com ODP, atingindo 90 % de acurácia logo na 15^o época de treinamento, em contraste com o aumento gradual do modelo CNN simples (Figura 5.4), que atingiu 90 % de acurácia somente na 40^o época de treinamento. Observa-se também, nos gráficos do modelo CNN com ODP, uma maior aproximação entre as curvas de acurácia de treinamento e validação, quando da estabilização do sistema. Os mesmos comportamentos citados ocorrem com as curvas de erro de treinamento e validação (à direita da Figura 5.7). Estes resultados evidenciam o aumento da capacidade e velocidade de aprendizagem dos modelos CNN quando precedidos pela aplicação da reordenação dos segmentos de canais, no pré-processamento, com a técnica de ODP.

Verdadeiro	Classe 0 - TEA	416	0
	Classe 1 - TD	3	413
		Classe 0 - TEA	Classe 1 - TD
		Predito	

Figura 5.8: Matriz de confusão do experimento n° 9 com CNN e técnica de ODP.

A matriz de confusão da Figura 5.8 possibilita a visualização do desempenho de classificação do melhor modelo proposto para o sistema CNN com ODP. Dentre as 832 amostras (416, TEA e 416, TD) da base de dados separadas para testes (20 % do total), constata-se que o algoritmo classificou corretamente, como TEA, 100 % das amostras disponíveis. Além disso, o mesmo algoritmo classificou 413

amostras corretamente como TD, falhando apenas para 3. Em concordância com o modelo CNN simples, estes resultados revelam a habilidade do modelo CNN com ODP em classificar, com baixa margem de erro, as duas classes de indivíduos, sendo superior para indivíduos com o TEA. Por fim, a curva ROC da Figura 5.9 para o melhor modelo proposto traça as taxas de verdadeiros positivos (eixo y) versus falsos positivos (eixo x), onde é possível notar uma curva ideal, confirmado pela área da curva ROC igual a 1,00.

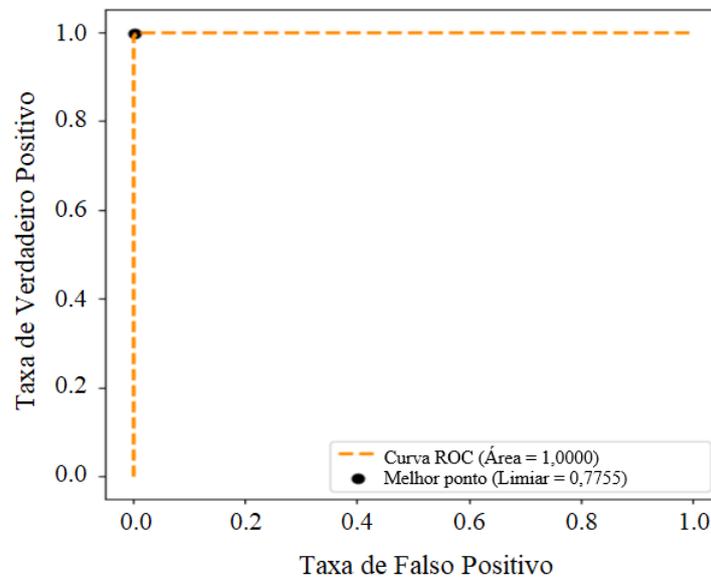


Figura 5.9: Curva ROC do experimento n° 9 com CNN e técnica de ODP.

Tabela 5.3: Resultados do sistema de classificação TEA com CNN e técnica de ODP, no conjunto de teste.

Experimento N°	RESULTADOS						
	Acurácia	AUC	Erro	Precisão	Recall	F1-score	Tempo de Treino (s)
1	0,9933	0,9998	0,0182	0,99	1	0,99	129,79
2	0,9900	0,9997	0,0211	0,98	1	0,99	112,22
3	0,9956	0,9999	0,0161	0,99	1	1	131,23
4	0,9752	0,9973	0,1114	0,96	0,99	0,98	112,95
5	0,9777	0,9990	0,0731	0,98	0,98	0,98	111,23
6	0,9939	0,9998	0,0221	0,99	1	0,99	225,12
7	0,9915	0,9998	0,0187	0,99	0,99	0,99	221,42
8	0,9921	0,9998	0,0032	0,98	1	0,99	258,74
9	0,9963	0,9999	0,0091	0,99	1	1	220,91
10	0,9934	0,9998	0,0055	0,99	1	1	202,79

Tabela 5.4: Parâmetros do sistema de classificação TEA com CNN e técnica de ODP.

Experi- mento N°	PARÂMETROS										
	Dimensão de entrada X	Frequência de Amostra gem	Labels Y	Treina- mento (%)	Teste (%)	Vali- dação (%)	Função de custo	Otimi- zador	Taxa de aprendi- zagem	Batch Size	Épo- cas
1	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	64	200
2	64x1000x2020	256 Hz	2020x1	60	20	20	CC	Adam	0,001	64	200
3	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	64	200
4	64x1000x2020	256 Hz	2020x1	60	20	20	BC	SGD	0,01	64	200
5	64x1000x2020	256 Hz	2020x1	60	20	20	BC	RMSprop	0,001	64	200
6	64x1000x4158	512 Hz	4158x1	60	20	20	BC	SGD	0,01	64	200
7	64x1000x4158	512 Hz	4158x1	60	20	20	BC	RMSprop	0,01	64	200
8	64x1000x4158	512 Hz	4158x1	70	15	15	CC	Adam	0,001	64	200
9	64x1000x4158	512 Hz	4158x1	60	20	20	CC	Adam	0,001	64	200
10	64x1000x4158	512 Hz	4158x1	50	25	25	CC	Adam	0,001	64	200

CC: Categorical crossentropy / BC: Binary crossentropy

5.5.3 Sistema de classificação com CNN sem pré-processamento

Nesta subseção, o bloco de pré-processamento descrito na Seção 3.2 é retirado do sistema de classificação com CNN para análise dos efeitos de sua ausência. As Tabelas 5.5, 5.6 e Apêndice B.3, apresentam, respectivamente, os resultados, parâmetros e arquiteturas para 5 experimentações distintas, também escolhidas com base no desempenho dos melhores modelos da subseção anterior. Nota-se, a partir das principais métricas de desempenho para avaliação, que o experimento de nº 3, em destaque, alcançou os melhores resultados dentre as configurações experimentadas para este sistema, com acurácia de classificação de 94,37%, AUC de 98,68%, erro de 0,0742, precisão de 94%, *recall* de 99% e *f1-score* de 97% (Tabela 5.5), resultados notadamente inferiores aos do sistema CNN com o bloco de pré-processamento (Tabela 5.1), indicando a importância do tratamento dos dados antes da entrada na rede neural. Além disso, é evidente o considerável aumento no tempo de treinamento dos modelos sem pré-processamento (para uma mesma quantidade de épocas), podendo chegar a aproximadamente o dobro do tempo do melhor modelo do sistema CNN com pré-processamento (Tabela 5.1), que possui parâmetros e arquitetura semelhantes.

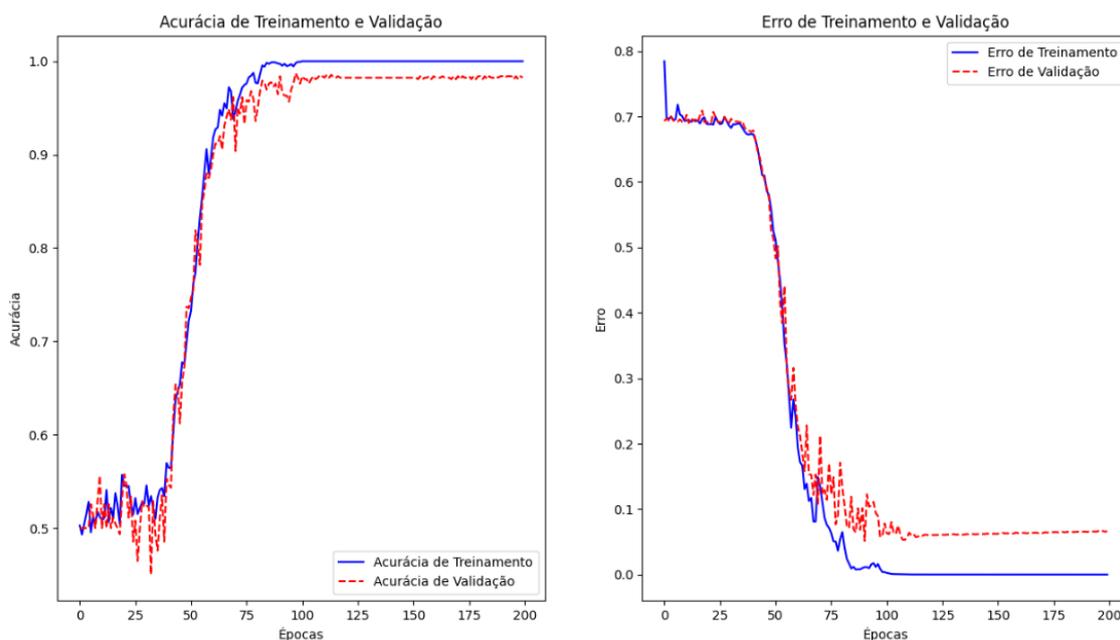


Figura 5.10: Curvas de acurácia e erro de treinamento e validação do experimento nº 3 com CNN sem pré-processamento.

A dificuldade inicial do sistema em trabalhar com dados brutos é confirmada pelas curvas de acurácia e erro de treinamento e validação da Figura 5.10, plotadas para o melhor modelo obtido do sistema CNN sem pré-processamento (experimento de n° 3), durante as 200 épocas. Observa-se, a partir dos gráficos, um retardamento no aumento da acurácia de treinamento e validação do modelo, iniciando mais significativamente somente na 40^o época de treinamento, em contraste com o sistema CNN com pré-processamento da Figura 5.4, que inicia com aumento gradativo logo nas primeiras iterações. Nota-se ainda, nos gráficos do modelo CNN sem pré-processamento, o mesmo retardamento no decaimento do erro de treinamento e um maior distanciamento entre as linhas de erro de treinamento e validação, quando da estabilização do sistema. Estes resultados evidenciam a diminuição da capacidade e velocidade de aprendizagem dos modelos CNN quando retirado o bloco de pré-processamento.

Verdadeiro	Classe 0 - TEA	298	14
	Classe 1 - TD	18	294
		Classe 0 - TEA	Classe 1 - TD
		Predito	

Figura 5.11: Matriz de confusão do experimento n° 3 com CNN sem pré-processamento.

Na análise da matriz de confusão da Figura 5.11, para o sistema CNN sem pré-processamento, é possível notar uma maior quantidade de erros de classificação de indivíduos TD e TEA, quando comparados à matriz de confusão do sistema CNN completo da Figura 5.5. Visualiza-se que, dentre as 624 amostras (312, TEA e 312,

TD) da base de dados separadas para testes (15% do total), o algoritmo classificou incorretamente 14 amostras de TEA e 18 amostras TD, ressaltando o prejuízo da retirada do bloco de pré-processamento. Apesar da diminuição da quantidade de acertos deste sistema, a curva ROC da Figura 5.12 se mostrou positiva, uma vez que esta se baseia em taxas comparativas entre verdadeiros positivos (eixo y) versus falsos positivos (eixo x).

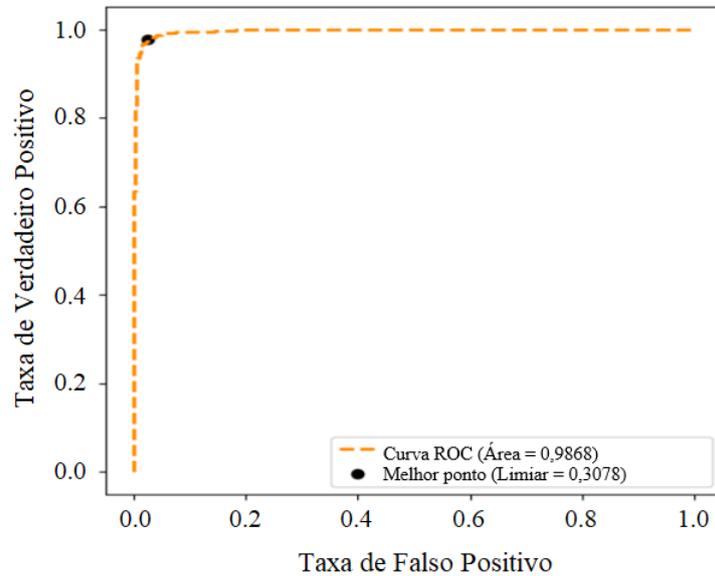


Figura 5.12: Curva ROC do experimento n° 3 com CNN sem pré-processamento.

Tabela 5.5: Resultados do sistema de classificação TEA com CNN sem pré-processamento, no conjunto de teste.

Experimento N°	RESULTADOS						
	Acurácia	AUC	Erro	Precisão	Recall	F1-score	Tempo de Treino (s)
1	0,8882	0,9574	0,2824	0,92	0,85	0,88	444,63
2	0,9259	0,9776	0,0912	0,97	0,96	0,97	410,69
3	0,9437	0,9868	0,0742	0,94	0,99	0,97	461,18
4	0,9307	0,9812	0,0894	0,98	0,96	0,97	413,99
5	0,8973	0,9612	0,1456	0,93	0,89	0,91	424,92

Tabela 5.6: Parâmetros do sistema de classificação TEA com CNN sem pré-processamento.

Experi- mento N°	PARÂMETROS										
	Dimensão de entrada X	Frequência de Amostragem	Labels Y	Treina- mento (%)	Teste (%)	Vali- dação (%)	Função de custo	Otimi- zador	Taxa de aprendi- zagem	Batch Size	Épo- cás
1	64x1000x4158	512 Hz	4158x1	60	20	20	BC	SGD	0,01	64	200
2	64x1000x4158	512 Hz	4158x1	60	20	20	BC	RMSprop	0,01	64	200
3	64x1000x4158	512 Hz	4158x1	70	15	15	CC	Adam	0,001	64	200
4	64x1000x4158	512 Hz	4158x1	60	20	20	CC	Adam	0,001	64	200
5	64x1000x4158	512 Hz	4158x1	50	25	25	CC	Adam	0,001	64	200

CC: Categorical crossentropy / BC: Binary crossentropy

5.5.4 Sistema de classificação com RNN

De maneira equivalente, as tabelas 5.7, 5.8 e Apêndice B.4, apresentam, respectivamente, os resultados, parâmetros e arquiteturas para 20 experimentações distintas, utilizando a abordagem de redes neurais recorrentes (com arquitetura LSTM) para a tarefa de classificação do transtorno do espectro autista. As referidas tabelas são complementares e resultantes de um único sistema de classificação. Dentre as 20 experimentações, destaca-se, em negrito, o melhor resultado alcançado, após diversas modificações de parâmetros e arquiteturas, tais como, frequência de amostragem, função de custo, otimizador, quantidades de camadas e quantidade de neurônios.

Observa-se, a partir das mesmas métricas de desempenho, que o experimento de n° 9 alcançou o melhor resultado global, com acurácia de classificação de 97,77 %, AUC de 98,23 %, erro de 0,236 e 98 % de precisão, *recall* e *f1-score* (Tabela 5.7), seguido pelos experimentos n° 18 e 19, com desempenhos semelhantes. O melhor resultado possui uma arquitetura com três camadas LSTM, de 128, 64 e 32 neurônios, com *dropout* de 20 % entre elas, e finalizada com uma camada densa binária, conforme detalhado na tabela do Apêndice B.4. Em termos de parâmetros (Tabela 5.8), utilizo os dados com frequência de amostragem de 256 Hz, função de custo *binary crossentropy*, otimizador RMSprop com taxa de aprendizagem de 0,001 e divisões percentuais de treinamento, validação e teste de 60, 20 e 20 %, respectivamente.

Comparando o melhor desempenho com os dois subsequentes (experimentos n° 18 e 19), verifica-se que o aumento da frequência de amostragem de dados para 512 Hz e ainda o aumento de camadas LSTM da rede (+1 de 256 neurônios) não resultou em desempenhos superiores nas métricas utilizadas, indicando que o modelo do experimento de n° 9 é suficiente para a complexidade dos dados de entrada em questão.

A Figura 5.13 apresenta as curvas de acurácia e erro no processo de treinamento e validação do melhor modelo RNN obtido, durante as 150 épocas. Verifica-se, a partir dos gráficos, o aumento abrupto da acurácia de treinamento da RNN proposta, chegando-se ao valor máximo de 100 % aproximadamente na 15^o época de treinamento, diferentemente da curva de acurácia na base de validação que apresentou instabilidade nos instantes iniciais e um pico negativo durante a iteração, a meio caminho. Após isso, a curva mostra aparente estabilização entre 96 e 97 %

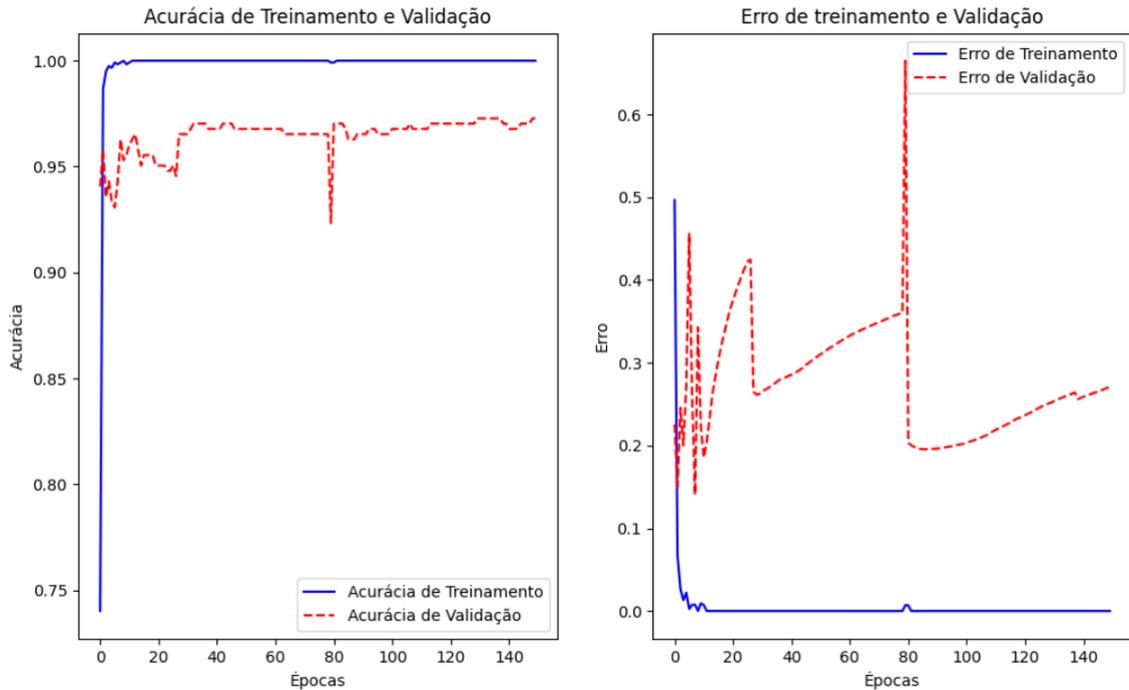


Figura 5.13: Curvas de acurácia e erro de treinamento e validação do experimento n° 9 com RNN.

de acurácia. À direita da mesma figura, percebe-se a diminuição abrupta do erro de treinamento, mas não da validação, que se apresenta grandemente instável em várias épocas. Ainda assim, o modelo possui potencial para reconhecer padrões e classificar corretamente os sinais EEG, através da arquitetura RNN.

A Figura 5.14 permite a visualização do desempenho de classificação do melhor modelo proposto com RNN. Reiterando, esta matriz relaciona as classificações reais ou verdadeiras (eixo y) com as preditas pela rede (eixo x). As classes 0 e 1 referem-se às classes TEA e TD, respectivamente. Dentre as 404 amostras (202, TEA e 202, TD) da base de dados separadas para testes, constata-se que o algoritmo classificou corretamente, como TEA, 198 amostras das 202 disponíveis, taxando erroneamente como negativo (ou TD) 4 amostras. Observa-se ainda que o mesmo algoritmo classificou 197 amostras corretamente como TD, falhando para 5 amostras. Estes resultados revelam a habilidade do modelo RNN em classificar, com baixa margem de erro, para as duas classes de indivíduos.

Por fim, a curva ROC da Figura 5.15 para o melhor modelo proposto (n° 9), traça as taxas de verdadeiros positivos (eixo y) versus falsos positivos (eixo x). De maneira semelhante ao modelo CNN, ainda que inferior, é possível notar também



Figura 5.14: Matriz de confusão do experimento n° 9 RNN

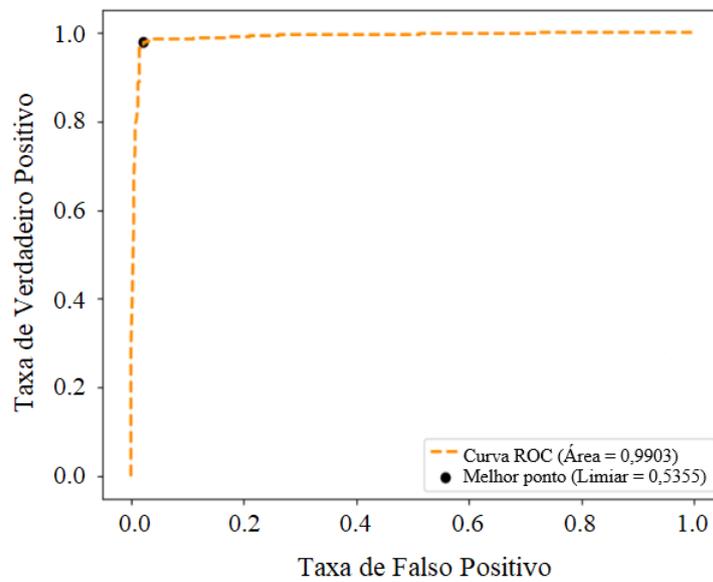


Figura 5.15: Curva ROC do experimento n° 9 RNN

uma curva muito próxima da ideal, confirmado pelo valor da AUC de 0,9903.

Tabela 5.7: Resultados do sistema de classificação TEA com RNN, no conjunto de teste.

Experimento N°	RESULTADOS						Tempo de Treino (s)
	Acurácia	AUC	Erro	Precisão	Recall	F1-score	
1	0,915	0,969	0,4382	0,92	0,91	0,91	88,92
2	0,8866	0,9262	0,3224	0,9	0,91	0,91	101,72
3	0,9066	0,9564	0,5787	0,91	0,91	0,91	102,04
4	0,8976	0,9122	0,3156	0,9	0,92	0,92	134,14
5	0,9191	0,9693	0,345	0,9	0,94	0,92	215,25
6	0,9356	0,9681	0,312	0,98	0,89	0,93	198,13
7	0,8943	0,9366	0,408	0,89	0,90	0,89	202,44
8	0,9339	0,9712	0,30495	0,97	0,89	0,93	345,5
9	0,9777	0,9823	0,236	0,98	0,98	0,98	305,62
10	0,9628	0,9674	0,4057	0,96	0,97	0,96	391,95
11	0,9471	0,9693	0,3227	0,96	0,93	0,95	257,15
12	0,9207	0,9554	0,4405	0,94	0,90	0,92	267,98
13	0,9554	0,9794	0,344	0,96	0,95	0,96	208,52
14	0,9166	0,958	0,4789	0,90	0,94	0,92	312,95
15	0,899	0,9531	0,4311	0,92	0,87	0,90	304,62
16	0,9567	0,991	0,1303	0,96	0,97	0,96	458,9
17	0,935	0,971	0,2821	0,95	0,92	0,93	409,94
18	0,9759	0,9878	0,2435	0,98	0,97	0,98	442,57
19	0,9727	0,9825	0,2937	0,97	0,97	0,97	545,18
20	0,9102	0,9621	0,3947	0,91	0,9	0,91	443,75

Tabela 5.8: Parâmetros do sistema de classificação TEA com RNN.

Experi- mento N°	PARÂMETROS										
	Dimensão de entrada X	Frequência de Amostra gem	Labels Y	Treina- mento (%)	Teste (%)	Vali- dação (%)	Função de custo	Otimi- zador	Taxa de aprendi- zagem	Batch Size	Épo- cas
1	64x1000x1002	128 Hz	1002x1	60	20	20	BC	Adam	0,001	8	100
2	64x1000x1002	128 Hz	1002x1	50	25	25	CC	SGD	0,01	8	100
3	64x1000x1002	128 Hz	1002x1	70	15	15	CC	RMSprop	0,001	8	100
4	64x1000x1002	128 Hz	1002x1	70	15	15	CC	RMSprop	0,001	8	100
5	64x1000x2020	256 Hz	2020x1	70	15	15	BC	Adam	0,001	8	100
6	64x1000x2020	256 Hz	2020x1	70	15	15	BC	Adam	0,001	8	100
7	64x1000x2020	256 Hz	2020x1	60	20	20	BC	Adam	0,001	8	100
8	64x1000x2020	256 Hz	2020x1	70	15	15	BC	Adam	0,001	8	150
9	64x1000x2020	256 Hz	2020x1	60	20	20	BC	RMSprop	0,001	8	150
10	64x1000x2020	256 Hz	2020x1	60	20	20	BC	RMSprop	0,001	8	150
11	64x1000x2020	256 Hz	2020x1	70	15	15	BC	Adam	0,001	8	100
12	64x1000x2020	256 Hz	2020x1	70	15	15	CC	Adam	0,001	8	100
13	64x1000x2020	256 Hz	2020x1	60	20	20	CC	RMSprop	0,001	8	100
14	64x1000x4158	512 Hz	4158x1	70	15	15	BC	SGD	0,001	8	100
15	64x1000x4158	512 Hz	4158x1	60	20	20	BC	RMSprop	0,001	8	100
16	64x1000x4158	512 Hz	4158x1	70	15	15	BC	Adam	0,001	8	100
17	64x1000x4158	512 Hz	4158x1	60	20	20	BC	Adam	0,001	8	100
18	64x1000x4158	512 Hz	4158x1	70	15	15	CC	RMSprop	0,001	8	100
19	64x1000x4158	512 Hz	4158x1	70	15	15	CC	RMSprop	0,001	8	100
20	64x1000x4158	512 Hz	4158x1	70	15	15	CC	SGD	0,01	8	100

CC: Categorical crossentropy / BC: Binary crossentropy

5.5.5 Sistema de classificação com RNN e técnica de ODP

Neste sistema, a técnica de ODP é adicionada ao bloco de pré-processamento descrito na Seção 3.2, utilizando o mesmo sistema de classificação com RNN. As Tabelas 5.9, 5.10 e Apêndice B.5, apresentam, respectivamente, os resultados, parâmetros e arquiteturas para 10 experimentações distintas, escolhidas com base no desempenho dos melhores modelos da subseção anterior.

Dentre as configurações experimentadas com ODP, nota-se, a partir das principais métricas de desempenho para avaliação, que o experimento de n° 7, em destaque, alcançou os melhores resultados na base de testes, com acurácia de classificação de 97,43%, AUC de 98,98%, erro de 0,2005, *recall* de 98% e 97% para precisão e *f1-score* (Tabela 5.9). Estes resultados, de modo geral, aproximam-se dos resultados do sistema RNN sem a técnica de ODP (Tabela 5.7), sendo inferior na métrica acurácia e superior nas métricas AUC e erro. Diferentemente do ocorrido com os modelos CNN, não houve redução significativa no tempo de treinamento dos modelos em geral (para uma mesma quantidade de épocas).

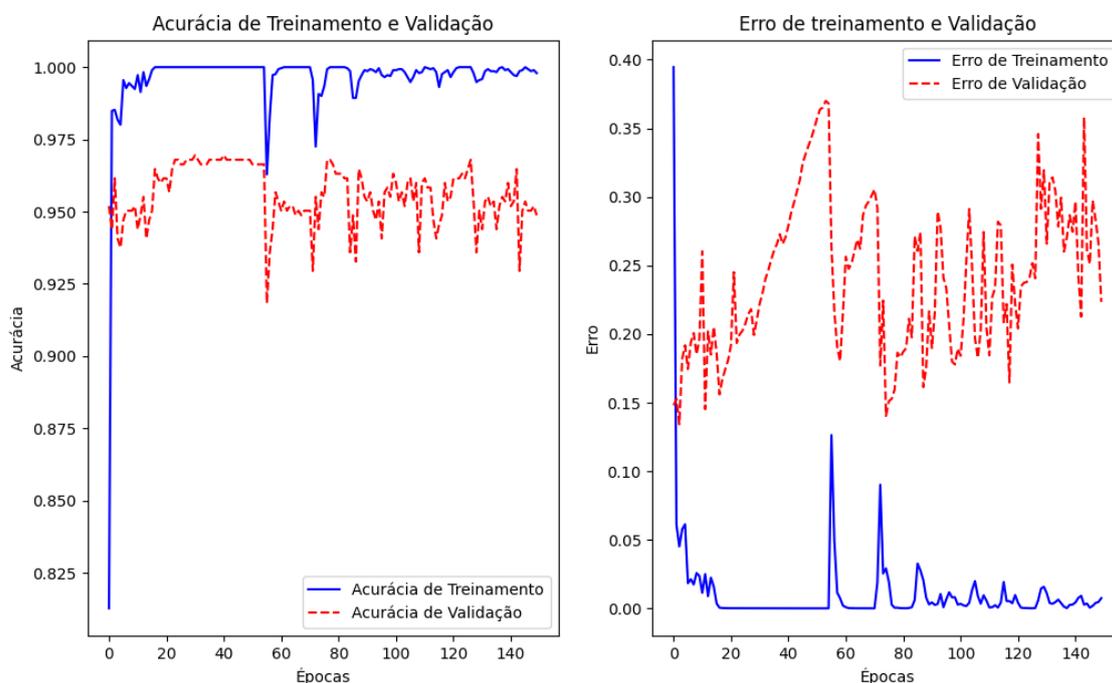


Figura 5.16: Curvas de acurácia e erro de treinamento e validação do experimento n° 7 com RNN e técnica de ODP.

A Figura 5.16 apresenta as curvas de acurácia e erro no processo de treinamento e validação do melhor modelo obtido para o sistema RNN com ODP (experi-

mento de n° 7), durante as 150 épocas. Verifica-se, a partir dos gráficos, o aumento abrupto da acurácia de treinamento do sistema RNN com ODP, não diferindo do modelo sem ODP. Entretanto, o modelo apresentou maior instabilidade durante as épocas de treinamento, quando comparado ao melhor modelo sem ODP. À direita da mesma figura, percebe-se a diminuição abrupta do erro de treinamento, mas não da validação, que se apresenta grandemente instável em várias épocas, junto à curva de treinamento. Estes resultados podem sugerir certa incompatibilidade dos modelos RNN quando precedidos pela aplicação da reordenação dos segmentos de canais, no pré-processamento, com a técnica de ODP.

Verdadeiro	Classe 0 - TEA	307	5
	Classe 1 - TD	11	301
		Classe 0 - TEA	Classe 1 - TD
		Predito	

Figura 5.17: Matriz de confusão do experimento n° 7 com RNN e técnica de ODP.

A matriz de confusão da Figura 5.17 permite a visualização do desempenho de classificação do melhor modelo proposto para o sistema RNN com ODP. Dentre as 624 amostras (312, TEA e 312, TD) da base de dados separadas para testes, constata-se que o algoritmo classificou corretamente, como TEA, 307 amostras das 312 disponíveis, taxando erroneamente como negativo (ou TD) 5 amostras. Observa-se ainda que o mesmo algoritmo classificou 301 amostras corretamente como TD, falhando para 11 amostras. Em que pese a instabilidade das curvas de treinamento e validação, o sistema RNN com ODP, na base de testes, possui a habilidade de

classificar entre as duas classes de indivíduos. Por fim, na Figura 5.18 é possível notar uma curva ROC próxima da ideal, confirmado pelo valor da AUC de 0,9898.

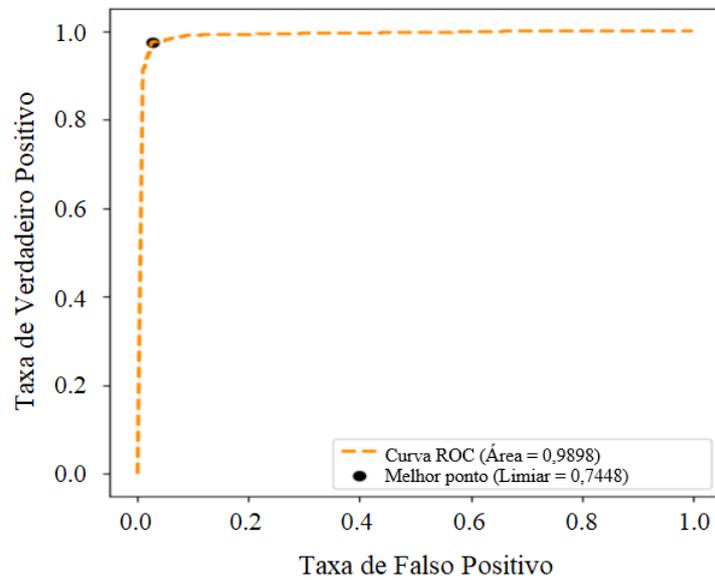


Figura 5.18: Curva ROC do experimento n° 7 com RNN e técnica de ODP.

Tabela 5.9: Resultados do sistema de classificação TEA com RNN e técnica de ODP, no conjunto de teste.

Experimento N°	RESULTADOS						
	Acurácia	AUC	Erro	Precisão	Recall	F1-score	Tempo de Treino (s)
1	0,9339	0,9683	0,2710	0,92	0,95	0,93	228,75
2	0,9603	0,9693	0,3425	0,96	0,96	0,96	204,44
3	0,9455	0,9558	0,3692	0,92	0,97	0,95	252,80
4	0,9207	0,9520	0,4349	0,93	0,91	0,92	121,17
5	0,9331	0,9440	0,6985	0,94	0,94	0,93	205,91
6	0,7836	0,8535	0,5538	0,76	0,83	0,79	264,13
7	0,9743	0,9898	0,2005	0,97	0,98	0,97	345,86
8	0,9579	0,9808	0,1887	0,95	0,96	0,96	317,51
9	0,9647	0,9897	0,2264	0,96	0,97	0,96	340,74
10	0,9679	0,9942	0,1626	0,98	0,96	0,97	370,62

Tabela 5.10: Parâmetros do sistema de classificação TEA com RNN e técnica de ODP.

Experi- mento N°	PARÂMETROS										
	Dimensão de entrada X	Frequência de Amostra gem	Labels Y	Treina- mento (%)	Teste (%)	Vali- dação (%)	Função de custo	Otimi- zador	Taxa de aprendi- zagem	Batch Size	Épo- cas
1	64x1000x2020	256 Hz	2020x1	70	15	15	BC	Adam	0,001	8	150
2	64x1000x2020	256 Hz	2020x1	60	20	20	BC	RMSprop	0,001	8	150
3	64x1000x2020	256 Hz	2020x1	60	20	20	BC	RMSprop	0,001	8	150
4	64x1000x2020	256 Hz	2020x1	70	15	15	BC	Adam	0,001	8	100
5	64x1000x2020	256 Hz	2020x1	60	20	20	CC	RMSprop	0,001	8	100
6	64x1000x4158	512 Hz	4158x1	70	15	15	BC	SGD	0,001	8	100
7	64x1000x4158	512 Hz	4158x1	70	15	15	BC	Adam	0,001	8	150
8	64x1000x4158	512 Hz	4158x1	60	20	20	BC	Adam	0,001	8	100
9	64x1000x4158	512 Hz	4158x1	70	15	15	CC	RMSprop	0,001	8	100
10	64x1000x4158	512 Hz	4158x1	70	15	15	CC	RMSprop	0,001	8	100

CC: Categorical crossentropy / BC: Binary crossentropy

5.5.6 Sistema de classificação com RNN sem pré-processamento

De forma equivalente ao sistema da Subseção 5.5.3, o bloco de pré-processamento descrito na Seção 3.2 é retirado do sistema de classificação com RNN para análise dos efeitos de sua ausência. As Tabelas 5.11, 5.12 e Apêndice B.6, apresentam, respectivamente, os resultados, parâmetros e arquiteturas para 5 experimentações distintas, também escolhidas com base no desempenho dos melhores modelos da subseção anterior. Nota-se, a partir das principais métricas de desempenho para avaliação, que o experimento de n° 2, em destaque, alcançou os melhores resultados dentre as configurações experimentadas para este sistema, com acurácia de classificação de 95,99 %, AUC de 97,37 %, erro de 0,2135, e 96 % para precisão, *recall* e *f1-score* (Tabela 5.11), resultados também notadamente inferiores aos do sistema RNN com o bloco de pré-processamento (Tabela 5.7), ressaltando a importância do tratamento dos dados antes da entrada na rede neural. Com a retirada do bloco em questão, é notório o considerável aumento no tempo de treinamento dos modelos sem pré-processamento (para uma mesma quantidade de épocas), alcançando em alguns casos o dobro do tempo do melhor modelo do sistema RNN com pré-processamento (Tabela 5.7), que possui parâmetros e arquitetura semelhantes.

A Figura 5.19, plotada para o melhor modelo obtido do sistema RNN sem pré-processamento (experimento de n° 2), durante as 200 épocas, apresenta as curvas de acurácia e erro de treinamento e validação, onde é possível observar a instabilidade que o sistema sofreu ao trabalhar com dados brutos, comparado ao modelo com pré-processamento da Figura 5.13. A mesma oscilação é observada no gráfico do erro de treinamento e validação (à direita da Figura 5.19). Estes resultados evidenciam a diminuição da capacidade e velocidade de aprendizagem dos modelos RNN quando retirado o bloco de pré-processamento.

Na análise da matriz de confusão da Figura 5.20, para o sistema RNN sem pré-processamento, é possível notar uma maior quantidade de erros de classificação de indivíduos TD e TEA, quando comparados à matriz de confusão do sistema CNN completo da Figura 5.14. Visualiza-se que, dentre as 624 amostras (312, TEA e 312, TD) da base de dados separadas para testes (15 % do total), o algoritmo classificou incorretamente 13 amostras de TEA e 12 amostras TD, confirmando o prejuízo da

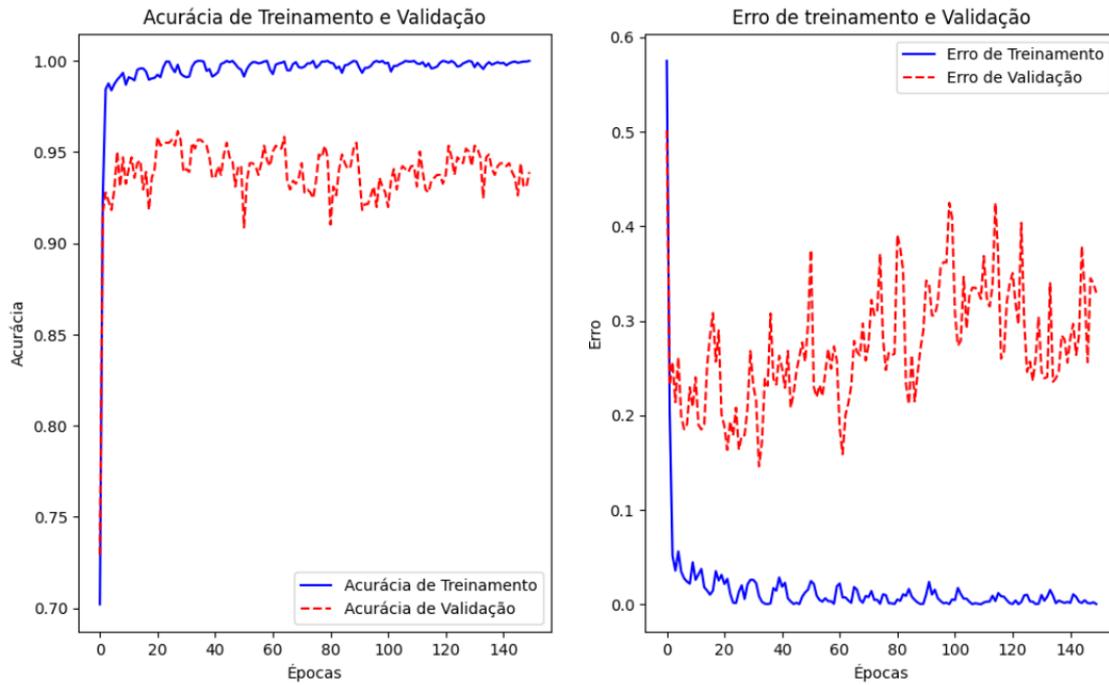


Figura 5.19: Curvas de acurácia e erro de treinamento e validação do experimento n° 2 com RNN sem pré-processamento.

Verdadeiro	Classe 0 - TEA	299	13
	Classe 1 - TD	12	300
		Classe 0 - TEA	Classe 1 - TD
		Predito	

Figura 5.20: Matriz de confusão do experimento n° 2 com RNN sem pré-processamento.

retirada do bloco de pré-processamento. Apesar da diminuição da quantidade de acertos deste sistema, a curva ROC da Figura 5.21 se mostrou positiva, uma vez que esta se baseia em taxas comparativas entre verdadeiros positivos (eixo y) versus

falsos positivos (eixo x).

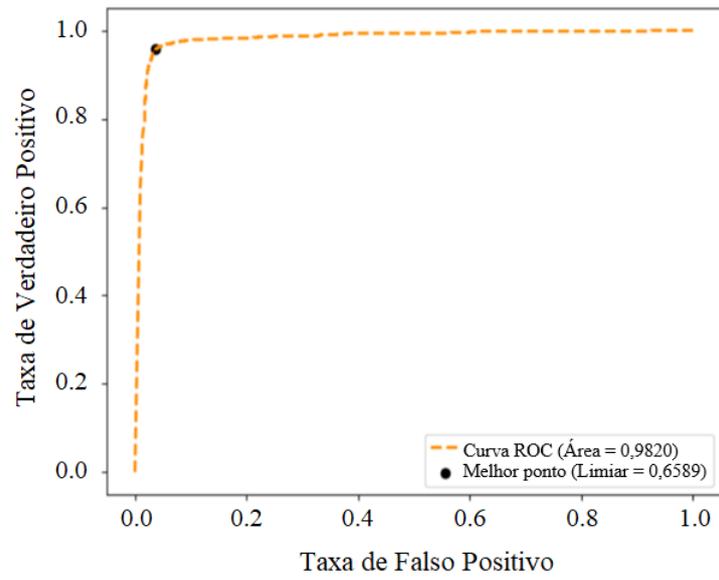


Figura 5.21: Curva ROC do experimento n° 2 com RNN sem pré-processamento.

Tabela 5.11: Resultados do sistema de classificação TEA com RNN sem pré-processamento, no conjunto de teste.

Experimento N°	RESULTADOS						
	Acurácia	AUC	Erro	Precisão	Recall	F1-score	Tempo de Treino (s)
1	0,8213	0,8856	0,7145	0,87	0,88	0,87	614,67
2	0,9599	0,9737	0,2135	0,96	0,96	0,96	802,34
3	0,9327	0,9425	0,3414	0,94	0,95	0,93	654,78
4	0,9121	0,9254	0,4216	0,93	0,94	0,92	634,15
5	0,8976	0,9176	0,5616	0,91	0,90	0,91	628,44

Tabela 5.12: Parâmetros do sistema de classificação TEA com RNN sem pré-processamento.

Experi- mento N°	PARÂMETROS										
	Dimensão de entrada X	Frequência de Amostragem	Labels Y	Treina- mento (%)	Teste (%)	Vali- dação (%)	Função de custo	Otimi- zador	Taxa de aprendi- zagem	Batch Size	Épo- cás
1	64x1000x4158	512 Hz	4158x1	70	15	15	BC	SGD	0,001	8	100
2	64x1000x4158	512 Hz	4158x1	70	15	15	BC	Adam	0,001	8	150
3	64x1000x4158	512 Hz	4158x1	60	20	20	BC	Adam	0,001	8	100
4	64x1000x4158	512 Hz	4158x1	70	15	15	CC	RMSprop	0,001	8	150
5	64x1000x4158	512 Hz	4158x1	70	15	15	CC	RMSprop	0,001	8	100

CC: Categorical crossentropy / BC: Binary crossentropy

5.5.7 Comparação de resultados com estudos anteriores de diagnóstico de TEA

Na Tabela 5.13, apresenta-se uma comparação dos melhores sistemas propostos com os estudos abordados na Seção 2.7, para a tarefa de auxílio ao diagnóstico do TEA. Na referida tabela, inclui-se ainda nosso artigo publicado no simpósio brasileiro de telecomunicações e processamento de sinais (SBrT2023) (Neto *et al.* [95]), como referência de sistema 2D-CNN antes da aplicação da técnica de ODP. Destaca-se que, com a mesma base de dados nos experimentos, nosso melhor sistema (ODP + 2D-CNN) superou em 17,83 pontos percentuais o melhor resultado de acurácia apresentado no trabalho de Milne *et al.* [62]. Para os demais sistemas com bases de dados distintas, a diferença percentual de acurácia é ainda expressiva, no entanto, possíveis diferenças na forma de aquisição e nos dispositivos de coleta de sinais, usados nos experimentos, podem prejudicar a comparação fidedigna das métricas utilizadas, servindo apenas de referência aproximada para outros estudos. Portanto, outras bases de dados são necessárias para verificação da capacidade de generalização dos sistemas propostos neste trabalho.

Tabela 5.13: Comparação com estudos anteriores de diagnóstico de TEA

Autor	Base de dados (BD)	Método	Acurácia
Bhaskarachary <i>et al.</i> [48]	Child Mind Institute	Extra Trees	67,70 %
Pistorius <i>et al.</i> [52]	Univers. de Califórnia	PCA + RQA + LDA/KNN	83,30 %
Kang <i>et al.</i> [2]	BD Kang <i>et al.</i> [2]	MRMR + SVM	85,44 %
Jayawardana <i>et al.</i> [3]	BD Jayawardana <i>et al.</i> [3]	1D-CNN	95,00 %
Ibrahim <i>et al.</i> [23]	Bonn, MIT, King A.	DWT + SE + kNN	94,62 %
Milne <i>et al.</i> [62]	BD Milne <i>et al.</i> [62]	DFA + ITC + MSE	81,80 %
Neto <i>et al.</i> [95]	BD Milne <i>et al.</i> [62]	2D-CNN	99,51 %
Sistema proposto 1	BD Milne <i>et al.</i> [62]	ODP + 2D-CNN	99,63 %
Sistema proposto 2	BD Milne <i>et al.</i> [62]	ODP + RNN	97,43 %

Capítulo 6

Conclusões

Nesta dissertação, concebemos um sistema de classificação de sinais EEG para auxílio ao diagnóstico do transtorno do espectro autista. Os resultados apresentados mostram que a estratégia de concatenação de sinais unidimensionais, transformando-os em bidimensionais, pode ser usada para a extração de características que possibilitem a detecção do TEA, a partir de modelos de redes neurais que utilizem imagens como entrada. A técnica de ordenação por diferença percentual se mostrou eficaz na tarefa de reordenação dos segmentos de canais dos sinais EEG. Quantitativamente, verificou-se que o melhor modelo proposto (2D-CNN com ODP), pôde distinguir, com acurácia de 99,63%, indivíduos diagnosticados com TEA daqueles tipicamente desenvolvidos, resultado ligeiramente superior ao do sistema 2D-CNN sem a técnica de ODP. Além disso, notou-se uma considerável redução no tempo de treinamento, quando comparado aos modelos sem ODP, e um crescimento acelerado da acurácia de treinamento, evidenciando o aumento da capacidade e velocidade de aprendizagem dos modelos CNN com ODP.

Quanto à utilização das técnicas de *transfer learning* e *autoencoder*, os resultados mostraram que os modelos de aprendizagem por transferência, em geral treinados a partir da base de dados *imagenet*, podem não conferir benefícios significativos aos modelos treinados com imagens formadas a partir de sinais EEG e, em relação aos modelos com *autoencoder*, pode haver perda de características importantes para o treinamento no processo da compactação. Por fim, a partir da observação dos resultados das principais métricas de desempenho das redes propostas, concluiu-se que os melhores modelos apresentados podem ser utilizados como base para o

treinamento de outras bases de dados semelhantes. Entretanto, não é possível afirmar se os modelos concebidos possuem a capacidade de generalização para outros sinais EEG coletados de indivíduos diagnosticados com o mesmo transtorno, uma vez que são necessários novos experimentos com diferentes bases de dados.

6.1 Proposta para Trabalhos Futuros

Nesta seção, apresentam-se algumas propostas para desenvolvimentos futuros relacionados aos métodos de classificação de sinais de eletroencefalograma abordados nesta dissertação. Assim, as propostas futuras são:

- Experimentar os modelos propostos em outras bases de dados públicas para verificação da capacidade de generalização dos algoritmos e, se necessário, treinar os modelos com a nova base de dados, em uma espécie de *transfer learning*.
- Realizar a aquisição de sinais EEG em pacientes diagnosticados previamente com TEA e em indivíduos TD, através de dispositivo de eletrodos de 64 canais ou mais, com aprovação por comitê de ética.
- Após validação do modelo em diferentes bases de dados e comprovada a eficácia do método para as demais bases, elaborar um software dedicado que integre todas as etapas do sistema, iniciando pela coleta dos sinais EEG em tempo real, depois realizando o pré-processamento com as técnicas apresentadas e utilizando o modelo treinado para calcular a probabilidade de presença do transtorno do espectro autista em pacientes.
- Identificar, através de novos experimentos, a quantidade mínima possível de eletrodos (canais) capazes de auxiliar no diagnóstico garantindo os mesmos níveis de acurácias encontrados.
- Identificar quais regiões cerebrais estão mais propícias a contribuir para o diagnóstico do transtorno, realizando experimentos com segmentação de canais e avaliação das métricas para cada configuração.

- Realizar experimentos com outras técnicas de ordenação por correlação na etapa de pré-processamento, inclusive com a retirada de níveis DC, e observar os resultados. Levantar as estatísticas e usar classificador com extrator de características.

Referências Bibliográficas

- [1] DEBORAH KERCHES, “Novo relatório do CDC”, <https://dradeborahkerches.com.br/novo-relatorio-aponta-1-a-cada-36-criancas-esta-no-espectro-do-autismo/>, 2023.
- [2] KANG, J., HAN, X., SONG, J., et al., “The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data”, *Computers in biology and medicine*, v. 120, pp. 103722, 2020.
- [3] JAYAWARDANA, Y., JAIME, M., JAYARATHNA, S., “Analysis of temporal relationships between ASD and brain activity through EEG and machine learning”. In: *2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI)*, pp. 151–158, 2019.
- [4] CABRAL, R., CABRAL, G., OTHERS, “Importance of early diagnosis and intervention in per-son with TEA for its inclusion in society”, *Brazilian Journal of Health Review*, v. 5, n. 2, pp. 7754–7765, 2022.
- [5] NAJARIAN, K., SPLINTER, R., *Biomedical signal and image processing*. Taylor & Francis, 2012.
- [6] CLANCY, E., MORIN, E., MERLETTI, R., “Sampling, noise-reduction and amplitude estimation issues in surface electromyography”, *Journal of Electromyography and Kinesiology*, v. 12, n. 1, pp. 1–16, 2002.
- [7] LEITE, N., PEREIRA, E., GURJÃO, E., et al., “Deep Convolutional Autoencoder for EEG Noise Filtering”. In: *Proc. IEEE Int. Conf. on Bioinformatics and Biomedicine*, pp. 2605–2612, 2018.

- [8] ROBERTO, M., DARIO, F., *Surface Electromyography: Physiology, Engineering, and Applications*. John Wiley & Sons, 2016.
- [9] GACEK, A., PEDRYCZ, W., *ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence*. Springer Science & Business Media, 2011.
- [10] SANEI, S., CHAMBERS, J., *EEG signal processing*. John Wiley & Sons, 2013.
- [11] MELO, W., FILHO, E., SILVA JR., W., “Electroencephalographic Signal Compression Based on Adaptive Segmentation and Video Encoders”. In: *Proc. XXXV Simp. Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*, pp. 1–5, 09 2017.
- [12] MELO, W., FILHO, E., SILVA JR., W., “SEMG signal compression based on two-dimensional techniques”, *BioMedical Engineering OnLine*, v. 15, pp. 1–31, 04 2016.
- [13] SOUZA, P., SILVA JR, W., FILHO, E., et al., “Compressão de Sinais de Eletrocardiograma Utilizando Técnicas de Codificação Distribuída”. In: *Proc. XXXVII Simp. Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*, 01 2019.
- [14] MELISSANT, C., YPMA, A., FRIETMAN, E., et al., “A method for detection of Alzheimer’s disease using ICA-enhanced EEG measurements”, *Artificial Intelligence in Medicine*, v. 33, n. 3, pp. 209–222, 2005.
- [15] JIANNAN, K., XIAOYA, H., OTHERS, “The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data”, *Computers in Biology and Medicine*, v. 120, pp. 103722, 2020.
- [16] SHEN, J., ZHANG, X., LI, J., et al., “Depression Detection from Electroencephalogram Signals Induced by Affective Auditory Stimuli”. In: *Proc. Int. Conf. on Affective Computing and Intelligent Interaction*, pp. 76–82, 2019.
- [17] ABDI, H., WILLIAMS, L., “Principal component analysis”, *Wiley interdisciplinary reviews: computational statistics*, v. 2, n. 4, pp. 433–459, 2010.

- [18] SILVA JR, W., ARAUJO, G., DA SILVA, E., et al., “Facial fiducial points detection using discriminative filtering on principal components”. In: *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, pp. 2681 – 2684, 10 2010.
- [19] JUNIOR, W. S. S., DA SILVA, E., GOLDENSTEIN, S., *Reconhecimento de Padrões utilizando Filtros de Correlação com Análise de Componentes Principais*, Tese de D.Sc., Universidade Federal do Rio de Janeiro (UFRJ), Agosto de 2010.
- [20] DE LATHAUWER, L., DE MOOR, B., VANDEWALLE, J., “An introduction to independent component analysis”, *Journal of Chemometrics: A Journal of the Chemometrics Society*, v. 14, n. 3, pp. 123–149, 2000.
- [21] LECUN, Y., BENGIO, Y., HINTON, G., “Deep learning”, *Nature*, v. 521, n. 7553, pp. 436–444, 2015.
- [22] BREIMAN, L., “Random forests”, *Machine learning*, v. 45, pp. 5–32, 2001.
- [23] IBRAHIM, S., DJEMAL, R., ALSUWAILEM, A., “Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis”, *Biocybernetics and Biomedical Engineering*, v. 38, n. 1, pp. 16–26, 2018.
- [24] MONTENEGRO, M. A., CENDES, F., GUERREIRO, M. M., et al., *EEG na prática clínica*. Thieme Revinter Publicações LTDA, 2018.
- [25] SANEI, S., CHAMBERS, J. A., *EEG signal processing*. John Wiley & Sons, 2013.
- [26] HU, L., ZHANG, Z., *EEG signal processing and feature extraction*. Springer, 2019.
- [27] BIASIUCCI, A., FRANCESCHIELLO, B., MURRAY, M. M., “Electroencephalography”, *Current Biology*, v. 29, n. 3, pp. R80–R85, 2019.
- [28] LYALL, K., CROEN, L., DANIELS, J., et al., “The changing epidemiology of autism spectrum disorders”, *Annual review of public health*, v. 38, pp. 81–102, 2017.

- [29] LORD, C., BRUGHA, T. S., CHARMAN, T., et al., “Autism spectrum disorder”, *Nature reviews Disease primers*, v. 6, n. 1, pp. 1–23, 2020.
- [30] WHITMAN, T. L., *O desenvolvimento do autismo*. M. Books Editora, 2019.
- [31] SIQUEIRA, C. C., FERREIRA, E. O., BITTENCOURT, R. G., et al., “221 O CÉREBRO AUTISTA: A BIOLOGIA DA MENTE E SUA IMPLICAÇÃO NO COMPROMETIMENTO SOCIAL”, *Revista Transformar*, v. 8, n. 8, pp. 221–237, 2016.
- [32] MANNING-COURTNEY, P., MURRAY, D., CURRANS, K., et al., “Autism spectrum disorders”, *Current problems in pediatric and adolescent health care*, v. 43, n. 1, pp. 2–11, 2013.
- [33] GRAF, W. D., MILLER, G., EPSTEIN, L. G., et al., “The autism “epidemic”: ethical, legal, and social issues in a developmental spectrum disorder”, *Neurology*, v. 88, n. 14, pp. 1371–1380, 2017.
- [34] KERCHES, D., *Autismo: ao longo da vida*. 1. Ed. São Paulo. Literare Books International, 2022.
- [35] FREITAS, D. F. C. L. D., GUIMARÃES, D. S., “Care relationships with people diagnosed with Autism Spectrum Disorder”, *Psicologia USP*, v. 32, 2021.
- [36] ROGERS, S. J., DAWSON, G., “Intervenção precoce em crianças com Autismo”, *Lisboa: Lidel*, 2010.
- [37] NILSSON, N. J., *Artificial intelligence: a new synthesis*. Morgan Kaufmann, 1998.
- [38] RUSSEL, S., NORVIG, P., OTHERS, *Artificial intelligence: a modern approach*. v. 256. Pearson Education Limited London, 2013.
- [39] HASTIE, T., TIBSHIRANI, R., FRIEDMAN, J. H., et al., *The elements of statistical learning: data mining, inference, and prediction*. v. 2. Springer, 2009.
- [40] GOODFELLOW, I., BENGIO, Y., COURVILLE, A., *Deep learning*. MIT press, 2016.

- [41] ESTEVA, A., ROBICQUET, A., RAMSUNDAR, B., et al., “A guide to deep learning in healthcare”, *Nature medicine*, v. 25, n. 1, pp. 24–29, 2019.
- [42] CHAR, D. S., SHAH, N. H., MAGNUS, D., “Implementing machine learning in health care—addressing ethical challenges”, *The New England journal of medicine*, v. 378, n. 11, pp. 981, 2018.
- [43] AGGARWAL, C. C., OTHERS, “Neural networks and deep learning”, *Springer*, v. 10, n. 978, pp. 3, 2018.
- [44] LECUN, Y., BOTTOU, L., BENGIO, Y., et al., “Gradient-based learning applied to document recognition”, *Proceedings of the IEEE*, v. 86, n. 11, pp. 2278–2324, 1998.
- [45] YANG, S., “Research on network behavior anomaly analysis based on bidirectional LSTM”. In: *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 798–802, 2019.
- [46] RASCHKA, S., MIRJALILI, V., “Python machine learning: Machine learning and deep learning with python”, *Scikit-Learn, and TensorFlow. Second edition ed*, v. 3, 2017.
- [47] JUNIOR, W., FILHO, E., MELO, W., “Compressão de Sinais Eletromiográficos Baseada em Técnicas Bidimensionais”. 01 2015.
- [48] BHASKARACHARY, C., NAJAFABADI, A. J., GODDE, B., “Machine Learning Supervised Classification Methodology for Autism Spectrum Disorder Based on Resting-State Electroencephalography (EEG) Signals”. In: *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–4, 2020.
- [49] CHILD MIND INSTITUTE BIOBANK, “Dataset for Autism Spectrum Disorder”, Disponível em: <https://childmind.org/blog/healthy-brain-network-biobank-bringing-big-data-to-child-mental-health>. Acesso em: 12 de outubro, 2022.

- [50] GEURTS, P., ERNST, D., WEHENKEL, L., “Extremely randomized trees”, *Machine learning*, v. 63, pp. 3–42, 2006.
- [51] LAURA, I., SANTI, S., “Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications”, 2017.
- [52] PISTORIUS, T., ALDRICH, C., AURET, L., et al., “Early detection of risk of autism spectrum disorder based on recurrence quantification analysis of electroencephalographic signals”. In: *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 198–201, 2013.
- [53] DE MELO, W. C., DE LIMA FILHO, E. B., DA SILVA JÚNIOR, W. S., “Electroencephalographic Signal Compression Based on Adaptive Segmentation and Video Encoders”, .
- [54] ACHARYA, U. R., SREE, S. V., CHATTOPADHYAY, S., et al., “Application of recurrence quantification analysis for the automated identification of epileptic EEG signals”, *International journal of neural systems*, v. 21, n. 03, pp. 199–211, 2011.
- [55] LOTTE, F., CONGEDO, M., LÉCUYER, A., et al., “A review of classification algorithms for EEG-based brain–computer interfaces”, *Journal of neural engineering*, v. 4, n. 2, pp. R1, 2007.
- [56] DELORME & MAKEIG, “EEGLAB Wiki”, Disponível em: <https://eeglab.org/>, 2013.
- [57] THE MATHWORKS, INC, “MATLAB for Artificial Intelligence”, Disponível em: <https://www.mathworks.com/>, 2022.
- [58] GOTHAM, K., RISI, S., PICKLES, A., et al., “The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity”, *Journal of autism and developmental disorders*, v. 37, pp. 613–627, 2007.
- [59] ANDRZEJAK, R. G., LEHNERTZ, K., MORMANN, F., et al., “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state”, *Physical Review E*, v. 64, n. 6, pp. 061907, 2001.

- [60] GOLDBERGER, A. L., AMARAL, L. A., GLASS, L., et al., “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals”, *circulation*, v. 101, n. 23, pp. e215–e220, 2000.
- [61] ALHADDAD, M. J., KAMEL, M. I., MALIBARY, H. M., et al., “Diagnosis autism by fisher linear discriminant analysis FLDA via EEG”, *International Journal of Bio-Science and Bio-Technology*, v. 4, n. 2, pp. 45–54, 2012.
- [62] MILNE, E., GOMEZ, R., GIANNADOU, A., et al., “Atypical EEG in autism spectrum disorder: Comparing a dimensional and a categorical approach.” *Journal of abnormal psychology*, v. 128, n. 5, pp. 442, 2019.
- [63] MILNE, E., GOMEZ, R., GIANNADOU, A., et al., “Atypical EEG in autism spectrum disorder: Comparing a dimensional and a categorical approach”, Disponível em: https://figshare.shef.ac.uk/articles/dataset/EEG_Data_for_Electrophysiological_signatures_of_brain_aging_in_autism_spectrum_disorder_/16840351. Acesso em: 22 de novembro, 2022.
- [64] CHANG, C.-Y., HSU, S.-H., PION-TONACHINI, L., et al., “Evaluation of artifact subspace reconstruction for automatic EEG artifact removal”. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1242–1245, 2018.
- [65] ALBAWI, S., MOHAMMED, T. A., AL-ZAWI, S., “Understanding of a convolutional neural network”. In: *2017 international conference on engineering and technology (ICET)*, pp. 1–6, 2017.
- [66] MASTROMICHALAKIS, S., “ALReLU: A different approach on Leaky ReLU activation function to improve Neural Networks Performance”, *arXiv preprint arXiv:2012.07564*, 2020.
- [67] MURRAY, N., PERRONNIN, F., “Generalized max pooling”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2473–2480, 2014.

- [68] GRAVES, A., GRAVES, A., “Long short-term memory”, *Supervised sequence labelling with recurrent neural networks*, pp. 37–45, 2012.
- [69] BUZUTI, L. F., THOMAZ, C. E., “RECONHECIMENTO DE FACES POR APRENDIZADO PROFUNDO: DROPOUT EM AUTOENCODERS”, .
- [70] TECHNOLOGY FOUNDATION STW, “Biosemi Active Two EEG System”, Disponível em: <https://www.biosemi.com/products.htm>, 2018.
- [71] BRAINARD, D. H., VISION, S., “The psychophysics toolbox”, *Spatial vision*, v. 10, n. 4, pp. 433–436, 1997.
- [72] DICKINSON, A., JESTE, S., MILNE, E., “Electrophysiological signatures of brain aging in autism spectrum disorder”, *Cortex*, v. 148, pp. 139–151, 2022.
- [73] NHS, “Health Research Authority”, Disponível em: <https://www.hra.nhs.uk/>, 2022.
- [74] INDI, “Autism Brain Imaging Data Exchange (ABIDE)”, Disponível em: https://fcon_1000.projects.nitrc.org/indi/abide/, 2017.
- [75] NIH, “National Database for Autism Research (NDAR)”, Disponível em: <https://www.re3data.org/repository/r3d100010717>, 2018.
- [76] TEAM, C., “Google Colaboratory”, Disponível em: <https://colab.research.google.com/>, 2018.
- [77] TEAM, P. C., “Python: A dynamic, open source programming language”, Disponível em: <https://www.python.org/>, 2019.
- [78] CHOLLET, F., OTHERS, “Keras”, Disponível em: <https://keras.io/>, 2015.
- [79] ABADI, M., OTHERS, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems”, Disponível em: <http://tensorflow.org/>, 2015.
- [80] HARRIS, C. R., MILLMAN, K. J., VAN DER WALT, S. J., et al., “Array programming with NumPy”, *Nature*, v. 585, n. 7825, pp. 357–362, 2020.

- [81] THE MATPLOTLIB DEVELOPMENT TEAM, “Matplotlib: Visualization with Python”, Disponível em: <https://matplotlib.org/>, 2017.
- [82] COURNAPEAU, D., “Scikit-learn: Machine Learning in Python”, Disponível em: <https://scikit-learn.org/>, 2007.
- [83] HOSSIN, M., SULAIMAN, M. N., “A review on evaluation metrics for data classification evaluations”, *International journal of data mining & knowledge management process*, v. 5, n. 2, pp. 1, 2015.
- [84] BADDELEY, A. J., “An error metric for binary images”, *Robust computer vision*, v. 5978, 1992.
- [85] KYNKÄÄNNIEMI, T., KARRAS, T., LAINE, S., et al., “Improved precision and recall metric for assessing generative models”, *Advances in Neural Information Processing Systems*, v. 32, 2019.
- [86] LAVIE, A., SAGAE, K., JAYARAMAN, S., “The significance of recall in automatic metrics for MT evaluation”. In: *Machine Translation: From Real Users to Research: 6th Conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC, USA, September 28-October 2, 2004. Proceedings 6*, pp. 134–143, 2004.
- [87] YACOUBY, R., AXMAN, D., “Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models”. In: *Proceedings of the first workshop on evaluation and comparison of NLP systems*, pp. 79–91, 2020.
- [88] WU, S., FLACH, P., “A scored AUC metric for classifier evaluation and selection”. In: *Second workshop on ROC analysis in ML, bonn, Germany, 2005*.
- [89] HAN, S., YUAN, B., LIU, W., “Rare class mining: progress and prospect”. In: *2009 Chinese Conference on Pattern Recognition*, pp. 1–5, 2009.
- [90] WANTO, A., WINDARTO, A. P., HARTAMA, D., et al., “Use of binary sigmoid function and linear identity in artificial neural networks for forecas-

ting population density”, *IJISTECH (International Journal of Information System and Technology)*, v. 1, n. 1, pp. 43–54, 2017.

- [91] KINGMA, D. P., BA, J., “Adam: A method for stochastic optimization”, *arXiv preprint arXiv:1412.6980*, 2014.
- [92] YANG, Q., ZHANG, Y., DAI, W., et al., *Transfer learning*. Cambridge University Press, 2020.
- [93] ZHANG, S., CHEN, D., TANG, Y., et al., “Children ASD evaluation through joint analysis of EEG and eye-tracking recordings with graph convolution network”, *Frontiers in Human Neuroscience*, v. 15, pp. 651349, 2021.
- [94] BANK, D., KOENIGSTEIN, N., GIRYES, R., “Autoencoders”, *arXiv preprint arXiv:2003.05991*, 2020.
- [95] NETO, J. L., FILHO, E. B. L., FURTADO, S. C., et al., “Classificação de Sinais de EEG para Diagnóstico do Transtorno do Espectro Autista utilizando CNN e RNN”, *XLI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*, 2023.
- [96] SYAFEEZA, A., KHALIL-HANI, M., LIEW, S. S., et al., “Convolutional neural networks with fused layers applied to face recognition”, *International Journal of Computational Intelligence and Applications*, v. 14, n. 03, pp. 1550014, 2015.

Apêndice A

Camadas de CNN

Camada Leaky ReLU

Após ser realizada a convolução é comum o uso de uma função de ativação para cada neurônio. Esta função assegura que o valor de ativação seja passado para a próxima camada a partir da sua entrada. Nos últimos anos, o uso da função de ativação *ReLU* substituiu as funções de ativação sigmóide e tangente hiperbólica no projeto de uma rede neural convolucional, pois a ReLU tem grandes vantagens sobre essas funções de ativação, tanto em termos de velocidade quanto de precisão. Uma das desvantagens dessa função é que, durante o processo de treinamento, alguns neurônios efetivamente morrem, especialmente sob circunstâncias na qual o valor empregado para a taxa de aprendizado é muito alto [66]. Deste modo, a função *leakyReLU* apresenta-se como uma alternativa para contornar esse problema, removendo o gradiente zero para os casos em que $x < 0$ por meio da adição de um pequeno componente linear a , conforme equação (A.1).

$$\begin{cases} f(x) = ax, x < 0 \\ f(x) = x, x \geq 0 \end{cases} \quad (\text{A.1})$$

Camada *Pooling*

A função da camada *pooling* é substituir determinado local de saída da rede por uma estatística resumida das saídas próximas. A camada de *pooling* permite diminuir a sensibilidade da rede com relação a pequenas alterações na imagem [67].

Durante o processo de *pooling* alguns neurônios são deixados de lado, o que torna a rede menos complexa, uma vez que o processo reduz sua dimensionalidade. Esta camada não possui pesos. O filtro é uma aplicação direta de uma operação, tal como o máximo (*maxpooling*) ou média, em uma sub-região da imagem.

Camada *Batch*

A mudança de covariância interna é um problema decorrente do treinamento de redes neurais convolucionais, ocorre quando os parâmetros da rede mudam no treinamento, juntamente com as ativações das camadas ocultas, o que resulta em uma convergência mais lenta de treinamento. Para reduzir esse efeito, utiliza-se a camada *batch* nos treinamentos. Esta camada é responsável por adicionar uma operação, pouco antes da função de ativação, em cada camada de convolução. Pesquisadores mostraram que o uso desta camada reduz o problema de desvanecimento do gradiente [96], ao ponto de permitir a utilização de funções de ativação de tangente hiperbólica ou sigmoide, melhorando assim a acurácia em diversos trabalhos.

Camada Inteiramente Conectada

Conhecida também como camada densa, a camada inteiramente conectada é idêntica as redes neurais da geração anterior, em que todos os neurônios de uma camada estão conectados a todos os neurônios de uma camada anterior, essa camada tem a função de vetorizar a camada anterior para conectá-la à camada final de classificação [43]. Em algumas linguagens, como o *Matlab*, a camada de classificação já faz automaticamente o processo de *flattening*, ou seja, converte uma matriz em um vetor para que com esse vetor seja realizada a classificação. Já outras linguagens, como o *Python*, necessitam que haja uma camada de *flatten* antes da camada de classificação.

Camada de Classificação

Esta camada está presente na maioria das arquiteturas que almejam a classificação. Ela tem por objetivo calcular a diferença entre a saída estimada e o valor desejado, para que assim os valores de perda possam ser minimizados. Em casos

em que há apenas duas classes, costuma-se utilizar a função de ativação *sigmoid*, e nos casos de duas ou mais classes, a função *softmax*, em razão da atribuição de probabilidades a cada uma das classes existentes no problema.

Apêndice B

Tabelas de arquiteturas dos sistemas de classificação

Tabela B.1: Arquiteturas dos sistemas de classificação TEA com CNN.

Experi- mento N°	CAMADAS							
	1°	2°	3°	4°	5°	6°	7°	8°
1	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
2	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
3	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
4	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
5	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
6	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
7	Conv2D = 32 LR = 0,1	Conv2D = 64 LR = 0,1	Conv2D = 128 LR = 0,1	Conv2D = 64 LR = 0,1	Dense = 128 FA = Linear	Dense = 2 FA = Sigmoid	-	-

	MP = (2,2)	MP = (2,2)	MP = (2,2)	MP = (2,2)	LR = 0,1			
8	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
9	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
10	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
11	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
12	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
13	Conv2D = 3 LR = 0,1	ResNet50 imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
14	Conv2D = 3 LR = 0,1	ResNet50 imagenet'	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
15	Conv2D = 3 LR = 0,1	MobileNet imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear	Dense = 2 FA = Sigmoid

16	Conv2D = 3 LR = 0,1	MobileNet imagenet'	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	LR = 0,1 Dense = 2 FA = Sigmoid	
17	Conv2D = 3 LR = 0,1	MobileNetV2 imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
18	Conv2D = 3 LR = 0,1	MobileNetV2 imagenet'	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-
19	Conv2D = 3 LR = 0,1	ResNet50 imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
20	Conv2D = 3 LR = 0,1	MobileNet imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
21	Conv2D = 3 LR = 0,1	MobileNetV2 imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
22	Conv2D = 3 LR = 0,1	ResNet50 imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
23	Conv2D = 3 LR = 0,1	MobileNet imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear	Dense = 2 FA = Sigmoid

							LR = 0,1	
24	Conv2D = 3 LR = 0,1	MobileNetV2 imagenet'	Conv2D = 1024 LR = 0,1	Conv2D = 512 LR = 0,1	Conv2D = 256 LR = 0,1	Conv2D = 128 LR = 0,1	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid
25	Autoenc Dense = 3200 FA: Linear	Autoenc Dense = 640 FA: Linear	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
26	Autoenc Dense = 6400 FA: Linear	Autoenc Dense = 640 FA: Linear	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
27	Autoenc Dense = 3200 FA: Linear	Autoenc Dense = 640 FA: Linear	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
28	Autoenc Dense = 6400 FA: Linear	Autoenc Dense = 640 FA: Linear	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
29	Autoenc Dense = 3200 FA: Linear	Autoenc Dense = 640 FA: Linear	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
30	Autoenc Dense = 6400 FA: Linear	Autoenc Dense = 640 FA: Linear	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-

1) Em todos os experimentos, na camada de convolução utiliza-se como padrão kernel = (3,3) e padding = "same".
LR: LeakyReLU; MP: MaxPooling2D; FA: Função de Ativação

Tabela B.2: Arquiteturas dos sistemas de classificação TEA com CNN e técnica de ODP.

Experi- mento N°	CAMADAS							
	1°	2°	3°	4°	5°	6°	7°	8°
1	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
2	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
3	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Softmax	-	-
4	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Softmax	-	-
5	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Softmax	-	-
6	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
7	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
8	Conv2D = 32 LR = 0,1	Conv2D = 64 LR = 0,1	Conv2D = 128 LR = 0,1	Conv2D = 64 LR = 0,1	Dense = 128 FA = Linear	Dense = 2 FA = Sigmoid	-	-

	MP = (2,2)	MP = (2,2)	MP = (2,2)	MP = (2,2)	LR = 0,1			
9	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
10	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-

1) Em todos os experimentos, na camada de convolução utiliza-se como padrão $kernel = (3,3)$ e $padding = "same"$.
 LR: LeakyReLU; MP: MaxPooling2D; FA: Função de Ativação

Tabela B.3: Arquiteturas dos sistemas de classificação TEA com CNN sem Pré-processamento.

Experi- mento N°	CAMADAS							
	1°	2°	3°	4°	5°	6°	7°	8°
1	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
2	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
3	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
4	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-
5	Conv2D = 32 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Conv2D = 128 LR = 0,1 MP = (2,2)	Conv2D = 64 LR = 0,1 MP = (2,2)	Dense = 128 FA = Linear LR = 0,1	Dense = 2 FA = Sigmoid	-	-

1) Em todos os experimentos, na camada de convolução utiliza-se como padrão kernel = (3,3) e padding = "same".
LR: LeakyReLU; MP: MaxPooling2D; FA: Função de Ativação

Tabela B.4: Arquiteturas dos sistemas de classificação TEA com RNN.

Experi- mento	CAMADAS								
	1°	2°	3°	4°	5°	6°	7°	8°	9°
1	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
2	LSTM = 5	Dropout = 0,1	LSTM = 5	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
3	LSTM = 5	Dropout = 0,1	LSTM = 5	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
4	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
5	LSTM = 8	Dropout = 0,1	LSTM = 8	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
6	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
7	LSTM = 5	Dropout = 0,1	LSTM = 5	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
8	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
9	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
10	LSTM = 256	Dropout = 0,2	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid
11	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
12	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-

13	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
14	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Sigmoid	-	-	-	-
15	LSTM = 10	Dropout = 0,1	LSTM = 10	Dropout = 0,1	Dense = 2 FA = Sigmoid	-	-	-	-
16	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
17	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
18	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-
19	LSTM = 256	Dropout = 0,2	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid
20	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid	-	-

FA: Função de Ativação

Tabela B.5: Arquiteturas dos sistemas de classificação TEA com RNN e técnica de ODP.

Experi- mento N°	CAMADAS								
	1°	2°	3°	4°	5°	6°	7°	8°	9°
1	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid		
2	LSTM = 128	Dropout = 0,2	LSTM = 128	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid		
3	LSTM = 256	Dropout = 0,2	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid
4	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Sigmoid				
5	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Softmax		
6	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Softmax				
7	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Softmax		
8	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid		
9	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid		
10	LSTM = 256	Dropout = 0,2	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid

FA: Função de Ativação

Tabela B.6: Arquiteturas dos sistemas de classificação TEA com RNN sem pré-processamento

Experi- mento N°	CAMADAS								
	1°	2°	3°	4°	5°	6°	7°	8°	9°
1	LSTM = 15	Dropout = 0,1	LSTM = 15	Dropout = 0,7	Dense = 2 FA = Softmax				
2	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Softmax		
3	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid		
4	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid		
5	LSTM = 256	Dropout = 0,2	LSTM = 128	Dropout = 0,2	LSTM = 64	Dropout = 0,2	LSTM = 32	Dropout = 0,2	Dense = 2 FA = Sigmoid

FA: Função de Ativação

Apêndice C

Artigos Publicados

Neste apêndice, apresentamos os artigos desenvolvidos nesta Dissertação.

C.1 Artigos Diretamente Relacionados ao Tema

1. NETO, J. L., FILHO, E. B. L., FURTADO, S. C., et al., "Classificação de Sinais de EEG para Diagnóstico do Transtorno do Espectro Autista utilizando CNN e RNN". In: *Anais do XLI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, Outubro de 2023, São José dos Campos - SP, SBrT 2023.