



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
DE COMPUTAÇÃO



Solução de Saúde Digital para a Comunicação Alternativa de Pessoas com Esclerose Lateral Amiotrófica

Felipe Ricardo dos Santos Fernandes

Orientador: Prof. Dr. Ricardo Alexsandro de Medeiros Valentim

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação da UFRN (área de concentração: Engenharia de Computação) como parte dos requisitos para obtenção do título de Doutor em Ciências.

Natal, RN, Agosto de 2023

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Central Zila Mamede

Fernandes, Felipe Ricardo dos Santos.

Solução de saúde digital para a comunicação alternativa de pessoas com esclerose lateral amiotrófica / Felipe Ricardo dos Santos Fernandes. – 2023.

124 f. : il.

Orientação: Prof. Dr. Ricardo Alexandro de Medeiros Valentim.

Tese (doutorado) – Universidade Federal do Rio Grande do Norte, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica e de Computação. Natal, RN, 2023.

1. Doença Neurodegenerativa – Tese. 2. Visão Computacional – Tese. 3. Machine Learning – Tese. 4. Processamento de Imagem – Tese. 5. Doença do Neurônio Motor – Tese. I. Valentim, Ricardo Alexandro de Medeiros. II. Título.

RN/UF/BCZM

CDU 616.8-003.98:004

Elaborado por Fernanda de Medeiros Ferreira Aquino - CRB-15/301

*À Maria das Dores dos Santos
Fernandes (in memoriam), minha
mãe.*

Agradecimentos

Início expressando minha gratidão à minha mãe, Maria das Dores dos Santos Fernandes (*in memoriam*). Com todo amor do mundo, ela cuidou de seus dois filhos e não mediu esforços para desobstruir obstáculos para que pudéssemos caminhar e conquistar nossos objetivos. Sou igualmente grato a meu pai, Sivanildo Fernandes, pelo suporte inestimável e pela energia positiva sempre presente. Ao meu irmão, Bruno Fernandes, minha imensa gratidão pelo companheirismo constante. E com especial carinho, agradeço aos meus avós, Dona Rita e Seu Antônio.

À minha querida esposa, Ingridy Barbalho, minha gratidão é tão profunda quanto o amor que tenho por você. Sua paciência, companhia, cuidado e amor foram os pilares que me sustentaram durante este percurso. Sua existência foi e é elementar em minha vida, especialmente para me manter bem, orientado e feliz. Juntos, superamos muitos desafios ao longo deste percurso acadêmico. Com todo o meu coração, obrigado!

Ao meu orientador, Professor Ricardo Valentim, sou imensamente grato pelos conhecimentos generosamente compartilhados e pelas oportunidades proporcionadas. Me sinto honrado primeiramente por ter sido acolhido como seu aluno/orientando, e ainda mais por ter tido a chance de contribuir para a sociedade através do Laboratório de Inovação Tecnológica em Saúde (LAIS) da Universidade Federal do Rio Grande do Norte (UFRN). Ser parte de sua equipe de pesquisadores no LAIS/UFRN é um privilégio e motivo de orgulho.

Aos meus estimados colegas pesquisadores do Laboratório de Inovação Tecnológica em Saúde (LAIS), minha sincera gratidão por todos os momentos enriquecedores que compartilhamos e pelo vasto aprendizado que me proporcionaram.

Resumo

A Esclerose Lateral Amiotrófica (ELA) é uma doença neurodegenerativa que compromete o sistema motor e as habilidades funcionais da pessoa de forma irreversível, causando, inclusive, a perda progressiva da capacidade de comunicação e da autonomia. Recursos tecnológicos baseados em Comunicação Aumentativa e Alternativa, Visão Computacional e Machine Learning são imprescindíveis para o desenvolvimento de soluções de saúde digital para a viabilização do processo comunicativo e da autonomia que, consequentemente, promovem melhorias na qualidade de vida e sobrevida da pessoa com ELA. Focado em uma abordagem de Interação Humano-Computador (IHC) baseada em imagens dos olhos provenientes de uma câmera simples e livre do corpo, este trabalho apresenta um recurso de tecnologia assistiva para a Comunicação Aumentativa e Alternativa de pessoas com ELA. A abordagem proposta neste trabalho consiste em um modelo algorítmico capaz de reconhecer o estado do olho (aberto ou fechado) em tempo real e de interoperar com o Autonomus, uma solução de saúde digital concebida pelo Laboratório de Inovação Tecnológica em Saúde da Universidade Federal do Rio Grande do Norte (LAIS/UFRN) para a comunicação de pessoas com ELA. O modelo é constituído por quatro processos metodológicos: (i) aquisição das imagens; (ii) detecção da Face; (iii) detecção dos olhos; e (iv) classificação do estado do olho, sendo esta a principal etapa para a IHC. Um estudo algorítmico com um grupo de controle foi conduzido para avaliar a performance geral do modelo e a capacidade de classificação da Rede Neural Convolutacional (CNN). Os resultados relacionados ao modelo proposto para a tarefa de classificação do estado do olho em tempo real são promissores e alcançam valores significativos de acurácia e f1-score, acima de 92%. Os resultados também apontam para a viabilidade do desenvolvimento de recursos de tecnologias assistivas de baixo custo que garantem o acesso universal, promoção de saúde, bem estar e redução das desigualdades, aspectos que vão além das melhorias no processo comunicativo de pessoas com ELA. Portanto, o objeto de estudo deste trabalho também possibilita e promove o exercício dos direitos, da cidadania, das liberdades fundamentais e do cuidado assistencial à saúde da pessoa com ELA.

Palavras-chave: Visão Computacional, Inteligência Artificial, Machine Learning, Processamento de Imagem, Doença do Neurônio Motor, Doença Neurodegenerativa, Saúde Pública

Abstract

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disease that irreversibly impairs an individual's motor system and functional abilities, even causing the progressive loss of communication skills and autonomy. Technological resources based on Augmentative and Alternative Communication, Computer Vision, and Machine Learning are essential for developing digital health solutions to enable the communicative process and autonomy that, consequently, promote improvements in the quality of life and survival of people with ALS. Focused on a Human-Computer Interaction (HCI) approach based on images of the eyes from a simple camera not mounted on the body, this work presents an assistive technology resource for Augmentative and Alternative Communication for people with ALS. The approach proposed in this work consists of an algorithmic model capable of recognizing the state of the eye (open or closed) in real-time and interoperating with Autonomus, a digital health solution designed by the Laboratory of Technological Innovation in Health at the Federal University of Rio Grande do Norte (LAIS/UFRN) for the communication of people with ALS. The model consists of four methodological processes: (i) image acquisition; (ii) Face detection; (iii) eye detection; and (iv) classifying the state of the eye, which is the foremost step for Human-Computer Interaction. An algorithmic study with a control group was conducted to evaluate the model's overall performance and the Convolutional Neural Network (CNN) classification ability. The results related to the proposed model for classifying the state of the eye in real-time are promising and reach significant values of accuracy and f1-score above 92%. The results also point to the viability of developing low-cost assistive technology resources that guarantee universal access, health promotion, well-being, and reduced inequalities, which go beyond improvements in the communicative process of people with ALS. Therefore, the object of study of this work is also to enable and promotes the exercise of rights, citizenship, fundamental freedoms, and health care for people with ALS.

Keywords: Computer Vision, Artificial intelligence, Machine Learning, Image Processing, Motor Neurone Disease, Neurodegenerative Diseases, Public health.

Sumário

Sumário	i
Lista de Figuras	iii
Lista de Tabelas	v
Lista de Algoritmos	vi
Lista de Símbolos e Abreviaturas	vi
1 Introdução	1
1.1 Problemática	3
1.2 Hipótese	6
1.3 Objetivos	6
1.3.1 Objetivo Geral	6
1.3.2 Objetivos Específicos	6
1.4 Contribuições	7
1.4.1 Produções Relacionadas à Tese e ao projeto revELA	7
1.4.2 Produções Relacionadas ao Curso Acadêmico	10
1.5 Organização do Trabalho	13
2 Fundamentos Teóricos	15
2.1 Aprendizado de Máquina	15
2.2 Visão Computacional	16
2.3 Projeto revELA	18
3 Sinais Biomédicos e Aprendizado de Máquina para ELA	21
3.1 Protocolo de Busca	21
3.2 Resultados	24
3.2.1 Estudos para o Diagnóstico	28
3.2.2 Estudos para a Comunicação	33
3.2.3 Estudos para a Predição de Sobrevida	35
3.3 Discussões	35
3.4 Conclusões	38

4	Comunicação Alternativa para ELA	40
4.1	Protocolo de Busca	40
4.2	Resultados	43
4.2.1	Questão de Pesquisa 01	46
4.2.2	Questão de Pesquisa 02	47
4.2.3	Questão de Pesquisa 03	49
4.2.4	Questão de Pesquisa 04	50
4.2.5	Questão de Pesquisa 05	51
4.3	Discussões	52
4.4	Considerações	54
5	Eye-Blink em Tempo Real para ELA	56
5.1	Desenho do Estudo	56
5.1.1	Definições	57
5.1.2	Compensação de Rotação Sequencial	59
5.1.3	Detecção de Face	60
5.1.4	Alinhamento Facial	61
5.1.5	<i>Patches</i> Oculares e Extração de Coordenadas	63
5.1.6	Classificação do Estado do Olho	65
5.1.7	Aumentação de Dados	66
5.1.8	Filtragem <i>Sliding-window</i>	67
5.1.9	Detecção do <i>Eye-blink</i>	67
5.1.10	Conjuntos de Dados Propostos	70
5.2	Experimentos e Resultados	71
5.2.1	Experimento de Classificação do Estado do Olho	73
5.2.2	Experimentos em Vídeos	75
5.3	Considerações Finais	83
6	Ecossistema Autonomus	85
6.1	Arquitetura do Ecossistema Autonomus	85
6.1.1	Módulo de Interação Humano-Computador	86
6.1.2	Módulo de Interface do Usuário	86
6.2	Resultados	91
7	Discussões	95
8	Considerações Finais	99
8.1	Limitações	99
8.2	Perspectivas	100
	Referências bibliográficas	101

Lista de Figuras

1.1	Cenário brasileiro sobre a população com deficiência (ou dificuldades funcionais).	5
2.1	Visão geral hierárquica dos tipos e algoritmos clássicos de aprendizado de máquina.	16
2.2	Visão geral do fluxo de trabalho (<i>pipeline</i>) básico no campo de visão computacional.	17
2.3	Ilustração dos procedimentos de pré-processamento.	18
2.4	Dimensões do projeto revELA.	19
3.1	Passos metodológicos.	22
3.2	Resultado dos procedimentos de busca e seleção de estudos primários para esta revisão sistemática.	25
3.3	Sumarização dos sinais utilizados e seus objetivos.	27
3.4	Número de indivíduos utilizados nos estudos.	28
3.5	Quantitativo de trabalhos por tipo de sinais biomédicos abordados na classe diagnóstico.	29
3.6	Quantitativo dos melhores modelos algorítmicos e os respectivos sinais biomédicos.	36
3.7	Pipeline genérico: esquema generalizado para resolver problemas de classificação.	36
4.1	Resultado do processo de busca e triagem dos estudos primários para esta revisão sistemática.	44
4.2	Estratégias para Interação Humano-Computador baseadas em imagens dos olhos.	46
4.3	Modelo genérico de fluxo de trabalho (<i>pipeline</i>).	47
4.4	Número de indivíduos submetidos aos testes nos estudos primários.	52
5.1	Fluxograma do Sistema.	57
5.2	Movimentos de rotação da cabeça ao redor dos eixos x , y e z	59
5.3	Compensador de Rotação.	59
5.4	Marcos faciais usados no conjunto de dados IBUG.	62
5.5	Performance do avaliador ocular em diferentes situações.. . . .	63
5.6	Arquitetura para as CNNs utilizadas neste trabalho.	64
5.7	Fluxo Geral do Sistema.	68

5.8	Pontuações da tarefa de classificação do estado do olho em função da variação do limiar.	74
5.9	Análise do alcance do piscar de olhos.	77
5.10	Análises de filtro.	81
6.1	Visão geral da Arquitetura do Autonomus.	86
6.2	<i>Persona</i> Antônio Pereira (pessoa com ELA).	88
6.3	<i>Persona</i> Maria Pereira (acompanhante/cuidadora de pessoa com ELA). . .	89
6.4	<i>Persona</i> João Paulo (pessoa com ELA).	89
6.5	Protótipo de alta fidelidade do Autonomus.	90
6.6	Protótipo de alta fidelidade do Autonomus com o módulo de automação residencial (Autonomus Casa).	91
6.7	Interação de um usuário saudável com a Plataforma Autonomus <i>Desktop</i> . . .	92
6.8	Interação de um usuário saudável com a Plataforma Autonomus <i>Mobile</i> . . .	93
6.9	Conjunto de dados do grupo de controle (pessoas saudáveis) para a tarefa de rastreamento da pupila.	94
6.10	Conjunto de dados do grupo de pessoas com ELA para a tarefa de rastreamento da pupila.	94
7.1	Visão geral dos Objetivos de Desenvolvimento Sustentável aos quais o Ecossistema Autonomus alinha-se.	96

Lista de Tabelas

1.1	População com deficiência (ou dificuldades funcionais) no Brasil em 2022.	4
3.1	Questões de Pesquisa.	22
3.2	Critérios de Inclusão.	23
3.3	Critérios de Exclusão.	23
3.4	Critérios de Avaliação de Qualidade.	24
3.5	Conjunto de artigos selecionados e suas principais características.	26
4.1	Questões de Pesquisa.	41
4.2	Critérios de Inclusão e Exclusão	42
4.3	Critérios de Avaliação de Qualidade	43
4.4	Resumo das principais características dos artigos incluídos nesta revisão sistemática da literatura.	45
5.1	Pontuações AUC e melhor acurácia (BA) para modelos ECC e ECS em combinação com o modelo EE para classificação do estado do olho.	75
5.2	Resultados para ECS e ECC combinados com modelos auxiliares para a tarefa de detecção de piscar de olhos em todos os conjuntos de dados.	78
5.3	Resultados para o ECS e modelos auxiliares na tarefa de detecção voluntária de piscar de olhos.	79
5.4	Valores de média e desvio padrão (Std.) das latências para cada etapa e para todo o modelo. Algumas etapas são apresentadas juntamente com suas siglas.	82
5.5	Comparação dos melhores <i>F1-Scores</i> entre diferentes estudos sob diferentes configurações e critérios experimentais.	83
6.1	Perfil dos especialistas envolvidos no processo de desenvolvimento da interface do usuário.	87

Lista de Símbolos e Abreviaturas

ABD	Conjunto de Dados <i>Autonomus Blink Detection</i>
Acc	Acurácia
BCI	<i>Brain Computer Interface</i>
CAA	Comunicação Aumentativa e Alternativa
CAQ	Critérios de Avaliação de Qualidade
CE	Critérios de Exclusão
CI	Critérios de Inclusão
CNN	<i>Convolutional Neural Network</i>
DWT	Transformada Discreta de Wavelet
ECC	<i>Eye Classifier Convolutional Neural Network</i>
ECS	<i>Eye Classifier Support Vector Machine</i>
EE	<i>Eye Evaluator</i>
EEG	Eletroencefalografia
ELA	Esclerose Lateral Amiotrófica
EMG	Eletromiografia
FHMM	Modelo Fatorial Oculto de Markov
GC	Grupo de Controle
GR	Ritmo da Marcha
HMM	Modelo Oculto de Markov
IHC	Interação Humano-Computador
IR	Infravermelho
LAIS	Laboratório de Inovação Tecnológica em Saúde

LDA	Análise Discriminante Linear
MLP	<i>Multilayer Perceptron</i>
MRI	Imagem de Ressonância Magnética
MS	Ministério da Saúde
OD	Outras Doenças
ODS	Objetivos de Desenvolvimento Sustentável
ONU	Organizações das Nações Unidas
PCA	Análise de Componentes Principais
Pop	População
QP	Questões de Pesquisa
RC	<i>Rotation Compensator</i>
RSL	Revisão Sistemática da Literatura
Sen	Sensibilidade
Spe	Especificidade
ST	Transformada de Stockwell
STR	<i>Strings</i> de Busca
SVM	<i>Support Vector Machine</i>
UFRN	Universidade Federal do Rio Grande do Norte

Capítulo 1

Introdução

A Esclerose Lateral Amiotrófica (ELA) é uma doença neurodegenerativa progressiva e irreversível que afeta os neurônios motores - tanto superiores quanto inferiores - do indivíduo (Barbalho et al. 2022, Fernandes et al. 2021). Essas características patológicas da doença resultam em uma perda gradual dos neurônios motores e, conseqüentemente, da funcionalidade, comprometendo assim os movimentos voluntários, a função respiratória e o processo comunicativo da pessoa com ELA (Fernandes et al. 2023, Goutman et al. 2022, Saadeh et al. 2017, Hardiman et al. 2017, van Es et al. 2017). Após a descrição formal da ELA há mais de 150 anos, por Jean-Martin Charcot, muitas pesquisas foram desenvolvidas e avanços são evidenciados em diversos aspectos relacionados ao diagnóstico (novos critérios, como o *Gold Coast criteria*), ao monitoramento ou telessaúde (Fidelix et al. 2023) e a tratamentos terapêuticos (Goetz 2000, Corcia et al. 2023).

Apesar dos avanços representativos no contexto da ELA, ainda não há cura para a doença ou formas de prevenção. Entretanto, para aliviar a carga da doença e os sintomas relacionados à saúde física, mental e social nessas pessoas, o acesso a um ecossistema que engloba a assistência/intervenção multiprofissional e tecnologias assistivas é imprescindível, e tem se mostrado essencial para a manutenção do bem-estar, da capacidade de comunicação, da autonomia e da interação ou participação social à medida que a doença progride (Zizzi et al. 2023, de Lima Medeiros et al. 2022, Papaiz et al. 2022, Silva et al. 2021, Silva et al. 2013). Aspectos como estes, que constituem o ecossistema, são elementares durante a trajetória de vida da pessoa com ELA e promovem o cuidado à saúde, o exercício dos direitos e das liberdades fundamentais, a inclusão, a melhora da qualidade de vida e, conseqüentemente, o aumento da sobrevivência (Mercadante e Al-Husinat 2023, Howard e Burgess 2021, Gillespie et al. 2021, Rosa Silva et al. 2020, Linse, Rüger, Joos, Schmitz-Peiffer, Storch e Hermann 2018, Linse, Aust, Joos e Hermann 2018, Londral et al. 2015, Rooney et al. 2015).

Um dos pontos de fundamental importância ao longo do percurso da ELA é a tecnologia assistiva, que surge com o objetivo de compensar as perdas relacionadas às habilidades funcionais e motoras causadas pela doença (de Lima Medeiros et al. 2022, Silva et al. 2021). O termo “tecnologia assistiva” abrange um subconjunto de recursos e serviços assistivos de saúde, aos quais, no Brasil, os cidadãos têm acesso garantido por lei, com base na Lei Nº 13.146, de 6 de Julho de 2015 (Lei Brasileira de Inclusão da Pessoa com Deficiência ou Estatuto da Pessoa com Deficiência) (Brasil 2015), e no Decreto

Nº 10.645, de 11 de Março de 2021 (que dispõe sobre o Plano Nacional de Tecnologia Assistiva) (Brasil 2021a). No campo das tecnologias assistivas e especialmente para promover a funcionalidade relacionada ao processo comunicativo, criou-se uma importante área, chamada Comunicação Aumentativa e Alternativa (CAA), dedicada à pesquisa, ao desenvolvimento e a implementação de recursos e estratégias que permitem acontecer a comunicação expressiva e a interação (Gutmann e Koul 2023, Fernandes et al. 2023).

Nesse contexto da CAA, existem diferentes mecanismos e paradigmas para a Interação Humano-Computador (IHC) e o controle de interfaces. Os sinais bioelétricos, por exemplo, são mecanismos amplamente investigados e utilizados em áreas como neurociência e reabilitação. A chamada Interface Cérebro-Computador (BCI, do Inglês *Brain-Computer Interface*), que utiliza os sinais cerebrais e dispositivos de eletroencefalografia para a IHC, é o mecanismo mais adequado para os casos onde as pessoas com ELA se encontram no estado denominado *locked-in*, isto é, sem movimentos e sem a capacidade de controlar voluntariamente os olhos (Ramakrishnan et al. 2020, Miao et al. 2020, Sorbello et al. 2018, Liu et al. 2017, Vansteensel et al. 2016, Mainsah et al. 2015, McCane et al. 2014).

Jaramillo-Yáñez et al. (2020), em uma revisão sistemática, mostram estudos que investigam os sinais de eletromiografia para a IHC por meio de contrações musculares e o reconhecimento de gestos. Outros estudos também mostraram que, utilizando-se recursos baseados em eletrooculograma, é possível controlar interfaces por meio da captura de movimentos oculares predefinidos (cima, baixo, esquerda e direita) ou piscar de olhos (Fernandes et al. 2023, Tonin et al. 2020, Zhang et al. 2019, Chang et al. 2017, Larson et al. 2017, Lingegowda et al. 2017, Pinheiro et al. 2011).

Mesmo com o desenvolvimento de recursos e tecnologias na área de Comunicação Aumentativa e Alternativa (CAA) utilizando os sinais bioelétricos, introduzir dispositivos dessa natureza em ambiente domiciliar para o uso por pessoas com deficiências motoras graves, como é o caso das pessoas com ELA, é um grande desafio. Essas limitações estão relacionadas à usabilidade doméstica, a habilidade de manusear o instrumento/artefato e ao tempo de seleção de caracteres ou itens em uma interface de CAA. A necessidade de usar eletrodos que precisam estar conectados ao usuário também é um fator limitador que, além disso, causa fadiga, desconforto e desmotivação para a utilização do recurso (Chaudhary et al. 2022, Fernandes et al. 2021, Singh e Singh 2019a, Chaudhary et al. 2016).

Outras abordagens fundamentadas em conceitos e métodos computacionais têm recebido crescente destaque na Interação Humano-Computador (IHC), especialmente nos campos de visão computacional e aprendizado de máquina. Estudos recentes indicam resultados promissores no uso de técnicas algorítmicas para o reconhecimento de padrões de piscadas dos olhos ou movimentos da pupila de pessoas com deficiência motora grave ou com patologias semelhantes a da ELA (Kayadibi et al. 2022, Fathi e Abdali-Mohammadi 2015, Mu et al. 2022, Hwang et al. 2020, Bignaut 2017, Chareonsuk et al. 2016, Liu et al. 2012, Liu et al. 2019). Operacionalmente, essas técnicas servem para processar, extrair características e estabelecer os padrões predefinidos *on time* (no momento em que os eventos ocorrem) a partir de um conjunto de imagens dos olhos, obtidas por meio de uma ou múltiplas câmeras. O objetivo principal desse conjunto sequencial de tarefas é a

criação de um modelo computacional capaz de realizar o reconhecimento de padrões para a IHC e, conseqüentemente, permitir o controle de interfaces e ambientes automatizados.

Para trabalhar com métodos computacionais baseados em imagens, Fathi e Abdali-Mohammadi (2015) apresentaram duas categorias: com e sem infravermelho. Modelos computacionais baseados em imagens com infravermelho são mais eficazes, entretanto, a exposição prolongada ao infravermelho pode causar desconforto e danos à saúde ocular dos usuários. Além disso, utilizar infravermelho exige *hardware* específico que, geralmente, é acoplado em um protótipo montado na cabeça do usuário. Os métodos com câmeras sem infravermelho são mais simples, entretanto, os aspectos de detecção dos movimentos oculares ou do piscar dos olhos são mais complexos, portanto, podem comprometer a precisão do sistema de IHC em relação ao objeto alvo de seleção na interface (Hwang et al. 2020, Holmqvist et al. 2022). É nesta perspectiva que surgem os desafios relacionados ao desenvolvimento de estratégias eficientes de IHC baseada em imagens dos olhos para a Comunicação Aumentativa e Alternativa em ambiente domiciliar.

Nesta tese, é apresentado um recurso de tecnologia assistiva de fundamental importância para o ecossistema da Esclerose Lateral Amiotrófica (ELA) no Brasil que pode ser incorporado e distribuído gratuitamente pelo Sistema Único de Saúde. Este recurso é uma solução de saúde digital acessível, baseada em inteligência artificial, capaz de promover a funcionalidade relacionada à comunicação e proporcionar autonomia, inclusão e participação social, interação com o ambiente e, principalmente, qualidade de vida às pessoas que vivem com a ELA. Os resultados desta pesquisa podem ajudar as pessoas que vivem com a ELA no Brasil ou em outros países de língua portuguesa, especialmente por ser uma tecnologia assistiva gratuita e de baixo custo, além de ser uma ferramenta de inclusão social. O objeto de estudo desta tese envolve uma arquitetura tecnológica de Interação Humano-Computador (IHC) que usa imagens dos olhos e abrange conceitos e tecnologias baseadas em métodos de visão computacional e aprendizado de máquina.

1.1 Problemática

O cuidado da pessoa com Esclerose Lateral Amiotrófica (ELA) abrange e necessita de vários aspectos, desde da perspectiva humana quanto tecnológica. Além disso, envolve também fatores associados às políticas públicas de saúde e questões socioeconômicas. No mundo, a incidência da ELA é de 1,68 casos por 100.000 habitantes (Erazo et al. 2022), enquanto a prevalência varia entre 4,1 a 8,4 casos por 100.000 habitantes (Barbalho et al. 2021, Genge e Chio 2023). No Brasil, não há registros de incidência e prevalência populacional baseados em dados coesos da ELA, pois, além da descentralização e não compartilhamento dos dados, não há a exigência de notificar novos casos, por isso há no país uma vazios de dados que prejudica a condução da política pública de saúde (Barbalho et al. 2023).

O estudo mais recente sobre a ELA no Brasil analisou certidões de óbito durante o período de 2004 a 2013 e apontou incidência média de 0,46 casos por 100.000 habitantes, com tendência de aumento de casos (Moura et al. 2016). Em relação à prevalência, o dado é de 1998 e os autores Dietrich-Neto et al. (2000), por meio de correspondências enviadas aos centros de referência do Brasil, estimaram uma taxa que varia de 0,9 a 1,5

por 100.000 habitantes. O Estado do Rio Grande do Norte, por sua vez, possui uma taxa estimada de incidência abaixo da nacional, com 0,37 casos por 100.000 habitantes (Oliveira et al. 2023).

Com a provável subnotificação de casos da ELA no Brasil, é de fundamental importância analisar os dados das pessoas com deficiência. Isso ajudará a delimitar as fronteiras desta pesquisa e a compreender a importância das tecnologias assistivas, bem como da necessidade de um ecossistema alinhado aos Objetivos de Desenvolvimento Sustentável, para garantir que ninguém seja deixado para trás, conforme a Agenda 2030 das Organizações das Nações Unidas (ONU) (United Nations 2015).

A Tabela 1.1, ilustrada pela Figura 1.1, mostra que o Brasil tem aproximadamente 18,6 milhões de pessoas com dificuldades funcionais, incluindo aquelas com ELA. Nota-se ainda que cerca de 64% desse grupo populacional vive sem renda ou com rendimento mensal per capita de até 1 salário mínimo. Essa condição financeira limitada certamente compromete a qualidade de vida e restringe o acesso a recursos assistenciais, como as tecnologias assistivas.

Tabela 1.1: População com deficiência (ou dificuldades funcionais) no Brasil em 2022.

Região	População (Pop)	Pop-deficiência	%^a	Taxa^b	Rendimento^c
Norte	17.349.619	1.532.000	8,4%	8.830	1.176.000
Nordeste	54.644.582	5.787.000	10,3%	10.590	4.478.000
Sudeste	84.847.187	7.217.000	8,2%	8.506	3.986.000
Sul	29.933.315	2.622.000	8,8%	8.759	1.361.000
Centro-Oeste	16.287.809	1.397.000	8,6%	8.577	769.000
Brasil	203.062.512	18.555.000	8,9%	9.138	11.771.000

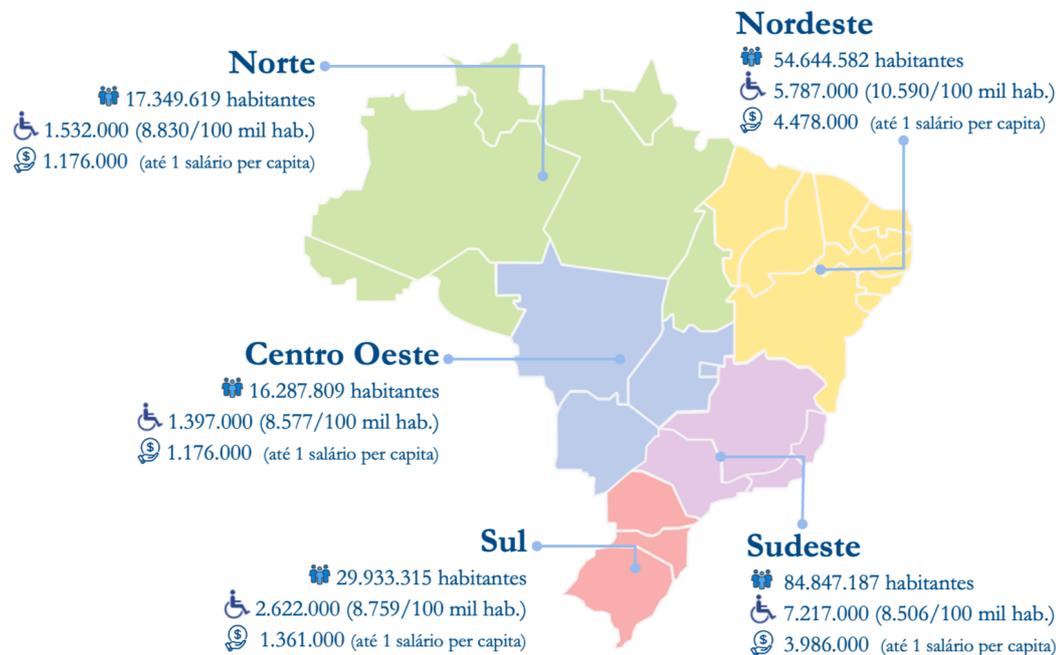
^a Porcentagem ajustada ao grupo etário de pessoas de 2 anos ou mais de idade do país/região que tinham algum tipo de deficiência.

^b $Taxa = \left(\frac{\text{população com deficiência}}{\text{população 2022}} \right) 100.000 \text{ habitantes.}$

^c Número de pessoas com deficiência sem rendimento ou com rendimento mensal domiciliar per capita de até 1 salário mínimo.

Fonte: Adaptado de Brasil (2023a) e Brasil (2023b).

Figura 1.1: Cenário brasileiro sobre a população com deficiência (ou dificuldades funcionais).



Nesse contexto, ao observar os variados desafios enfrentados por essas pessoas com deficiências ou dificuldades funcionais mais severas, principalmente no contexto da ELA, esta pesquisa decide investigar e buscar uma solução de saúde digital viável e acessível para o problema funcional relacionado ao processo comunicativo. Nesta perspectiva, a investigação concentra-se no desenvolvimento de um recurso de Interação Humano-Computador baseado em imagens dos olhos para a Comunicação Aumentativa e Alternativa de pessoas com a ELA. Para isso, conceitos e modelos algorítmicos de visão computacional e aprendizado de máquina são explorados e se apresentam como alternativas eficientes e de baixo custo para o desenvolvimento da solução.

Assim, surgem as seguintes questões de pesquisa:

- Q1: É possível desenvolver tecnologias assistivas de baixo custo e com código fonte livre e aberto, diferentemente dos que estão disponíveis no mercado, para pessoas que vivem com a Esclerose Lateral Amiotrófica, com o objetivo de lhes garantir a comunicação e lhes devolver a autonomia?
- Q2: É possível utilizar técnicas de visão computacional e aprendizado de máquina para desenvolver tecnologias que possam prover Comunicação Aumentativa e Alternativa para pessoas que vivem com a Esclerose Lateral Amiotrófica, considerando os dispositivos móveis neste contexto?

1.2 Hipótese

Diversas abordagens computacionais para a interação humano-computador baseadas em imagens dos olhos são exploradas na literatura, tanto para o reconhecimento do piscar dos olhos quanto para o rastreamento da pupila. Entretanto, são trabalhos que ficam no campo teórico e não chegam, de forma direta, a quem de fato necessita (Fernandes et al. 2023), e os que chegam são caros e inviáveis para que o Sistema Único de Saúde do Brasil, o SUS, possa absorver e distribuir a todos que realmente necessitam. Há neste sentido, os dispositivos comerciais para a Comunicação Aumentativa e Alternativa baseada em imagens dos olhos, que são providos de alta tecnologia e possuem bom desempenho na Interação Humano-Computador. Contudo, o seu custo ultrapassa 12 mil reais¹, o que dificulta o acesso e a possibilidade de escalar o produto no contexto do Brasil. Com isso, foi levantada a hipótese de que é possível desenvolver tecnologias assistivas baseadas em métodos computacionais para garantir formas de Comunicação Aumentativa e Alternativa para pessoas que vivem com Esclerose Lateral Amiotrófica (ELA).

1.3 Objetivos

1.3.1 Objetivo Geral

Desenvolver um ecossistema de Interação Humano-Computador acessível, baseado em imagens dos olhos e técnicas de visão computacional e aprendizado de máquina, para a Comunicação Aumentativa e Alternativa de pessoas com Esclerose Lateral Amiotrófica (ELA).

1.3.2 Objetivos Específicos

1. Revisar na literatura, a partir da definição e execução de um protocolo de revisão sistemática, abordagens computacionais de comunicação baseadas em imagens dos olhos;
2. Descrever uma arquitetura de Comunicação Aumentativa e Alternativa para pessoas que vivem com a ELA;
3. Construir um conjunto de dados (dataset) de imagens oculares de pessoas saudáveis e que vivem com a ELA;
4. Projetar um módulo de Interação Humano-Computador baseado em visão computacional e aprendizado de máquina para detecção do piscar dos olhos;
5. Integrar o módulo de Interação Humano-Computador à arquitetura de Comunicação Aumentativa e Alternativa para pessoas com ELA.

¹Tablet Indi Tobii Dynavox: <https://bit.ly/tablet-indi-tobii-dynavox>. Acesso em 20 de ago de 2023.

1.4 Contribuições

1.4.1 Produções Relacionadas à Tese e ao projeto revELA

- 1. Título:** *Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review*
Periódico: *BioMedical Engineering OnLine*;
status: Publicado.
Fator de Impacto: 3,9
Citescore: 6,1
Qualis CAPES: A3
Como citar:
Fernandes, Felipe, Ingridy Barbalho, Daniele Barros, Ricardo Valentim, César Teixeira, Jorge Henriques, Paulo Gil e Mário Dourado Júnior (2021), 'Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review', *BioMedical Engineering OnLine* **20**(1), 61. **URL:** <https://doi.org/10.1186/s12938-021-00896-2>.
- 2. Título:** *Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam*
Periódico: *Expert Systems with Applications*;
status: Publicado.
Fator de Impacto: 8,5
Citescore: 12,6
Qualis CAPES: A1
Como citar:
de Lima Medeiros, Paulo Augusto, Gabriel Vinícius Souza da Silva, Felipe Ricardo dos Santos Fernandes, Ignacio Sánchez-Gendriz, Hertz Wilton Castro Lins, Daniele Montenegro da Silva Barros, Danilo Alves Pinto Nagem e Ricardo Alexandro de Medeiros Valentim (2022), 'Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam', *Expert Systems with Applications* **188**, 116073. **URL:** <https://doi.org/10.1016/j.eswa.2021.116073>.
- 3. Título:** *Digital Alternative Communication for Individuals with Amyotrophic Lateral Sclerosis: what we have*
Periódico: *Journal of Clinical Medicine | Clinical Neurology*;
status: Publicado.
Fator de Impacto: 3,9
Citescore: 5,4
Qualis CAPES: A2
Como citar:
Fernandes, Felipe, Ingridy Barbalho, Arnaldo Bispo Júnior, Luca Alves, Danilo Nagem, Hertz Lins, Ernano Arrais Júnior, Karilany D. Coutinho, Antônio H. F. Moraes, João Paulo Q. Santos, Guilherme Medeiros Machado, Jorge Henriques, César Teixeira, Mário E. T. Dourado Júnior, Ana R. R. Lindquist e Ricardo A. M. Valen-

tim (2023), 'Digital alternative communication for individuals with amyotrophic lateral sclerosis: What we have', *Journal of Clinical Medicine* **12**(16), 5235. **URL:** <https://doi.org/10.3390/jcm12165235>.

4. **Título:** *National registry for amyotrophic lateral sclerosis: a systematic review for structuring population registries of motor neuron diseases*
Periódico: *BMC Neurology*;
status: Publicado.
Fator de Impacto: 2,6
Citescore: 3,2
Qualis CAPES: A4
Como citar:
Barbalho, Ingridy, Ricardo Valentim, Mário Dourado Júnior, Daniele Barros, Hércules Pedrosa Júnior, Felipe Fernandes, César Teixeira, Thaísa Lima, Jailton Paiva e Danilo Nagem (2021) 'National registry for amyotrophic lateral sclerosis: a systematic review for structuring population registries of motor neuron diseases', *BMC Neurology* **21**(1), 269. **URL:** <https://doi.org/10.1186/s12883-021-02298-2>.
5. **Título:** *Electronic Health Records in Brazil: prospects and technological challenges*
Periódico: *Frontiers in Public Health | Digital Public Health*;
status: Publicado.
Fator de Impacto: 5,2
Citescore: 3,8
Qualis CAPES: A1
Como citar:
Barbalho, Ingridy M. P., Felipe Fernandes, Daniele M. S. Barros, Jailton C. Paiva, Jorge Henriques, Antônio H. F. Morais, Karilany D. Coutinho, Giliate C. Coelho Neto, Arthur Chioro e Ricardo A. M. Valentim (2022), 'Electronic health records in brazil: Prospects and technological challenges', *Frontiers in Public Health* **10**, 963841. **URL:** <https://doi.org/10.3389/fpubh.2022.963841>.
6. **Título:** *Translational Research in Health Technologies: A Scoping Review*
Periódico: *Frontiers in Digital Health | Health Technology Implementation*;
status: Publicado.
Fator de Impacto: 2,3
Citescore: 2,2
Qualis CAPES: -
Como citar:
Mayrink, Nadja N. V., Luís Alcoforado, Arthur Chioro, Felipe Fernandes, Thaisa S. Lima, Erika B. Camargo e Ricardo A. M. Valentim (2022), 'Translational research in health technologies: A scoping review', *Frontiers in Digital Health* **4**, 957367. **URL:** <https://doi.org/10.3389/fdgth.2022.957367>.
7. **Título:** *Digital Health Solution for Monitoring and Surveillance of Amyotrophic Lateral Sclerosis in Brazil*

Periódico: *Frontiers in Public Health | Digital Public Health*;

status: Aceito para publicação.

Fator de Impacto: 5,2

Citescore: 3,8

Qualis CAPES: A1

Como citar:

Barbalho, Ingridy M. P., Aleika L. Fonseca, Felipe Fernandes, Jorge Henriques, Paulo Gil, Danilo Nagem, Raquel Lindquist, Thaísa Santos-Lima, João Paulo Q. Santos, Jailton C. Paiva, Antônio H. F. Morais e Ricardo A. M. Valentim (2023), 'Digital health solution for monitoring and surveillance of amyotrophic lateral sclerosis in brazil', *Frontiers in Public Health* **11**, 1209633. **URL:** <https://doi.org/10.3389/fpubh.2023.1209633>

8. **Título:** *Computer Vision and AI Applied to Blink Detection for Communication Interface for ALS Patients*

Conferência: *Intelligenza Artificiale: per una governance umana. Prospettive educative e sociali*;

status: Publicado. **Como citar:**

de Lima Medeiros, Paulo Augusto, Ricardo Alexsandro de Medeiros Valentim, Gabriel Vinícius Souza da Silva, Daniele Montenegro da Silva Barros, Hertz Wilton Castro Lins, Felipe Ricardo dos Santos Fernandes, Danilo Alves Pinto Nagem (2021), 'Computer Vision and AI Applied to Blink Detection for Communication Interface for ALS Patients', em *Intelligenza Artificiale: per una governance umana. Prospettive educative e sociali*. **URL:** <https://bit.ly/cv-ai-communication-for-als>.

9. **Título:** *An alternative keyboard for communication of patients with amyotrophic lateral sclerosis*

Conferência: *Intelligenza Artificiale: per una governance umana. Prospettive educative e sociali*;

status: Publicado.

Como citar:

Bezerra, André Herman Freire, Ricardo Alexsandro de Medeiros Valentim, Daniele Montenegro da Silva Barros, John Victor Alves Luiz, Hertz Wilton Castro Lins, Felipe Ricardo dos Santos Fernandes, Danilo Alves Pinto Nagem (2021), 'An alternative keyboard for communication of patients with amyotrophic lateral sclerosis', em *Intelligenza Artificiale: per una governance umana. Prospettive educative e sociali*. **URL:** <https://bit.ly/alternative-keyboard-als>.

10. **Título:** *Artificial Intelligence applied in a Human-Machine Interface for an alternative communication system to ALS patients*

Conferência: *Intelligenza Artificiale: per una governance umana. Prospettive educative e sociali*;

status: Publicado.

Como citar:

Pires de Souza, Maria Luiza, Ricardo Alexsandro de Medeiros Valentim, Eulália Costa Ribeiro, John Victor Alves Luiz, Daniele Montenegro da Silva Barros, Gabriel Vinicius Sousa da Silva, Hertz Wilton Castro Lins, Felipe Ricardo dos Santos Fernandes, Danilo Alves Pinto Nagem (2021), ‘Artificial Intelligence applied in a Human-Machine Interface for an alternative communication system to ALS patients’, em *Intelligenza Artificiale: per una governance umana. Prospettive educative e sociali*. **URL:** <https://bit.ly/ai-hci-als>.

11. **Registro de Software:** Autonomus
Processo N°: BR512023001526-1.
status: Expedido em 2023.

12. **Registro de Software:** Autonomus Mobile
Processo N°: BR512023001573-3.
status: Expedido em 2023.

13. **Registro de Software:** Observatório Nacional da ELA
Processo N°: BR512023002161-0.
status: Expedido em 2023.

14. **Registro de Software:** Registro Nacional da ELA
Processo N°: BR512023002160-1.
status: Expedido em 2023.

1.4.2 Produções Relacionadas ao Percorso Acadêmico

1. **Título:** *Data Report: “Health care of Persons Deprived of Liberty” Course from Brazil’s Unified Health System Virtual Learning Environment*
Periódico: *Frontiers in Medicine | Healthcare Professions Education*;
status: Publicado;
Fator de Impacto: 3,9
Citescore: 3,6
Qualis CAPES: A1
Como citar:
Valentim, Janaína, Eloiza da S. G. Oliveira, Ricardo A. de M. Valentim, Sara Dias-Trindade, Aline de Pinho Dias, Aliete Cunha-Oliveira, Ingridy Barbalho, Felipe Fernandes, Rodrigo Dantas da Silva, Manoel Honorio Romão, César Teixeira e Jorge Henriques (2021), ‘Data report: “health care of persons deprived of liberty” course from brazil’s unified health system virtual learning environment’, *Frontiers in Medicine* **8**, 742071. **URL:** <https://doi.org/10.3389/fmed.2021.742071>.

2. **Título:** *Virtual Learning Environment of the Brazilian Health System (AVASUS): Efficiency of Results, Impacts, and Contributions*
Periódico: *Frontiers in Medicine | Healthcare Professions Education*;
status: Publicado;
Fator de Impacto: 3,9
Citescore: 3,6
Qualis CAPES: A1
Como citar:
Valentim, Ricardo, Carlos A.P. Oliveira, Eloiza Oliveira, Eduardo Ribeiro, Soneide Costa, Ione Morais, Felipe Fernandes, Alexandre Caitano, Cristine Gusmão, Aliete Cunha-Oliveira, Maria C. Rêgo, Karilany Coutinho, Daniele Barros e Ricardo Cecim (2022), ‘Virtual learning environment of the brazilian health system (avasus): Efficiency of results, impacts, and contributions’, *Frontiers in Medicine* **9**, 896208. **URL:** <https://doi.org/10.3389/fmed.2022.896208>.
3. **Título:** *The Relevancy of Massive Health Education in The Brazilian Prison System: The Course “Health Care for People Deprived of Freedom” and Its Impacts*
Periódico: *Frontiers in Public Health | Life-Course Epidemiology and Social Inequalities in Health*;
status: Publicado.
Fator de Impacto: 5,2
Citescore: 3,8
Qualis CAPES: A1
Como citar:
Valentim, Janaína L. R. S., Sara Dias-Trindade, Eloiza S. G. Oliveira, José A. M. Moreira, Felipe Fernandes, Manoel H. Romão, Philippi S. G. Morais, Alexandre R. Caitano, Aline P. Dias, Carlos A. P. Oliveira, Karilany D. Coutinho, Ricardo B. Ceccim e Ricardo A. M. Valentim (2022), ‘The relevancy of massive health education in the brazilian prison system: The course “health care for people deprived of freedom” and its impacts’, *Frontiers in Public Health* **10**, 935389. **URL:** <https://doi.org/10.3389/fpubh.2022.935389>.
4. **Título:** *Effectiveness of COVID-19 Vaccination on Reduction of Hospitalizations and Deaths in Elderly Patients in Rio Grande do Norte, Brazil*
Periódico: *International Journal of Environmental Research and Public Health*;
status: Publicado.
Fator de Impacto: 4,6
Citescore: 5,4
Qualis CAPES: A1
Como citar:
Sales-Moioli, Ana Isabela L., Leonardo J. Galvão-Lima, Talita K. B. Pinto, Pablo H. Cardoso, Rodrigo D. Silva, Felipe Fernandes, Ingridy M. P. Barbalho, Fernando L. O. Farias, Nicolas V. R. Veras, Gustavo F. Souza, Agnaldo S. Cruz, Ion G. M. Andrade, Lúcio Gama e Ricardo A. M. Valentim (2022), ‘Effectiveness of covid-19 vaccination on reduction of hospitalizations and deaths in elderly pati-

ents in rio grande do norte, brazil’, *International Journal of Environmental Research and Public Health* **19**(21), 13902. **URL:** <https://doi.org/10.3390/ijerph192113902>.

5. **Título:** *Development of a Cyclic Voltammetry-Based Method for the Detection of Antigens and Antibodies as a Novel Strategy for Syphilis Diagnosis*
Periódico: *International Journal of Environmental Research and Public Health*;
status: Publicado.
Fator de Impacto: 4,6
Citescore: 5,4
Qualis CAPES: A1
Como citar:
Barros, Gabriel M. C., Dionísio D. A. Carvalho, Agnaldo S. Cruz, Ellen K. L. Moraes, Ana Isabela L. Sales-Moioli, Talita K. B. Pinto, Melise C. D. Almeida, Ignacio Sanchez-Gendriz, Felipe Fernandes, Ingridy M. P. Barbalho, João P. Q. Santos, Jorge M. O. Henriques, César A. D. Teixeira, Paulo Gil, Lúcio Gama, Angélica E. Miranda, Karilany D. Coutinho, Leonardo J. Galvão-Lima e Ricardo A. M. Valentim (2022), ‘Development of a cyclic voltammetry-based method for the detection of antigens and antibodies as a novel strategy for syphilis diagnosis’, *International Journal of Environmental Research and Public Health* **19**(23), 16206. **URL:** <https://doi.org/10.3390/ijerph192316206>.
6. **Título:** *Osteoporosis screening using machine learning and electromagnetic waves*
Periódico: *Scientific Reports*;
status: Publicado.
Fator de Impacto: 4,6
Citescore: 7,1
Qualis CAPES: A1
Como citar:
Albuquerque, Gabriela A., Dionísio D. A. Carvalho, Agnaldo S. Cruz, João P. Q. Santos, Guilherme M. Machado, Ignácio S. Gendriz, Felipe R. S. Fernandes, Ingridy M. P. Barbalho, Marquiony M. Santos, César A. D. Teixeira, Jorge M. O. Henriques, Paulo Gil, Adrião D. D. Neto, Antonio L. P. S. Campos, Josivan G. Lima, Jailton C. Paiva, Antonio H. F. Moraes, Thaisa Santos Lima e Ricardo A. M. Valentim (2023), ‘Osteoporosis screening using machine learning and electromagnetic waves’, *Scientific Reports* **13**(1), 12865. **URL:** <https://doi.org/10.1038/s41598-023-40104-w>.
7. **Título:** *Evaluation of Massive Education in Prison Health: a perspective of health care for the person deprived of freedom in Brazil*
Periódico: *Frontiers in Public Health | Public Health Education and Promotion*;
status: Aceito para publicação.
Fator de Impacto: 5,2
Citescore: 3,8
Qualis CAPES: A1

Como citar:

Valentim, Janaína L. R. S., Sara Dias-Trindade, Eloiza S. G. Oliveira, Manoel H. Romão, Felipe Fernandes, Alexandre R. Caitano, Marilyn A. A. Bonfim, Aline P. Dias, Cristine M. G. Gusmão, Philippi S. G. Morais, Ronaldo S. Melo, Gustavo Fontoura de Souza, Kelson C. Medeiros, Maria C. F. D. Rêgo, Ricardo B. Ceccim e Ricardo A. M. Valentim, 'Evaluation of Massive Education in Prison Health: a perspective of health care for the person deprived of freedom in Brazil' *Frontiers in Public Health*, **11**, 1239769. **URL:** <https://doi.org/10.3389/fpubh.2023.1239769>.

8. **Título:** *Computational Methods Applied to Syphilis: Where are we, and Where are we Going?*

Periódico: *Frontiers in Public Health | Digital Public Health*;

status: Aceito para publicação.

Fator de Impacto: 5,2

Citescore: 3,8

Qualis CAPES: A1

Como citar:

Albuquerque, Gabriela, Fernandes, Felipe, Ingridy M. P. Barbalho, Daniele M. Barros, Philippi S. G. Morais, Antônio H. F. Morais, Marquiony M. Santos, Leonardo J. Galvão-Lima, Ana Isabela L. Sales-Moioli, João Paulo Q. Santos, Paulo Gil, Jorge Henriques, César Teixeira, Thaisa Santos Lima, Karilany D. Coutinho, Talita K. B. Pinto e Ricardo A. M. Valentim, 'Computational Methods Applied to Syphilis: Where are we, and Where are we Going?' *Frontiers in Public Health*, **11**, 1201725. **URL:** <https://doi.org/10.3389/fpubh.2023.1201725>.

9. **Registro de Software:** SALUS 2.0

Processo N°: BR512022002956-1.

status: Expedido em 2022.

1.5 Organização do Trabalho

Esta tese está organizada em 8 capítulos, somado o presente Capítulo 1 de Introdução, da forma como segue:

Capítulo 2: apresentam-se os aspectos teóricos básicos que fundamentam e orientam o desenvolvimento da abordagem de Interação Humano-Computador baseada em imagens dos olhos;

Capítulo 3: é dedicado à apresentação do artigo intitulado "*Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review*", que reúne evidências relacionadas ao uso de sinais biomédicos e abordagens computacionais no contexto da Esclerose Lateral Amiotrófica;

Capítulo 4: apresenta o artigo intitulado “*Digital Alternative Communication for Individuals with Amyotrophic Lateral Sclerosis: what we have*”, que fornece evidências e identifica lacunas relacionadas a abordagens de Interação Humano-Computador baseadas em imagens dos olhos para a comunicação aumentativa e alternativa de pessoas que vivem com Esclerose Lateral Amiotrófica;

Capítulo 5: apresenta o módulo de Interação Humano-Computador baseado em imagens dos olhos como uma solução de comunicação viável e acessível para as pessoas que vivem com Esclerose Lateral Amiotrófica. Este capítulo é a versão publicada do artigo intitulado “*Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam*”;

Capítulo 6: é dedicado à apresentação do Ecossistema Autonomus, uma solução de saúde digital para a Comunicação Aumentativa e Alternativa de Pessoas com Esclerose Lateral Amiotrófica, bem como os resultados alcançados durante o seu desenvolvimento;

Capítulo 7: apresentam-se as discussões desta tese, que abordam o objeto de estudo tanto na perspectiva tecnológica quanto na humana, pensando nas pessoas que vivem com Esclerose Lateral Amiotrófica ou outras dificuldades funcionais relacionadas ao processo comunicativo;

Capítulo 8: é dedicado à apresentação das considerações finais, das limitações identificadas e das perspectivas para trabalhos futuros.

Capítulo 2

Fundamentos Teóricos

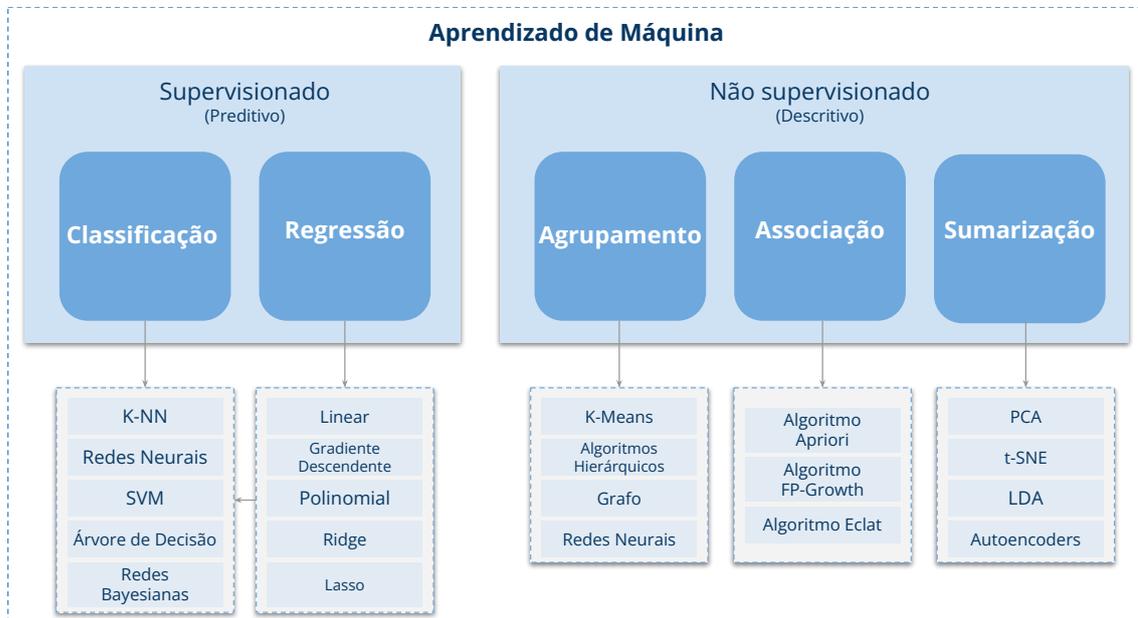
Este capítulo apresenta conceitos preliminares e essenciais para a compreensão do objeto de estudo. Apresenta-se uma visão geral sobre os princípios básicos das áreas de aprendizado de máquina e visão computacional. Além disso, é apresentado o projeto revELA, ao qual esta tese está inserida. Os conceitos relevantes sobre a Esclerose Lateral Amiotrófica foram abordados resumidamente no Capítulo 1.

2.1 Aprendizado de Máquina

O termo *machine learning*, ou simplesmente aprendizado de máquina, apresenta-se como uma área de estudo, emergente do campo da Inteligência Artificial, centrada no desenvolvimento de algoritmos capazes de realizar tarefas complexas, de diferentes domínios, com base em dados (Mitchell 1997, Faceli et al. 2021, Russell e Norving 2022). Os algoritmos computacionais de aprendizado de máquina são amplamente utilizados e de forma bem-sucedida em problemas que se assemelham ao desta tese, tais como reconhecimento facial e processamento de imagens (Paguirigan et al. 2023, Rakhmatulin e Duchowski 2020, Singh 2019, Chen e Jenkins 2017).

Faceli et al. (2021) apresentam duas categorias principais de tarefas em que os algoritmos de aprendizado de máquina têm sido aplicados: preditivas e descritivas. Os algoritmos desenvolvidos para cada categoria seguem paradigmas diferentes. Esses paradigmas estão relacionados ao uso do conhecimento ou informação do conjunto de dados no processo de aprendizagem. Na tarefa de predição, os algoritmos fazem o uso do conhecimento dos dados no processo de treinamento (aprendizagem), isto é, o conjunto de dados inclui os rótulos (as soluções desejadas). Neste caso, os algoritmos seguem o paradigma de aprendizado supervisionado. No caso da tarefa de descrição, não há dados rotulados e os modelos algorítmicos aprendem a partir da extração de padrões do conjunto de dados. Esta abordagem segue o paradigma de aprendizado não supervisionado (Faceli et al. 2021, Russell e Norving 2022, Géron 2022). A Figura 2.1 mostra uma visão geral das tarefas de predição e descrição, e apresenta exemplos de algoritmos clássicos de cada uma das categorias.

Figura 2.1: Visão geral hierárquica dos tipos e algoritmos clássicos de aprendizado de máquina.



Fonte: Adaptada de Faceli et al. (2021) e Géron (2022).

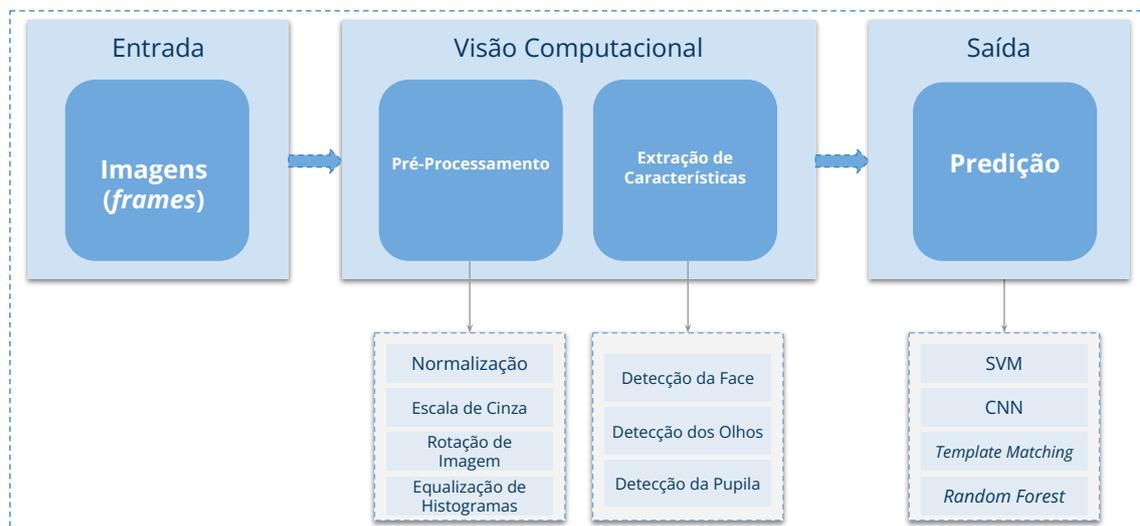
No contexto recente do campo de aprendizado de máquina, devido ao grande volume de dados disponível e à evolução do poder computacional, em termos de processamento, muitas pesquisas têm sido desenvolvidas utilizando redes neurais profundas, dando origem a uma nova subárea de aprendizado de máquina chamada de aprendizado profundo ou *deep learning*. Algoritmos de *deep learning* vêm sendo amplamente utilizados e têm mostrado resultados promissores na resolução de aplicações práticas complexas em diversos domínios, como em tarefas de processamento de imagem, de voz e de linguagem natural (Goodfellow et al. 2016, Faceli et al. 2021). Nesta tese, por exemplo, um algoritmo de *deep learning*, baseado em uma rede neural convolucional (CNN, do Inglês *Convolutional Neural Networks*), é projetado para o reconhecimento do piscar voluntário de olhos em imagens de pessoas. Esses algoritmos baseados em CNN são, em sua essência, redes neurais especializadas, inspiradas no córtex visual do cérebro que, diferentemente das redes neurais artificiais convencionais, empregam em sua arquitetura pelo menos uma operação matemática chamada convolução (Goodfellow et al. 2016, Géron 2022). Maiores detalhes sobre a arquitetura CNN projetada para o recurso de Comunicação Aumentativa e Alternativa são apresentados no Capítulo 5.

2.2 Visão Computacional

Visão Computacional é um campo de estudo que busca desenvolver um conjunto de técnicas algorítmicas capazes de extrair descrições de objetos a partir de imagens (Forsyth e Ponce 2011, Kumar e Renuka 2023). Um modelo de visão computacional, cujo obje-

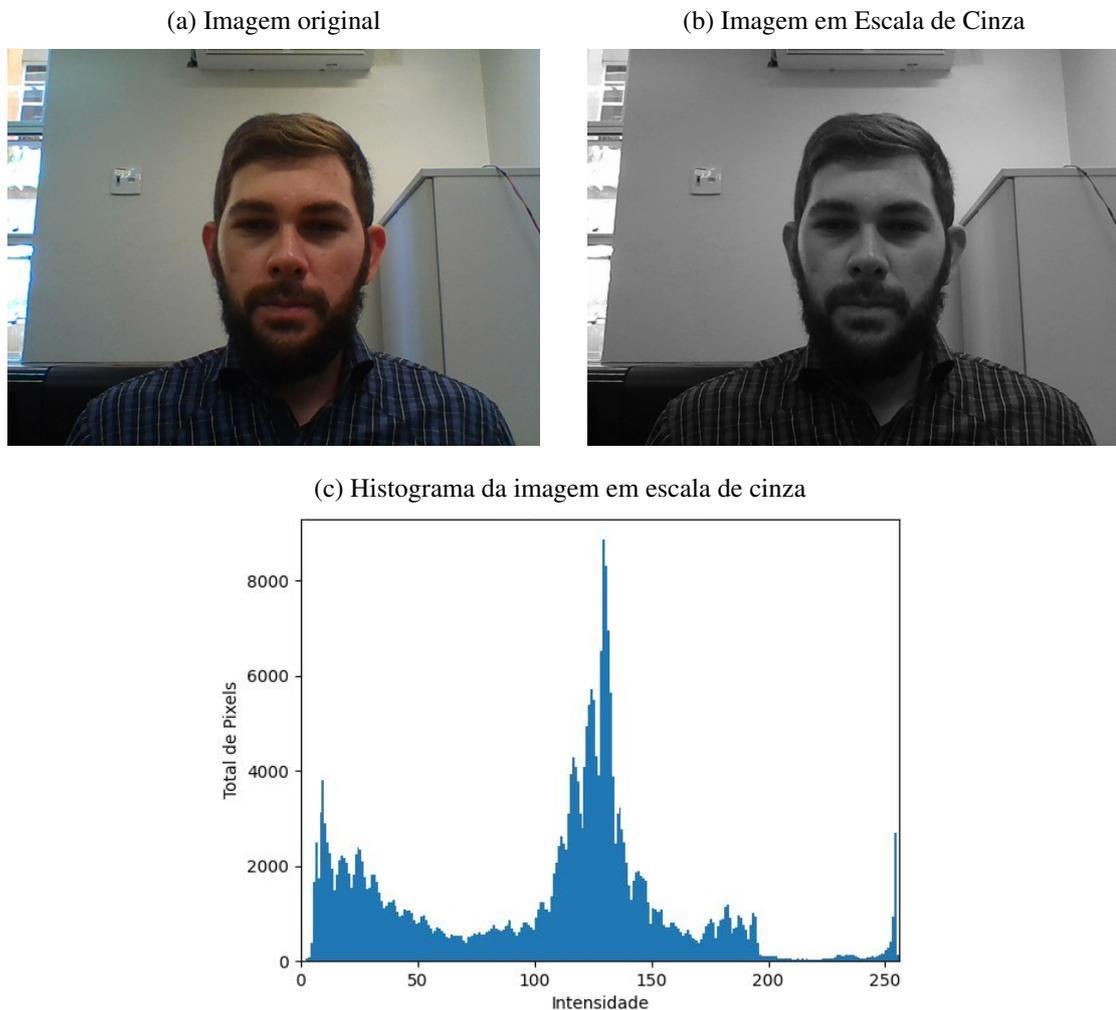
tivo é a interpretação ou reconhecimento de objetos em imagens, envolve conhecimentos e técnicas de outras áreas, como por exemplo inteligência artificial, aprendizado de máquina, matemática ou estatística. Este modelo, normalmente, segue um fluxo de trabalho organizado e sequencial para que haja a compreensão e percepção das informações do campo visual. A Figura 2.2 apresenta um fluxo genérico que se inicia com a aquisição de uma ou mais imagens do ambiente (entrada), passa pela etapa de manipulação das imagens e preparação dos dados (visão computacional: pré-processamento e extração de características, respectivamente) e chega na predição, isto é, no reconhecimento do objeto (saída) (Kumar e Renuka 2023).

Figura 2.2: Visão geral do fluxo de trabalho (*pipeline*) básico no campo de visão computacional.



A visão computacional atua, principalmente, no tratamento, preparação e enriquecimento dos dados provenientes das imagens para servir modelos de aprendizado de máquina. A Figura 2.3, por exemplo, ilustra o processo simples de tratamento de uma imagem, por meio da aplicação das técnicas de escala de cinza (Figura 2.3b) e de equalização de histograma (Figura 2.3c) na imagem original (Figura 2.3a). Outros recursos mais robustos de visão computacional são abordados nos Capítulos 4 e 5.

Figura 2.3: Ilustração dos procedimentos de pré-processamento.



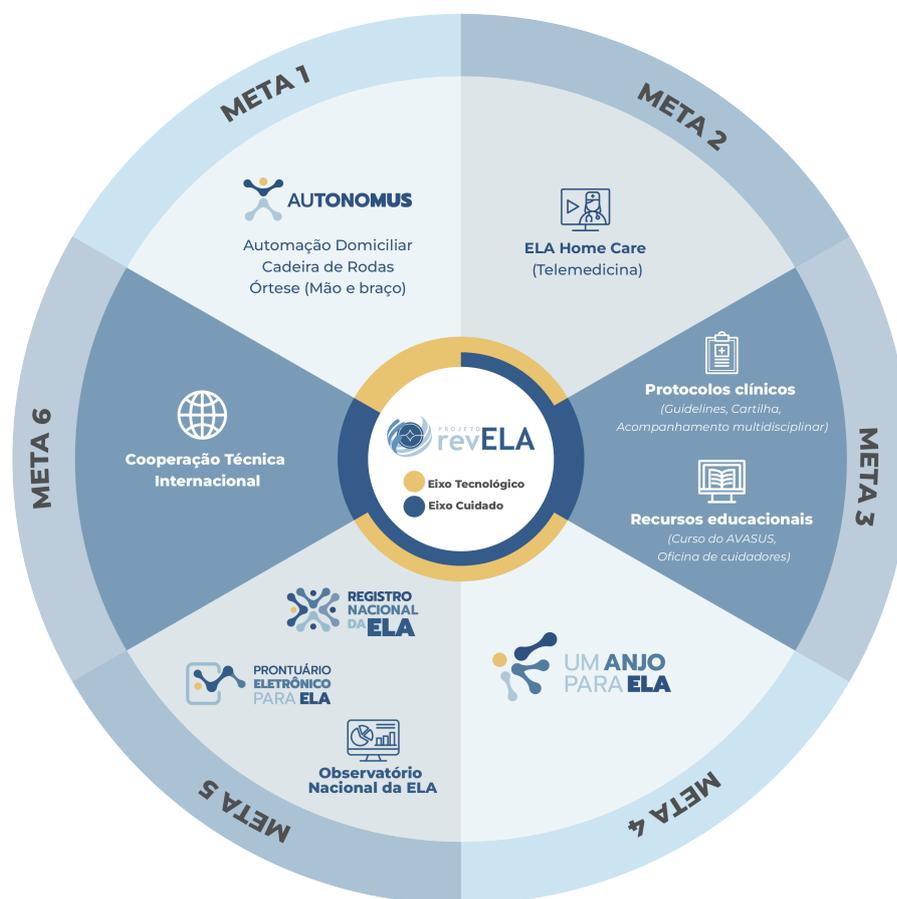
2.3 Projeto revELA

O projeto revELA refere-se à atividade de Desenvolvimento Científico e Tecnológico Aplicado a Esclerose Lateral Amiotrófica, relativa a TED 132/2018 - um acordo de cooperação técnica entre a Universidade Federal do Rio Grande do Norte (UFRN), por intermédio do Laboratório de Inovação Tecnológica em Saúde (LAIS), e o Ministério da Saúde do Brasil (MS). Com atuação em âmbito nacional e Internacional e centrado no fortalecimento do Sistema Único de Saúde (SUS), o projeto revELA tem como principal objetivo o desenvolvimento de pesquisas e estudos voltados para a produção de tecnologias e inovações para o tratamento, monitoramento e a definição de protocolos para o acompanhamento adequado das pessoas com Esclerose Lateral Amiotrófica (Nagem et al. 2023).

Conforme ilustrado na Figura 2.4, o projeto revELA é estruturado em seis dimensões, chamadas de metas. Cada meta desempenha ações específicas, tanto do eixo tecnológico

quanto do cuidado, para o alcance dos objetivos do projeto. Esta tese, em especial, está situada na meta 1, no qual envolve o desenvolvimento de uma plataforma tecnológica, solução de saúde digital, para auxílio à comunicação e interação com o ambiente domiciliar de pessoas com ELA. O produto proveniente desta meta chama-se Autonomus, um recurso de tecnologia assistiva acessível para a comunicação alternativa de pessoas com ELA. Os capítulos 5 e 6 desta tese descrevem cuidadosamente o desenvolvimento do módulo de Interação Humano-Computador e a arquitetura do Autonomus, respectivamente, componentes de fundamental importância que sustentam o Ecossistema Autonomus. Ambos os capítulos apresentam resultados pertencentes à meta 1.

Figura 2.4: Dimensões do projeto revELA.



As outras dimensões buscam atender aspectos relacionados a outras necessidades no contexto da ELA. Na meta 2, por exemplo, são desenvolvidas atividades e produtos (tele-saúde) para a prestação de cuidados multidisciplinares de pessoas que vivem com a ELA (Fidelix et al. 2023). Na meta 3, também com vistas ao cuidado, protocolos clínicos das áreas de fisioterapia e nutrição foram desenvolvidos (de Souza et al. 2022, Pondofe et al. 2022, Silva et al. 2022). Ainda nesta meta, foram desenvolvidos recursos educacionais de acesso gratuito, no Ambiente Virtual de Aprendizagem do Sistema Único de

Saúde (AVASUS) (Valentim et al. 2022), para a capacitação de profissionais de saúde bem como para orientar a população sobre diversos assuntos sobre a ELA. A trilha educativa para ELA no AVASUS ¹ conta com 21 cursos atualmente e mais de 32 mil alunos inscritos, onde cada aluno, em média, se matricula em dois cursos (AVASUS).

A meta 4 desenvolve ações voltadas para o acompanhamento remoto on-line das pessoas com ELA com foco, principalmente, na análise de sinais vitais e biomédicos, tais como a eletromiografia para medir os níveis de fasciculações ou perdas dos movimentos musculares durante o progresso da doença. Buscando resolver problemas relacionados à ausência de dados sobre a ELA no Brasil. Um fator a ser considerado é que parte desta pesquisa é continuidade do projeto Um Anjo para ELA que surgiu em 2013 (Silva et al. 2013), portanto, trata-se de um estudo que ganha maturidade por meio do Projeto revELA. A meta 5, que faz interface com a meta 4, desenvolve ações para realizar o mapeamento contínuo e online sobre a ELA. Além disso, a meta 5 é responsável pelo desenvolvimento de um prontuário eletrônico para pessoas com ELA, o que contribui diretamente para o acompanhamento clínico pela equipe multidisciplinar (Barbalho et al. 2023).

A dimensão relacionada a meta 6 busca fortalecer as redes de cooperação nacional e internacional, para a realização de estudos, avaliações e homologações de tecnologias desenvolvidas no âmbito do projeto revELA. Aspecto considerado importante para o fortalecimento da pesquisa no campo da Esclerose Lateral Amiotrófica no contexto da Saúde Global, especialmente, por contribuir também com a implementação da Meta 17 ² dos Objetivos de Desenvolvimento Sustentável da Agenda 2030.

¹Acesse a Trilha: <https://avasus.ufrn.br/local/avasplugin/cursos/raras.php>

²Parcerias e meios de implementação: Reforçar os meios de implementação e revitalizar a parceria global para o desenvolvimento sustentável. Acesse: <https://brasil.un.org/pt-br/sdgs/17>

Capítulo 3

Sinais Biomédicos e Aprendizado de Máquina para ELA

Este capítulo apresenta o artigo intitulado “*Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review*”, em sua versão publicada no periódico internacional *BioMedical Engineering OnLine*. O artigo reúne, avalia e fornece evidências, bem como lacunas, relacionadas a sinais biomédicos e abordagens computacionais, baseadas em técnicas de aprendizado de máquina, aplicadas a desafios no contexto de pessoas com Esclerose Lateral Amiotrófica. Além disso, este artigo representa o início das investigações no contexto da Esclerose Lateral Amiotrófica que, conseqüentemente, contribuiu e endossou o objeto de estudo desta tese.

Fernandes, Felipe, Ingridy Barbalho, Daniele Barros, Ricardo Valentim, César Teixeira, Jorge Henriques, Paulo Gil e Mário Dourado Júnior (2021), ‘Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review’, *BioMedical Engineering OnLine* **20**(1), 61. URL: <https://doi.org/10.1186/s12938-021-00896-2>

3.1 Protocolo de Busca

Esta pesquisa foi desenvolvida baseando-se no *guideline* para revisões sistemáticas proposto por Kitchenham (2004) e, na perspectiva de investigar aplicações tecnológicas na ELA, objetiva: (i) identificar quais são os sinais biomédicos mais utilizados; (ii) identificar a finalidade da utilização dos sinais biomédicos; (iii) e verificar a utilização das técnicas de aprendizado de máquina ou abordagens inteligentes no processamento dos sinais biomédicos. Com essa premissa, foram levantadas as questões de pesquisa (QP) contidas na Tabela 3.1.

O processo de busca e seleção de estudos primários nas bases de dados bibliográficas digitais foi dividido em quatro etapas, conforme mostra a Figura 3.1. Na primeira etapa, um conjunto inicial de artigos foi selecionado a partir dos resultados retornados das buscas realizadas nas bases de dados *IEEE Xplore*, *Web of Science*, *Science Direct*, *Springer* e *PubMed*. Para as buscas, as seguintes *Strings* de busca (STR) foram utilizadas nesta primeira etapa:

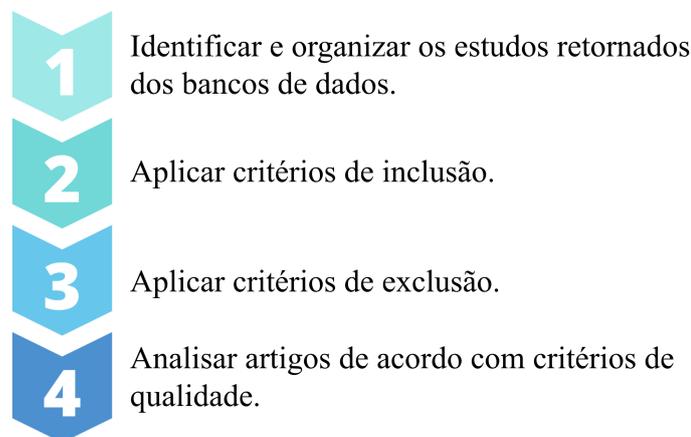
CAPÍTULO 3. SINAIS BIOMÉDICOS E APRENDIZADO DE MÁQUINA PARA ELA22

- STR01: (((“signals processing” OR “biomedical signals”) OR (“smart systems” OR “machine learning” OR “artificial intelligence” OR “computational intelligence” OR “algorithm” OR “algorithms”)) AND (“amyotrophic lateral sclerosis” OR “als”));
- STR02: (((“signals processing” OR “biomedical signals”) OR (“intelligent systems” OR “machine learning” OR “artificial intelligence” OR “algorithms” OR “Computational Intelligence”)) AND (“amyotrophic lateral sclerosis” OR “als”)).

Tabela 3.1: Questões de Pesquisa.

QP	Descrição
01	O processamento de sinais é utilizado com qual finalidade?
02	Quais são os tipos de sinais analisados pelo estudo?
03	Quais as técnicas inteligentes são utilizadas no estudo?
04	Qual a performance das técnicas analisadas?
05	Quantos pacientes são utilizados para testar ou validar o estudo?

Figura 3.1: Passos metodológicos.



Na segunda etapa, os critérios de inclusão (CI) predefinidos, apresentados na Tabela 3.2, foram aplicados ao conjunto inicial de artigos da etapa anterior. Os CI delimitam, de forma primária, as fronteiras ou o escopo da investigação e possibilitam formar um novo subconjunto de artigos com maior probabilidade de conter respostas às questões de pesquisa. Neste caso, no subconjunto estão incluídos artigos dos últimos dez anos, que são publicados em *journals* e que estão relacionados diretamente à grande área de interesse desta revisão sistemática.

Tabela 3.2: Critérios de Inclusão.

CI	Descrição
01	Artigos publicados entre 2009 e 2019.
02	Artigos de pesquisa publicados em <i>Journals</i> .
03	Artigos das áreas de tecnologia, engenharias ou ciência da computação.

Após a seleção dos artigos pelos CI, nesta terceira etapa acontece a verificação e remoção de artigos duplicados e um procedimento de filtragem, através da leitura do título, resumo e palavras-chave é aplicado para remover artigos que não apresentam os termos específicos associados ao tema de investigação desta revisão sistemática. Esta inspeção é norteada pelos critérios de exclusão (CE), apresentados na Tabela 3.3, e foi realizada com o auxílio da aplicação web *Rayyan* (Ouzzani et al. 2016).

Tabela 3.3: Critérios de Exclusão.

CE	Descrição
01	Artigos duplicados.
02	Estudos não relacionados à processamento de sinais biomédicos, aprendizado de máquina, sistemas inteligentes e análise de dados de pessoas com ELA.

Na quarta etapa, foi realizada a leitura completa dos artigos filtrados do passo anterior. Assim, foi executado o protocolo baseado nos Critérios de Avaliação de Qualidade (CAQ), apresentados na Tabela 3.4. Neste procedimento, foi atribuída para cada critério uma pontuação que mensura a relevância do artigo no assunto alvo desta pesquisa. A pontuação é distribuída no formato de pesos para possíveis respostas aos CAQ, contidas nos estudos primários, sendo 1.0 o peso de maior relevância e 0 o de menor (Equação 3.1).

$$w_{CAQ} = \begin{cases} 1.0, & \text{sim, descreve totalmente,} \\ 0.5, & \text{sim, descreve parcialmente,} \\ 0, & \text{não descreve.} \end{cases} \quad (3.1)$$

Um *score* (Equação 3.2), a média aritmética das pontuações dos critérios da CAQ (Equação 3.1), foi gerado para cada artigo. Neste caso, todos os artigos que obtiveram *score* maior ou igual a 0.5 ($0.5 \leq \text{score} \leq 1$) foram selecionados para esta pesquisa e constituem o conjunto final de artigos.

$$\text{score} = \frac{1}{n_{CAQ}} \sum_{i=1}^{n_{CAQ}} w_{CAQ_i} \quad (3.2)$$

Tabela 3.4: Critérios de Avaliação de Qualidade.

CAQ	Descrição
01	O estudo descreve claramente os tipos de sinais biomédicos?
02	O estudo descreve como é realizado o processamento do sinal (por exemplo, as técnicas algorítmicas)?
03	O estudo descreve o processo da aplicação para pessoas com ELA (detalha o experimento)?
04	O estudo descreve claramente suas contribuições científicas para a evolução das pesquisas relacionadas à ELA?

Os registros pertinentes a cada etapa, bem como os dados extraídos dos artigos, foram armazenados adequadamente em planilhas e na aplicação web *Rayyan* (Ouzzani et al. 2016) para extração dos dados. Dados como ano de publicação, autores e possíveis respostas às questões de pesquisa foram extraídos do conjunto de artigos provenientes da última etapa e permitirá a análise final para cumprimento dos objetivos desta revisão sistemática.

3.2 Resultados

O resultado do processo de busca e seleção de estudos primários é sintetizado na Figura 3.2. Na primeira etapa, 10128 artigos candidatos foram identificados após busca nas bases de dados usando as STR. Na fase seguinte, três procedimentos de refinamento baseados nos CI (Tabela 3.2) foram aplicados e 9914 artigos foram excluídos por não satisfazerem os CI. Após o refinamento, 214 artigos foram considerados aptos e selecionados para a terceira etapa. Nesta nova etapa, os filtros baseados nos CE (Tabela 3.3) excluíram 186 artigos duplicados e com ausência de termos alvos da pesquisa, restando 28 artigos para leitura completa e avaliação através dos critérios de CAQ (Tabela 3.4). Após procedimento de CAQ, quarta e última etapa, 18 artigos superaram o *score* mínimo preestabelecido, conforme resultado apresentado na respectiva coluna na Tabela 3.5, e foram incluídos para análise e investigação final desta revisão.

CAPÍTULO 3. SINAIS BIOMÉDICOS E APRENDIZADO DE MÁQUINA PARA ELA25

Figura 3.2: Resultado dos procedimentos de busca e seleção de estudos primários para esta revisão sistemática.

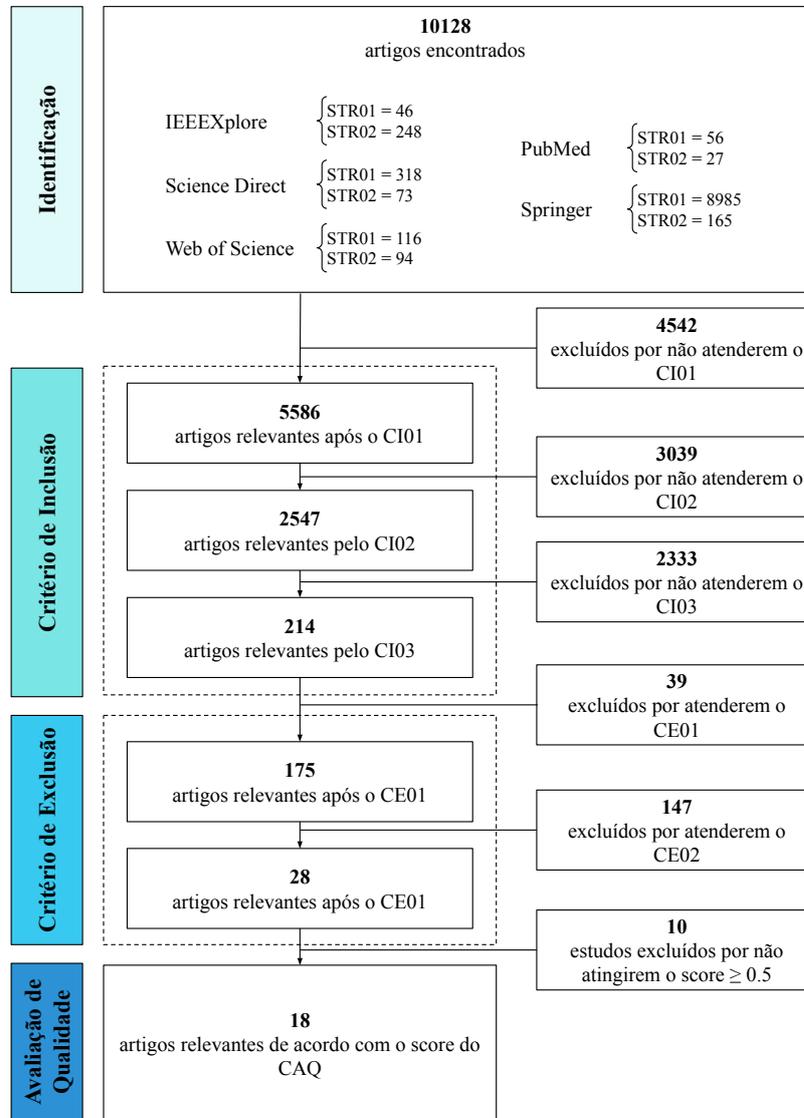


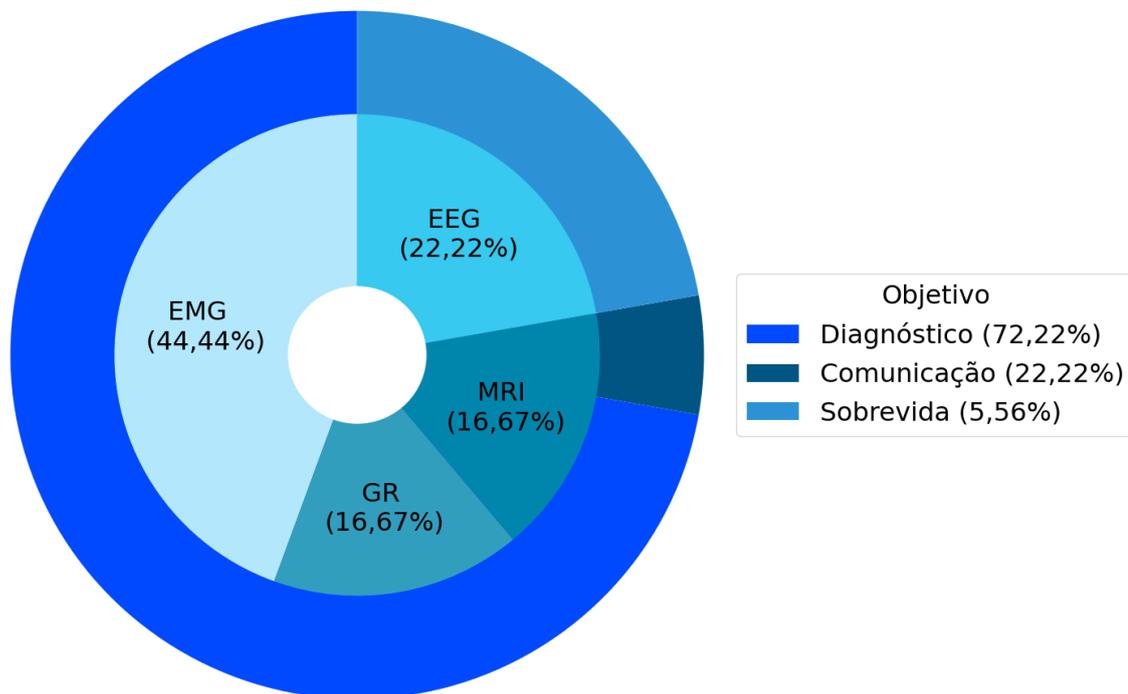
Tabela 3.5: Conjunto de artigos selecionados e suas principais características.

ID	Estudo	Ano	Score	Objetivo	Sinal	Dataset	Indivíduos GC/ELA/OD	Melhor Modelo	Performance (%)		
									Acc	Spe	Sen
01	Chatterjee et al. (2019)	2019	1.0	Diagnóstico	EMG	Público	-8/7	SVM	98.58	99.5	97.66
02	Zhang et al. (2014)	2014	1.0	Diagnóstico	EMG	Local	11/10/-	LDA	-	100	90
03	Hazarika et al. (2019)*	2019	1.0	Diagnóstico	EMG	Público	10/8/7 e 4/4/4	QDC e QDC	99.03 e 100	99.58 e 100	96 e 100
04	Gokgoz e Subasi (2014)	2014	1.0	Diagnóstico	EMG	Local	10/8/7	SVM	92.55	90.3	96.33
05	Ambikapathy et al. (2018)	2018	0.875	Diagnóstico	EMG	Público	-	RNA	96.6	100	93.7
06	Doulah et al. (2014)	2014	1.0	Diagnóstico	EMG	Público	10/8/7	KNN	98.8	100	98
07	Vallejo et al. (2018)	2018	1.0	Diagnóstico	EMG	Público	10/8/7	RNA	98	97.5	100
08	Gokgoz e Subasi (2015)	2015	1.0	Diagnóstico	EMG	Público	10/8/7	RF	96.67	94.75	99.58
09	Xia et al. (2015)	2015	0.875	Diagnóstico	GR	Público	16/13/35	SVM	96.55	94	100
10	Ren et al. (2017)	2017	1.0	Diagnóstico	GR	Público	16/13/35	MLP	-	-	-
11	Khorasani et al. (2016)	2016	0.875	Diagnóstico	GR	Público	16/13/-	FHMM	93.1	93.75	92.31
12	Welsh et al. (2013)	2013	1.0	Diagnóstico	MRI	Local	31/32/-	SVM	71.5	-	-
12	Ferraro et al. (2017)	2017	0.75	Diagnóstico	MRI	Local	78/123/64	RF	91	92	91
14	Miao et al. (2020)	2020	1.0	Comunicação	EEG	Local	-/18/-	BLDA	90	-	-
15	Liu et al. (2017)	2017	0.875	Comunicação	EEG	Local	-/5/-	KNN e LDA	95.25	-	-
16	Sorbello et al. (2018)	2018	0.75	Comunicação	EEG	Local	4/4/-	LDA	-	-	-
17	Mainsah et al. (2015)	2015	0.875	Comunicação	EEG	Local	-/10/-	DSLML	76.39	-	-
18	van der Burgh et al. (2017)	2017	1.0	Sobrevida	MRI	Local	-/135/-	DL	84.4	-	-

Abreviação: GC = Grupo de Controle, ELA = Esclerose Lateral Amiotrófica, OD = Outras Doenças, Acc = Acurácia, Spe = Especificidade, Sen = Sensibilidade, EMG = Eletromiografia, EEG = Eletroencefalograma, MRI = Imagem de Ressonância Magnética, GR = Ritmo da Marcha, SVM = *Support Vector Machine*, RF = *Random Forest*, LDA = Análise Discriminante Linear, QDC = Classificador Quadrático, RNA = Rede Neural Artificial, KNN = *k-Nearest Neighbor*, MLP = *Multilayer Perceptron*, FHMM = Modelo Fatorial Oculto de Markov, BLDA = Análise Discriminante Linear Bayesiana, DSLM = *Dynamic Stopping with Language Model*, DL = *Deep Learning*, * O estudo conduziu experimentos em dois conjuntos de dados.

Em suma, considerando os 18 artigos primários incluídos nesta pesquisa, os resultados sintetizados na Figura 3.3 apontam para três grandes classes de possíveis aplicações práticas no processamento de sinais biomédicos e aprendizado de máquina no contexto da doença ELA: diagnóstico (ou classificação), comunicação e previsão de sobrevida. Além da categorização das finalidades dos estudos em questão, a Figura 3.3 expressa ainda o quantitativo dos sinais biomédicos utilizados e as respectivas classes que os usaram. Quatro diferentes tipos de sinais foram identificados: EMG, EEG, GR e MRI.

Figura 3.3: Sumarização dos sinais utilizados e seus objetivos.



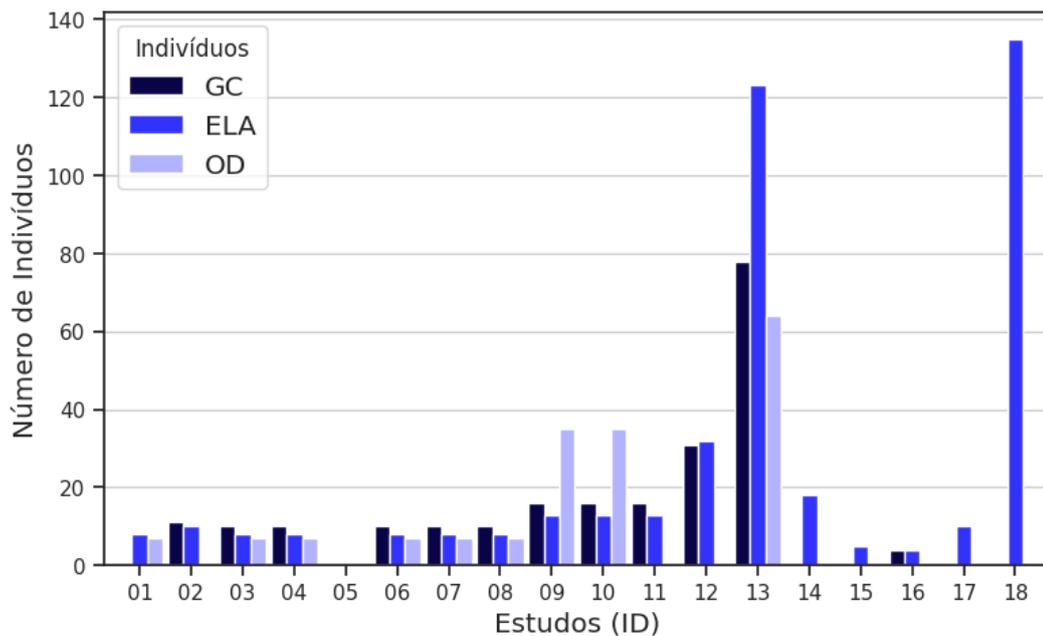
Dos estudos analisados, 44,44% deles concentram-se no processamento do sinal EMG, sendo o sinal biomédico mais utilizado (Figura 3.3) e, exclusivamente para a tarefa de classificação. Isso significa dizer que o sinal EMG foi utilizado como um potencial biomarcador para identificar (diagnóstico) indivíduos saudáveis (grupo de controle), ELA e, em alguns casos, outras doenças. Com o mesmo objetivo, especialmente para classificação, 16,67% dos estudos utilizam GR e 11,11% MRI. O sinal MRI também foi usado, em um único artigo, para a previsão de sobrevida dos indivíduos com ELA, o que representa 5,56%. Na categoria de comunicação, 22,22% dos estudos focam exclusivamente na abordagem através do processamento do sinal EEG, sendo o único sinal apresentado para este fim. Esta primeira análise geral dos estudos, identificando os objetivos dos artigos e os sinais utilizados, responde às questões de pesquisa 01 e 02.

Outras características de fundamental importância para esta pesquisa, extraídas dos 18 artigos primários incluídos, são sintetizadas na Tabela 3.5 para subsidiar a análise e as respostas às questões de pesquisa 03, 04 e 05. Independentemente das classes observadas, diagnóstico, comunicação ou previsão de sobrevida, todos os trabalhos utilizam

algoritmos de aprendizado de máquina. São diferentes modelos algorítmicos abordados e, de acordo com a análise da performance do algoritmo em relação as métricas de avaliação acurácia (Acc), especificidade (Spe) ou sensibilidade (Sen), o melhor ou o único modelo proposto de cada trabalho é apresentado na Tabela 3.5, bem como sua respectiva performance.

Como parte do processo de teste, validação e avaliação das abordagens propostas nos trabalhos, as técnicas algorítmicas foram aplicadas em um conjunto de dados de indivíduos, distribuídos em diferentes combinações entre grupos de controle (saudáveis), ELA e/ou miopatia ou outras doenças neurológicas. O quantitativo de indivíduos e tipo dos grupos participantes nos experimentos de cada trabalho são especificados na Tabela 3.5 e resumidos na Figura 3.4. Além disso, a Tabela 3.5 mostra ainda a origem do conjunto de dados (*Dataset*) e se são provenientes de repositórios públicos ou locais (privados).

Figura 3.4: Número de indivíduos utilizados nos estudos.



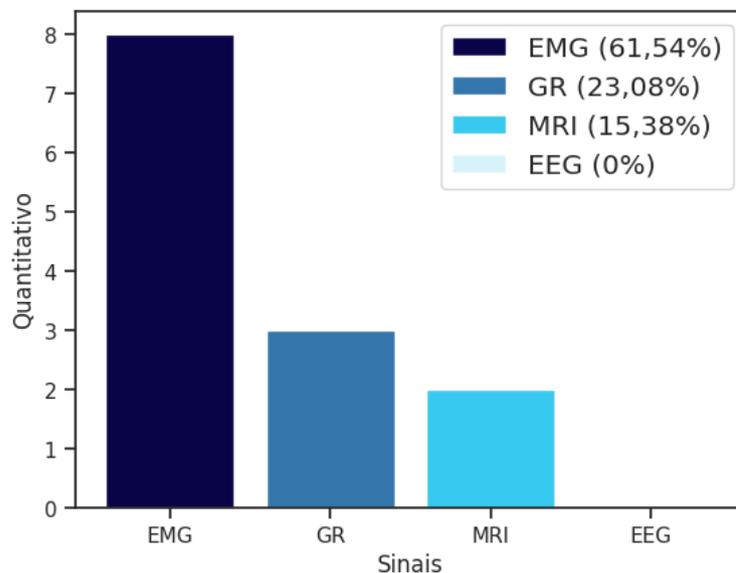
3.2.1 Estudos para o Diagnóstico

A classe de diagnóstico para pessoas com ELA é a mais numerosa e conta com o total de 72,22% dos estudos incluídos nesta revisão sistemática. Considerando apenas esta grande classe, a Figura 3.5 apresenta uma visão geral do quantitativo dos sinais biomédicos utilizados. Destaca-se o uso do sinal EMG, abordado em 61,54% dos estudos analisados para diagnóstico (Chatterjee et al. 2019, Zhang et al. 2014, Hazarika et al. 2019, Gokgoz e Subasi 2014, Ambikapathy et al. 2018, Doulah et al. 2014, Vallejo et al. 2018, Gokgoz e Subasi 2015). GR é utilizado em 23,08% dos estudos (Xia et al. 2015, Ren et al. 2017, Khorasani et al. 2016) e MRI em 15,38% (Welsh et al. 2013, Ferraro et al. 2017). O Sinal biomédico EEG não é utilizado para esta finalidade.

CAPÍTULO 3. SINAIS BIOMÉDICOS E APRENDIZADO DE MÁQUINA PARA ELA29

Nos estudos é evidenciado o foco no desenvolvimento de estratégias capazes de reduzir os ruídos dos sinais biomédicos coletados, cujo objetivo é encontrar melhores características de tal modo que potencialize a performance dos modelos algorítmicos e, conseqüentemente, que isto reflita na classificação dos indivíduos corretamente. No caso do sinal EMG, Gokgoz e Subasi (2014) mostram o ganho de performance na classificação de GC, ELA e OD em diferentes algoritmos de aprendizado de máquina, como KNN, RNA e SVM, após a utilização da técnica de remoção de ruídos Análise de Componentes Principais Multiescala (MSPCA) aliada à técnica de Classificação Única Múltipla (MUSIC), para a extração de características. Neste estudo, a performance mais satisfatória foi obtida pelo modelo SVM, com 92,55% de Acc.

Figura 3.5: Quantitativo de trabalhos por tipo de sinais biomédicos abordados na classe diagnóstica.



No ano seguinte, também abordando o sinal EMG, Gokgoz e Subasi (2015) propuseram outro estudo com foco na redução de ruídos e seleção de características. Neste estudo, os autores desenvolveram uma estrutura para remover ruídos, também utilizando a técnica MSPCA, com uma nova estratégia baseada em Transformada Wavelet Discreta (DWT) para a extração de características. Neste estudo, os autores realizaram testes experimentais, com e sem a estrutura de remoção de ruídos e extração de características, em três diferentes modelos algorítmicos baseados em Árvore de Decisão: CART, C4.5 e *Random Forest* (RF). Da mesma forma que em seu estudo anterior, Gokgoz e Subasi (2015) obtiveram resultados satisfatórios na classificação de GC, ELA e OD e demonstraram que o uso do método de remoção de ruídos MSPCA em conjunto com a estratégia de extração de características DWT melhorou a performance do modelo algorítmico RF, que obteve no melhor caso 96,67% de Acc utilizando o sinal EMG. Em ambos os estudos, três grupos distintos de indivíduos foram considerados: GC = 10, ELA = 8 e OD = 7.

Vallejo et al. (2018) adotaram a abordagem DWT para decompor os sinais EMG e

CAPÍTULO 3. SINAIS BIOMÉDICOS E APRENDIZADO DE MÁQUINA PARA ELA30

gerar um hiperespaço de características. Com esse hiperespaço bem definido, os autores selecionaram as características mais relevantes, através da técnica de entropia *fuzzy*, e serviram de entrada para um algoritmo de aprendizado de máquina baseado em Rede Neural Artificial (RNA). Para verificar e validar o método proposto, os autores utilizaram uma RNA *feedforward* com quatro camadas de sete neurônios cada, com exceção da camada de saída, e a função de ativação log-sigmoide. A RNA, atuando em conjunto com as abordagens baseadas em DWT e entropia *fuzzy*, obteve 98% de Acc no melhor caso. Os resultados indicam que o método proposto para extração e seleção de características melhoram o processo de classificação dos indivíduos em três diferentes grupos, incluindo GC, ELA e/ou miopatia. Nos experimentos, os autores incluíram dez indivíduos saudáveis, oito com ELA e sete com miopatia.

Em Doulah et al. (2014), a abordagem DWT é utilizada para decompor, filtrar e extrair características relevantes de sinais EMG pré-processados. O pré-processamento do EMG abrangeu dois métodos propostos para a classificação dos indivíduos em GC e ELA e/ou miopatia. No método EMG com DWT, as características são extraídas após uma análise sequencial de *frame a frame*. No segundo método, um conjunto de características é selecionado a partir do procedimento chamado Potencial de Ação da Unidade Motora Dominante (D-MUAP) e, baseado na decomposição DWT, as principais características para classificação são extraídas. Doulah et al. (2014) verificaram ambos os métodos propostos em um classificador KNN. Os autores avaliaram a performance do algoritmo com variadas configurações paramétricas e em três conjuntos de dados correspondentes a indivíduos saudáveis, com ELA e outras doenças (10, 8 e 7 indivíduos, respectivamente). Ambos os métodos apresentaram resultados significativos na tarefa de classificação de duas ou três classes (GC, ELA e OD). O método D-MUAP pode ser destacado por seu desempenho consistente, atingindo 98,8% de Acc na tarefa de identificar três classes.

Chatterjee et al. (2019) apresentam uma generalização da Transformada de Stockwell (ST), chamada de Janela Modificada ST (MWST), para pré-processar os sinais EMG e gerar uma matriz (ou, um plano de tempo-frequência) de características mais representativa. O conjunto de parâmetros da MWST (α , β , γ , e δ), os quais controlam a forma da janela ou a concentração de energia no plano tempo-frequência, são definidos de forma otimizada através da metaheurística baseada em enxame de partículas. Após a aplicação da técnica MWST nos sinais EMG de indivíduos com ELA e miopatia, quatro características são extraídas da matriz para servir de entrada para quatro modelos de aprendizado de máquina: SVM, kNN, Naïve Bayes e Árvore de Decisão. A abordagem proposta pelos autores apresentou resultados significativamente melhores na classificação de indivíduos entre ELA e miopatia quando comparada ao método ST convencional. O modelo SVM apresentou a performance mais satisfatória, com 98.28% de Acc.

Quatro métodos distintos para a extração de características de sinais EMG no domínio tempo-frequência foram abordados em Ambikapathy et al. (2018): ST, Transformada Síncronas, Distribuição Wigner-Ville e Transformada de Fourier de Tempo Curto. Os métodos aplicados aos sinais geraram imagens e, usando a técnica de Matriz de Coocorrência de Nível de Cinza, cada método extraiu 20 características, originando um conjunto de 80 características. Para otimizar a performance de uma Rede Neural Artificial (com uma camada oculta de 10 neurônios e função de ativação tan sigmoid), um subconjunto

contendo a combinação mais adequada de 15 características (de um conjunto finito de possibilidades) foi definido através de um Algoritmo Genético (AG), implementado com o classificador KNN para realizar a avaliação da *fitness* das diferentes combinações (população). A estratégia proposta por Ambikapathy et al. (2018) para auxiliar no diagnóstico obteve resultados promissores e com significância estatística no processo de classificação de indivíduos entre saudáveis e ELA (86,6% de Acc, 86,6% de Spe, 86,6% de Sen), entre saudáveis, ELA e/ou miopatia (82,2% Acc, 81,89% Sen, 91,31% Spe), e entre ELA e miopatia (96,6% Acc, 93,7% Sen, 100% Spe).

Zhang et al. (2014) desenvolveram um método para caracterizar padrões em sinais EMG de superfície com baseado em três marcadores/características para suporte ao diagnóstico: índice de agrupamento; a curtose do histograma de amplitude do sinal EMG; e a curtose da expansão da taxa de cruzamento do sinal EMG. Além disso, um algoritmo de aprendizado de máquina, chamado Análise Discriminante Linear (LDA, do Inglês *Linear Discriminant Analysis*), foi aplicado no processo de classificação de indivíduos entre saudáveis e ELA. O algoritmo implementado recebeu como entrada de dados um vetor de características, resultado da concatenação dos vetores construídos pelos respectivos marcadores mencionados. Os experimentos realizados neste estudo basearam-se na análise de dados de 10 pessoas com ELA e 11 saudáveis. Cada um dos três marcadores apresentaram resultados promissores. Quando usados de forma sinérgica em uma abordagem mais robusta que os combina com o algoritmo LDA, apresenta resultados ainda mais significativos, com Spe de 100% e Sen de 90%. Tal aspecto indica perspectivas favoráveis para ser utilizado no apoio ao diagnóstico da ELA.

Hazarika et al. (2019) propuseram um novo sistema de inferência e avaliação de processos com uma estrutura robusta de pré-processar e extrair características de sinais EMG. Essa estrutura é composta por procedimentos que, inicialmente, avaliam as características do sinal EMG através de uma abordagem envolvendo estratégias de particionamento dos dados de entrada, chamada de multivisualização direta, e de decomposição pela DWT, e em seguida aplicam o método de Análise de Correlação Discriminante de Multivisão Multidomínio (mmDCA). O mmDCA analisa a correlação das características, verifica redundância, elimina características irrelevantes e sincroniza em um único vetor as características provenientes das partições. Este vetor é incorporado em diferentes algoritmos de aprendizado de máquina, como KNN, RNA (rede *feedforward back-propagation* com duas camadas ocultas de 8 neurônios cada), Classificador Linear Baseado em Densidade Normal (LDC), Classificador Quadrático (QDC) e SVM. A performance de cada algoritmo foi avaliada positivamente na classificação de pessoas saudáveis, com ELA e miopatia em dois conjuntos de dados (de sinais EMG) em tempo real. Em ambos conjuntos de dados o modelo QDC apresentou os resultados mais robustos: 99,03% de Acc no primeiro (10 pessoas saudáveis, 8 com ELA e 7 com miopatia) e 100% de Acc no segundo (4 pessoas saudáveis, 4 com ELA e 4 com miopatia).

Os estudos mencionados até aqui mostram o potencial e a importância do processamento do sinal biomédico EMG. Como resultado desta tarefa, há fornecimento de um conjunto de características ou dados consistentes e elementares para o processo de aprendizado dos algoritmos inteligentes e, conseqüentemente, para o auxílio no diagnóstico de pessoas com ELA. Entretanto, como alternativa aos sinais EMG para o diagnóstico, outros

estudos investigaram a performance dos modelos de aprendizado de máquina com estratégias que processam dados provenientes dos sinais biomédicos GR (Xia et al. 2015, Ren et al. 2017, Khorasani et al. 2016).

Xia et al. (2015) realizaram a extração de características de sinais GR em cinco registros de séries temporais com base em análise estatística. Características como valor médio, desvio padrão, valores máximo e mínimo, assimetria e curtose foram definidas. Os autores também propuseram a extração de mais três características através dos métodos estatísticos: complexidade de Lempel-Ziv, entropia fuzzy e do operador de energia Teager-Kaiser. Xia et al. (2015) executaram ainda uma abordagem de seleção de características baseada em três procedimentos. O primeiro procedimento analisou estatisticamente as características, o segundo consistiu na avaliação de desempenho dos modelos algorítmicos de classificação e, por último, um algoritmo de otimização chamado hill climbing foi utilizado para encontrar e definir um subconjunto ótimo de características.

Após seleção do subconjunto ótimo de características, uma série de experimentos com algoritmos de aprendizado de máquina foi realizada, considerando quatro classificadores diferentes: SVM, RF, RNA (*feedforward* com função de ativação sigmoid) e KNN. O conjunto de dados usado para os experimentos incluiu dados de 16 pessoas saudáveis, 13 com ELA e 35 com outras doenças. Todos os classificadores apresentaram performances satisfatórias na classificação binária entre pessoas saudáveis e com ELA, bem como pessoas saudáveis outras doenças neurodegenerativas além de ELA. A melhor performance foi obtida pelo algoritmo SVM, com 96,55% de Acc.

Ren et al. (2017) também apresentam uma estratégia para extrair e selecionar características de sinais GR em cinco registros de séries temporais. Para isso, os autores propuseram uma abordagem que utiliza o método de Decomposição de Modo Empírico (EMD) para extrair características do particionamento das séries temporais do sinal GR. Assim, produzindo seis componentes - dos quais cinco são usados e um descartado - que foram prontamente submetidos à análise estatística, pelo método do Coeficiente de Concordância de Kendall, para mensurar a significância e a relação das características. Em seguida, é realizado um cálculo nas amplitudes através do procedimento de razão para mudança de energia e, finalmente, a técnica de redução de dimensionalidade PCA (do Inglês *Principal Component Analysis*) é aplicada para definir o conjunto final de características.

A estratégia proposta pelos autores foi avaliada em cinco classificadores: Naïve Bayes, SVM, RF, *Multilayer Perceptron* (MLP) e Regressão Logística Simples. Diferentemente da avaliação da performance dos modelos de aprendizado de máquina considerada na Tabela 3.5, neste estudo a área sob a curva ROC (AUC) é usada para medir a performance. O resultado mais significativo para a tarefa de classificação de indivíduos entre saudáveis e com ELA foi apresentado pelo modelo MLP *feedforward*, com valor AUC = 0.934. O valor médio da AUC, considerando os cinco classificadores, foi de 0.898 e, no geral, a abordagem mostrou resultados promissores.

Diferentemente das outras pesquisas desenvolvidas para o auxílio ao diagnóstico, em Khorasani et al. (2016) uma generalização do Modelo Oculto de Markov (HMM), chamado de Modelo Fatorial Oculto de Markov (FHMM), foi proposta para reconhecer padrões e classificar indivíduos em saudáveis ou com ELA a partir de séries temporais de sinais GR. Depois de pré-processar o sinal GR para remoção de *outliers*, o sinal é seg-

mentado em 50 séries temporais e algumas características são extraídas, como média ou variância, para fornecerem os dados de entrada do FHMM. O método foi avaliado com dados de 16 indivíduos saudáveis e 13 com ELA. Além disso, seu desempenho foi comparado ao modelo tradicional HMM e ao algoritmo *Least-Squares SVM* (LS-SVM) com kernels gaussianos. Com um valor de 93,1% de Acc, o modelo de classificação FHMM apresentou o melhor desempenho na tarefa de reconhecimento de pessoas saudáveis ou com ELA.

Pouco exploradas no contexto do diagnóstico de pessoas com ELA, as estratégias baseadas em neuroimagens e aprendizado de máquina ainda são desafiadoras. Em 2013, Welsh et al. (2013) propuseram uma abordagem baseada em SVM para identificar pessoas saudáveis ou com ELA através da análise de dados de ressonância magnética funcional (fMRI). Os dados de séries temporais da fMRI foram pré-processados e submetidos a um conjunto de procedimentos robustos de extração e seleção de características, incluindo a execução consecutiva de estratégias como PCA e Análise de Componentes Independentes (ICA), bem como a criação de mapas (ou vetores) de coeficientes de correlação de diferentes regiões do cérebro. Na sequência, os dados foram fornecidos ao algoritmo SVM de kernel linear, e sua performance foi avaliada. Para os experimentos, foram fornecidos dados de 32 pessoas diagnosticadas com ELA e 31 saudáveis. Welsh et al. (2013) classificou os resultados como modestos (71,5% de Acc) para o reconhecimento de pessoas com a doença ELA utilizando fMRI em estado de repouso.

Alguns anos após o estudo de Welsh et al. (2013), utilizando fMRI, Ferraro et al. (2017) desenvolveram um método para classificar indivíduos com doenças do neurônio motor, incluindo a ELA, baseado em ressonância magnética estrutural multimodal e um algoritmo de aprendizado de máquina. Os dados da MRI foram divididos em regiões de interesse distintas e submetidas a análise estatística através de softwares da literatura. A avaliação da abordagem proposta para o diagnóstico foi realizada em um modelo algorítmico RF. A performance do algoritmo na classificação de indivíduos entre saudáveis e com ELA, especificamente com o modelo de MRI combinado, foi expressiva. Para os experimentos, foram utilizados dados de 78 indivíduos saudáveis, 123 com ELA e 64 com outras doenças. O modelo desenvolvido apresentou Acc de 91% quanto à classificação de pessoas saudáveis e com ELA. Os resultados de ambos os estudos (Welsh et al. 2013, Ferraro et al. 2017), no geral, apontam que somente os dados de estratégias baseadas em MRI não são suficientes para obter uma boa performance, sendo necessário considerar outros tipos de dados da pessoa. Os estudos baseados em neuroimagens e aprendizado de máquina para ELA ainda são poucos. Entretanto, os resultados são relevantes e mostram potencial para novas pesquisas e para o desenvolvimento de um sistema capaz de auxiliar no diagnóstico da ELA.

3.2.2 Estudos para a Comunicação

Devido a degeneração progressiva dos neurônios motores superiores e/ou inferiores, às pessoas diagnosticadas com ELA perdem a capacidade de falar e interagir com o ambiente. Os recursos tecnológicos desenvolvidos para intermediar e promover a comunicação são cruciais para garantir a qualidade de vida e bem-estar dessas pessoas. Apro-

ximadamente 22% dos estudos incluídos nesta revisão sistemática desenvolveram recursos que promovem melhorias no processo comunicativo da pessoa com ELA, exclusivamente utilizando o sinal biomédico EEG (Sorbello et al. 2018, Liu et al. 2017, Mainsah et al. 2015, Miao et al. 2020).

No estudo conduzido por Sorbello et al. (2018), foi proposta uma arquitetura que, através de um sistema baseado em Interface Cérebro-Computador (BCI), fosse possível controlar um robô humanoide e promover o mínimo de autonomia à pessoa com ELA. De forma geral, a estrutura do sistema, chamada de *Brain Computer Robotic Interface* (BCRI), foi composta por um sistema BCI, dispositivos de EEG e *eye-tracking*, e uma rede local para fazer a conexão entre os sistemas BCI e robótico. O algoritmo LDA foi utilizado, após o pré-processamento e extração de características do sinal EEG, para classificar e traduzir corretamente a ação do usuário em comandos de controle do robô humanoide. Os autores avaliaram a proposta através da realização de experimentos em quatro pessoas saudáveis e quatro com ELA. Os resultados foram satisfatórios e a arquitetura proposta para promover a comunicação e autonomia de pessoas com ELA foi validada após todos os participantes dos experimentos conseguirem controlar o robô humanoide.

Liu et al. (2017) desenvolveram uma abordagem, aplicando os conceitos de Dimensão Fractal (FD) e do critério de Fisher, para otimizar a seleção de canais de EEG e a caracterização dos dados obtidos do sinal. Com isso, os autores visaram melhorar a capacidade de classificação de um algoritmo de aprendizado de máquina em um sistema de BCI para pessoas com ELA. Os métodos Grassberger-Procaccia (GPF) e Higuchi (HFD) foram implementados para estimar a Dimensão Fractal e as principais características, provenientes de 30 canais de EEG, foram extraídas e concatenadas em um único vetor para servir a dois modelos algorítmicos: KNN e LDA. Após testes realizados em cinco pessoas com ELA, os resultados mostraram-se satisfatórios e o método GPF superou o HFD. As performances dos algoritmos KNN e LDA foram semelhantes, com 95,25% de Acc. Os resultados foram significativos quando comparados com o desempenho dos modelos utilizando os dados de entrada contendo os 30 canais de EEG.

A existência de uma interface simples com uma taxa de transferência de informações precisa e rápida é essencial para manter a eficiência da comunicação em um sistema BCI baseado em sinais de EEG para pessoas com ELA. Com foco nessa questão, Mainsah et al. (2015) desenvolveram um classificador Bayesiano (de parada antecipada baseado em dados), chamado de DS, para otimizar o processo de seleção de caracteres de uma BCI baseada em ERP (do inglês, *event-related potentials*) do tipo P300 *spellers*. Além disso, uma variação do DS (DSL) foi proposta com a aplicação de modelagem estatística por inferência Bayesiana para a previsibilidade da linguagem. Características correlacionadas com o interesse do usuário foram extraídas do sinal EEG para treinar o classificador LDA. No estudo, os denominados testes online foram realizados em 10 pessoas com a doença ELA. Ambos os algoritmos DS e DSL mostraram-se eficientes na tarefa de minimizar o tempo de seleção de caracteres, obtendo valor médio de acurácia em 75,40% e 76,39%, respectivamente. Não houve diferença estatística entre os dois algoritmos.

Na mesma perspectiva de Mainsah et al. (2015), Miao et al. (2020) propuseram uma abordagem de BCI baseada em ERP usando como estratégia um paradigma de organização ortográfica periférica com possibilidade de *feedback* no centro do *display*. Os sinais

de EEG foram gravados e analisados através de um software da plataforma BCI, já existente na literatura. As características foram extraídas de dados obtidos de 16 eletrodos *offline* para treinar o classificador LDA Bayesiana (BLDA) que, posteriormente, foi utilizado nos testes do sistema online. A abordagem proposta foi avaliada em relação ao paradigma ortográfico convencional, em matriz. Os experimentos foram realizados em 18 pessoas com ELA e os resultados apresentados não mostram diferença significativa do algoritmo BLDA em relação à abordagem convencional, mesmo obtendo 90% de Acc. Entretanto, as pessoas com ELA foram capazes de utilizar o sistema de forma eficaz.

3.2.3 Estudos para a Predição de Sobrevida

A predição de sobrevida das pessoas com ELA é definida empiricamente baseando-se, geralmente, na análise dos dados clínicos. Apenas um dos estudos incluídos nesta revisão dedica-se à predição de sobrevida. No estudo desenvolvido por van der Burgh et al. (2017), foi proposto um modelo para predição de sobrevida (curta, média ou longa) de pessoas com ELA combinando dados clínicos, neuroimagens e *deep learning* (DL). Quatro cenários foram definidos para a aplicação do algoritmo DL. O primeiro, foi baseado apenas em dados clínicos. O segundo e terceiro cenários utilizaram imagens de MRI, com conectividade estrutural e morfológica, respectivamente. O quarto cenário foi construído a partir da combinação dos cenários anteriores. Para cada cenário, um modelo foi implementado e sua performance avaliada por meio de um banco de dados que continha dados de 135 pessoas com ELA. O modelo que combinou dados clínicos e de ressonância magnética obteve o melhor desempenho (com 84,4% de Acc) e apresenta-se como uma estratégia viável para predição de sobrevida de pessoas com ELA. Os demais modelos apresentaram resultados intermediários, embora indicassem abordagens promissoras.

3.3 Discussões

Nesta Revisão Sistemática de Literatura foi investigada abordagens baseadas em inteligência computacional que atuam de forma sinérgica e complementar ao processamento de sinais biomédicos no contexto específico da doença ELA. Um conjunto de 18 artigos foi incluído e revisado, e três grandes classes de aplicações foram encontradas: auxílio ao diagnóstico, facilitação da comunicação e previsão de sobrevivência. Foram identificados e quantificados os modelos algorítmicos mais adequados e os respectivos sinais biomédicos responsáveis pelo fornecimento dos dados (ver Figura 3.6).

Baseado na análise dos 13 artigos para auxílio ao diagnóstico de pacientes com ELA, independentemente do sinal biomédico ou algoritmo de aprendizado de máquina utilizado, é possível definir um esquema (*pipeline*) metodológico padrão e comum a todos os estudos, o qual de forma genérica pode ser visto na Figura 3.7. Com exceção de Khorasani et al. (2016), que investigou um novo algoritmo de classificação, os estudos sugerem abordagens ou métodos para o processo de tratamento de dados que podem aprimorar a etapa de treinamento e, conseqüentemente, a etapa de classificação. Esse processo de tratamento de dados, que inclui as fases de extração e seleção de atributos, por exemplo, é importante para eliminar ruídos, redundâncias e reduzir a dimensionalidade dos dados,

CAPÍTULO 3. SINAIS BIOMÉDICOS E APRENDIZADO DE MÁQUINA PARA ELA36

além de maximizar o desempenho dos algoritmos por meio do fornecimento de dados refinados e consistentes (Phinyomark et al. 2012). Os diversos modelos de aprendizado de máquina implementados foram apresentados como técnicas de avaliação e validação das propostas dos estudos. No entanto, foram técnicas elementares no processo de diagnóstico que estão presentes em todos os artigos.

Figura 3.6: Quantitativo dos melhores modelos algorítmicos e os respectivos sinais biomédicos.

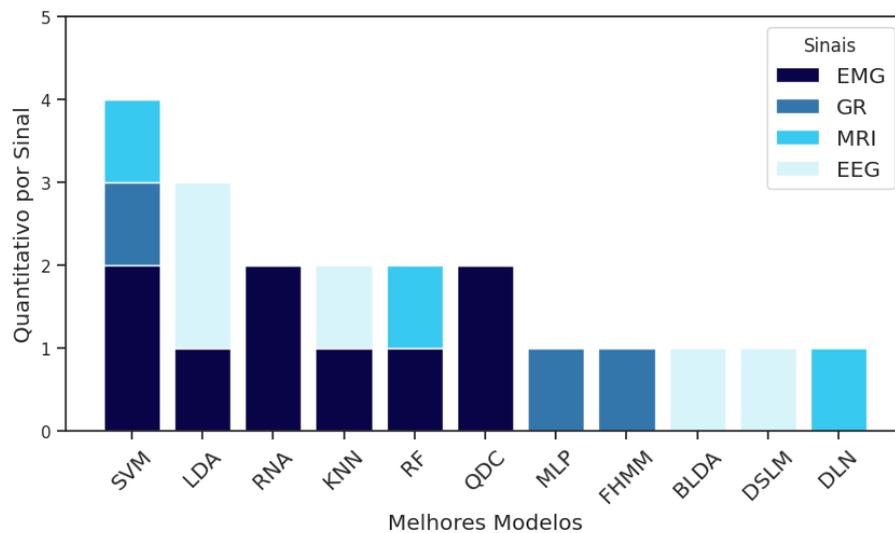
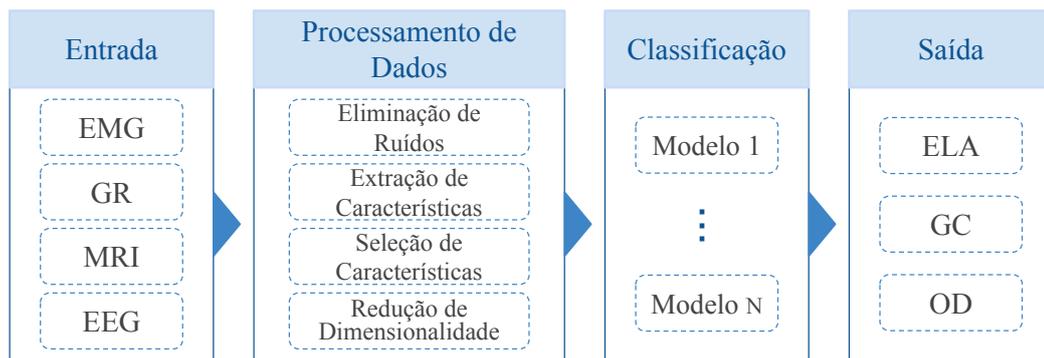


Figura 3.7: Pipeline genérico: esquema generalizado para resolver problemas de classificação.



Os estudos (Chatterjee et al. 2019, Hazarika et al. 2019, Gokgoz e Subasi 2014, Ambikapathy et al. 2018, Doulah et al. 2014, Vallejo et al. 2018, Gokgoz e Subasi 2015, Xia et al. 2015, Ren et al. 2017, Ferraro et al. 2017) para o auxílio ao diagnóstico, além da ELA, também testaram abordagens para classificação binária ou multirrotulo considerando outras condições neurológicas, como miopatia, doença de Parkinson, doença de Huntington, predominantemente doença do neurônio motor superior e distúrbios que mimetizam a ELA.

Em relação aos quatro artigos pertencentes à classe de comunicação, duas categorias distintas podem ser observadas nos estudos analisados. A primeira é a do estudo realizado por Sorbello et al. (2018), que visa complementar e adaptar um sistema BCI com um robô humanóide para fornecer não apenas comunicação, mas também um mínimo de autonomia. Na segunda categoria identificada, os estudos (Liu et al. 2017, Mainsah et al. 2015, Miao et al. 2020) sugerem abordagens alternativas que incluem aprendizado de máquina para otimizar o tempo de seleção de caracteres em um sistema BCI. Essas abordagens vão desde a otimização de eletrodos de EEG até a personalização inteligente da interface. A importância dos sistemas BCI na promoção da comunicação é evidente. Esses sistemas são amplamente utilizados em pesquisas para estabelecer uma via de comunicação entre o cérebro humano e dispositivos externos, reconhecendo mudanças voluntárias na atividade cerebral de seus usuários (Grimann et al. 2009, Herff et al. 2015, Anumanchipalli et al. 2019, Cooney et al. 2019, Dash, Ferrari e Wang 2020, Dash, Ferrari, Hernandez, Heitzman, Austin e Wang 2020).

Apesar das pesquisas focarem no desenvolvimento de sistemas BCI, existem limitações quanto ao seu uso doméstico. Uma das principais razões pelas quais os BCI ainda não foram introduzidos no ambiente doméstico, é a relação do tempo de seleção de caracteres, que ainda são considerados lentos e imprecisos, em comparação com outras abordagens sem a utilização de sinais originados do cérebro, e da necessidade de eletrodos conectados à cabeça do paciente (Tamura et al. 2016, Mainsah et al. 2015, Xia et al. 2015). Outras abordagens de sistemas de Interação Humano-Computador que não envolvem necessariamente sinais cerebrais por EEG podem ser vistas em Pinheiro et al. (2011), Hori et al. (2004), Fathi e Abdali-Mohammadi (2015), Harezlak e Kasprowski (2018), Villanueva et al. (2009), Królak e Strumiłło (2012), Zhao et al. (2015), Liu et al. (2019) e Aharonson et al. (2020).

O único estudo de previsão de sobrevida com pacientes com ELA analisa o quão desafiador é desenvolver sistemas para tal finalidade. O estudo (van der Burgh et al. 2017) indica que a ressonância magnética e a técnica DL são promissoras para a previsão de sobrevida e sugere uma exploração mais significativa no campo da neuroimagem. Além disso, a pesquisa revela a importância e os benefícios dos dados clínicos dos pacientes no processo de prever a sobrevida nos três níveis de ELA. Esta observação, aliada à análise feita até agora, revela tanto a ausência como a possibilidade de utilização de dados clínicos para o diagnóstico. Correlacionado ao aspecto de sobrevivência, estudos recentes indicam que é possível aplicar abordagens de aprendizado de máquina com biomarcadores digitais usando o sinal de fala para monitorar a progressão da ELA (Wang et al. 2018), incluindo aplicações para classificação automática da Escala Funcional de Esclerose Lateral Amiotrófica (ALSFRS) (Wisler et al. 2019, Cedarbaum et al. 1999).

Em relação aos algoritmos de aprendizado de máquina, observa-se que eles são especificamente supervisionados em todos os estudos. O tipo de sinal biomédico varia apenas nos estudos de diagnóstico, sendo a EMG o sinal mais utilizado, seguida da GR e da MRI. O sinal EEG é aplicado apenas para aplicações de habilitação de comunicação. O sinal biomédico baseado em ressonância magnética é usado tanto em aplicações de diagnóstico quanto de previsão de sobrevivência. Schuster et al. (2016) afirmam que os biomarcadores baseados em ressonância magnética são raramente usados para auxiliar na identificação

da ELA. Esta observação é complementada pelos resultados apresentados nesta SLR, que também relata o número limitado de estudos baseados em neuroimagem voltados para aplicações de suporte diagnóstico e previsão de sobrevida da doença de ELA, apesar do potencial mencionado por van der Burgh et al. (2017). Além desses sinais biomédicos mencionados até agora, estudos mostram a viabilidade do uso do biosinal da fala para o diagnóstico precoce da ELA, conforme indicam Wang et al. (2016), Suhas et al. (2020), An et al. (2018), Vieira et al. (2019) e Wisler et al. (n.d.), e acompanhamento de alterações em indivíduos com ELA bulbar (Stegmann et al. 2020).

Os 18 estudos realizaram testes experimentais com conjuntos de dados de indivíduos saudáveis e indivíduos com ELA ou outras doenças neurológicas. 50% dos estudos usaram *datasets* locais ou próprios. Os outros 50% das investigações coletaram dados de repositórios públicos online. Em alguns casos, como os de diagnóstico e comunicação, exceto no estudo realizado por Ferraro et al. (2017), é evidente a limitação no número de pacientes com ELA (ver Figura 3.4). Esses resultados sugerem que ainda é desafiador desenvolver e validar um estudo robusto com um número mais considerável de sujeitos com ELA ou em ambiente ambulatorial.

3.4 Conclusões

Este artigo apresenta um protocolo RSL para investigar estudos relevantes dos últimos dez anos (2009-2019) que abordam técnicas de aprendizado de máquina e processamento de sinais biomédicos que contribuem para o avanço de pesquisas no contexto da ELA. Com base em 18 estudos primários, os resultados apresentam estratégias para minimizar problemas e/ou promover meios de suporte diagnóstico, comunicação e predição de sobrevida. Considerando os estudos analisados, 88,89% deles relatam a importância do tratamento de sinais biomédicos para fornecer dados robustos e consistentes para modelos algorítmicos de aprendizado de máquina.

Além disso, pode-se observar que há predominância no tipo de sinais biomédicos utilizados pelos estudos nas categorias de comunicação e predição de sobrevida, sendo exclusivamente e respectivamente os sinais de EEG e imagens de MRI. Para a classe de diagnóstico, em particular, são reportados três tipos de dados brutos, nomeadamente EMG (61,54%), GR (23,08%) e MRI (15,38%). Em relação aos modelos algorítmicos de ML e analisando as melhores performances, o SVM foi o mais utilizado, seguido das técnicas LDA e RNA. Apesar dos 18 estudos utilizarem aprendizado de máquina, com exceção de um estudo que propôs um novo algoritmo, o tema não é o principal foco. De forma geral, limitada aos objetivos desta revisão, a literatura sugere e dedica-se ao tratamento dos sinais biomédicos.

Os estudos são promissores, mas ainda há oportunidades para serem exploradas no futuro. Para diagnóstico, por exemplo, estudos desenvolvidos podem ser testados em ambulatórios médicos para assistência prática em casos ainda não confirmados de ELA ou em estágios iniciais da doença. Além disso, incluindo também o campo de predição de sobrevida de pessoas com ELA, a utilização do conceito de *big data* em conjunto aos dados clínicos do paciente pode contribuir no resultado final e é passível de investigação. Em relação à comunicação, há questões em aberto quanto ao uso de BCI no ambiente

CAPÍTULO 3. SINAIS BIOMÉDICOS E APRENDIZADO DE MÁQUINA PARA ELA39

domiciliar, tendo em vista aspectos do ambiente, custos, interfaces eficientes que evitam fadiga e otimização dos eletrodos dos sinais EEG.

Capítulo 4

Comunicação Alternativa para ELA

Este capítulo apresenta o artigo intitulado “*Digital Alternative Communication for Individuals with Amyotrophic Lateral Sclerosis: what we have*”, em sua versão publicada no *Journal of Clinical Medicine*, na seção de neurologia clínica. O artigo reúne, avalia e fornece evidências, bem como lacunas, relacionadas à abordagens de Interação Humano-Computador baseadas em imagens dos olhos para a comunicação aumentativa e alternativa de pessoas com Esclerose Lateral Amiotrófica. A investigação e as buscas por estudos primários na literatura foram baseadas em recursos metodológicos de uma Revisão Sistemática da Literatura (Kitchenham et al. 2016, Keele 2007, Kitchenham 2004, Brereton et al. 2007, Snyder 2019).

Fernandes, Felipe, Ingridy Barbalho, Arnaldo Bispo Júnior, Luca Alves, Danilo Nagem, Hertz Lins, Ernano Arrais Júnior, Karilany D. Coutinho, Antônio H. F. Morais, João Paulo Q. Santos, Guilherme Medeiros Machado, Jorge Henriques, César Teixeira, Mário E. T. Dourado Júnior, Ana R. R. Lindquist e Ricardo A. M. Valentim (2023), ‘Digital alternative communication for individuals with amyotrophic lateral sclerosis: What we have’, *Journal of Clinical Medicine* **12**(16), 5235. URL: <https://www.mdpi.com/2077-0383/12/16/5235>

4.1 Protocolo de Busca

Esta pesquisa foi desenvolvida com base nas diretrizes de revisão sistemática proposta por Kitchenham (2004) e guiada pelo *checklist* PRISMA (do inglês, *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) (Page et al. 2021). O protocolo para condução desta revisão foi registrado no banco de dados internacional de revisões sistemáticas PROSPERO (registro no. CRD42021230721) (Fernandes e Barbalho 2021). Inicialmente, e como parte fundamental do protocolo, foram elaboradas cinco Questões de Pesquisa (QP) (ver Tabela 4.1 a seguir).

Tabela 4.1: Questões de Pesquisa.

QP	Descrição
01	Qual estratégia é usada para a Interação Humano-Computador baseada em imagens dos olhos?
02	Qual técnica computacional é usada para processar e classificar as imagens dos olhos (visão computacional ou aprendizado de máquina, por exemplo)?
03	Qual é a performance das técnicas computacionais exploradas (avaliada através da acurácia, precisão, sensibilidade, especificidade, erro)?
04	Qual é o suporte de hardware para a aquisição das imagens?
05	Qual é o perfil do grupo de indivíduos submetido aos testes experimentais do estudo (grupo de controle, ELA ou outras doenças)?

O processo de identificação dos estudos primários relacionados ao objeto de investigação desta Revisão Sistemática da Literatura foi constituído por buscas em seis repositórios: *Science Direct*, *Web of Science*, *Springer*, *IEEE Xplore*, *ACM Digital Library*, and *PubMed*. As buscas em todas as bases foram realizadas em 18 novembro de 2021. Com exceção do repositório *PubMed*, duas *strings* de busca (STR01 e STR02) foram utilizadas nas buscas. Especificamente para o repositório *PubMed*, uma terceira *string* de busca (STR03) foi considerada, a qual foi definida a partir do vocabulário do *Medical Subject Headings* (MeSH) *thesaurus*. As *strings* de busca são apresentadas a seguir:

- STR01: (eye) AND (track OR gaze OR blink OR localization) AND (camera OR webcam) AND (“amyotrophic lateral sclerosis” OR als);
- STR02: (eye) AND (track OR gaze OR blink OR localization) AND (camera OR webcam) AND (“neuromuscular disease” OR “motor neuron disease”);
- STR03: ((((((Eye Tracking Technology) OR (Eye-Tracking Technologies) OR (Technology, Eye-Tracking) OR (Eyetracking Technology) OR (Eyetracking Technologies) OR (Gaze-Tracking Technology) OR (Gaze Tracking Technology) OR (Gaze-Tracking Technologies) OR (Eye-Tracking System) OR (Eye Tracking System) OR (Eye-Tracking Systems) OR (Eyetracking System) OR (Eyetracking Systems) OR (Eye Movement Data Analysis) OR (Gaze-Tracking) OR (Gaze Tracking) OR (Gaze-Tracking System) OR (Gaze Tracking System) OR (Gaze-Tracking Systems) OR (Gazetracking System) OR (Gazetracking Systems) OR (Eye-Tracking) OR (Eye Tracking)) OR ((Eye Movement) OR (Movement, Eye) OR (Movements, Eye))) OR ((Eye Movement Measurement) OR (Measurement, Eye Movement) OR (Measurements, Eye Movement))) OR ((Focusing, Ocular) OR (Ocular Focusing) OR (Ocular Fixation) OR (Eye Gaze) OR (Eye Gazes) OR (Gaze, Eye) OR (Gazes, Eye))) OR ((Saccade) OR (Saccadic Eye Movements) OR (Eye Movement, Saccadic) OR (Eye Movements, Saccadic) OR (Movement, Saccadic Eye) OR (Movements, Saccadic Eye) OR (Saccadic Eye Movement) OR (Pursuit, Saccadic) OR (Pursuits, Saccadic) OR (Saccadic Pursuit) OR (Saccadic Pursuits)))

AND ((Sclerosis, Amyotrophic Lateral) OR (Gehrig's Disease) OR (Gehrig Disease) OR (Gehrigs Disease) OR (Charcot Disease) OR (Motor Neuron Disease, Amyotrophic Lateral Sclerosis) OR (Lou Gehrig's Disease) OR (Lou-Gehrigs Disease) OR (Disease, Lou-Gehrigs) OR (ALS - Amyotrophic Lateral Sclerosis) OR (ALS Amyotrophic Lateral Sclerosis) OR (Lou Gehrig Disease) OR (Amyotrophic Lateral Sclerosis, Guam Form) OR (Amyotrophic Lateral Sclerosis-Parkinsonism-Dementia Complex 1) OR (Amyotrophic Lateral Sclerosis Parkinsonism Dementia Complex 1) OR (Guam Form of Amyotrophic Lateral Sclerosis) OR (Guam Disease) OR (Disease, Guam) OR (Amyotrophic Lateral Sclerosis, Parkinsonism-Dementia Complex of Guam) OR (Amyotrophic Lateral Sclerosis, Parkinsonism Dementia Complex of Guam) OR (Amyotrophic Lateral Sclerosis With Dementia) OR (Dementia With Amyotrophic Lateral Sclerosis)).

Após a identificação e definição do conjunto inicial de registros, o processo de triagem foi executado para selecionar um subconjunto de estudos primários elegíveis. Este processo foi organizado e executado a partir da aplicação de três procedimentos elementares: (i) Critério de Inclusão – CI; (ii) Critério de Exclusão – CE; e, (iii) Critério de Avaliação de Qualidade – CAQ.

No primeiro procedimento (i), um subconjunto de estudos primários foi definido a partir dos Critérios de Inclusão (Tabela 4.2), aplicados através dos filtros disponibilizados nos repositórios. Na etapa posterior (ii), uma triagem orientada pelos Critérios de Exclusão (Tabela 4.2) e com base na leitura do título, resumo e palavras-chave foi realizada no subconjunto de estudos primários. O *Rayyan* (Ouzzani et al. 2016), aplicativo *web* para revisões sistemáticas, auxiliou na realização desta etapa (ii).

Tabela 4.2: Critérios de Inclusão e Exclusão

ID	Critério de Inclusão	Critério de Exclusão
01	Artigos publicados entre 2010 e 18/11/2021.	Artigos duplicados.
02	Artigos de pesquisa originais e completos publicados em periódicos ou conferências.	Artigos de revisão.
03	Artigos das áreas de tecnologia, engenharia ou ciência da computação.	Artigos não relacionados à estratégias baseadas em imagens dos olhos para a Interação Humano-Computador a partir de câmeras genéricas.

Objetivando determinar o conjunto final de artigos elegíveis e buscar respostas às Questões de Pesquisa (ver Tabela 4.1), uma triagem, guiada pelos Critérios de Avaliação de Qualidade (Tabela 4.3), foi realizada a partir da leitura completa dos estudos primários. Para qualificação e classificação dos estudos, uma condição eliminatória (CAQ01) e uma métrica de avaliação, chamada *score* (ver Equação 4.1), foram utilizadas. O *score* é a média aritmética dos pesos (w) atribuídos para cada Critério de Avaliação de Qualidade. O peso (w), que pode variar entre 0, 0.5 e 1.0, mensura o quão satisfatória é a resposta

daquele artigo para um determinado Critério de Avaliação de Qualidade, conforme mostra a Equação 4.2. Os artigos primários que obtiveram *score* igual ou superior a 0.5 (isto é, $0.5 \leq score \leq 1.0$) foram considerados elegíveis para esta Revisão Sistemática da Literatura. Dois revisores atribuíram os *scores* e os dados elementares do conjunto final de estudos elegíveis, extraídos com base nas Questões de pesquisa, foram resumidos na Tabela 4.4.

$$score = \frac{1}{n_{CAQ}} \sum_{i=1}^{n_{CAQ}} w_{CAQ_i} \quad (4.1)$$

onde:

- n_{CAQ} : variável usada para representar o total de CAQ
- w_{CAQ} : variável usada para determinar o valor referente ao peso w atribuído ao CAQ analisado (ver os valores possíveis na Equação 4.2).

$$w_{CAQ} = \begin{cases} 1.0, & \text{sim, descreve totalmente,} \\ 0.5, & \text{sim, descreve parcialmente,} \\ 0, & \text{não descreve.} \end{cases} \quad (4.2)$$

Tabela 4.3: Critérios de Avaliação de Qualidade

CAQ	Descrição	Requisito
01	O estudo tem como objeto de investigação uma abordagem de Interação Humano-Computador baseada em imagens dos olhos para pessoas com ELA ou doença do neurônio motor?	Sim
02	O estudo descreve a abordagem para o processamento de imagens?	Não
03	O estudo descreve a técnica algorítmica e sua performance (acurácia, precisão, sensibilidade, especificidade)?	Não
04	O estudo descreve o hardware utilizado para a aquisição das imagens?	Não
05	O estudo realiza experimentos em grupos de controle (pessoas saudáveis), pessoas com ELA ou outras doenças?	Não

4.2 Resultados

Os resultados quantitativos detalhados da execução do protocolo desta Revisão Sistemática da Literatura foram sumarizados na Figura 4.1. Após a identificação de 9084 registros e a realização da triagem, constituída pela aplicação dos Critérios de Inclusão (8586

estudos excluídos), Critérios de Exclusão (449 estudos excluídos) e Critérios de Avaliação de Qualidade (24 estudos excluídos), um conjunto de 25 estudos primários foram considerados elegíveis e incluídos nesta Revisão Sistemática da Literatura para responder às Questões de Pesquisa (Tabela 4.1). Os resultados mais abrangentes da investigação foram organizados e apresentados na mesma sequência das Questões de Pesquisa. Toda a análise é embasada nos dados extraídos dos 25 artigos elegíveis, resumidamente descritos na Tabela 4.4. Esta tabela foi organizada em grupos de Interação Humano-Computador, *Score* e ano de publicação dos artigos, dispostos em ordem decrescente.

Figura 4.1: Resultado do processo de busca e triagem dos estudos primários para esta revisão sistemática.

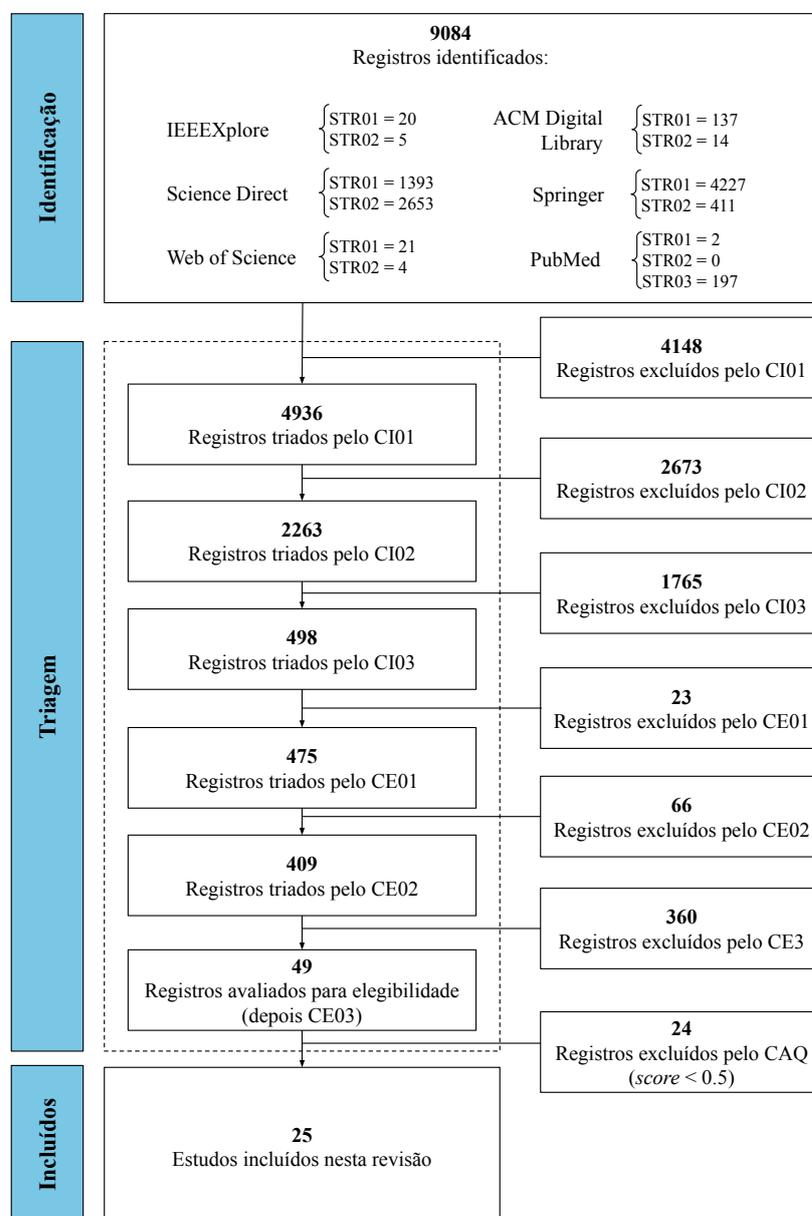


Tabela 4.4: Resumo das principais características dos artigos incluídos nesta revisão sistemática da literatura.

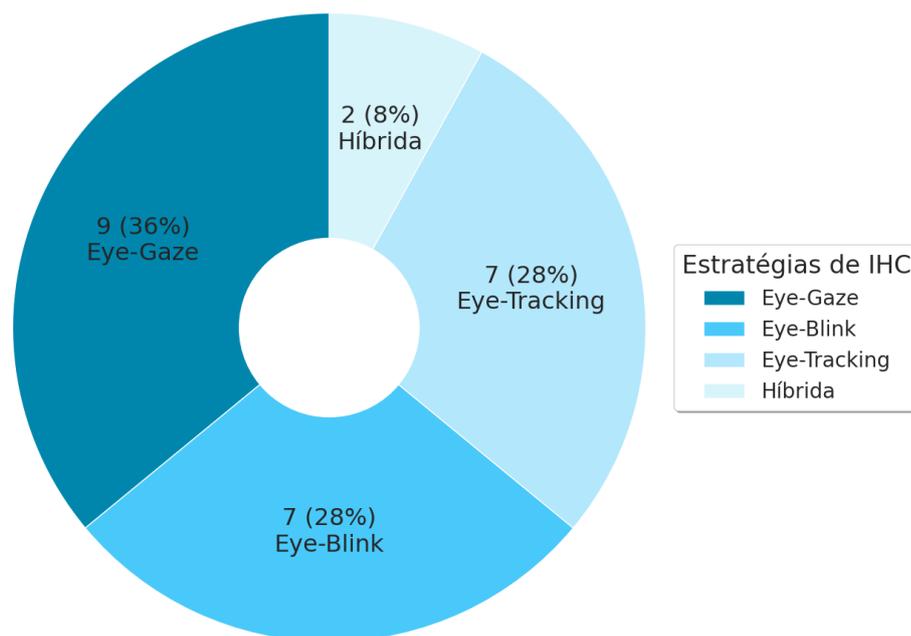
ID	Estudo	Score	IHC	Hardware	Indivíduos GC/ELA/OD	Técnicas	Performance(%)		
							Acc	Recall	Precisão
01	Eom et al. (2019)	0.8	<i>Eye-Gaze</i>	Câmara	6/0/0	<i>Haar-like/Binarização/grayscale/RNA</i>	86%	Outra Abordagem de Avaliação	-
02	Zhang et al. (2017)	0.8	<i>Eye-Gaze</i>	iPhone e iPad	12/0/0	Fast face alignment/GD/TM	100%	-	-
03	Aslam et al. (2019)	0.7	<i>Eye-Gaze</i>	Câmara	3/0/0	<i>Haar-like/CHT</i>	-	-	-
04	Abe et al. (2011)	0.7	<i>Eye-Gaze</i>	Câmara	5/0/0	<i>Limbus Tracking</i>	-	-	0.56° / 1.09°
05	Rahnama-ye Moqaddam e Vahdat-Nejad (2015)	0.6	<i>Eye-Gaze</i>	Câmara	4/1/0	<i>Haar cascade/GVM/TM</i>	-	-	5.68%
06	Rozado, Rodriguez e Varona (2012)	0.6	<i>Eye-Gaze</i>	Câmara com IR	15/0/0	<i>ITU Gaze Tracker/E-HTM</i>	98%	-	-
07	Yıldiz e Yorulmaz (2019)	0.5	<i>Eye-Gaze</i>	CAC	1/0/0	CHT/KNN	-	-	0.98%
08	Nakazawa et al. (2018)	0.5	<i>Eye-Gaze</i>	CAC com IR	5/0/0	CHT	93.32%	-	-
09	Rozado, Agustin, Rodriguez e Varona (2012)	0.5	<i>Eye-Gaze</i>	CAC com IR	20/0/0	HTM/Needleman-Wunsch	95%	-	-
10	Królak e Strumillo (2012)	0.8	<i>Eye-Blink</i>	Câmara	37/0/12	Viola-Jones/GD/TM	95.17%	96.91%	98.13%
11	Singh e Singh (2019b)	0.7	<i>Eye-Blink</i>	Câmara com fonte de luz	10/0/0	Viola-Jones/PMA	90%	-	-
12	Singh e Singh (2018)	0.7	<i>Eye-Blink</i>	Câmara com fonte de luz	10/0/0	Viola-Jones/PMA	91.2%	-	94.11%
13	Missimer e Betke (2010)	0.7	<i>Eye-Blink</i>	Câmara	20/0/0	TM/Optical flow algorithm	96.6%	-	-
14	Rupanagudi et al. (2018)	0.6	<i>Eye-Blink</i>	Câmara com IR	50/0/0	<i>grayscale/SBT/2PVM</i>	-	Outra Abordagem de Avaliação	-
15	Rakshita (2018)	0.5	<i>Eye-Blink</i>	Câmara	1/0/0	<i>grayscale/FLD/EAR</i>	-	Outra Abordagem de Avaliação	-
16	Krapic et al. (2015)	0.5	<i>Eye-Blink</i>	Câmara	12/0/0	e Viacam software	-	Outra Abordagem de Avaliação	-
17	Park e Park (2016)	0.8	<i>Eye-Tracking</i>	Câmara com IR	4/0/0	Reflexão da Córnea e Centro da Pupila	1-2°	-	-
18	Saleh e Tarek (2021)	0.7	<i>Eye-Tracking</i>	CAC com IR	5/0/0	<i>grayscale/CHT/GD</i>	-	-	-
19	Atasoy et al. (2016)	0.7	<i>Eye-Tracking</i>	Câmara	30/0/0	Viola-Jones/ <i>grayscale/CHT/GD</i>	90%	Outra Abordagem de Avaliação	-
20	Aharanson et al. (2020)	0.6	<i>Eye-Tracking</i>	CAC	4/0/0	OpenCV/Polynomial/Projection	-	Outra Abordagem de Avaliação	-
21	Oyabu et al. (2012)	0.6	<i>Eye-Tracking</i>	Câmara com IR	5/0/0	Binarização/CMUPL	-	Outra Abordagem de Avaliação	-
22	Kaushik et al. (2018)	0.5	<i>Eye-Tracking</i>	CAC com IR	1/0/0	Eyeecan software	95%	-	-
23	Kavale et al. (2018)	0.5	<i>Eye-Tracking</i>	Câmara com IR	1/0/0	Binarização/GD	-	Outra Abordagem de Avaliação	-
24	Zhao et al. (2015)	0.8	Híbrida	Câmara com IR	7/0/0	Binarização/GD	92.69%	Outra Abordagem de Avaliação	-
25	Xu e Lin (2017)	0.7	Híbrida	Câmara com IR	1/0/0	FLD/GD	100%	-	-

Abreviações: IHC Interação Humano-Computador, GC Grupo de Controle, ELA Esclerose Lateral Amiotrófica, OD Outras Doenças, Acc Acurácia, CAC Câmara Ajustada na Cabeça, IR Infravermelho, HTM Memória Temporal Hierárquica, E-HTM Extensão do HTM, RNA Rede Neural Artificial, GVM Método do vetor de gradientes, TM *Template Matching*, GD Dependências Geométricas, CHT Transformada de Hough, KNN *K-Nearest Neighbor*, PMA Análise de Movimento de Pixels, SBT Segmentação Baseada em *Thresholding*, 2PVM Metodologia de Verificação de 2 Pixels, FLD Detector de Pontos Facial, EAR *Eye Aspect Ratio*, POLYNOMIAL Algoritmo Paramétrico Baseado em Interpolação, PROJECTION Algoritmo Baseado em Modelo, CMUPL *Clustering Method of Unbroken Pixel Lines*.

4.2.1 Questão de Pesquisa 01

Com base nos estudos primários e conforme mostra a Figura 4.2, as estratégias evidenciadas para a Interação Humano-Computador através de imagens dos olhos são divididas em quatro categorias: *Eye-Gaze*; *Eye-Blink*; *Eye-Tracking*; e estratégias híbridas, que combinam algumas das categorias anteriores. A maior representação encontrada entre as categorias foi a abordagem de Interação Humano-Computador baseada em *Eye-Gaze* (36% dos estudos) que, geralmente, busca estimar a direção do olhar com base na movimentação da pupila nos eixos horizontal (esquerda e direita) e vertical (cima e baixo) para selecionar o objeto alvo na interface (Eom et al. 2019, Zhang et al. 2017, Aslam et al. 2019, Abe et al. 2011, Rahnama-ye Moqaddam e Vahdat-Nejad 2015, Rozado, Rodriguez e Varona 2012, Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012).

Figura 4.2: Estratégias para Interação Humano-Computador baseadas em imagens dos olhos.



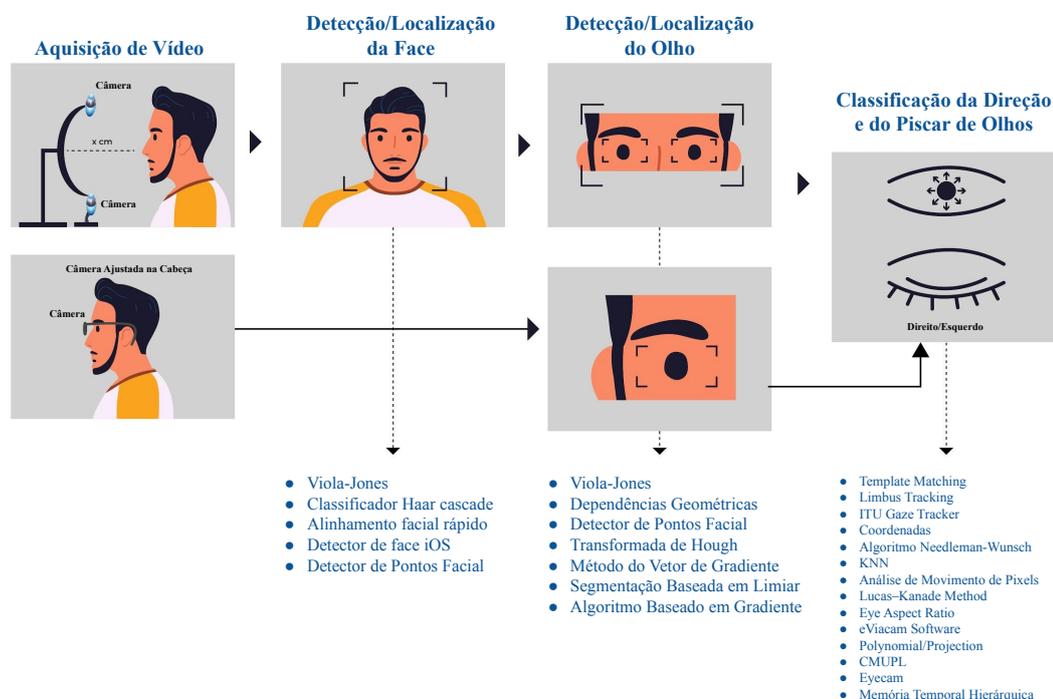
A estratégia *Eye-Blink* foi evidenciada em 28% do conjunto de estudos primários (Królak e Strumiłło 2012, Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Rupanagudi et al. 2018, Rakshita 2018, Krapic et al. 2015). Nesta categoria, a abordagem de selecionar o alvo na interface varia e pode ser baseada na detecção/identificação do piscar voluntário dos olhos (piscar e manter os olhos fechados por um período curto de tempo) (Królak e Strumiłło 2012), na simulação de cliques análogos ao mouse (piscar olho direito ou esquerdo) (Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Rakshita 2018, Krapic et al. 2015) e na combinação sequencial de piscadas em um espaço temporal (Rupanagudi et al. 2018). Com a mesma porcentagem, 28% dos

estudos primários fornecem abordagens para Interação Humano-Computador baseadas em *Eye-Tracking*, isto é, que buscam identificar e classificar a direção efetiva da pupila (Park e Park 2016, Saleh e Tarek 2021, Atasoy et al. 2016, Aharonson et al. 2020, Oyabu et al. 2012, Kaushik et al. 2018, Kavale et al. 2018). Apesar da semelhança com o *Eye-Gaze*, essa estratégia busca estimar, geralmente, a direção do olhar em relação aos pixels da imagem, o que a torna mais precisa, pois vai além das linhas horizontais e verticais. Os estudos de Zhao et al. (2015) e Xu e Lin (2017), que pertencem à categoria de estratégias híbridas, combinam as estratégias de *Eye-Gaze* e *Eye-Blink*.

4.2.2 Questão de Pesquisa 02

As técnicas algorítmicas exploradas nos estudos primários variam em razão das abordagens de Interação Humano-Computador, vistas na seção anterior (Questão de Pesquisa 01), e da configuração do ambiente em que a câmera foi alocada para a aquisição das imagens/vídeos do usuário, como mostra a Figura 4.3, na etapa chamada Aquisição de Vídeo. A Figura 4.3 mostra ainda um fluxo de trabalho genérico dos procedimentos (tarefas) e das respectivas técnicas computacionais de visão computacional ou aprendizado de máquina frequentemente utilizadas para solucionar os desafios que compreendem a Interação Humano-Computador através do rastreamento da pupila ou detecção do piscar dos olhos.

Figura 4.3: Modelo genérico de fluxo de trabalho (*pipeline*).



No conjunto de estudos primários, principalmente para as etapas de Detecção/Localização da Face e Detecção/Localização do Olho, foram observadas evidências da utili-

zação de recursos algorítmicos de visão computacional da Biblioteca de Visão Computacional de Código Aberto (OpenCV, do Inglês *Open Source Computer Vision Library*) (Bradski 1998) e Dlib (King 2009). Para a tarefa de Detecção/Localização da Face, por exemplo, os autores mencionaram o uso de técnicas baseadas no algoritmo Viola-Jones ou classificador *Haar cascade*, disponíveis na OpenCV, e Detector de Pontos Facial da biblioteca Dlib (Eom et al. 2019, Aslam et al. 2019, Rahnema-ye Moqaddam e Vahdat-Nejad 2015, Królak e Strumiłło 2012, Rakshita 2018, Atasoy et al. 2016, Xu e Lin 2017). Outras fontes de ferramentas computacionais para processamento de imagem e visão computacional também foram exploradas na etapa de Detecção/Localização da Face, como por exemplo, Singh e Singh (2019b) e Singh e Singh (2018) utilizaram o algoritmo Viola-Jones da empresa *MathWorks*. Os autores Zhang et al. (2017) exploraram a detecção de rostos utilizando ferramentas da biblioteca de aprendizado de máquina do iOS e a detecção de pontos com Dlib.

Os desafios encontrados na tarefa de Detecção/Localização do Olho estão relacionados, principalmente, ao recorte dos olhos e realce da pupila ou da íris. As técnicas computacionais para a tarefa de recorte dos olhos estão intrinsecamente vinculadas à abordagem de aquisição de vídeo que não dispõe de câmeras montadas na cabeça do usuário e objetivam delimitar ou extrair a região de interesse (o olho). São técnicas computacionais comumente exploradas: algoritmo Viola-Jones (Singh e Singh 2019b, Singh e Singh 2018, Atasoy et al. 2016); modelos algorítmicos baseados em Detecção de Pontos Facial (Zhang et al. 2017, Rakshita 2018); e dependências geométricas aliadas a estratégias de Binarização/*grayscale*/OpenCV (Eom et al. 2019, Nakazawa et al. 2018, Królak e Strumiłło 2012, Aharonson et al. 2020, Kavale et al. 2018, Zhao et al. 2015).

Para a tarefa de realçar ou de destacar a pupila ou a íris em uma imagem do olho, é evidenciado o uso de uma técnica robusta e versátil para a detecção de círculos, chamada de Transformada Circular de Hough (Aslam et al. 2019, Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Saleh e Tarek 2021, Atasoy et al. 2016). Outras técnicas que abrangem a tarefa de Detecção/Localização do Olho são mencionadas, como o método *Limbus tracking* (Abe et al. 2011), algoritmos baseados em gradiente (Rahnema-ye Moqaddam e Vahdat-Nejad 2015, Xu e Lin 2017), modelos de aprendizado de máquina baseados em Memória Temporal Hierárquica (Rozado, Rodriguez e Varona 2012, Rozado, Agustin, Rodriguez e Varona 2012), Erosão com um elemento de estrutura em forma de cruz (Missimer e Betke 2010), segmentação baseada em limiar (Rupanagudi et al. 2018) e o método de agrupamento de linhas de pixel ininterruptas (CMUPL) (Oyabu et al. 2012).

Na perspectiva de detectar o piscar dos olhos, os autores exploraram modelos computacionais baseados em correspondência (*Template Matching*) (Królak e Strumiłło 2012, Missimer e Betke 2010), algoritmo de fluxo óptico e Análises de Movimentos de Pixels (Singh e Singh 2019b, Singh e Singh 2018), funções ou fórmulas matemáticas para calcular o *Eye Aspect Ratio* (EAR) (Rakshita 2018, Xu e Lin 2017), bem como a altura e largura da íris (Zhao et al. 2015), e a metodologia de Verificação de 2 Pixels (preto e branco: olho aberto; preto e preto: olho fechado) (Rupanagudi et al. 2018). Além do uso da técnica baseada em *Template Matching*, Missimer e Betke (2010) incorporaram ao modelo proposto o método de Lucas-Kanade (algoritmo de fluxo óptico) e máquinas de estado finito. Krapic et al. (2015) utilizaram um *software*, chamado eViacam, com implemen-

tações integradas de Análises de Movimentos, dos autores (Grauman et al. 2001, Chau e Betke 2005), e técnicas computacionais (não especificadas) da biblioteca de código aberto OpenCV.

No conjunto de estudos em que o objetivo consistiu em desenvolver uma estratégia de Interação Humano-Computador baseada na direção do olhar, é evidente que os autores seguem o fluxo de forma a identificarem antecipadamente a pupila/íris nas imagens, utilizando métodos variados como a Transformada Circular de Hough (Aslam et al. 2019, Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Saleh e Tarek 2021, Atasoy et al. 2016) ou algoritmo baseado em gradiente (Rahnama-ye Moqaddam e Vahdat-Nejad 2015, Xu e Lin 2017), e extraem valores referentes às coordenadas que servem de entrada para modelos matemáticos (chamados neste trabalho de dependências geométricas) que calculam e identificam a direção do olhar (Aslam et al. 2019, Yildiz e Yorulmaz 2019, Rahnama-ye Moqaddam e Vahdat-Nejad 2015, Nakazawa et al. 2018, Saleh e Tarek 2021, Atasoy et al. 2016, Kavale et al. 2018, Zhao et al. 2015, Xu e Lin 2017). Os autores Eom et al. (2019) e Yildiz e Yorulmaz (2019) utilizaram as dependências geométricas para treinar e criar modelos de aprendizado de máquina através de redes neurais e do algoritmo KNN (*K-Nearest Neighbor*), respectivamente.

Outras abordagens para classificar a direção do olhar também foram mencionadas, Zhang et al. (2017) utilizaram *Template Matching*, Abe et al. (2011) exploram o método de detecção vertical do *Eye-Gaze*, que também é baseado no método *Limbus Tracking*, e os autores Rozado201202 e Rozado, Agustin, Rodriguez e Varona (2012) aliam aos modelos de aprendizado de máquina baseados em Memória Temporal Hierárquica o *ITU Gaze Tracker* (*software* de código aberto) e o algoritmo de Needleman-Wunsch, respectivamente. No estudo de Oyabu et al. (2012) a posição da pupila é definida usando o método de agrupamento de linhas de pixel ininterruptas. Também realizando operações matemáticas, Aharonson et al. (2020) calculam a posição da pupila usando dois algoritmos diferentes: (i) algoritmo paramétrico baseado em interpolação (chamado de *polynomial*); e o algoritmo baseado em modelo (chamado de *projection*). Park e Park (2016) constroem um sistema embarcado especialista, chamado de “*Pupil Center Corneal Reflection*”, capaz de realizar o rastreamento da pupila através de um *hardware* com luzes adaptáveis acopladas e um programa baseado em modelo matemático. Kaushik et al. (2018) utilizam o *software* *eyecan*.

4.2.3 Questão de Pesquisa 03

A avaliação relacionada à performance das técnicas computacionais exploradas nos estudos primários aponta resultados promissores em testes em grupo de controle. A análise dos 13 estudos que reportaram as performances das técnicas em termos percentuais mostra que a acurácia (Acc) média alcança o valor de 94,12% (std = 4,14; mediana = 95%) (Zhang et al. 2017, Aslam et al. 2019, Rozado, Rodriguez e Varona 2012, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012, Królak e Strumiłło 2012, Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Atasoy et al. 2016, Kaushik et al. 2018, Zhao et al. 2015, Xu e Lin 2017). A abordagem proposta pelos autores Park e Park (2016) obteve uma acurácia de 1-2°. Além da Acc, de 95,17%, Królak e Strumiłło

(2012) mensuram o *Recall* e a Precisão, que obtiveram os valores de 96,91% e 98,13%, respectivamente. Singh e Singh (2018) também mensuram a performance em mais de uma métrica, apresentando 91,2% de Acc e 94,11% de Precisão. Abe et al. (2011) apresentam o erro médio em duas perspectivas de detecção do *Eye-Gaze*: detecção vertical (0,56°); a detecção horizontal (1,09°). Os autores Rahnama-ye Moqaddam e Vahdat-Nejad (2015) reportam a taxa de erro médio 5,68% e, agora observando a melhor taxa de erro obtida pelo sistema, Yildiz e Yorulmaz (2019) apresentam como resultado 0,98%. Todas as técnicas computacionais exploradas podem ser identificadas na Tabela 4.4.

Outras abordagens para avaliar a performance da tecnologia para a Interação Humano-Computador são utilizadas. Eom et al. (2019) conduzem experimentos computacionais com um grupo de controle e somam o erro do movimento do olhar (vertical e horizontal) dos participantes individualmente. Zhang et al. (2017) avaliam, além do sistema de *Eye-Gaze*, a usabilidade do software de Comunicação Aumentativa e Alternativa através de um questionário com questões baseadas na escala de Likert (Likert 1932). De forma semelhante, Krapic et al. (2015) realizam a avaliação através de testes de usabilidade. Rupanagudi et al. (2018) avaliam e comparam a velocidade do algoritmo proposto com outra abordagem da literatura. Com um sistema de avaliação baseado em reconhecimento de padrão, Rakshita (2018) reporta a eficiência da abordagem (sem quantificar). Saleh e Tarek (2021) avaliam a proposta com base em uma interface com 6 alvos que representam as necessidades dos usuários. Aharonson et al. (2020) constroem uma tabela contendo o desvio médio em graus para cada usuário. Os resultados experimentais em Oyabu et al. (2012) são apresentados através do cálculo do tempo a partir do uso de um sistema chamado “click experiment screenshot”. Kavale et al. (2018) mostram a performance das técnicas utilizadas através de imagens.

4.2.4 Questão de Pesquisa 04

Com base na etapa de Aquisição de Vídeo apresentada na Figura 4.3, 76% dos estudos primários (Eom et al. 2019, Zhang et al. 2017, Aslam et al. 2019, Abe et al. 2011, Rahnama-ye Moqaddam e Vahdat-Nejad 2015, Rozado, Rodriguez e Varona 2012, Królak e Strumiňo 2012, Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Rupanagudi et al. 2018, Rakshita 2018, Krapic et al. 2015, Park e Park 2016, Atasoy et al. 2016, Oyabu et al. 2012, Kavale et al. 2018, Zhao et al. 2015, Xu e Lin 2017) realizaram experimentos computacionais utilizando dispositivos para a coleta de imagens alocados em uma mesa ou integrados ao próprio computador, como no caso de notebooks ou *smartphones* com câmeras integradas, o que caracteriza uma abordagem de Interação Humano-Computador onde os usuários ficam livres de dispositivos em seu corpo. De modo alternativo, 24% dos estudos exploraram uma abordagem de Interação Humano-Computador onde o protótipo para coleta de imagens, a câmera, é montado na cabeça do usuário (Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012, Saleh e Tarek 2021, Aharonson et al. 2020, Kaushik et al. 2018).

Ainda em uma perspectiva geral, 52% dos estudos primários propuseram dispositivos de Interação Humano-Computador equipados com algum tipo de fonte de luz projetada para o olho ou face do usuário, sendo luzes de infravermelho (Rozado, Rodriguez

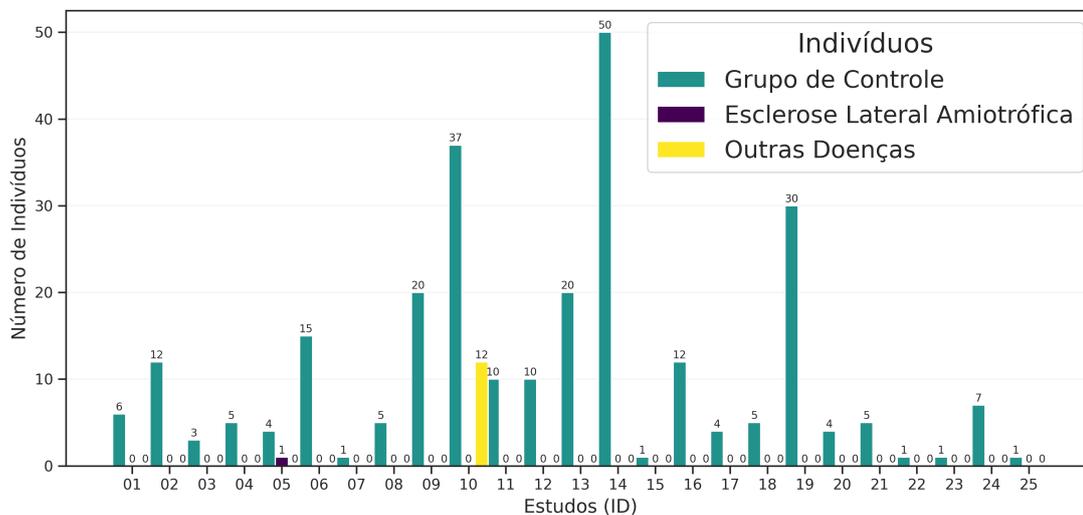
e Varona 2012, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012, Rupanagudi et al. 2018, Park e Park 2016, Saleh e Tarek 2021, Oyabu et al. 2012, Kaushik et al. 2018, Kavale et al. 2018, Zhao et al. 2015, Xu e Lin 2017) e lâmpadas (Singh e Singh 2019b, Singh e Singh 2018). Na categoria de estratégias para Interação Humano-Computador baseadas em *Eye-Gaze*, que acumulou o maior número de estudos (nove), cinco propostas de abordagens de Comunicação Aumentativa e Alternativa (aproximadamente 55,6%) utilizam câmeras livres de algum tipo de recurso adicional ou do corpo (Eom et al. 2019, Zhang et al. 2017, Aslam et al. 2019, Abe et al. 2011, Rahnama-ye Moqaddam e Vahdat-Nejad 2015). Dos outros quatro estudos da mesma categoria, três exploram abordagens com câmeras montadas na cabeça do usuário (Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012), sendo duas propostas com infravermelho (Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012), e os autores Rozado, Rodriguez e Varona (2012) adicionam à câmera luzes e infravermelho.

Na categoria de *Eye-Blink*, todos os sete estudos exploram a captura de imagens através de câmeras livres do corpo do usuário (Królak e Strumiłło 2012, Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Rupanagudi et al. 2018, Rakshita 2018, Krapic et al. 2015). Três estudos equiparam as câmeras com algum tipo de fonte de luz projetada para o olho ou face do usuário, sendo um com luzes de infravermelho (Rupanagudi et al. 2018) e dois com lâmpadas (Singh e Singh 2019b, Singh e Singh 2018). A Aquisição de Vídeo nos estudos pertencentes à categoria de *Eye-Tracking* varia entre as abordagens com câmera ajustada na cabeça (Saleh e Tarek 2021, Aharonson et al. 2020, Kaushik et al. 2018), sendo dois dispositivos com infravermelho (Saleh e Tarek 2021, Kaushik et al. 2018), e com câmeras não montadas na cabeça equipadas com (Park e Park 2016, Oyabu et al. 2012, Kavale et al. 2018) e sem infravermelho (Atasoy et al. 2016). Os autores Zhao et al. (2015) e Xu e Lin (2017), da categoria de estratégias híbridas, investigam as técnicas para a Interação Humano-Computador a partir de imagens coletadas de câmeras com infravermelho.

4.2.5 Questão de Pesquisa 05

Os dados extraídos dos estudos primários para responder a esta Research Questions foram sumarizados na Figura 4.4. A Figura 4.4 mostra claramente que apenas um estudo, dos autores Rahnama-ye Moqaddam e Vahdat-Nejad (2015), para *Eye-Gaze*, realizou testes experimentais em uma pessoa com Esclerose Lateral Amiotrófica. Um segundo estudo considerou participantes com outras doenças (não especificadas). Królak e Strumiłło (2012) incluíram 12 pessoas nos testes experimentais para avaliar a abordagem de Interação Humano-Computador através de *Eye-Blink*. Todos os estudos primários realizam testes em pessoas saudáveis (grupo de controle), com uma média de 10,76 participantes por estudo (std = 12,3; mediana = 5).

Figura 4.4: Número de indivíduos submetidos aos testes nos estudos primários.



4.3 Discussões

Esta revisão sistemática da literatura investigou um conjunto de 25 estudos primários relacionados à abordagens de Interação Humano-Computador baseadas em imagens através de câmeras simples e de baixo custo para a Comunicação Aumentativa e Alternativa de pessoas com Esclerose Lateral Amiotrófica. Inicialmente, como resposta à questão problema, os resultados em sua essência apontam para a possibilidade e viabilidade do desenvolvimento de tecnologias de baixo custo para a Interação Humano-Computador por meio de imagens dos olhos. Entretanto, ainda há desafios a serem explorados nas grandes áreas de visão computacional, aprendizado de máquina e Comunicação Aumentativa e Alternativa relacionados não somente com o custo, mais também a eficiência e usabilidade das tecnologias de Interação Humano-Computador que são utilizadas por meio dos olhos, particularmente, no contexto das pessoas com Esclerose Lateral Amiotrófica. É nesta perspectiva que esta revisão sistemática da literatura organizou e discutiu, na sequência, os principais achados.

A primeira análise, relacionada às estratégias de Interação Humano-Computador, evidenciou quatro estratégias abordadas nos estudos primários: *Eye-Gaze* (36%) (Eom et al. 2019, Zhang et al. 2017, Aslam et al. 2019, Abe et al. 2011, Rahnama-ye Moqaddam e Vahdat-Nejad 2015, Rozado, Rodriguez e Varona 2012, Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012); *Eye-Blink* (28%) (Królak e Strumiłło 2012, Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Rupanagudi et al. 2018, Rakshita 2018, Krapic et al. 2015); *Eye-Tracking* (28%) (Park e Park 2016, Saleh e Tarek 2021, Atasoy et al. 2016, Aharonson et al. 2020, Oyabu et al. 2012, Kaushik et al. 2018, Kavale et al. 2018); and Hybrid strategies (8%) (Zhao et al. 2015, Xu e Lin 2017). Foi observado também que 52% dos estudos adotaram recursos adicionais para controlar a luz do ambiente ou que incidem no olho ou face do usuário. Deste grupo de estudos, 11 recorreram ao uso de infravermelho (Rozado, Rodriguez

e Varona 2012, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012, Rupanagudi et al. 2018, Park e Park 2016, Saleh e Tarek 2021, Oyabu et al. 2012, Kaushik et al. 2018, Kavale et al. 2018, Zhao et al. 2015, Xu e Lin 2017) e 2 à luz fluorescente (Singh e Singh 2019b, Singh e Singh 2018). O direcionamento de feixes de luz para o olho objetivou, em termos práticos, criar efeitos reflexivos na região da pupila, no caso de luzes infravermelho, ou pontos de referência na pupila/íris/esclera, tal que facilitasse o processamento da imagem, e, conseqüentemente, a detecção e classificação da direção do olhar ou estado do olho (aberto ou fechado). Ainda na perspectiva de melhorar as condições de processamento de imagem, detecção do movimento do olhar e performance do dispositivo de Interação Humano-Computador, seis estudos (24%) realizaram experimentos com a câmera ajustada na cabeça dos usuários (Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012, Saleh e Tarek 2021, Aharonson et al. 2020, Kaushik et al. 2018).

As pesquisas destacadas neste estudo podem ser importantes se forem alinhadas ao compartilhamento do conhecimento e do próprio recurso de Comunicação Aumentativa e Alternativa (produto final). Portanto, o objetivo deve ser o de melhorar as habilidades funcionais de pessoas com deficiências motoras, ou seja, a autonomia destes indivíduos. Deste modo, é possível mitigar os efeitos do isolamento social. Além disso, promove também o exercício dos direitos, a cidadania, das liberdades fundamentais e do cuidado assistencial à saúde da pessoa com Esclerose Lateral Amiotrófica. Esse aspecto é muito significativo, por atuar diretamente na promoção de saúde, bem estar e na redução das desigualdades, o que poderá refletir, essencialmente, na promoção da equidade. Esses fatores, estão inclusive previstos nos Objetivos de Desenvolvimento Sustentável (ODS), que fazem parte da chamada Agenda 2030 das Organizações das Nações Unidas (ONU), particularmente, os (3.8) e 10 (United Nations 2015). Logo, essa temática não se trata, somente, de desenvolver novas tecnologias, ou, simplesmente, desenvolver tecnologias de baixo custo, mas de atuar por meio da mediação tecnológica como pesquisas indutoras da inclusão social, da redução das desigualdades e na promoção de equidade - valores muitas vezes não mensuráveis nas pesquisas científicas de cunho mais tecnológicos. Apesar de grande parte dos resultados terem ocorrido em grupo de controle, eles demonstram resultados que sugerem a viabilidade do investimento em pesquisas neste campo. Todavia, é importante que a priorização de pesquisas nesta área seja dada pelos formuladores das políticas públicas de saúde, especialmente, com o objetivo de garantir aos pacientes mais pobres, diagnosticados com a Esclerose Lateral Amiotrófica, acesso à tecnologias assistivas que possam melhorar a qualidade de vida deles. Não adianta, somente, desenvolver novas tecnologias é preciso garantir que os pacientes com Esclerose Lateral Amiotrófica tenham acesso a elas, independente de suas condições sociais. A garantia de investimentos em pesquisa neste campo é importante não apenas para dar acesso aos pacientes, algo que é fundamental, mas também para garantir a sustentabilidade e os avanços dos estudos nesta área, que muitas vezes é negligenciada pela indústria, pois o mercado é muito restrito.

A Esclerose Lateral Amiotrófica é considerada rara e, apesar dos esforços para buscar soluções digitais de saúde, ainda existem desafios significativos a serem enfrentados. Isso inclui a necessidade de mais dados, estudos e evidências sobre a incidência e prevalência

de doenças, que são informações essenciais, mas escassas, no contexto da saúde global (Fernandes et al. 2021, Papaiz et al. 2022, Gromicho et al. 2022, Tavazzi et al. 2022, Gordon e Lerner 2019, Ahangaran e Chiò 2020, Özden O. Dalgıç et al. 2021, Ahangaran et al. 2022, Bede et al. 2022, Iadanza et al. 2022, Thome et al. 2022, Greco et al. 2021, Kocar, Behler, Ludolph, Müller e Kassubek 2021, Leão et al. 2021, Grollemund et al. 2021, Kocar, Müller, Ludolph e Kassubek 2021, Grollemund et al. 2020, Myszczyńska et al. 2020, Chen et al. 2020, Grollemund et al. 2019, Pinto et al. 2020, Barbalho et al. 2021). Existem poucos registros ou estudos epidemiológicos no Brasil, sendo apenas dois estudos de nível nacional citados na literatura. Em 1998, Dietrich-Neto et al. (2000) realizaram uma pesquisa nacional e relataram taxas de incidência e prevalência de 0,4 e 1,2 por 100.000 habitantes, respectivamente. Mais recentemente, ao analisar o período de 2004 a 2013, os pesquisadores Moura et al. (2016) estimaram a incidência média em 0,461 casos por 100.000 habitantes (com tendência de aumento da incidência ao longo dos anos), taxa semelhante à de Dietrich-Neto et al. (2000). Vale ressaltar que, no Brasil, até 2019, não havia um sistema de notificação compulsória ou registro nacional da Esclerose Lateral Amiotrófica, o que leva a uma provável subnotificação.

Para tentar resolver esse problema de subnotificação no Brasil, Barbalho et al. (2023) destaca o Registro Nacional de Esclerose Lateral Amiotrófica, projeto de pesquisa aplicada apoiado pelo Ministério da Saúde do Brasil. Segundo Barbalho et al. (2023), esse Registro Nacional é um projeto ainda em andamento e em implantação em todo o Brasil, cujo objetivo é fazer um mapeamento online de todas as pessoas com Esclerose Lateral Amiotrófica no país de forma contínua. A partir deste Registro Nacional será possível, por exemplo, desenvolver estudos epidemiológicos e análises que poderão apoiar as decisões das autoridades públicas na formulação das políticas de saúde no contexto da Esclerose Lateral Amiotrófica no Brasil. Há, neste sentido, o Projeto de Lei nº 4691, de 2019 (Brasil 2019), que visa tornar obrigatório a notificação das doenças raras no Brasil, portanto, o Registro Nacional de Esclerose Lateral Amiotrófica é parte estruturante deste Projeto de Lei. Destaca-se no Brasil o Estado do Rio Grande do Norte, que fica na Região Nordeste do país, pois foi o primeiro Estado brasileiro a publicar a Lei nº 10.924, de 10 de Junho de 2021 (Brasil 2021*b*), que tornou compulsória a notificação da Esclerose Lateral Amiotrófica.

Os desafios no contexto da Esclerose Lateral Amiotrófica são vários e é evidente que estes perpassam os campos das ciências da saúde e tecnológicas. Entretanto, a compreensão da transdisciplinaridade e o uso adequado dessas tecnologias ou soluções de saúde digital são fatores que podem contribuir significativamente no acesso à saúde de qualidade, na redução das desigualdades e na melhoria da qualidade de vida, principalmente, das pessoas com Esclerose Lateral Amiotrófica. Para isso, é preciso ainda pensar nas tecnologias também como ferramentas para o desenvolvimento social e sustentável das sociedades.

4.4 Considerações

Este artigo, através da execução de um protocolo de revisão sistemática da literatura, investigou estudos primários na literatura e evidenciou cinco pontos relevantes que podem

contribuir diretamente para o desenvolvimento e efetividade tecnológica no provimento de estratégias de Interação Humano-Computador baseadas em imagens dos olhos para a Comunicação Aumentativa e Alternativa de pessoas com Esclerose Lateral Amiotrófica. O primeiro ponto mostrou as abordagens de Interação Humano-Computador com os olhos: *Eye-Gaze* (36%) (Eom et al. 2019, Zhang et al. 2017, Aslam et al. 2019, Abe et al. 2011, Rahnama-ye Moqaddam e Vahdat-Nejad 2015, Rozado, Rodriguez e Varona 2012, Yildiz e Yorulmaz 2019, Nakazawa et al. 2018, Rozado, Agustin, Rodriguez e Varona 2012); *Eye-Blink* (28%) (Królak e Strumiłło 2012, Singh e Singh 2019b, Singh e Singh 2018, Missimer e Betke 2010, Rupanagudi et al. 2018, Rakshita 2018, Krapic et al. 2015); *Eye-Tracking* (28%) (Park e Park 2016, Saleh e Tarek 2021, Atasoy et al. 2016, Aharonson et al. 2020, Oyabu et al. 2012, Kaushik et al. 2018, Kavale et al. 2018); e estratégias híbridas (8%) (Zhao et al. 2015, Xu e Lin 2017). Estas abordagens de Interação Humano-Computador são resultados de esforços da comunidade científica no desenvolvimento de soluções de baixo custo e sinalizam a viabilidade para o uso como tecnologias assistivas para a Comunicação Aumentativa e Alternativa de pessoas com Esclerose Lateral Amiotrófica ou outras doenças que comprometem as habilidades funcionais. Os recursos computacionais relacionados às técnicas de visão computacional/aprendizado de máquina e ao suporte de hardwares para aquisição e aperfeiçoamento das imagens também foram examinados e descritos na Tabela 4.4, que sumariza as respostas sobre estes e os demais pontos investigados.

Os modelos computacionais identificados mostraram-se promissores nas tarefas de detecção da face, dos olhos e de rastreamento dos movimentos oculares ou classificação do estado do olho (aberto ou fechado), embora haja limitações quanto aos experimentos em pessoas com Esclerose Lateral Amiotrófica e, em alguns estudos, a densidade metodológica da estrutura e aplicação do modelo. Além dessas limitações, é importante destacar outra observação, a de que as técnicas computacionais alcançaram um limiar de eficiência (quanto a performance), isto é, que estão bem consolidadas para Interação Humano-Computador através dos olhos. Entretanto, vale lembrar que experimentos computacionais controlados com um número baixo e não diversificado de usuários podem mascarar os resultados reais, apresentando bons resultados nos testes, mas sem a capacidade de generalização do modelo. Estes aspectos podem ser explorados em novas pesquisas, bem como aspectos relacionados às abordagens sem o uso de câmera ajustada na cabeça do usuário ou infravermelho, o que pode direcionar novos testes considerando pessoas com Esclerose Lateral Amiotrófica sem causar desconfortos.

A proposta desta revisão sistemática da literatura de reunir os achados sobre abordagens de Interação Humano-Computador baseadas em imagens dos olhos para a Comunicação Aumentativa e Alternativa de pessoas com Esclerose Lateral Amiotrófica é essencialmente esperançosa e otimista quanto à inovação, ao desenvolvimento e disponibilização de tecnologias de baixo custo para o acesso universal e melhorias significativas na qualidade de vida de pessoas com Esclerose Lateral Amiotrófica ou outras deficiências motoras.

Capítulo 5

Eye-Blink em Tempo Real para ELA

Este capítulo apresenta o artigo intitulado “*Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam*”, em sua versão publicada no periódico internacional *Expert Systems with Applications*. O artigo apresenta uma solução acessível para a Interação Humano-Computador em tempo real baseada em imagens dos olhos, com especial atenção para o contexto de pessoas com Esclerose Lateral Amiotrófica (de Lima Medeiros et al. 2022).

de Lima Medeiros, Paulo Augusto, Gabriel Vinícius Souza da Silva, Felipe Ricardo dos Santos Fernandes, Ignacio Sánchez-Gendriz, Hertz Wilton Castro Lins, Daniele Montenegro da Silva Barros, Danilo Alves Pinto Nagem e Ricardo Alexandro de Medeiros Valentim (2022). ‘Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam’, *Expert Systems with Applications*, **188**, 116073. URL: <https://doi.org/10.1016/j.eswa.2021.116073>.

5.1 Desenho do Estudo

O sistema proposto foi projetado para lidar com tarefas de classificação do estado do olho e detecção de piscadas. A primeira tarefa é classificar o estado do olho na imagem do rosto como aberto ou fechado. As entradas nesta tarefa não necessariamente apresentam relação temporal ou sequencial e, para cada entrada, há uma saída esperada.

Por outro lado, a detecção de piscada é a tarefa de classificar o piscar dos olhos em um dado vídeo, ou seja, em uma sequência de *frames*. Como a entrada desta tarefa é um vídeo, os *frames* tem uma relação sequencial. Para cada entrada de vídeo, uma lista de ocorrências de piscadas é esperada como saída. A piscada é definida como uma sequência de *frames* com olhos fechados. Desta maneira, a piscada vai ser representada pelos limites da sua sequência correspondente. Nesse trabalho, nós tratamos a tarefa de detecção de *eye-blink* como uma extensão de classificação do estado do olho.

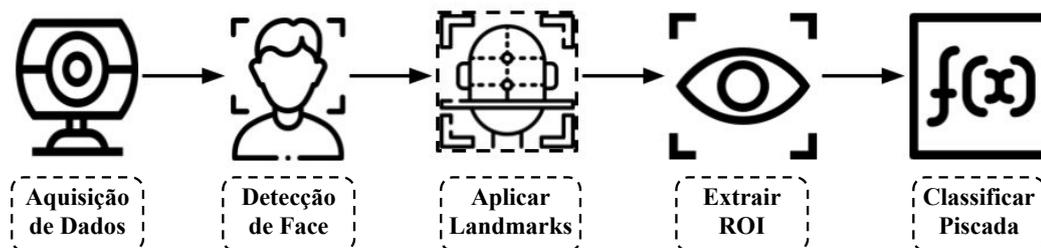
Primeiro, será apresentado o funcionamento do sistema para a tarefa de classificação do estado do olho e, em seguida, será mostrado como a arquitetura é estendida para detecção de *eye-blink*.

A arquitetura geral do sistema (Figura 5.1) pode ser dividida em etapas ou blocos sequenciais. O primeiro passo é detectar o rosto do usuário. Em seguida, os pontos de

referência são detectados na face atual. Esses pontos são usados para alinhamento facial e para extrair as coordenadas dos *patches* dos olhos e do olho. Como a extração do *patch* do olho pode falhar ao recuperar um *patch* ocular sem um olho real, um modelo de Redes Neurais Convolucionais (CNN, do Inglês *Convolutional Neural Network*) é usado para classificar essas extrações em imagens oculares ou imagens não oculares. A última etapa é a classificação do estado do olho e dois modelos foram desenvolvidos para esta tarefa de classificação. O primeiro modelo é baseado em uma CNN e classificará os *patches* dos olhos corretos em olhos abertos ou fechados, enquanto o segundo usa uma Máquina de Vetor de Suporte (SVM, do Inglês *Support Vector Machine*) para classificar as coordenadas oculares extraídas.

Esses dois modelos podem mapear, com baixa taxa de erro e baixa latência, uma imagem de *webcam* para uma classe (olhos abertos/fechados), resolvendo assim a tarefa de classificação do estado do olho. Além disso, o sistema proposto também será usado para detecção de piscar de olhos e, portanto, a dimensão temporal está presente. Dada essa nova dimensão e o alto desempenho do sistema proposto, alguns benefícios podem ser apontados. Essas características permitem o desenvolvimento de modelos auxiliares, como um filtro de janela deslizante a ser aplicado nos mapeamentos de quadros consecutivos, ou seja, saídas do classificador de estado do olho logo após sua computação, em vez de limitar as saídas dos classificadores de estado do olho logo após sua computação, a limitação é realizada somente após a filtragem. Esses filtros atuam como removedores de ruído e podem aprimorar o desempenho da classificação de cada quadro. Um segundo benefício diz respeito à resolução do problema de perder a detecção de faces rotacionadas. Um modelo auxiliar simples, mas eficaz, é apresentado para resolver esse problema como uma etapa de compensação de rotação.

Figura 5.1: Fluxograma do Sistema.



5.1.1 Definições

Nesta seção serão apresentados conceitos e terminologias que foram utilizadas ao longo do desenvolvimento do sistema.

5.1.1.1 Eye-Blink

Para este trabalho, uma piscada em um vídeo será definida como uma sequência de *frames* consecutivos com os olhos fechados. Portanto, os *frames* em que os olhos não estão totalmente fechados (*frames* em que os olhos estão fechando ou abrindo) não irão

compor o intervalo de piscadas. Esta definição é baseada na aplicação do sistema e é justificada a seguir.

Uma vez que uma piscada será usada como um sinal de comunicação e somente piscadas voluntárias serão consideradas (consulte a Seção 5.1.1.2), o usuário deve ter controle sobre a duração da piscada. Também foi consenso entre os autores que o usuário tem controle apenas sobre o tempo que fica com os olhos completamente fechados, não controlando, portanto, quanto tempo leva para fechar e abrir (transição entre olhos abertos-fechados) os olhos. Com isso, a escolha de definir o piscar conforme mencionado acima se adapta melhor ao aplicativo. Deve-se observar também que, seguindo esta definição, piscadas inacabadas não serão consideradas como estado do olho fechado, pois não há olhos fechados para esses eventos.

5.1.1.2 *Eye-Blink* Voluntário e Involuntário

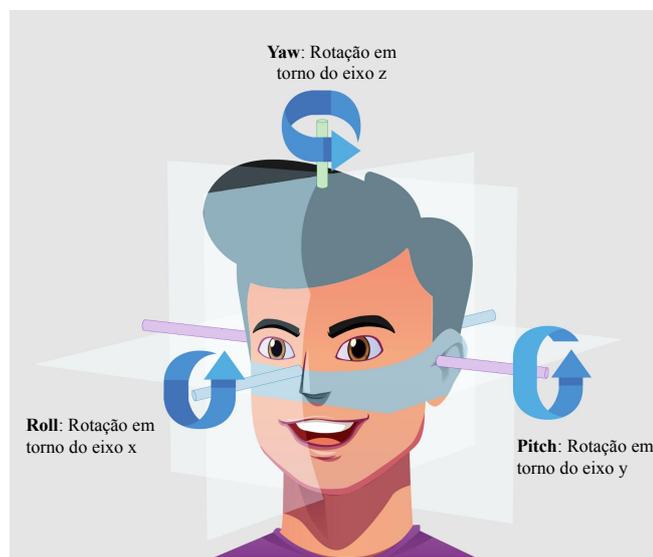
Na tarefa de detecção de piscadas, pode-se classificar as piscadas detectadas em uma de duas classes: voluntárias ou involuntárias. As abordagens encontradas para esta tarefa de classificação utilizaram apenas a duração da piscada como descritor (Singh e Singh 2018, Fathi e Abdali-Mohammadi 2015, Chau e Betke 2005). Se esta duração for superior a um limite predeterminado, a piscada é considerada voluntária. Caso contrário, é uma piscada involuntária. Portanto, as piscadas involuntárias e voluntárias também podem ser chamadas de piscadas curtas e longas, respectivamente. Em indivíduos saudáveis, a duração média do piscar de olhos voluntário foi relatada em $572 \pm 25\text{ms}$ (Kwon et al. 2013) e, portanto, períodos de tempo maiores que esse valor podem ser usados como indicadores de um piscar voluntário.

O sistema apresentado detecta um piscar de olhos e envia sua duração para o sistema de comunicação alternativa. Este, por sua vez, será então responsável por definir o limite de piscadas voluntárias. No entanto, recomenda-se que seja concedida uma margem de segurança ao valor escolhido. Essa margem é desejada porque, enquanto a falta de um sinal de ativação esperado requer apenas a repetição da ação de piscar por parte do usuário, levar em consideração um sinal indesejado pode desencadear uma ação indesejada do sistema.

5.1.1.3 Rotação da Cabeça

Em relação aos movimentos da cabeça, dado o espaço 3D, existem três eixos diferentes para rotação. Este trabalho usa a mesma terminologia exibida na Figura 5.2: a rotação em torno do eixo x é denominada *roll*; ao redor do eixo y é denominado *pitch*; ao redor do eixo z é chamado de *yaw* (guinada). Essas rotações afetam o sistema de maneira diferente e foram desenvolvidas abordagens para lidar com alguns desses impactos.

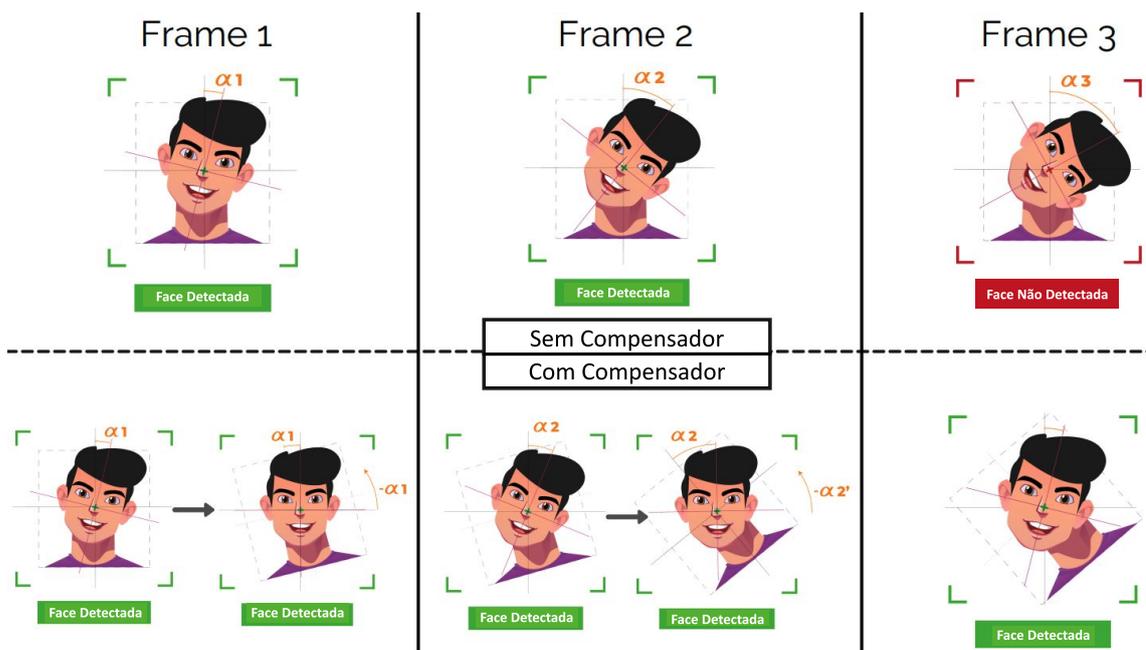
Figura 5.2: Movimentos de rotação da cabeça ao redor dos eixos x, y e z.



5.1.2 Compensação de Rotação Sequencial

Detectores faciais recentes são realmente robustos (Zhu et al. 2020). No entanto, falhas ainda ocorrem quando a face alvo é rotacionada exageradamente, já que esta não é uma entrada comumente esperada. Para diminuir o número de detecções perdidas devido a esse problema, um modelo auxiliar de compensação de rotação foi projetado para funcionar em uma sequência de quadros. Seu processo é representado na Figura 5.3.

Figura 5.3: Compensador de Rotação.



Dado que o i -ésimo *frame* f_i tem ângulo de rotação da face igual a α_i , pode-se compensar este ângulo de rotação em f_{i+1} girando-o em $-\alpha_i$. Assim, este novo ângulo de rotação da face é $\alpha'_{i+1} = \alpha_{i+1} - \alpha_i$. Dada a alta taxa de *frames* do sistema, espera-se que a variação do ângulo α'_{i+1} seja pequena. Assim, a chance de tal ângulo causar o problema relatado é menor. A rotação em f_{i+1} ocorre em torno do centro da face encontrado em f_i . Quando f_{i+2} for processado, ele será girado por $-\alpha_{i+1}$ antecipadamente, como uma etapa de compensação de rotação. Tal valor pode ser calculado usando $\alpha'_{i+1} = \alpha_{i+1} - \alpha_i$: α'_{i+1} é calculado pelo passo de alinhamento de face e α_i é o último valor usado pelo compensador de rotação.

Observe que, quando α_i e α_{i+1} representam ângulos de direções opostas, $|\alpha'_{i+1}| = |\alpha_{i+1}| + |\alpha_i|$, que pode ser prejudicial ao detector facial em valores mais altos. No entanto, dada a alta taxa de *frames* do sistema, ou seja, pequeno intervalo de tempo entre dois quadros, se quadros consecutivos tiverem ângulos de rotação de cabeça opostos, esses ângulos devem ser baixos.

Espera-se que este componente reduza a taxa de falhas do detector de face e também contribua para o alinhamento da face, fornecendo uma face menos enrolada para esta etapa. Se a etapa de detecção de rosto perder um quadro, não haverá compensação no próximo *frame* válido, ou seja, nenhuma compensação de rotação é feita.

5.1.3 Detecção de Face

A detecção de faces é uma área tradicional do campo de pesquisa em visão computacional e é um dos principais blocos do sistema proposto. Neste trabalho, a técnica chamada *Single Shot MultiBox Detector* (SSD) (Liu et al. 2015) é utilizada para esta tarefa. Este modelo é uma extensão das Redes Neurais Convolucionais.

Essas arquiteturas foram apresentadas pela primeira vez em LeCun et al. (1998), sendo posteriormente reintroduzidas em Krizhevsky et al. (2012), e são redes neurais especializadas que lidam com dados semelhantes a grade ou blocos (Goodfellow et al. 2016). Amplamente pesquisada e utilizada na literatura, é a arquitetura eleita da maioria dos modelos de ponta para tarefas de visão computacional, (Wang et al. 2019, Gu et al. 2019, Tan e Le 2019) já tendo superado o nível humano de desempenho em algumas tarefas (He et al. 2015).

Com um dado de entrada $x \in \mathbb{R}^{W \times H \times D}$, onde W , H e D representam a largura, altura e profundidade da imagem, respectivamente, as CNNs calcularão uma codificação $y \in \mathbb{R}^{D'}$ para x . Essas redes são construídas por uma sequência de blocos, composta por filtros convolucionais K , cujos pesos são aprendidos durante a fase de treinamento, seguidos por funções não lineares. Em cada um desses blocos, o objetivo é transformar a entrada reduzindo sua resolução espacial e expandindo sua profundidade. As representações intermediárias, saídas de tais blocos, são chamadas de mapas de características. Finalmente, essas redes podem ser pensadas como uma codificação não linear $f : \mathbb{R}^{W \times H \times D} \rightarrow \mathbb{R}^{D'}$ parametrizado por K .

Deve-se notar que as CNNs não podem detectar objetos por conta própria. Diferentes soluções, como YOLO (Redmon et al. 2016), *Faster R-CNN* (Ren et al. 2015) e SSD (Liu et al. 2015) foram propostas na literatura e, para este trabalho, a última está empregada.

Esta é uma arquitetura *single-short*, pois a localização e a classificação das detecções são feitas em uma única passagem direta. A arquitetura SSD usa uma CNN base e mapas de recursos específicos calculados por essa base são usados para localizar as caixas delimitadoras das detecções de candidatos. Um conjunto fixo de caixas delimitadoras está vinculado a cada célula nesses mapas de recursos. Os filtros convolucionais são aplicados a esses mapas de recursos específicos para calcular os deslocamentos das caixas e pontuações para cada classe. Como essas caixas delimitadoras se sobrepõem e o mesmo objeto pode ser detectado por mais de uma caixa delimitadora, uma supressão não máxima é executada para eliminar as caixas redundantes.

Para obter menor latência, convoluções separáveis em profundidade (Howard et al. 2017) são usadas neste trabalho. Esse tipo de bloco convolucional fornece um aprimoramento no desempenho do tempo e na contagem de parâmetros de modelos profundos ao custo de uma taxa de erro ligeiramente maior (Howard et al. 2017).

Além disso, Equalização de Histograma Adaptativo Limitada por Contraste (CLAHE, do Inglês *Contrast Limited Adaptive Histogram Equalization*) (Zuiderveld 1994) é aplicada a cada *frame* como uma etapa de pré-processamento, antes da detecção facial. Como são imagens RGB, o espaço de cores é primeiro convertido em YCbCr e, em seguida, essa técnica de equalização é empregada no componente *luma*. Por fim, o espaço de cores é reconvertido para o espaço de cores RGB.

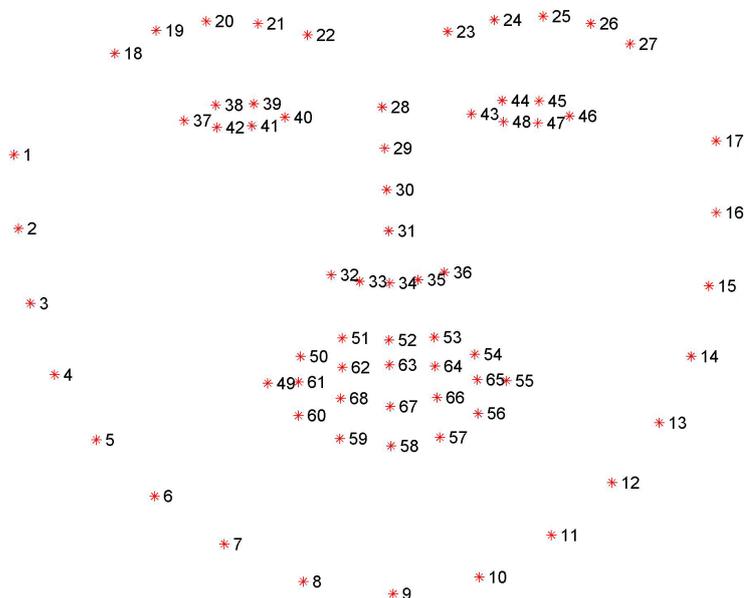
Quando nenhum rosto é encontrado na tarefa de classificação do estado do olho, o sistema é incapaz de produzir uma saída. O comportamento deste componente quando nenhum rosto é detectado na tarefa de detecção de piscar de olhos é descrito na Seção 5.1.9.

5.1.4 Alinhamento Facial

O alinhamento facial diz respeito à transformação de uma imagem facial para encaixá-la em um modelo facial geométrico. Seguindo outros trabalhos (Wang 2017, Li et al. 2018), o método usado no sistema apresentado também se baseia na extração de pontos de referência faciais frontais, recursos que também são usados em outras tarefas, como reconhecimento facial (Soltanpour et al. 2017), reconhecimento de gênero (Bekios-Calfa et al. 2014) e reconhecimento de emoção a partir da expressão facial (Martinez e Du 2012).

Os marcos faciais são pontos estratégicos cuidadosamente escolhidos para se assemelhar à estrutura facial. Assim, sua localização pode ser útil para o alinhamento da face e, como também utilizado neste sistema, para extração de coordenadas e *patch* da região ocular.

Figura 5.4: Marcos faciais usados no conjunto de dados IBUG.



Fonte: Sagonas et al. (2016).

Diferentes conjuntos de *landmarks* faciais podem ser encontrados na literatura, cada um deles definido por um conjunto de dados específico (Belhumeur et al. 2013, Sagonas et al. 2016, Le et al. 2012). Além disso, o mesmo conjunto de dados pode fornecer diferentes conjuntos de pontos, cada um deles com uma finalidade específica. Neste trabalho, foram utilizadas as anotações do conjunto de dados do IBUG (ver Figura 5.4) (Sagonas et al. 2016).

Para a extração de *landmarks*, foi utilizado o modelo apresentado por Kazemi e Sullivan (2014). Consiste em uma cascata de árvores de regressão, aprendidas com o algoritmo de aumento de gradiente. Para cada regressor na cascata, sua entrada é uma função da estimativa de forma do regressor passado, que gera melhores características para a árvore atual. Além disso, para reduzir a complexidade do espaço de busca, as estimativas de forma são assumidas como lineares e mantidas em um subespaço linear. Ao combinar esses pontos em seu trabalho, um novo preditor de forma em tempo ultra-real e com baixo erro foi concebido.

Para alinhar as faces em suas posições corretas, foi utilizado o algoritmo de Umeyama (1991). Este método apresenta uma solução para o problema de encontrar os parâmetros de transformação de similaridade que fornecem o menor erro quadrático médio entre dois conjuntos de pontos do mesmo tamanho. Um conjunto são os pontos de referência extraídos da imagem facial e o outro é a meta de alinhamento. O último foi recuperado calculando a média dos padrões de pontos extraídos do conjunto de dados YEC apresentado na Seção 5.1.10. Além disso, a técnica CLAHE foi aplicada a cada quadro antes de extrair os pontos de referência.

5.1.5 Patches Oculares e Extração de Coordenadas

Os pontos de referência extraídos também serão usados para coordenadas oculares e extração de *patches* oculares. O tamanho do *patch* ocular usado é 40×40 . Conforme representado na Figura 5.4, o olho esquerdo (*left*) e direito (*right*) são representados por $C_{left} = \{P_i \mid 43 \leq i \leq 48\}$ e $C_{right} = \{P_i \mid 37 \leq i \leq 40\}$, respectivamente. Estas são as coordenadas a serem extraídas. Este subconjunto também foi usado por Li et al. (2018) para extração do *patch* ocular.

Neste trabalho, os *patches* oculares podem ser extraídos de dois retângulos paralelos de eixos diferentes para cada olho: $R_{esquerda}^{min}$, $R_{esquerda}^{max}$, $R_{direita}^{min}$ e $R_{direita}^{max}$. Para o olho esquerdo, R_{left}^{max} é o menor retângulo contendo $C_{left} \cup \{P_{18}, P_{22}, P_{29}\}$, enquanto R_{left}^{min} usa apenas os pontos nos cantos dos olhos: P_{37}, P_{42} . Para este último, o menor retângulo que contém esses pontos é calculado e, em seguida, sua altura é nivelada à sua largura, mantendo o ponto médio original. Para o olho direito, as formas simétricas são extraídas: R_{right}^{max} usa $C_{right} \cup \{P_{23}, P_{27}, P_{29}\}$, enquanto R_{right}^{min} usa P_{43}, P_{46} . Esses retângulos foram projetados visando R_{left}^{min} para estar dentro de R_{left}^{max} e R_{right}^{min} dentro de R_{right}^{max} .

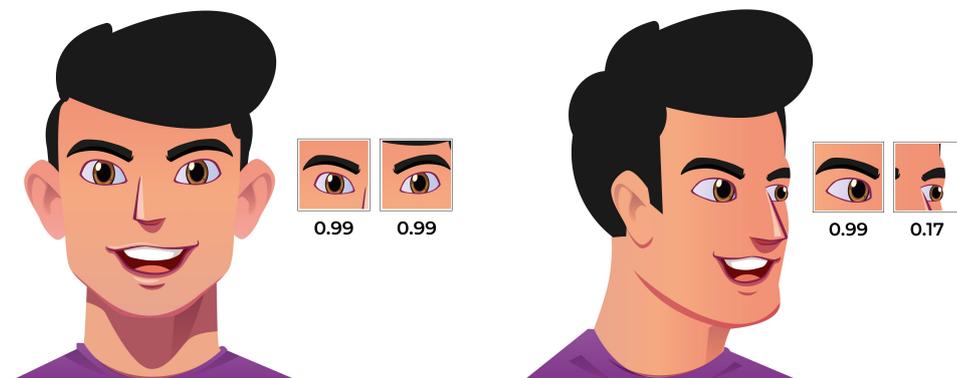
Ambos R_{right}^{max} e R_{left}^{max} apresentam uma área maior e, portanto, uma margem de segurança como *patche* ocular. Portanto, essas formas são mais interessantes para extração em cenários da vida real. Um uso mais atraente para R_{right}^{min} e R_{left}^{min} é mostrado na Seção 5.1.7.

Um modelo auxiliar baseado em uma CNN foi projetado para avaliar a qualidade (se um olho está realmente presente) dos *patches* oculares extraídos. Essa avaliação é crítica, pois os pontos de referência faciais são projetados para faces frontais e, portanto, *patches* oculares insatisfatórios podem ser extraídos de faces não frontais (resultantes de rolamento, guinada ou inclinação). Esses *patches* oculares provavelmente extraídos incorretamente sem olhos podem causar um comportamento inesperado, comprometendo o desempenho do sistema. A Figura 5.5 ilustra a utilização do modelo de avaliação.

Figura 5.5: Performance do avaliador ocular em diferentes situações..

(a) Exemplo de alta avaliação

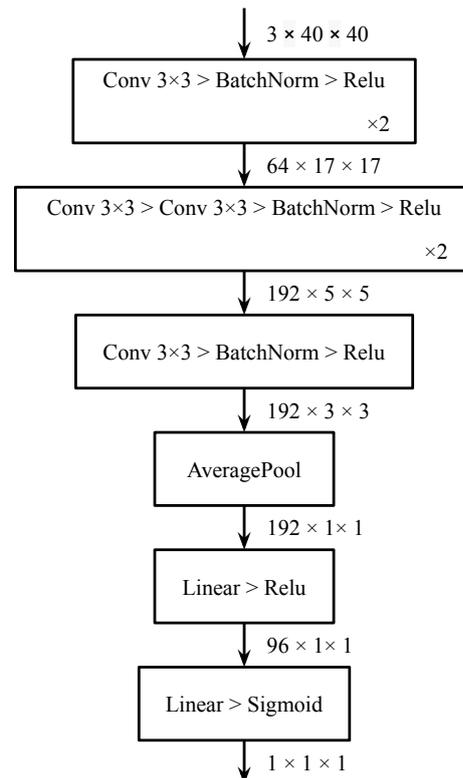
(b) Exemplo de baixa avaliação



Conforme apresentado, as CNNs são usadas como codificadores de imagem. Quando seguidas por uma sequência de camadas totalmente conectadas, essas redes podem ser usadas para classificação de imagens e alcançar resultados de ponta (Tan e Le 2019). O

modelo apresentado segue a arquitetura representada na Figura 5.6. Seu pequeno tamanho é crucial para manter a latência do sistema baixa. A entrada esperada são as manchas extraídas no canal de cor RGB (um tensor $3 \times 40 \times 40$) e o modelo mostra a probabilidade de tais manchas serem *patches* oculares adequados (um escalar para cada mancha).

Figura 5.6: Arquitetura para as CNNs utilizadas neste trabalho.



Um algoritmo simples (Algoritmo 1) foi projetado para este modelo auxiliar. Ele espera os *patches* oculares extraídos do *frame* atual como entrada e os filtra de acordo com dois limites de probabilidade fixos, *upper_thres* (limite superior) e *lower_thres* (limite inferior), *superior_thres* > *inferior_thres*. Observe que a saída do algoritmo é afetada pela tarefa que está executando. Como o *eye-blink* é uma tarefa de vídeo, o sistema pode optar por retornar None e, portanto, pular a etapa de classificação para o *frame* atual. A tarefa de classificação do estado do olho tem apenas uma imagem como entrada, portanto, nenhum salto é possível e pelo menos um *patch* deve ser retornado.

O Algoritmo 1 espera os *patches* extraídos como entrada e determina os melhores e os piores *patches* de acordo com suas avaliações (linhas 1-8). Posteriormente, é realizada uma análise de como as pontuações atribuídas a cada *patch* estão dentro dos intervalos induzidos pelos limiares (linhas 9-27).

- Se a melhor pontuação não for maior que *lower_thres*, a pior pontuação também será menor, devido à transitividade (linhas 9-16). Nesse caso, se a tarefa atual for a classificação do estado do olho, deve haver uma saída. Em seguida, o melhor

Algoritmo 1: Filter Frames

```

Entrada: left_patch, right_patch
Saída: valid patches for classification
1 best_patch ← left_patch
2 worst_patch ← right_patch
3 punct_best ← eval(best_patch)
4 punct_worst ← eval(worst_patch)
5 se punct_best < punct_worst então
6   | swap(best_patch, worst_patch)
7   | swap(punct_best, punct_worst)
8 fim
9 se punct_best ≤ lower_thres então
10  | se current_task = eye-state classification então
11  |   | retorna best_patch
12  |   fim
13  | senão se current_task = blink detection então
14  |   | retorna None
15  |   fim
16 fim
17 senão se punct_best ≤ upper_thres então
18  | retorna best_patch
19 fim
20 senão
21  | se punct_worst > upper_thres então
22  |   | retorna best_patch, worst_patch
23  |   fim
24  | senão
25  |   | retorna best_patch
26  |   fim
27 fim

```

patch é retornado. No entanto, se a tarefa for detecção de piscar de olhos, dada a propriedade temporal e a alta taxa de quadros do sistema, o quadro inválido pode ser facilmente ignorado e o possível piscar contínuo desconsiderado. Assim, None é retornado.

- Se a melhor pontuação estiver dentro do intervalo (*lower_thres*, *upper_thres*), apenas o melhor *patch* será retornado (linhas 17-18).
- Se apenas a melhor pontuação for maior que *upper_thres*, apenas esta pontuação será retornada. No entanto, se a pior pontuação também for maior que *upper_thres*, ambos os *patches* serão retornados (linhas 20-27).

5.1.6 Classificação do Estado do Olho

Dois classificadores diferentes foram usados neste trabalho para estimar o estado do olho a partir dos *patches* oculares previamente avaliados e possivelmente filtrados. O primeiro modelo é uma CNN que espera a figura de um olho como entrada e gera como

saída a probabilidade de ser um olho fechado. O segundo modelo é um SVM que toma como entrada as coordenadas X de um olho e também mostra a probabilidade de ser um olho fechado. Observe que, em cada quadro, dois *patches* oculares/coordenadas serão classificados, a menos que o passo apresentado na Seção 5.1.5 desqualifique um deles. Quando dois *patches*/coordenadas são usados como entrada, a média de suas saídas será usada para definir a saída do quadro, ou seja, sua probabilidade final.

O algoritmo SVM foi originalmente introduzido em 1963 por Vladimir N. Vapnik. Desde então, passou por atualizações que lhe permitiram classificar dados não lineares. O SVM visa encontrar o hiperplano que melhor discrimine as classes em um espaço N dimensional (N sendo o número de feições). Portanto, este algoritmo busca maximizar a distância entre o hiperplano e os pontos.

Os vetores de suporte são pontos de dados que estão mais próximos do hiperplano, sendo utilizados para maximizar a margem entre as classes e o hiperplano. Para a classificação não linear, o SVM utiliza o truque do *kernel*, método que aumenta a dimensão dos dados a classificar e separar as classes em um novo espaço.

Para construir um modelo SVM para classificação do estado do olho, 6 pontos de referência são usados para cada olho, resultando em 12 características no \mathbb{R}^2 para cada olho. A escala de dados foi modificada para variar de 0 a 1 usando os valores mínimo e máximo para cada variável, esta transformação é calculada como:

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (5.1)$$

Onde x representa um único vetor de recursos. Como as saídas do classificador representam a probabilidade do quadro conter olhos fechados, a classe correspondente ainda precisa ser atribuída se a tarefa for classificação do estado do olho. Para isso, um valor fixo T_{image} é usado para limiarização. Quando a tarefa é detecção de piscar de olhos, a limiarização pode ser feita somente após o processo descrito na Seção 5.1.8.

5.1.7 Aumentação de Dados

Para construir CNNs robustas, o aumento de dados foi fundamental durante a fase de treinamento desses modelos. Esta é uma técnica que tem sido amplamente utilizada devido aos seus impactos positivos (Shorten e Khoshgoftaar 2019). Como este sistema foi projetado para ser amplamente utilizado, diferentes condições de equipamentos e ambientes são esperadas durante a fase de aquisição de dados, ou seja, captura de *frames*. Assim, brilho, saturação e contraste foram modificados aleatoriamente para cada imagem de treinamento. Além disso, são esperadas várias configurações de exposição para captura de quadros, algo que está altamente relacionado ao ruído da imagem. Portanto, diferentes ruídos foram introduzidos aleatoriamente para simular essas variações.

A variação da proporção do quadro também foi usada para evitar enviesar o classificador para a mesma resolução. Este aumento é feito extraindo R_{left}^{min} e R_{right}^{min} e expandindo aleatoriamente essas formas independentemente em todas as quatro direções sem atravessar R_{left}^{max} e R_{right}^{max} .

5.1.8 Filtragem *Sliding-window*

Filtros de janela deslizante foram usados como modelos auxiliares para suprimir o ruído das saídas ainda sem limite dos classificadores, beneficiando-se da relação temporal entre *frames* consecutivos. Esperando que o processo de ruído tenha alta frequência e média zero (Rangayyan 2015), os filtros de média são, portanto, uma solução viável. São filtros passa-baixo, ou seja, preservam os sinais de baixa frequência e atenuam os de alta frequência. Além disso, a média do ruído amostrado tenderá a zero à medida que o tamanho da janela aumenta.

Para este trabalho, foram utilizadas duas versões ponderadas do filtro de média móvel: gaussiana e retangular. Além disso, o filtro de mediana também foi implementado devido à sua conhecida eficácia na remoção de ruído (Nussbaumer 1981). Os filtros foram aplicados às saídas dos classificadores para cada *frame*.

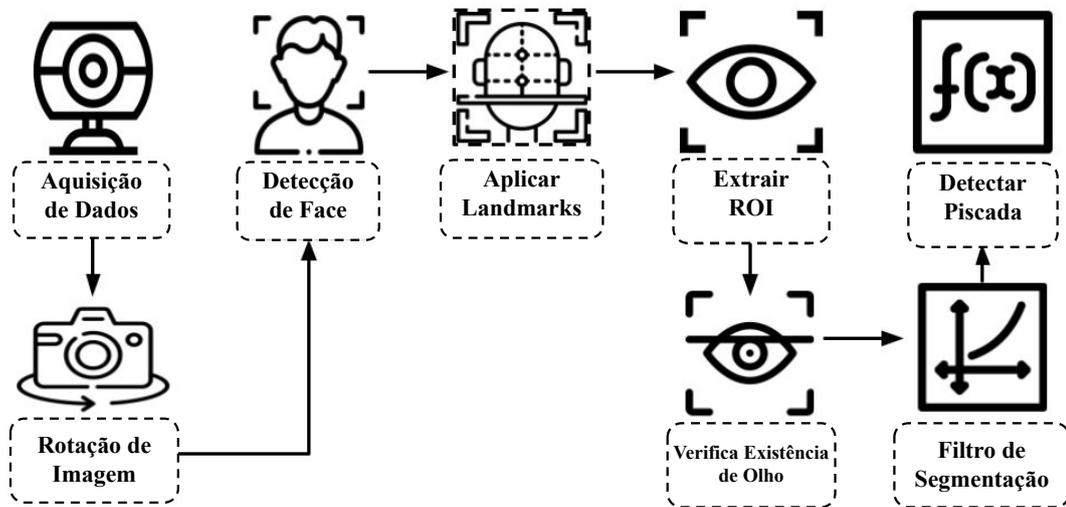
5.1.9 Detecção do *Eye-blink*

O modelo de detecção do *eye-blink* considerou as saídas de cada *frame*, já processadas pelos filtros propostos (Seção 5.1.8). Cada uma dessas saídas é um valor entre $[0; 1]$, que será binarizado por um limite T_{video} , implicando que o respectivo quadro apresenta um aberto (0) ou fechado (1) olho.

Seguindo as definições acima mencionadas (Seção 5.1.1), uma piscada é qualquer sequência de *frames* consecutivos cujas saídas são 1. Uma vez que apenas piscadas voluntárias são usadas como sinal de comunicação, um método para medir a duração da piscada precisa ser apresentado.

A duração do piscar pode ser medida de duas maneiras. Se uma taxa de transferência fixa (*frames* por segundo) for assumida, o número de *frames* de olhos fechados indicará a duração da piscada contínua. Portanto, um contador simples pode ser usado neste caso. Porém, em aplicações da vida real, o sistema estará exposto a variações de desempenho, inviabilizando a hipótese de *throughput* fixo. Portanto, os *timestamps* são uma maneira mais precisa de medir a duração de uma piscada. Marcar o início e o fim da piscada e depois calcular a diferença é uma maneira mais simples e precisa de medir a duração da piscada. Quando uma piscada é detectada, o sistema pode finalmente enviar sua duração para o sistema de comunicação alternativa. O sistema completo para detecção de *eye-blink* é representado na Figura 5.7.

Figura 5.7: Fluxo Geral do Sistema.



5.1.9.1 Falhas de Detecção da Face

As falhas de detecção de face são tratadas de maneira diferente do *eye-blink*, pois os *frames* podem ser pulados ou o posicionamento da cabeça do quadro anterior pode ser reutilizado. O Algoritmo 2 descreve o processo de detecção de face e o tratamento de detecções perdidas, que é sensível ao uso do compensador de rotação e do avaliador de *patches* oculares.

Este algoritmo depende do avaliador do *patch* do olho no sentido de que, quando ativo (*use_eye_eval* é definido como *true*), este modelo auxiliar pode repetir a posição facial detectada anteriormente (*last_face_pos*) e então será responsável por decidir se o quadro deve ser pulado (no caso de não extrair *patches* oculares válidos) conforme explicado na Seção 5.1.5.

Além disso, se o modelo auxiliar do compensador de rotação estiver ativo (*use_rc* é definido como *true*), a detecção perdida é seguida pela reexecução do detector facial no quadro não girado (*unrot_frame*). Se uma face for encontrada, ela será usada nas etapas seguintes e o compensador de rotação será zerado (*rc.reset()*). Caso contrário, se o avaliador de *patch* ocular estiver inativo, o *frame* é pulado, o que também acontece quando o compensador de rotação não é utilizado. Observe que o detector facial processa o *frame* não girado apenas quando necessário, não prejudicando assim as latências dos modelos.

Algoritmo 2: Erro de Detecção de Face

```

Entrada: last_face_pos
1 unrot_frame ← capture_frame()
2 frame ← unrot_frame
3 se use_rc então
4   | frame ← compensate(unrot_frame)
5 fim
6 curr_face_pos ← detect_face(frame)
7 se curr_face_pos == None então
8   | se use_rc and not rc.is_reseted() então
9     | unrot_face_pos ← detect_face(unrot_frame)
10    | se unrotated_face_pos == None então
11      | se use_eye_eval então
12        |   | retorna last_face_pos
13        |   | fim
14        |   | senão
15        |   |   | rc.reset()
16        |   |   | retorna None
17        |   |   | fim
18        |   | fim
19        |   | senão
20        |   |   | rc.reset()
21        |   |   | retorna unrot_face_pos
22        |   |   | fim
23        |   | fim
24        |   | senão
25        |   |   | se use_eye_eval então
26        |   |   |   | retorna last_face_pos
27        |   |   |   | fim
28        |   |   |   | senão
29        |   |   |   |   | retorna None
30        |   |   |   |   | fim
31        |   |   |   | fim
32 fim

```

5.1.9.2 Condições de Uso

Uma *webcam* genérica e um computador são os requisitos mínimos para usar o detector fornecido. A câmera deve ser posicionada na frente do rosto do usuário para que todo o rosto esteja presente no quadro extraído. As pessoas com Esclerose Lateral Amiotrófica passavam grande parte do dia sentadas (Karam et al. 2016, Güell et al. 2013, Soriani e Desnuelle 2017), de modo que a câmera poderia ser posicionada sobre uma mesa ou em um suporte preso a suas cadeiras. O único requisito real é que uma face frontal seja detectada. Na Seção 5.2.2, o sistema será avaliado em diferentes conjuntos de dados, sob diferentes resoluções, iluminação e condições do usuário.

5.1.10 Conjuntos de Dados Propostos

Diferentes conjuntos de dados já foram propostos para avaliar a tarefa de detecção de *eye-blink* (Drutarovsky e Fogelton 2014, Pan et al. 2007, Fogelton e Benesova 2016). No entanto, a maioria deles não atende aos propósitos deste trabalho no que diz respeito a questões de anotação e duração do piscamento voluntário. Portanto, para desenvolver e avaliar os modelos projetados neste trabalho, dois conjuntos de dados foram construídos: um conjunto de dados de treinamento e outro de teste.

O conjunto de dados de treinamento, denominado Classificação do Estado do Olho (YEC, *Youtube Eye-state Classification*), foi construído extraindo *frames* de vídeos do *youtube*, enquanto o conjunto de dados de teste, denominado *Autonomus Blink Detection* (ABD), consiste em gravações de vídeo de voluntários - com o devido consentimento das partes. O conjunto de dados YEC contém imagens faciais e é anotado com os estados oculares dessas faces. Assim, ele será usado para treinar a tarefa de classificação do estado do olho (que pode ser estendida para a tarefa de detecção do piscar de olhos). O conjunto de dados ABD contém vídeos em que cada *frame* possui as anotações do estado do olho e, portanto, será usado para avaliar a tarefa de detecção de piscar de olhos. As piscadas foram anotados neste conjunto de dados seguindo nossas definições apresentadas na Seção 5.1.1.

5.1.10.1 Conjunto de Dados YEC

O conjunto de dados YEC foi construído extraindo faces dos vídeos apresentados no conjunto de dados *AVSpeech* (Ephrat et al. 2018) e anotando o estado dos olhos desses rostos. O conjunto de dados *AVSpeech* consiste em uma coleção de links para vídeos do *Youtube* e em cada um dos vídeos há um segmento no qual uma única pessoa visível fala.

Para cada segmento de vídeo de interesse, foi realizada uma etapa de pré-anotação na qual, para cada *frame*, foram executadas a detecção e o alinhamento da face, foram extraídos *patches oculares* e anotados os estados dos olhos usando o modelo apresentado em Li et al. (2018). Posteriormente, foi realizada uma etapa final de anotação na qual pré-anotações incorretas foram corrigidas manualmente. Embora o próprio conjunto de dados *AVSpeech* forneça as anotações de posição facial, o detector de face *UltraLight*¹ foi usado.

Para construir um conjunto de dados balanceado, buscou-se extrair o mesmo número de *frames* para cada classe (olhos abertos/fechados) de cada vídeo. Além disso, como a diversidade e a imparcialidade são direcionadas, não foram usados mais de 20 *frames* de cada vídeo, o que aumenta a variabilidade.

Portanto, diferentes faces com diferentes resoluções foram extraídas para formar este conjunto de dados de treinamento. No total, foram extraídas 15.052 faces de olhos fechados e 28.998 faces de olhos abertos, totalizando 44.050 imagens faciais.

¹<https://github.com/Linzaer/Ultra-Light-Fast-Generic-Face-Detector-1MB>

5.1.10.2 Conjunto de Dados ABD

O protocolo experimental foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal do Rio Grande do Norte (UFRN), Natal, Brasil, por meio da carta CAAE No.25687819.3.0000.5537, e de acordo com os Acordos de Helsinki (modificado em 2004). O ABD foi construído gravando vídeos de voluntários saudáveis com segmentos de piscadas dos olhos, com uma webcam integrada a um *laptop*. Os sujeitos ficaram sentados a uma distância de aproximadamente 70cm em frente ao laptop, olhando para a câmera, simulando condições reais de uso. Os vídeos foram capturados em uma sala fechada com luz controlada.

A resolução de gravação foi de 640×480 pixels a 30 *frames* por segundo. Cada vídeo tem 40 segundos de duração e os voluntários foram instruídos a piscar voluntariamente exatamente duas vezes. Tais piscadas ocorreram em instantes aleatórios equiprováveis entre 10 e 35 segundos das gravações, com cada piscada voluntária durando 2 segundos. Não houve restrições quanto a piscadas involuntárias. As piscadelas foram anotadas seguindo nossa definição (Seção 5.1.1.1), ou seja, apenas *frames* com os olhos completamente fechados compuseram as piscadelas. Durante todo esse processo, os participantes foram orientados a evitar movimentos excessivos.

No total, 10 pessoas participaram do experimento. Todos os sujeitos gravaram vídeos sem óculos e 7 com óculos. Quanto às características dos sujeitos, 7 eram do sexo masculino e 3 do sexo feminino. Além disso, 5 eram de pele branca e 5 de pele parda. A idade média era de 23,1 anos.

Para anotar os vídeos gravados, foi desenvolvido um *pipeline* simples. Primeiro, uma etapa de pré-anotação foi executada para o conjunto de dados YEC (usando o mesmo modelo de classificação do estado do olho). Em seguida, para cada vídeo, foi realizada a anotação manual. Foi desenvolvido um roteiro composto por três etapas: (I) validação da pré-anotação em cada *frame*; (II) correção dos *frames* anotados errados; (III) análise de possíveis sequências inconsistentes, nas quais um pequeno número de *frames* é dissonante de seus arredores.

Todo o conjunto de dados consiste em 17 vídeos. Ao todo, são 16.921 imagens com olhos abertos e 3.054 imagens com olhos fechados, totalizando 19.975.

5.2 Experimentos e Resultados

Para avaliar o desempenho dos modelos apresentados nos conjuntos de dados propostos, diferentes experimentos de avaliação foram conduzidos.

Uma vez que a tarefa de detecção do piscar de olhos foi abordada como uma extensão da tarefa de classificação do estado do olho, surge a necessidade de avaliar nosso sistema na última tarefa (Seção 5.2.1). O desempenho do modelo base e o impacto causado pela adição do modelo de avaliação do patch ocular descrito na Seção 5.1.5 ao *pipeline*, ou seja, Avaliador Ocular (EE, do Inglês *Eye Evaluator*), foram avaliados tanto no conjunto de dados ABD quanto no conjunto de dados “Olhos Fechados na Natureza” (CeW, do Inglês *Closed Eyes in the Wild*) (Song et al. 2014).

Quanto aos experimentos com a tarefa de detecção de piscar de olhos, algumas considerações foram levantadas sobre a diferença entre as anotações de piscar de olhos nos diferentes conjuntos de dados usados. Esperava-se que tais diferenças impactassem os resultados.

Primeiramente, de forma semelhante à tarefa de classificação do estado do olho, foram avaliados o desempenho do modelo base e o impacto causado pela adição dos modelos auxiliares EE e Compensador de Rotação (RC, do Inglês *Rotation Compensator*) ao *pipeline*. Para cada combinação possível, foi encontrada a configuração do filtro móvel que maximiza a pontuação da avaliação. Em seguida, um experimento semelhante foi conduzido, mas considerando apenas as piscadas voluntárias. Além disso, foi realizado um estudo para analisar o impacto de diferentes configurações de filtro no desempenho do sistema. Por fim, foi realizada uma análise de latência do modelo final e uma comparação com trabalhos relacionados.

Conforme declarado nas Seções 5.1.3 e 5.1.4, o sistema proposto é construído com um detector de face e um extrator de marcos. O modelo *UltraLight* implementará a etapa de detecção facial. Este é um modelo de última geração em tempo real. Para a etapa de extração dos marcos, foi utilizado *Dlib* (King 2009).

Os classificadores CNN e SVM, apresentados na Seção 5.1.6, foram denominados *Eye Classifier* CNN (ECC) e *Eye Classifier* SVM (ECS), respectivamente. O RC, apresentado na Seção 5.1.2, e os filtros de janela (Seção 5.1.8) foram analisados em relação à tarefa de detecção do piscar de olhos. As combinações de tais modelos são descritas concatenando os nomes dos modelos com o símbolo '+’.

A implementação das arquiteturas varia dependendo do modelo. O modelo ECS foi implementado usando *Scikit Learning* (Pedregosa et al. 2011) com $C = 1$. O *kernel* utilizado foi `rbf` com $\gamma = 'scale'$. Ambos ECC e EE foram implementados e treinados usando a biblioteca *pytorch* (Paszke et al. 2017). O *batch-size* foi 32 e foi utilizado o otimizador *Adadelta* (Zeiler 2012). Os modelos foram treinados por 180 épocas cada. Os modelos ECS, ECC e EE foram treinados no conjunto de dados YEC (Seção 5.1.10). Este conjunto de dados não é usado para testes, portanto, nenhum viés pode afetar os resultados.

Dado que o conjunto de dados YEC é constituído por imagens faciais e que o modelo EE espera um candidato a *patch* ocular como entrada, a fase de treinamento desse modelo teve que ser conduzida de maneira diferente. Assim, em cada iteração de treinamento deste modelo, foi selecionado aleatoriamente e equiprovavelmente se um olho ou uma amostra negativa (sem *patch* ocular) deveria ser usado como entrada. Tais amostras negativas foram extraídas da metade superior da imagem sem interseção com a região dos olhos, sendo estas computadas pelo extrator de marcos (Seção 5.1.5). As amostras positivas foram retiradas da região dos olhos. O modelo de EE treinado alcançou uma pontuação de Área sob a Curva ROC (AUC) de 0,996 em um subconjunto do conjunto de dados YEC que não foi usado para treinamento. $Thres_A$ e $Thres_B$, os limites para filtragem válida de *patch* ocular (Seção 5.1.5), foram definidos para 0,54 e 0,2, respectivamente.

5.2.1 Experimento de Classificação do Estado do Olho

Neste experimento, todas as combinações possíveis de modelos ECC e ECS com o modelo auxiliar EE foram analisadas quanto à acurácia, precisão, *recall* e AUC na tarefa de classificação de imagem, ou seja, classificação do estado do olho. Dois conjuntos de dados foram usados: CeW (Song et al. 2014) e $ABD_{imagens}$, uma adaptação do conjunto de dados ABD.

O conjunto de dados CeW é desafiador para a classificação do estado do olho na natureza, tendo o mundo real irrestrito como cenário. Ambas as variações de iluminação e ângulo, bem como a oclusão, estão presentes. É um conjunto de dados balanceado contendo faces de pessoas com olhos abertos ou fechados. Há um total de 2.423 imagens.

O conjunto de dados ABD_{images} é um subconjunto de dados de imagens construído a partir do conjunto de dados ABD original. Para o vídeo i -ésimo no conjunto de dados ABD, considerou-se que ele possui c_i imagens de olhos fechados e o_i imagens de olhos abertos. Uma vez que $c_i < o_i$ e um conjunto de dados balanceado foi buscado, c_i imagens de olhos fechados e c_i imagens de olhos abertos são extraídas sem substituição. O conjunto de dados final é assim balanceado com amostras de 3.054 para cada classe. Comparado com o conjunto de dados CeW, este é um conjunto de dados muito menos desafiador.

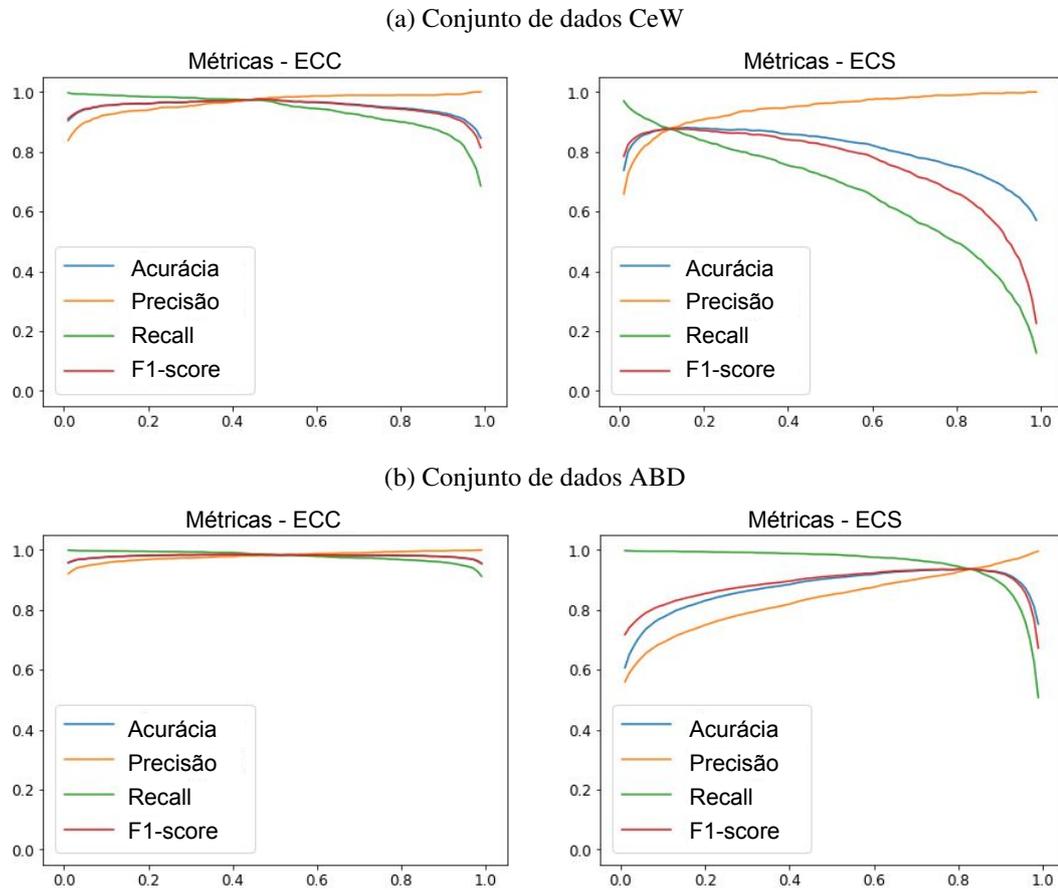
As métricas que foram usadas para este experimento de avaliação dependem dos possíveis resultados desta tarefa de classificação binária, que são:

- Verdadeiro Positivo (TP): Olhos fechados classificados como olhos fechados
- Falso Positivo (FP): Olhos abertos classificados como olhos fechados
- Verdadeiro Negativo (TN): Olhos abertos classificados como olhos abertos
- Falso Negativo (FN): Olhos fechados classificados como olhos abertos

5.2.1.1 Resultados

Acurácia, precisão, *recall* e *F1-Score* foram plotados em função do valor limite na Figura 5.8, variando dentro do intervalo de $(0, 1)$ para os classificadores do estado do olho. Os resultados dos modelos ECC+EE e ECS+EE são plotados para os conjuntos de dados CeW e ABD_{images} .

Figura 5.8: Pontuações da tarefa de classificação do estado do olho em função da variação do limiar.



Nota-se que as curvas do modelo ECC+EE são mais estáveis à variação do limiar e apresentam escores mais elevados do que as traçadas para o modelo ECS+EE. Essa maior estabilidade está relacionada à capacidade do classificador binário de separar amostras positivas de amostras negativas. Além disso, ao comparar as curvas da Figura 5.8a com as curvas da Figura 5.8b, é possível observar que ambos os modelos obtêm melhores resultados no $ABD_{imagens}$ conjunto de dados, o que confirma que a classificação do estado do olho neste conjunto de dados é uma tarefa menos desafiadora. Além disso, como esse conjunto de dados, conforme declarado na Seção 5.1.10.2, se aproxima das condições de uso esperadas do modelo, esses resultados indicam o desempenho do sistema nessas condições.

Tabela 5.1: Pontuações AUC e melhor acurácia (BA) para modelos ECC e ECS em combinação com o modelo EE para classificação do estado do olho.

Modelos	CeW		ABD _{images}	
	BA	AUC	BA	AUC
ECS	0.9050	0.9571	0.9484	0.9571
ECS+EE	0.9063	0.9569	0.9484	0.9862
ECC	<u>0.9744</u>	0.9966	<u>0.9805</u>	0.9966
ECC+EE	<u>0.9744</u>	<u>0.9967</u>	<u>0.9805</u>	<u>0.9982</u>

A Área sob a Curva ROC (AUC) e os melhores valores de acurácia obtidos são mostrados na Tabela 5.1. Apesar de todas as variações apresentarem ótimas pontuações, o modelo ECC apresentou melhores valores que os modelos ECS, apresentando seus melhores resultados quando combinado com o modelo EE. Para o conjunto de dados CeW, o modelo ECC+EE foi capaz de diminuir a taxa de erro em 72,7% em comparação com o modelo ECS+EE.

Os modelos ECC e ECS diferem tanto ao nível da extração de características como nas arquiteturas utilizadas. Enquanto o último depende muito do desempenho do extrator de marcos, o primeiro se beneficia de uma tolerância a erros que é introduzida quando os *patches* dos olhos são extraídos usando a margem proposta (Seção 5.1.5). Em relação às diferentes arquiteturas, as CNNs são conhecidas por sua superioridade no processamento de imagens. Além disso, o modelo ECC se beneficia do processo de aumento de dados apresentado na Seção 5.1.7. Tal técnica permitiu o desenvolvimento de um modelo menos suscetível à variância, podendo alcançar ótimos resultados mesmo no conjunto de dados CeW. Assim, o desempenho superior do modelo ECC em relação ao ECS já era esperado.

5.2.2 Experimentos em Vídeos

Esses experimentos se concentraram em testar a eficiência do sistema na tarefa de detecção de piscar de olhos, avaliando *precision*, *recall*, and *F1-score*. Quatro conjuntos de dados foram usados: o conjunto de dados ABD original, *Eyeblink8* (Fogelton e Benesova 2016), *ZJU* (Pan et al. 2007) e *Talking Face*². O conjunto de dados *Eyeblink8* consiste em gravações de 4 sujeitos caucasianos diferentes, cada um aparecendo em 2 vídeos. Os vídeos foram capturados em resolução 630×480 e contêm 70.992 *frames* no total. O conjunto de dados *ZJU* contém 80 vídeos de curta duração, capturados em resolução 320×240 a 30 FPS. 20 sujeitos gravaram 4 vídeos cada, variando a posição da câmera e o uso dos óculos. Os vídeos têm, em média, 136 frames. O conjunto de dados *Talking Face* consiste em um único vídeo do rosto de um homem caucasiano falando na frente de uma câmera. O vídeo é gravado a 25 FPS em resolução 720×576 , com duração de 200 segundos para um total de 5000 *frames*.

Em todos esses conjuntos de dados, as condições de luz, fundo e câmera são estáveis. Os sujeitos estão em sua maioria parados, apresentando poucos movimentos de cabeça,

²https://personalpages.manchester.ac.uk/staff/timothy.f.cootes/data/talking_face/talking_face.html

nenhum deles exagerado. Além disso, não há distinção entre o piscar de olhos voluntário e involuntário. As anotações das piscadas usadas para esses conjuntos de dados são as fornecidas em (Fogelton e Benesova 2016).

5.2.2.1 Calculando as Métricas

A *precisão* e o *recall* podem ser calculadas dentro de um determinado limite, analisando as piscadas corretamente detectadas, que podem ser encontrados calculando o valor de Interseção sobre a União (IoU) entre as piscadas detectadas e verdadeiras (Fogelton e Benesova 2016). A função IoU, bastante conhecida por seu uso em tarefas de detecção de objetos em imagens (Lin et al. 2014, Everingham et al. 2010), é calculada como a razão de duas áreas: a interseção entre uma detecção e uma anotação sobre a área da união entre esses elementos. No entanto, enquanto as imagens estão em duas dimensões, os intervalos de piscada estão em apenas uma dimensão. Portanto, as áreas devem ser substituídas por comprimentos.

Para calcular o valor IoU, vamos denotar a uma piscada anotada e d uma piscada detectada. Além disso, vamos definir $b, e : \text{blink} \rightarrow \mathbb{N}$ para recuperar suas posições inicial e final, respectivamente. Os comprimentos I e U da interseção e união entre a e d podem ser calculados da seguinte forma:

$$I_{a,d} = \max(\min(e(a), e(d)) - \max(b(a), b(d)), 0) \quad (5.2)$$

$$U_{a,d} = (e(a) - b(a)) + (e(d) - b(d)) - I_{a,d} \quad (5.3)$$

Agora, um algoritmo quadrático é capaz de combinar piscadas detectadas e anotadas usando essa pontuação. Para o `blink` detectado d , encontre $\text{match}_d = \arg \max_a \text{IoU}(a, d)$ ou seja, o `blink` anotado que maximiza as pontuações IoU para este `blink`.

Se $\text{IoU}(\text{match}_d, d) > 0,2$, d foi detectado corretamente, ou seja, é um TP. Caso contrário, é um FP. Finalmente, todos os `blinks` anotados que não foram associados a um TP são FN. Em outras palavras, TP são piscadas que ocorrem na saída de anotação e detecção e são correspondidas, FP são piscadas que são apenas detectadas, mas não correspondem à anotação e FN são piscadas anotadas que não foram combinadas com nenhuma detecção. *Precision*, *recall* e *F1-score* podem, portanto, ser calculados usando os valores encontrados de TP, FP e FN.

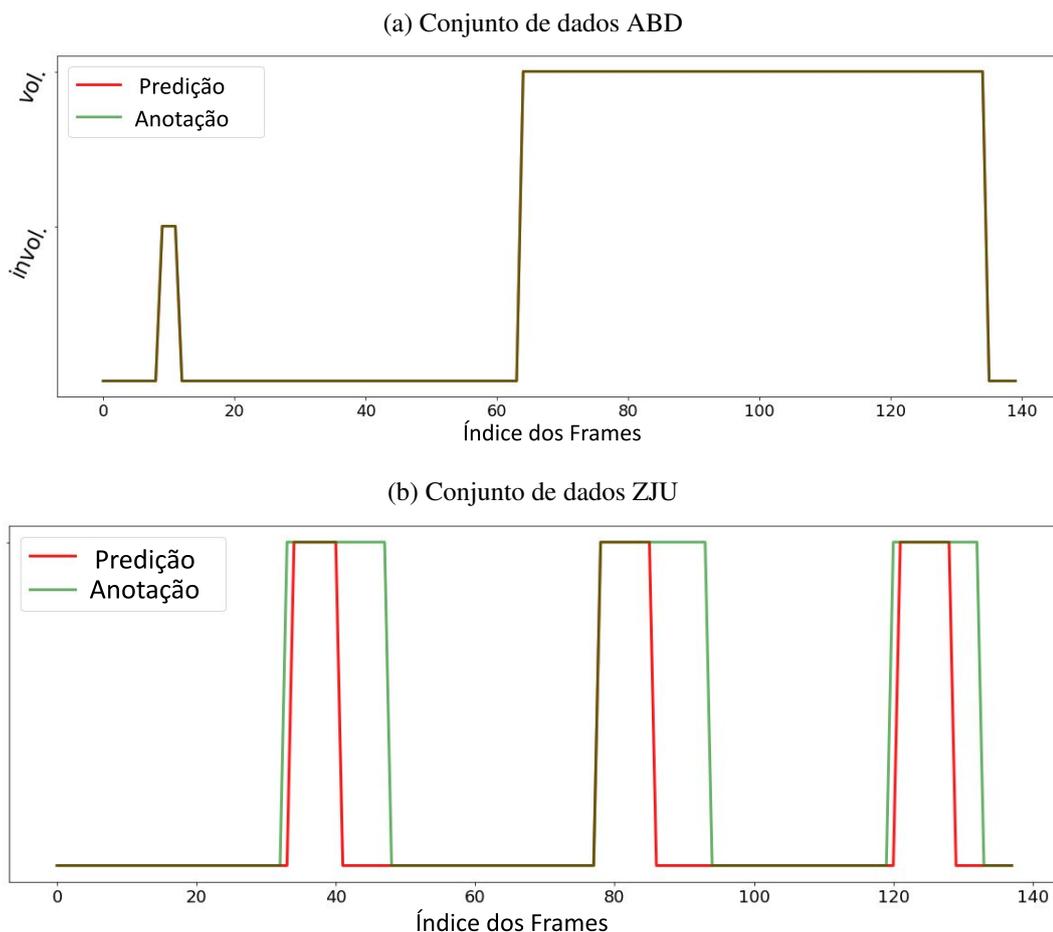
5.2.2.2 Análise do alcance do Piscar de Olhos

Antes de passar para os resultados principais, é preciso observar que os conjuntos de dados usados neste experimento têm origens diferentes. Portanto, as regras de anotação e as definições de piscadas não são necessariamente as mesmas. Percebeu-se que, embora as piscadas no conjunto de dados ABD estejam de acordo com nossa definição (Seção 5.1.10), os intervalos de piscadas em outros conjuntos de dados tendem a começar no início do fechamento dos olhos e terminar apenas quando eles são completamente abertos. Assim, as piscadas no conjunto de dados ABD têm, por definição, um alcance mais estreito do que aqueles anotados por (Fogelton e Benesova 2016).

Dado que nossos modelos também seguem nossa definição de piscada, espera-se que eles se ajustem melhor ao intervalo de piscadas no conjunto de dados ABD. Foi feita uma comparação entre o desempenho de ECC+EE+RC em partes dos conjuntos de dados ABD (Figura 5.9a) e ZJU (Figura 5.9b), conforme mostrado na Figura 5.9. Essas figuras foram construídas extraíndo segmentos de vídeo dos conjuntos de dados e exibindo valores de previsão e anotação nesses vídeos. Do ZJU foi utilizado um vídeo inteiro e, para manter a mesma escala, foi extraído do ABD um segmento de vídeo com a mesma contagem de quadros. Para o conjunto de dados ABD, a classificação como voluntária ou involuntária é mostrada no eixo y; o conjunto de dados ZJU não possui tais anotações.

O contraste entre o intervalo de piscadas adotado nas anotações ZJU e ABD é visível ao examinar a Figura 5.9, que destaca a diferença de alinhamento entre anotação e previsão entre os conjuntos de dados. O leitor deve lembrar que os modelos apresentados (incluindo ECC+EE+RC) foram treinados no conjunto de dados YEC, um conjunto de dados de imagem, não apresentando viés para o conjunto de dados ABD.

Figura 5.9: Análise do alcance do piscar de olhos.



5.2.2.3 Avaliação em Diferentes Conjuntos de Dados

Este experimento avalia o desempenho de todos os modelos propostos, explorando todas as possíveis configurações de filtro (assim como configurações sem filtro), em todos os conjuntos de dados de vídeo. Ela foi conduzida com o objetivo de explorar as melhores pontuações possíveis que podem ser alcançadas com as variações do modelo. Uma vez que todas essas combinações dependem de um classificador binário, limites dentro do intervalo de $[0,01;0,99]$ foram avaliados e apenas o melhor *FI-Score* alcançado a partir deles é mostrado e usado como parâmetro de comparação.

Como mais de uma configuração de filtro pode resultar na mesma pontuação máxima, apenas a mais simples é mostrada. Para tanto, o primeiro critério de simplicidade diz respeito à função janela (a ausência de filtros é mais simples que a retangular, que é mais simples que a gaussiana, que é mais simples que a mediana). E o segundo critério diz respeito ao tamanho da janela: quanto menor, mais simples. Os resultados são mostrados na Tabela 5.2.

Tabela 5.2: Resultados para ECS e ECC combinados com modelos auxiliares para a tarefa de detecção de piscar de olhos em todos os conjuntos de dados.

Modelos	Eyeblink			ZJU			Talking Face			ABD		
	<i>FI-Score</i>	$W_{function}$	W_{size}									
ECS	0.5582	Median	3	0.6940	Rect.	7	0.8640	Rect.	5	0.6429	Median	3
ECS+EE	0.5570	Median	3	0.6896	Rect.	7	0.8640	Rect.	5	0.6429	Median	3
ECS+RC	0.5507	Median	3	0.6905	Rect.	7	0.8760	Rect.	5	0.6460	Median	3
ECS+EE+RC	0.5532	Median	3	0.6905	Rect.	7	0.8760	Rect.	5	0.6460	Median	3
ECC	0.8656	Rect.	7	0.9219	Rect.	5	<u>0.9500</u>	Rect.	5	0.9256	-	-
ECC+EE	0.8624	Rect.	7	<u>0.9237</u>	Rect.	5	<u>0.9500</u>	Rect.	5	0.9256	-	-
ECC+RC	<u>0.8697</u>	Rect.	7	0.9140	Rect.	5	<u>0.9500</u>	Rect.	5	<u>0.9263</u>	Gaussian	3
ECC+EE+RC	0.8641	Rect.	7	0.9140	Rect.	5	<u>0.9500</u>	Rect.	5	0.9241	Gaussian	3

Inicialmente, pode-se perceber que os tamanhos das janelas que maximizam o *FI-Score* são pequenos (no máximo 7) e, também, funções de janela mais simples requerem um tamanho maior para atingir os escores máximos. Além disso, independentemente dos modelos auxiliares (EE, RC) utilizados, a mesma configuração de filtro maximizará a pontuação dos modelos em um determinado conjunto de dados. Assim, a configuração do filtro depende mais do modelo base (ECC, ECS) e conjunto de dados em questão do que da combinação dos modelos auxiliares analisados.

Além disso, como mostrado na Seção 5.2.1, o modelo ECC obteve melhores resultados do que o ECS. Como a tarefa de detecção do piscar de olhos é apenas uma extensão da tarefa de classificação do estado do olho, os modelos ECC e ECS são usados da mesma maneira. Assim, as razões apresentadas na Seção 5.1.5 (diferentes pipelines e arquiteturas de extração de características) para a superioridade do modelo ECC sobre o ECS também podem ser aplicadas aqui. As reduções de erro entre os conjuntos de dados obtidos comparando o modelo ECC com o ECS variaram entre 59,68% (conjunto de dados *Talking Face*) e 79,19% (conjunto de dados ABD), alcançando ótimos resultados em todos os quatro conjuntos de dados. Dado que os diferentes conjuntos de dados apresentam várias posições de câmera, luminância, resolução e etnia do usuário, os autores acreditam que o modelo ECC e suas variações apresentam uma solução robusta para a tarefa de detecção de piscar de olhos.

Além disso, as variações dos modelos ECC e ECS se comportaram de maneira semelhante entre os conjuntos de dados. Para ambos os modelos base, enquanto os melhores resultados foram alcançados no conjunto de dados *Talking Face*, os piores resultados ocorreram com o conjunto de dados *Eyeblink* e, para ambos os modelos, os resultados nos conjuntos de dados ZJU e ABD foram semelhantes. Esses resultados refletem a dificuldade dos conjuntos de dados. Enquanto o *Talking Face* é um conjunto de dados fácil para ambos os modelos, ZJU e ABD apresentam dificuldade média e o conjunto de dados mais desafiador foi o *Eyeblink*.

Os modelos auxiliares EE e RC, entretanto, não foram especificamente testados neste experimento. Conforme declarado na Seção 5.1, esses modelos são úteis quando o sujeito apresenta rotação da cabeça, oclusão ocular ou outras características desafiadoras de extração ocular. Eles não estão presentes nos conjuntos de dados usados neste experimento. No entanto, a introdução dos modelos auxiliares nos modelos ECC e ECS não prejudica significativamente seus desempenhos e pode até aumentá-los para algumas configurações.

Portanto, por se acreditar que tanto o modelo EE quanto o RC produzem variações mais adaptáveis e devido às ótimas pontuações alcançadas pelo modelo ECC, a variante ECC+EE+RC é considerada a principal contribuição deste trabalho para a tarefa de detecção do piscar de olhos.

5.2.2.4 Piscar de Olhos Voluntários

Como apenas piscadas voluntárias são consideradas um sinal de comunicação, um experimento foi conduzido para avaliar o desempenho dos modelos na tarefa de detecção de piscadas voluntárias. Como o conjunto de dados ABD contém piscadas voluntárias e involuntárias (de acordo com as definições na Seção 5.1.1), ele foi escolhido para este experimento. Todas as piscadas menores que 55 *frames* foram desconsideradas tanto das anotações quanto das saídas dos modelos, constituindo o conjunto de dados ABD_{vol} . Conforme a Seção 5.2.2.3, o *F1-Score* máximo alcançado foi descrito. Os resultados do modelo ECS e suas variações são apresentados na Tabela 5.3. O modelo ECC alcançou 1,0 *F1-Score* em todas as quatro variações, para quase todas as configurações de filtro e também quando nenhum filtro foi usado.

Tabela 5.3: Resultados para o ECS e modelos auxiliares na tarefa de detecção voluntária de piscar de olhos.

Modelos	ABD_{vol}		
	<i>F1-Score</i>	$W_{function}$	W_{size}
ECS	0.9032	-	-
ECS+EE	0.9032	-	-
ECS+RC	0.8955	Rect.	17
ECS+EE+RC	0.8955	Rect.	17

Comparando a pontuação das piscadas involuntárias e voluntárias, pode-se concluir diretamente que a tarefa referente apenas às piscadas voluntárias é mais fácil.

Como o sistema utilizará apenas piscadas voluntárias como sinal de comunicação, os modelos construídos em ECC parecem adequados ao objetivo proposto. Além disso, como já dito na Seção 5.2.2.3, os modelos auxiliares EE e RC produzem variações de modelo mais adaptáveis. Portanto, o modelo ECC+EE+RC é considerado o modelo final para a tarefa de detecção voluntária do piscar de olhos.

5.2.2.5 Analisando os Parâmetros do Filtro

Ambos os experimentos nas Seções 5.2.2.3 e 5.2.2.4 exploram todas as configurações de filtro possíveis (Seção 5.1.8). No entanto, um experimento sobre o comportamento dessas diferentes configurações de filtro no modelo final (ECC+EE+RC) deve ser analisado. Os resultados deste experimento podem determinar as configurações de filtro adequadas para serem usadas junto com este modelo em produção.

Para este experimento, todos os tamanhos de janela ímpares no intervalo [3;31] foram testados e diferentes limites dentro do intervalo de [0,01;0,99] foram analisados. Também foram consideradas as funções retangular, gaussiana e mediana dos filtros móveis. Este experimento analisa os resultados dos modelos no conjunto de dados ABD, devido à simulação das condições de uso pretendidas, e no conjunto de dados *Eyeblink*, devido às suas propriedades desafiadoras.

Uma abordagem geral foi seguida e levou em consideração piscadas involuntárias e voluntárias neste experimento. A restrição do conjunto de dados ABD apenas a piscadas voluntárias foi desconsiderada, pois diferentes cenários podem usar variados limites para piscadas voluntárias. A estabilidade pela variação do tamanho da janela e a melhor pontuação obtida com as configurações dos filtros são os critérios levados em consideração para esta análise. Além disso, os resultados obtidos com o conjunto de dados ABD são os considerados para a tomada de decisão final, uma vez que este conjunto de dados representa melhor as condições de uso propostas.

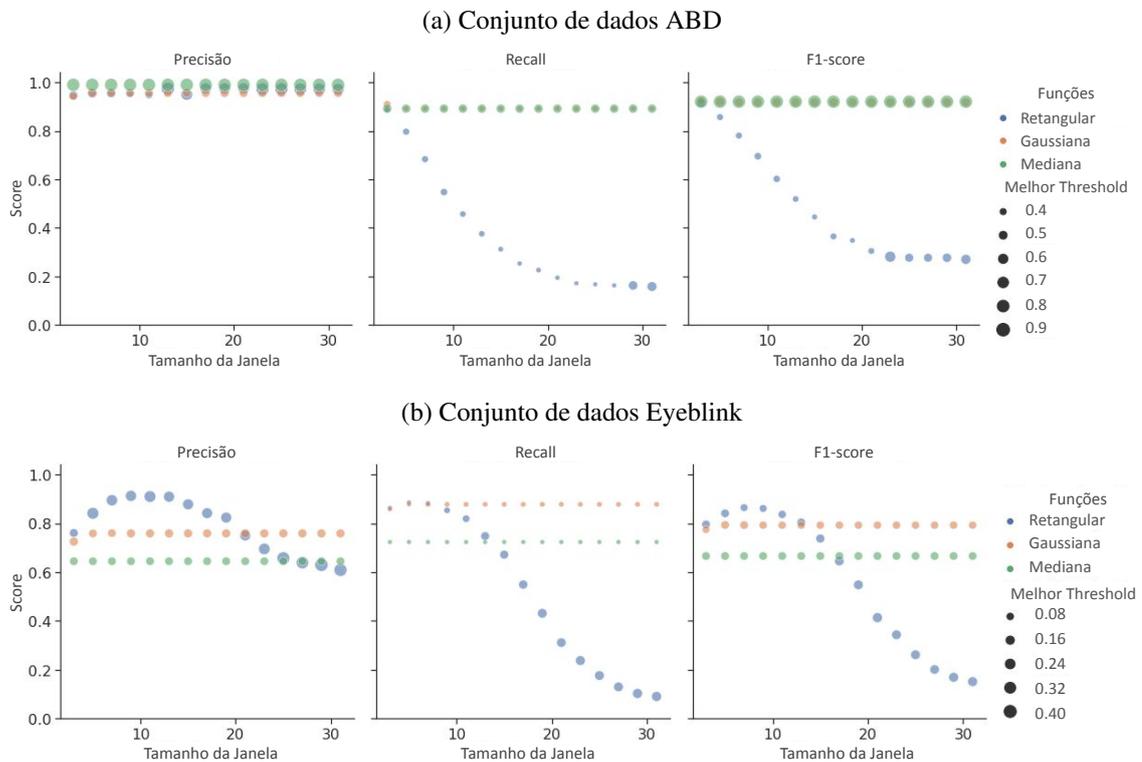
As Figuras 5.10a e 5.10b representam o experimento. Essas figuras possuem 3 colunas contendo cada uma um gráfico para uma das métricas propostas. Para cada um desses gráficos, foram plotadas 3 curvas, mostrando o impacto do tamanho da janela na respectiva pontuação para uma função de janela. O diâmetro de cada ponto nessas curvas é proporcional ao limite que maximiza a métrica correspondente.

As configurações podem ser inicialmente analisadas usando os critérios escolhidos. As funções mediana e gaussiana mostram estabilidade através de diferentes tamanhos de janela. Provavelmente pelo fato de que, enquanto a função mediana é capaz de ignorar *outliers*, a função gaussiana dará mais peso às saídas provenientes de quadros mais próximos da corrente, o que atribui a essas funções baixa sensibilidade ao aumento do tamanho da janela. A função retangular, no entanto, apresenta alta instabilidade do *F1-Score*, estimulada principalmente pela diminuição da recordação quando tamanhos maiores de janela são testados.

Em relação às melhores pontuações, todas as três funções apresentam resultados comparáveis. Como as funções gaussiana e mediana apresentam estabilidades maiores, seus F1-Scores são quase constantes. O retangular, no entanto, apresenta uma curva em forma de sino para o conjunto de dados *Eyeblink*, cujo máximo é a maior pontuação para este

conjunto de dados. No conjunto de dados ABD, no entanto, a curva para as funções retangulares apresenta uma pontuação máxima semelhante às alcançadas pelas outras funções.

Figura 5.10: Análises de filtro.



Dado que o conjunto de dados ABD é nosso foco e há instabilidade da função retangular, apenas as funções gaussiana e mediana foram consideradas como candidatas. As pontuações alcançadas por essas funções são extremamente semelhantes no conjunto de dados ABD. Portanto, os valores obtidos no conjunto de dados *Eyeblink* foram os considerados e a função Gaussiana é então eleita a mais adequada.

Devido à alta estabilidade da curva gaussiana, qualquer tamanho de janela apresenta pontuações valiosas. No entanto, janelas menores requerem um *buffer* menor e menos cálculos, o que as torna melhores candidatas. Intuitiva e arbitrariamente, o tamanho da janela de 7 foi escolhido.

5.2.2.6 Performance em Tempo Real

Um experimento de performance relacionada ao tempo também foi conduzido para avaliar a latência dos modelos. Foi utilizado o modelo ECC+EE+RC e uma webcam genérica capturou os *frames* ao vivo. O experimento durou 3 minutos e, para cada *frame* processada, foram computadas as latências de todo o modelo e de cada passo individual. As etapas listadas aqui são as da Seção 5.1. A média e o desvio padrão dos valores coletados são apresentados na Tabela 5.4. O experimento foi conduzido em uma máquina Ubuntu 20 com processador i5-7200U e 8GB de RAM.

Tabela 5.4: Valores de média e desvio padrão (Std.) das latências para cada etapa e para todo o modelo. Algumas etapas são apresentadas juntamente com suas siglas.

Etapa	Média (ms)	Std. (ms)
Aquisição de Dados	6.8	1.5
Compensador de Rotação (RC)	1.4	0.2
Deteção de Face	14.6	1.1
Alinhamento da Face	1.6	0.1
Extração da ROI	1.6	0.2
Avaliador Ocular (EE)	4.1	0.3
Classificador (ECC)	3.1	0.3
Filtro	<0.01	<0.01
Modelo Completo	33.5	1.7

A princípio, pode-se perceber que o sistema é estável e em tempo real, dados os baixos valores de desvio padrão e que seu *throughput* situa-se entre 28 e 31 *frames* por segundo ($33,5 \pm 1,7$ ms). Em segundo lugar, as etapas mais dispendiosas foram as etapas de aquisição de dados e detecção de faces, exigindo respectivamente 20,3% e 43,6% do tempo total. Enquanto o primeiro depende do sistema operacional, *drivers* e especificações de *hardware* e atualmente está fora do escopo de otimização, o segundo ainda é uma tarefa em aberto na literatura e pode ser otimizado.

Além disso, os modelos auxiliares propostos neste trabalho (filtros RC e janela deslizando) não comprometem a propriedade de tempo real do sistema. Juntos, esses modelos somam apenas 16,4% do tempo total de processamento e entregam um sistema mais robusto, sendo, portanto, uma atualização viável para o modelo básico.

Além disso, como as etapas de EE e ECC estão usando a mesma arquitetura CNN (conforme descrito na Seção 5.1.5), custos de tempo semelhantes podem ser encontrados para ambas as etapas. O ECC, no entanto, pode ser menos dispendioso do que o EE, uma vez que alguns candidatos a *patches* oculares podem ser desconsiderados (e, portanto, não usados como entrada para o modelo ECC) devido às suas baixas pontuações como *patches* oculares.

5.2.2.7 Comparação com Outros Estudos

Diferentes estudos testaram seus modelos nos mesmos conjuntos de dados. No entanto, conforme já exposto (Seção 5.2.2.2), esses trabalhos podem apresentar diferentes padrões de anotação, definições ou mesmo metodologias de avaliação – sendo tais diferenças motivadas principalmente por terem objetivos diferentes. Essas diferenças inibem uma comparação exata entre as obras. Com isso em mente, a Tabela 5.5 exhibe resultados de diferentes estudos que ainda podem ser úteis para o escopo deste artigo.

Como exemplos dessas diferenças, Fogelton e Benesova (2016) Fogelton e Benesova (2018) focaram no monitoramento da síndrome do olho seco, levando em consideração piscadas incompletas, e não focando piscadas voluntárias. No que diz respeito às metodologias de avaliação, diferentes técnicas do método IoU têm sido utilizadas, como o método baseado em picos (Al-gawwam e Benaissa 2018) ou os descritos em Soukupová

e Cech (2016) ou Drutarovsky e Fogelton (2014). Além disso, alguns estudos falharam em descrever seu procedimento de avaliação, base de dados e outros aspectos, conforme apontado por Fogelton e Benesova (2018).

Os experimentos aqui conduzidos, como mencionado anteriormente, usaram as anotações fornecidas em Fogelton e Benesova (2016) e também a metodologia de avaliação IoU com o mesmo limite que em Fogelton e Benesova (2016).

Finalmente, enquanto alguns trabalhos relatam apenas os resultados da detecção de piscar de olhos (Fogelton e Benesova 2016, Fogelton e Benesova 2018), os experimentos aqui conduzidos envolvem detecção e alinhamento de face, extração de *patch* do olho e validação dos modelos auxiliares propostos. Consequentemente, o sistema apresentado neste estudo é *end-to-end* e é avaliado como tal.

Tabela 5.5: Comparação dos melhores *F1-Scores* entre diferentes estudos sob diferentes configurações e critérios experimentais.

Trabalho	<i>Eyeblink8</i>	ZJU	<i>Talking Face</i>
(Anas et al. 2017)	-	0.937	1
(Soukupová e Cech 2016)	0.952	0.952	0.948
(Radlak e Smolka 2013)	-	0.992	-
(Fogelton e Benesova 2018)	0.913	0.976	0.971
Trabalho proposto	0.8697	0.9237	0.95

5.3 Considerações Finais

Foi apresentado um sistema robusto, em tempo real e de baixo custo para detecção do piscar de olhos voluntário. Este sistema pode ser integrado a um Sistema de Comunicação Alternativa, emitindo as piscadas detectadas e suas durações como sinais de comunicação. O sistema de detecção de piscar de olhos foi construído contendo um *pipeline* robusto, modelos auxiliares bem projetados (avaliação do olhar, compensador de rotação) e um filtro móvel.

O conjunto de dados de vídeo ABD usado para a tarefa de detecção do piscar de olhos (seguindo os critérios descritos no estudo) e o conjunto de dados de imagem YEC para a tarefa de classificação do estado ocular foram construídos e serão publicados a fim de contribuir para a comunidade acadêmica. Além disso, experimentos mostrando as propriedades de tempo real e de baixo erro do sistema foram conduzidos. Os resultados mostram que as extensões propostas contribuem positivamente para o sistema.

Concluimos assim que nosso sistema atende aos requisitos propostos dentro de seu contexto de aplicação. Ao associar alta qualidade, baixo custo e disponibilidade gratuita, espera-se que nossa proposta seja um *plugin* atraente para Sistemas de Comunicação Alternativa que utilizam blinks como sinal de ativação. Uma série de benefícios resulta da melhoria das habilidades de comunicação dos pacientes com ELA:

- Melhorar a qualidade de vida de pacientes e cuidadores
- Manutenção das habilidades cognitivas dos pacientes

- Contribuição com a independência de pacientes em estágios avançados de ELA

Com relação a trabalhos futuros, alguns pontos que podem ser considerados são:

- Diferentes alternativas para extração de patch ocular, evitando o uso de *landmarks*
- Melhoria do modelo RC
- Experimentos com pacientes de ELA
- Implementação de técnicas de concorrência com foco na otimização de latência

Capítulo 6

Ecosistema Autonomus

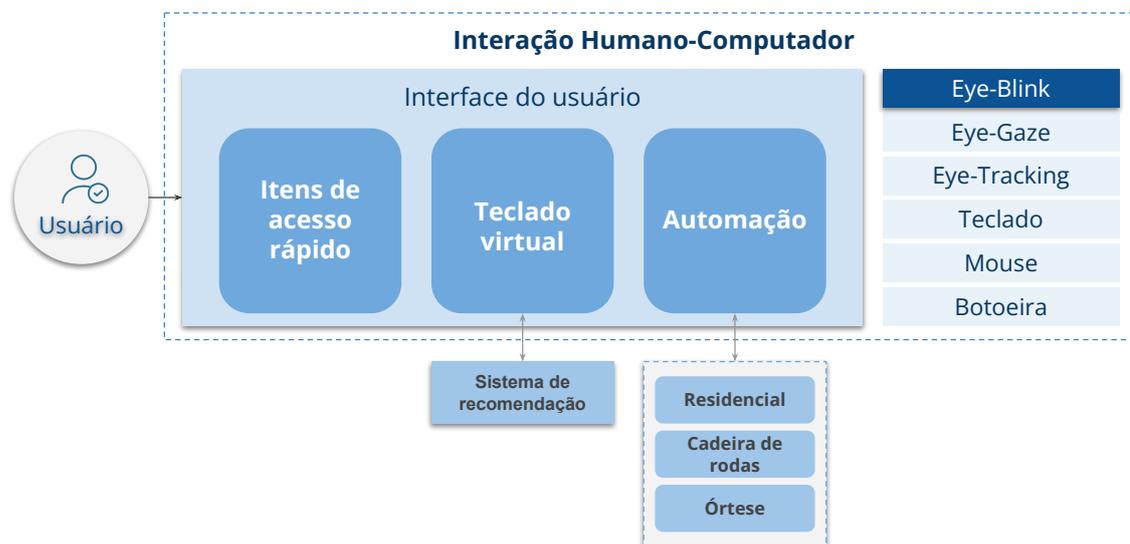
O Ecosistema Autonomus refere-se a um ambiente tecnológico colaborativo, planejado e desenvolvido com o propósito de integrar e interagir com diversos componentes de software e hardware. Seu principal objetivo é reunir recursos e produtos de tecnologias assistivas para o suporte e a promoção da autonomia, inclusão, participação social, interação com o ambiente e qualidade de vida para pessoas com ELA. Além disso, o Ecosistema Autonomus busca assegurar condições de igualdade ao exercício dos direitos e das liberdades fundamentais previstos no Estatuto da Pessoa com Deficiência (Brasil 2015).

Para manter a perenidade e o funcionamento pleno deste ecossistema, sua arquitetura foi projetada baseando-se em módulos (ou componentes), assim facilitando a dinâmica do desenvolvimento de novos recursos, da manutenção, evolução e implantação (Martin 2019). O Ecosistema Autonomus atua, resumidamente, como intermediário na comunicação e interação do usuário, portanto, é fundamental que o processo comunicativo aconteça eficientemente e não desmotive ou cause fadiga ao usuário. Neste trabalho, a investigação é direcionada especificamente para o módulo responsável pela forma de interação entre o usuário e o computador, o denominado módulo *eye-blink*.

6.1 Arquitetura do Ecosistema Autonomus

A arquitetura do Ecosistema Autonomus envolve módulos elementares para que o processo comunicativo aconteça efetivamente. Visto a complexidade do desenvolvimento de um sistema de comunicação alternativa acessível e inclusivo (acessibilidade) para pessoas que vivem com a ELA, da concepção ao produto final, foi definida a arquitetura ilustrada na Figura 6.1. Cada módulo presente nesta arquitetura é atribuída uma funcionalidade no ecossistema que, por sua vez, compreende uma lista de tarefas e desafios pertencentes a diferentes áreas do conhecimento. Nas seções a seguir serão apresentadas breves descrições dos módulos.

Figura 6.1: Visão geral da Arquitetura do Autonomus.



6.1.1 Módulo de Interação Humano-Computador

A Interação Humano-Computador é uma grande área multidisciplinar que envolve pesquisas voltadas para melhorar a experiência do usuário relacionada a sua interação com dispositivos de computação ou interfaces (Ebert et al. 2012). O respectivo módulo no Ecosystema Autonomus é a base da arquitetura, uma vez que é o meio de interação entre o usuário e todos os demais componentes. Existem diferentes estratégias para interagir com um sistema, tais como por meio das mãos (teclado, mouse, toque na tela), reconhecimento de gestos, da voz ou dos olhos. Esta última, como mostra Fernandes et al. (2023), é a mais adequada e eficiente para pessoas que vivem com a ELA em estágios mais avançados.

No contexto da ELA, quando se trata de estratégias de Interação Humano-Computador baseadas nos olhos, identificam-se quatro categorias segundo Fernandes et al. (2023): *eye-gaze*; *eye-blink*; *eye-tracking*; e estratégias híbridas, que combinam as anteriores.

Na Arquitetura do Autonomus, um estudo piloto foi desenvolvido centrado no módulo (estratégia) *eye-blink*. Este módulo, resumidamente, emprega técnicas de aprendizado de máquina e visão computacional com o objetivo principal de classificar o piscar de olhos como voluntários ou involuntários. Neste contexto, o piscar de olhos voluntário atua, a nível de interação entre usuário e sistema, como uma ação intencional do usuário para acionar funcionalidades ou fazer requisições ao sistema. O módulo *eye-blink* presente na Figura 6.1 representa uma abstração do que foi descrito no Capítulo 5, onde é detalhada toda a estrutura tecnológica subjacente.

6.1.2 Módulo de Interface do Usuário

O módulo de interface do usuário está intrinsecamente relacionado ao de Interação Humano-Computador. A interface do usuário serve como o canal pelo qual ocorre efeti-

vamente o processo de interação entre usuário e sistema computacional. É uma parte importante do sistema computacional, projetada para apresentar elementos gráficos que facilitam a visualização pelo usuário, o que permite a interação por meio desses elementos e o alcance dos objetivos ou funcionalidades propostas pelo sistema (Stone et al. 2005, Ebert et al. 2012, Stigberg 2020). Quando se fala em interação e interface, dois termos são rapidamente associados: usabilidade e design. A norma ISO 9241-11:1998 define a usabilidade como sendo “A extensão na qual um produto pode ser usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto de uso especificado” (ISO 1998).

No contexto da ELA, no qual há variadas especificidades patológicas da doença e diferentes perfis de usuário, a interface deve essencialmente ser projetada e desenvolvida com vistas às necessidades e limitações da pessoa que vive com a ELA. É nesta perspectiva que surge o termo design, mais especificamente, a abordagem design centrado no usuário, que busca entender o contexto dos usuários (social, físico e do ambiente, por exemplo) para desenvolver uma interface do usuário útil (ou utilizável) capaz de alcançar a usabilidade (Stone et al. 2005).

No Ecosistema Autonomus, especificamente no módulo de interface do usuário, foi utilizada a metodologia *Design Thinking* (Melo e Abelheira 2015) para o processo de design centrada no usuário. Um quadro de especialistas (Tabela 6.1) foi definido para entrevistar e observar as pessoas com ELA (usuários), etapa chamada de empatia.

Tabela 6.1: Perfil dos especialistas envolvidos no processo de desenvolvimento da interface do usuário.

Especialista	Descrição
01	Doutora, especialista em fonoaudiologia, linguística e psicologia. Atua no centro de referência em ELA do Rio Grande do Norte e desenvolve pesquisas na área de comunicação alternativa há mais de 10 anos.
02	Doutora em fisioterapia com mais de 10 anos de experiência em tecnologias assistivas e reabilitação.
03	Doutor em engenharia mecânica com mais de 10 anos de experiência em tecnologias assistivas.
04	Doutor em engenharia da computação com mais de 10 anos de experiência em tecnologias assistivas.
05	Doutor em neurologia que atua no centro de referência em ELA do Rio Grande do Norte há mais de 10 anos.
06	Designer gráfico que atua há mais de 10 anos em projetos envolvendo ilustração, design gráfico, motion design e edição de vídeo.
07	Designer, especialista em comunicação digital que atua há mais de 7 anos nos temas relacionados a experiência do usuário, usabilidade, design de interface, interação humano-computador, critérios de usabilidade e interação.

A etapa de empatia da metodologia *Design Thinking* foi fundamental para elucidar as necessidades, os desafios e a rotina dos usuários que vivem com a ELA, pois resultou na definição das personas, isto é, os perfis dos usuários. As Figura 6.2 e Figura 6.3 apresentam o contexto de um homem que vive com a ELA e sua esposa (cuidadora), respectivamente. Nessas figuras, é possível notar o problema central no contexto da família: a comunicação e autonomia. O segundo usuário que vive com a ELA, representado através da Figura 6.4, relata desafios e comportamentos semelhantes. Ambos os usuários utilizam tabelas de comunicação alternativa para a interação e participação nas atividades diárias.

Figura 6.2: *Persona* Antônio Pereira (pessoa com ELA).

Antônio Pereira

Biografia

Sr. Antônio é agricultor, tem um sítio em Caicó no RN, é analfabeto e foi diagnosticado com ELA há 1 ano e meio. Tem uma rotina já "programada", com idas frequentes a Natal para consultas. Usa o Bipap e já fez gastrectomia. Ultimamente não está mais conseguindo falar, o que está deixando-o frustrado, pois sente falta de conversar com sua família. Pela necessidade de se comunicar, foi aconselhado a utilizar métodos alternativos, como tabelas e recentemente aplicativos.

Referências




"Tem sido extremamente difícil, mas não posso perder a esperança!"

Dados demográficos:
 Idade: 52
 Profissão: Agricultor
 Casado com Dona Maria
 Tem três filhos

Dores

Sente falta de conversar com a família;
 Acabou de perder a fala;
 Tem dificuldades com tecnologias e como as utilizar;
 Frustra-se ao comunicar-se.

Necessidades

Gosta de conversar;
 Quer ter mais autonomia;
 Não sabe como chamar alguém quando precisa de ajuda;
 Quer uma solução personalizada para sua realidade.

Comportamentos

É impaciente;
 Tem receio em utilizar formas diferentes de comunicação;
 Tenta se comunicar com tabela de palavras e imagens do seu cotidiano.

Com a abordagem de Interação Humano-Computador e as personas bem definidas, foram desenvolvidos três tipos de protótipos de interfaces, os chamados protótipos de baixa, média e alta fidelidade. Ao pensar na inclusão de pessoas analfabetas, como o caso da pessoa que vive com ELA da Figura 6.2, foram organizados na interface do usuário itens customizáveis baseados em imagens e palavras.

A Figura 6.5, que representa o protótipo de alta fidelidade, agrupa uma sequência de retratos da interface do usuário do sistema computacional Autonomus. Os itens ou elementos da interface foram projetados para o acesso rápido (módulo de itens de acesso rápido da arquitetura), com no máximo quatro ações, para minimizar o esforço cognitivo da pessoa com ELA no alcance do objetivo. A exceção é o teclado virtual, que possibilita a escrita livre pelo usuário e representa um módulo específico na Arquitetura do Autonomus, pois são embarcados recursos tecnológicos baseados em inteligência artificial para otimizar o processo da escrita.

Figura 6.3: *Persona* Maria Pereira (acompanhante/cuidadora de pessoa com ELA).

Maria Pereira

Biografia

Dona Maria é agricultora, casada com Antônio, eles têm um sítio em Caicó no RN. é analfabeta e após o diagnóstico do marido, sua vida tem sido mais intensa. Tem uma rotina já 'programada', acorda muito cedo para realizar os afazeres da casa e do marido, o acompanha nas idas frequentes a Natal para consultas. Como ele não está mais falando, ela recebeu ajuda para usar um aplicativo e uma tabela de palavras com Sr. Antônio. Às vezes se sente frustrada por não conseguir entendê-lo.

Referências




“Sei que ele quer dizer algo, mas muitas vezes não o entendo”

Dados demográficos:
 Idade: 50
 Profissão: Agricultora
 Casada com Sr. Antônio
 Tem três filhos

Dores

Sente falta das conversas com o marido;
 Tem dificuldades com tecnologias e como as utilizar;
 Dos produtos que já viu, achou caro;
 Sente-se sobrecarregada.

Necessidades

Gosta de conversar;
 Quer que ele tenha mais autonomia;
 Poder saber quando Antônio precisa de ajuda;
 Poder personalizar uma tabela ou app para a realidade deles, sem ajuda de terceiros.

Comportamentos

Tem receio em utilizar formas diferentes de comunicação;
 Tenta usar uma tabela de palavras e imagens do seu cotidiano. Usa o aplicativo só para frases que tem no sistema;
 Deixou de trabalhar para cuidar do marido.

Figura 6.4: *Persona* João Paulo (pessoa com ELA).

João Paulo

Biografia

Sr. João é professor, mora em Natal no RN, é alfabetizado e foi diagnosticado com ELA há 2 anos. Tem uma rotina com fisioterapeutas, com idas frequentes a consultas e quando está em casa, gosta de ver filmes. Usa o Bipap e já fez gastrectomia. Não fala há cinco meses, utiliza uma tabela feita pela esposa para poder se comunicar, mas não gosta de usar aplicativos assistivos, ele acredita que não ajuda. Utilizava dispositivos móveis e computadores antes, mas hoje em dia só para leitura.

Referências




“Não gosto de sair da minha zona de conforto”

Dados demográficos:
 Idade: 55
 Profissão: Professor de escola pública
 Casado com Dona Noemi
 Tem dois filhos

Dores

Sente falta de conversar com a família;
 Esteve depressivo nos últimos meses;
 Não gosta de ser tratado de maneira infantil;
 Não consegue realizar tarefas com demandas cognitivas altas, por exemplo: ler.

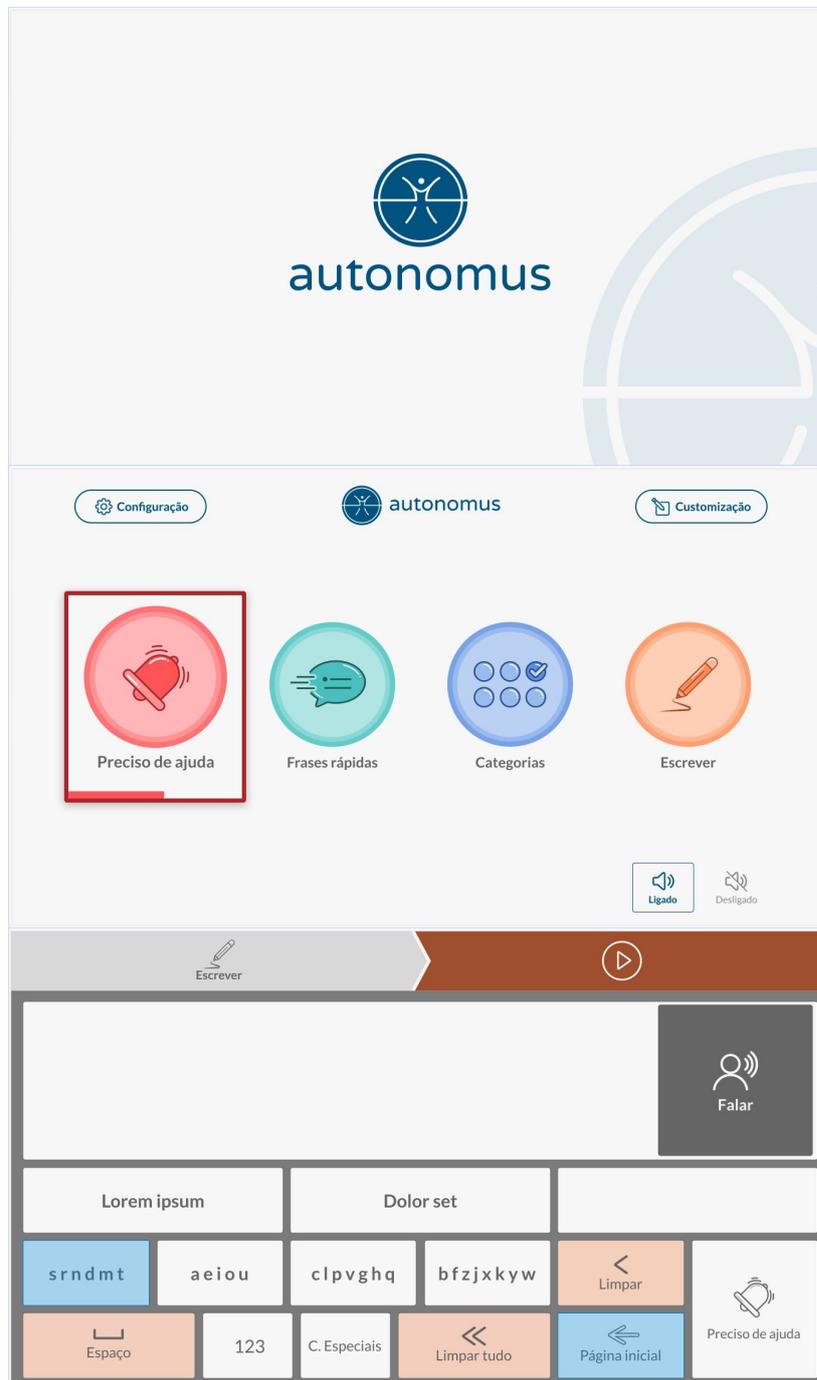
Necessidades

Gosta de conversar;
 Quer ter mais autonomia;
 Não sabe como chamar alguém quando precisa de ajuda;
 Precisa de uma customização para simplificar a comunicação.

Comportamentos

Sente-se desmotivado ultimamente;
 Tentou utilizar um aplicativo assistivo para comunicar-se, mas não gostou. O aplicativo não o ajudou no dia a dia;
 Não gosta de experimentar formas novas para comunicar-se;
 Se comunica com tabela de palavras e imagens. Fora isso, a esposa sabe o que ele pede..

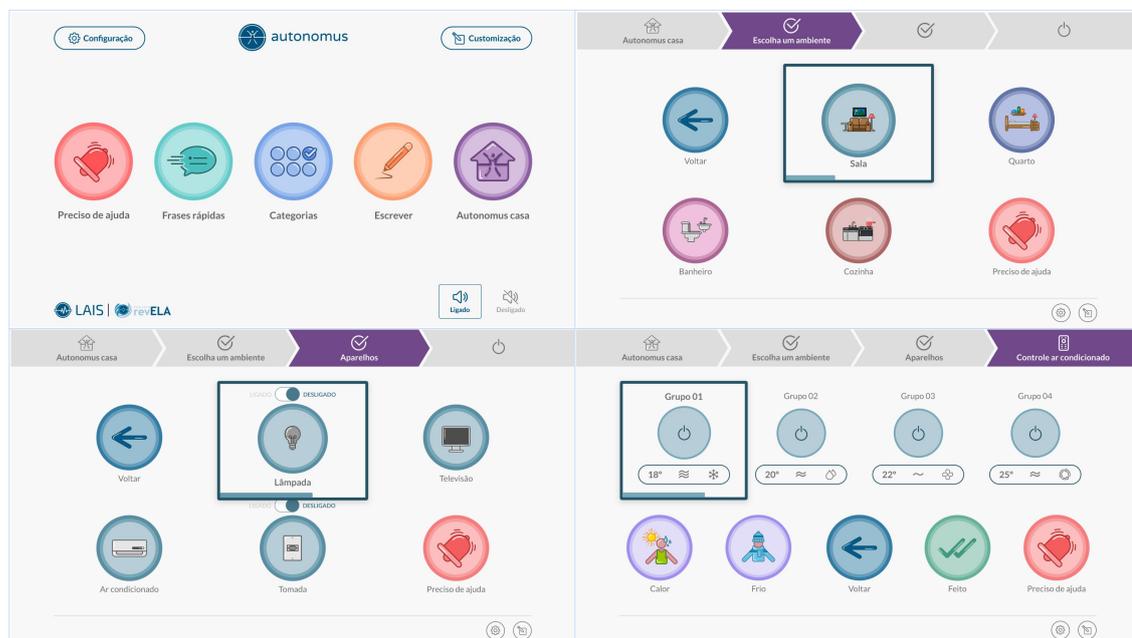
Figura 6.5: Protótipo de alta fidelidade do Autonomus.



Ao analisar a Arquitetura do Autonomus, conforme ilustrada na Figura 6.1, destaca-se também a presença do módulo de automação. Este módulo desempenha um papel importante para a autonomia da pessoa que vive com a ELA, pois permite a interação do usuário com dispositivos em ambiente domiciliar e intermedia o controle de tecnologias assistivas, tais como cadeira de rodas e órteses.

A Figura 6.6 mostra o protótipo de alta fidelidade deste módulo. Nela, é possível notar quatro ambientes automatizados: sala, quarto, banheiro e cozinha, bem como um conjunto de dispositivos, incluindo lâmpada, televisão, ar-condicionado e tomada, todos habilitados para a interação com o usuário. A interface deste módulo oferece ainda a opção de customização, o que possibilita inserir ou excluir dispositivos. É importante ressaltar que a infraestrutura tecnológica subjacente ao módulo ultrapassa os domínios da Interação Humano-Computador e da interface do usuário, pois é preciso projetar e preparar um ambiente automatizado, o que requer recursos físicos e especialistas com expertise no campo.

Figura 6.6: Protótipo de alta fidelidade do Autonomus com o módulo de automação residencial (Autonomus Casa).

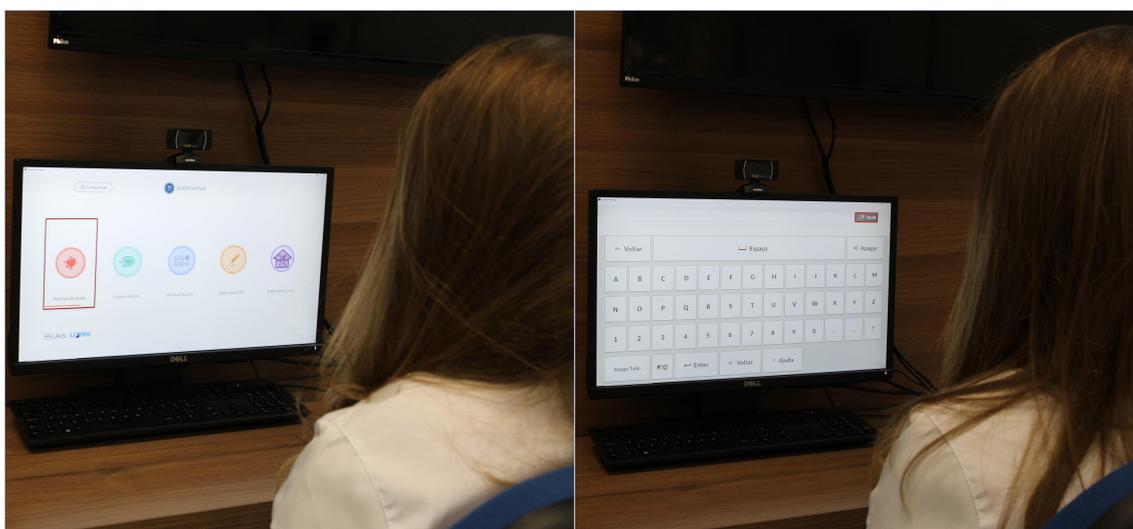


6.2 Resultados

A soma dos esforços no desenvolvimento do Ecosistema Autonomus resultou em contribuições importantes tanto para o público alvo do projeto, as pessoas que vivem com a Esclerose Lateral Amiotrófica ou com dificuldades funcionais relacionadas ao processo comunicativo, quanto para a comunidade que possui interesses em colaborar direta ou indiretamente com o *Software* de Comunicação Alternativa Autonomus. O principal alcance prático deste trabalho foi a disponibilização, de forma gratuita, do recurso de tecnologia assistiva para a comunicação alternativa de pessoas que vivem com a Esclerose Lateral Amiotrófica. O Autonomus, como é chamado o *software*, foi projetado e desenvolvido para computadores de mesa, ou notebooks, e dispositivos móveis, tais como celulares ou *tablets*.

O Autonomus, em sua versão 1.1 *desktop*, está disponível para *download* em <https://revela.lais.ufrn.br/projetos/autonomus/>. A Figura 6.7 ilustra a interação de um usuário saudável com a interface do Autonomus *desktop*. Por ser um *software* de código aberto, o que permite a colaboração e reprodução do recurso de tecnologia assistiva por pesquisadores e toda a comunidade interessada, foi também compartilhado um repositório no GitHub¹ com o respectivo código fonte. Para dar suporte aos usuários do Autonomus *Desktop* e aos cuidadores de pessoas que vivem com a Esclerose Lateral Amiotrófica, foram produzidos vídeos tutoriais que auxiliam no processo de instalação do *software*² e instruem sobre sua utilização³. É importante ressaltar que esses vídeos estão sendo avaliados por uma equipe, composta por profissionais da fisioterapia e fonoaudiologia, especializada em reabilitação e comunicação alternativa.

Figura 6.7: Interação de um usuário saudável com a Plataforma Autonomus *Desktop*.



A versão do Autonomus para dispositivos móveis, chamada de Autonomus *Mobile*, está em fase de testes de usabilidade pela equipe de desenvolvedores e pesquisadores da área da saúde envolvidos no projeto Autonomus. Entretanto, a versão de teste do Autonomus *Mobile*, especificamente para dispositivos providos do sistema operacional *Android*, está disponível para *download* no site: <https://bit.ly/autonomusmobile>. A Figura 6.8 ilustra a interação de um usuário saudável com a interface do Autonomus *Mobile*. A abordagem de Interação Humano-Computador em ambos os *softwares* é baseada no piscar voluntário dos olhos, como descrito no Capítulo 5. É importante destacar que, com o mesmo princípio do desenvolvimento colaborativo, o *software* Autonomus *Mobile* também possui código aberto e qualquer pessoa com potencial interesse em contribuir no desenvolvimento do *software* pode acessar o repositório no GitHub com o respectivo

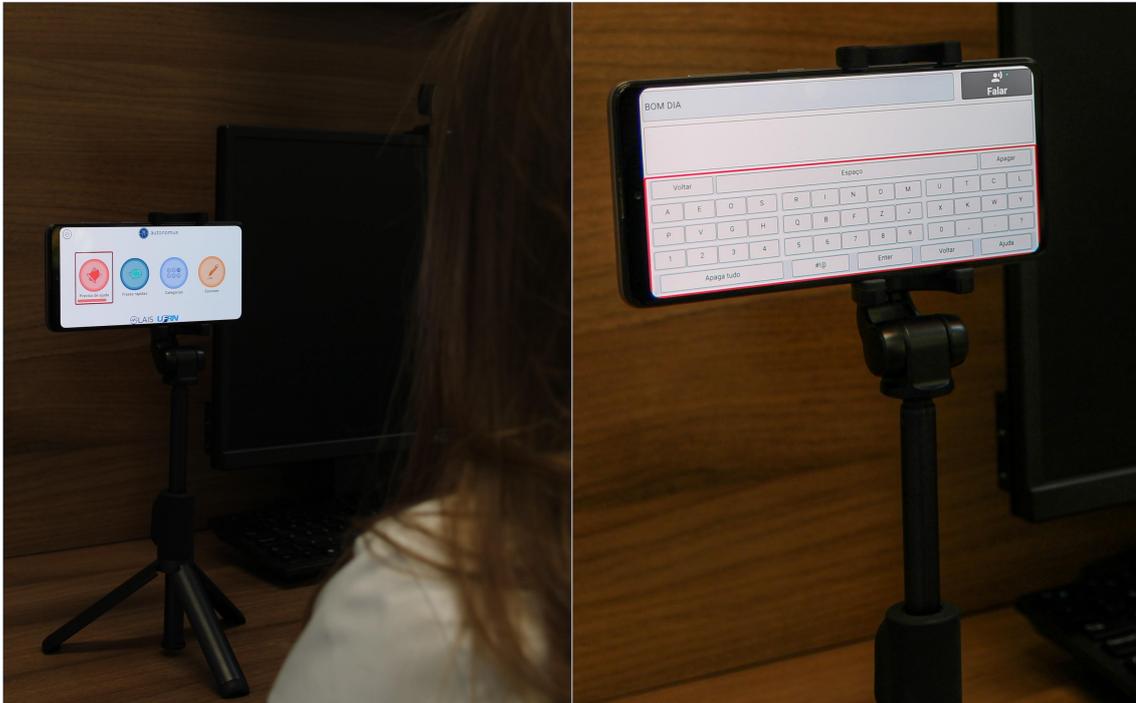
¹Autonomus *Desktop* - Licença Creative Commons 4.0: <https://github.com/feliiperiicardo/autonomus-app>

²Vídeo tutorial para instalação. Disponível em: <https://youtu.be/jkusaTdxudM>

³Vídeo tutorial para utilização do *software*. Disponível em: https://youtu.be/r80-M_H9o7I

código fonte⁴.

Figura 6.8: Interação de um usuário saudável com a Plataforma Autonomus *Mobile*.

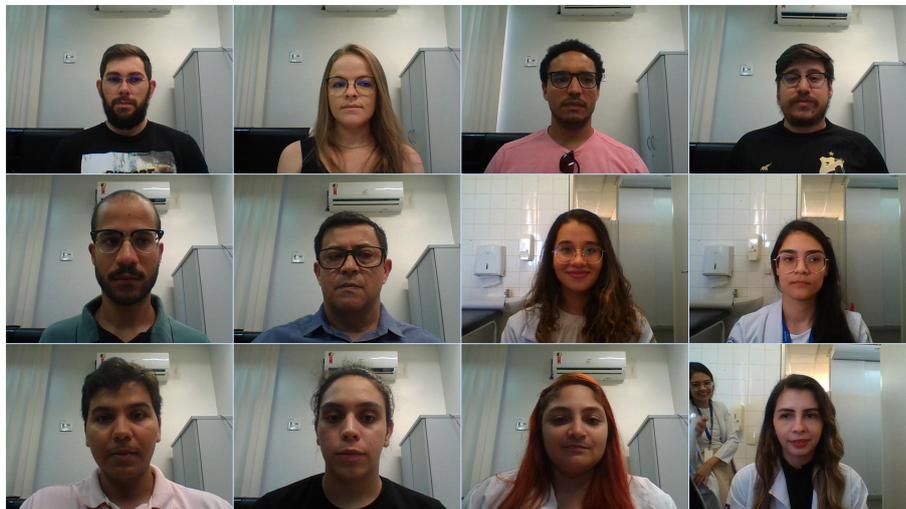


Na perspectiva de contribuições relevantes para o campo de pesquisa, o que pode impulsionar e encorajar novas investigações, foram destacadas duas importantes contribuições no Capítulo 5: (i) o módulo de Interação Humano-Computador para compor a Arquitetura do Autonomus, sendo a proposta centrada em modelos de *deep learning* e na estratégia de interação baseada na detecção do piscar voluntário dos olhos; (ii) e o conjunto de dados *Autonomus Blink Detection* (ABD), que disponibiliza 17 vídeos (19.975 imagens), provenientes de 10 pessoas saudáveis, com simulações do piscar de olhos voluntários e involuntários.

Como parte de um novo estudo, derivado desta tese e ainda de acesso restrito aos pesquisadores do projeto, foi construído um novo conjunto de dados para subsidiar o desenvolvimento do módulo de Interação Humano-Computador baseado no rastreamento da pupila. A Figura 6.9 apresenta uma amostra do conjunto de dados do grupo de controle (pessoas saudáveis). Esse conjunto de dados é composto por 18 vídeos (24.300 imagens), provenientes de 12 pessoas saudáveis com as seguintes características: 6 pessoas do sexo feminino e 6 do sexo masculino; idade média de 30,3 anos; e 7 pessoas de cor ou raça branca, 4 parda e uma preta.

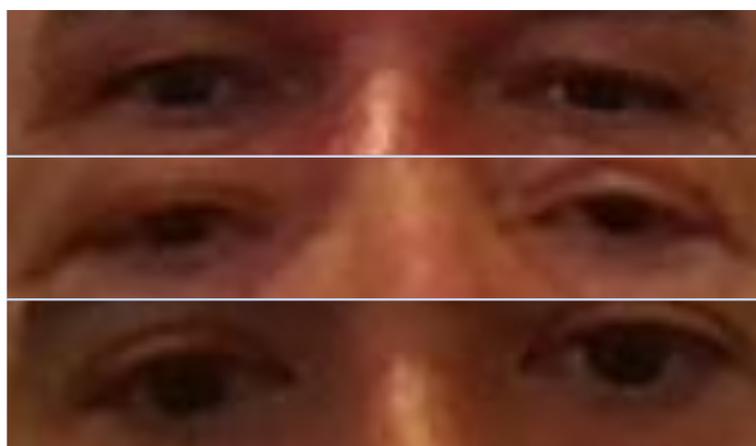
⁴Autonomus *Mobile* - Licença Creative Commons 4.0: <https://github.com/feliiperiicardo/autonomus-mobile>

Figura 6.9: Conjunto de dados do grupo de controle (pessoas saudáveis) para a tarefa de rastreamento da pupila.



Da mesma forma, foi construído também um conjunto de dados de pessoas que vivem com a Esclerose Lateral Amiotrófica. A Figura 6.10 apresenta uma amostra contendo recortes dos olhos do conjunto de dados de pessoas com a doença. Esse conjunto de dados é composto por 3 vídeos (4.050 imagens), obtidos de 3 pessoas com a Esclerose Lateral Amiotrófica com as seguintes características: 2 pessoas do sexo masculino e uma do sexo feminino; idade média de 59 anos; e 3 pessoas de cor ou raça branca. É de fundamental importância ressaltar que o protocolo experimental foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal do Rio Grande do Norte (UFRN), Natal, Brasil, por meio da carta CAAE No.25687819.3.0000.5537, e de acordo com os Acordos de Helsinki (modificado em 2004).

Figura 6.10: Conjunto de dados do grupo de pessoas com ELA para a tarefa de rastreamento da pupila.



Todos os resultados aqui apresentados estão vinculados à meta 1 do projeto RevELA.

Capítulo 7

Discussões

Segundo consta no relatório global sobre tecnologia assistiva, construído pela Organização Mundial de Saúde (OMS) em conjunto com o Fundo das Nações Unidas para a Infância (UNICEF), o acesso à tecnologias assistivas é um direito humano e um pré-requisito para participação e oportunidades iguais. Os Estados-Membros e as suas instituições são responsáveis por garantir que os seus cidadãos têm acesso a tecnologias de apoio seguras, eficazes e acessíveis. Ainda segundo o relatório, o acesso às tecnologias assistivas alinha-se ao ODS 3.8: Atingir a cobertura universal de saúde, incluindo a proteção do risco financeiro, o acesso a serviços de saúde essenciais de qualidade e o acesso a medicamentos e vacinas essenciais seguros, eficazes, de qualidade e a preços acessíveis para todos. O acesso às tecnologias assistivas, tais como ao Ecossistema Autonomus, alinha-se também ao ODS 10, para reduzir a desigualdade dentro dos países e entre eles (United Nations 2015, United Nations 2022).

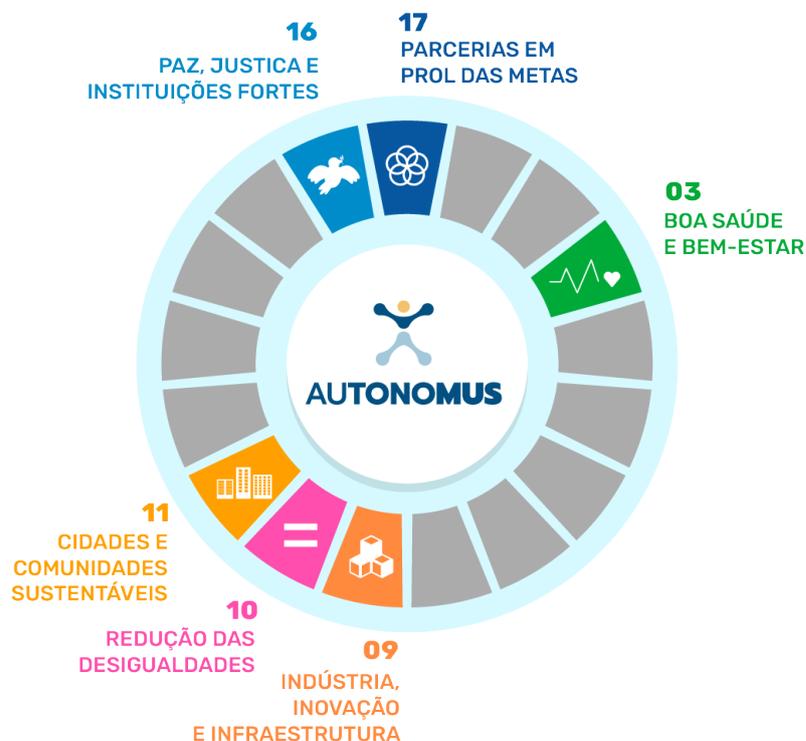
No Brasil, em uma linha do tempo mais recente e significativamente representativa, uma nova perspectiva relacionada à atenção e ao cuidado de pessoas com limitações funcionais foi evidenciada a partir da Lei Nº 13.146, de 6 de Julho de 2015 (Lei Brasileira de Inclusão da Pessoa com Deficiência ou Estatuto da Pessoa com Deficiência), destinada a assegurar e a promover, em condições de igualdade, o exercício dos direitos e das liberdades fundamentais por pessoa com deficiência, visando à sua inclusão social e cidadania (Brasil 2015).

Em virtude desta Lei, especialmente do Art. 75º, foi publicado o Decreto Nº 10.645, de 11 de Março de 2021 que define as diretrizes, os objetivos e os eixos do Plano Nacional de Tecnologia Assistiva (Brasil 2021a), que tornou-se público através da Portaria MCTI nº 5.366, de 02.12.2021 (Brasil 2021c). Ações como essas, bem como a criação do Sistema Nacional de Laboratórios de Tecnologias Assistivas (Portaria MCTI nº 6.033, de 24.06.2022) (Brasil 2022), incentivam, engajam e impulsionam o desenvolvimento de novas tecnologias assistivas e asseguram o acesso aos recursos no âmbito do Sistema Único de Saúde (SUS).

O Ecossistema Autonomus é um projeto inserido nesse contexto da legislação brasileira e atua em conformidade tanto com o Plano Nacional de Tecnologia Assistiva quanto com a Estratégia Nacional de Ciência, Tecnologia e Inovação 2016|2020 (Brasil 2018). Como uma solução de saúde digital, que dispõe de recursos de tecnologias assistivas para a Comunicação Aumentativa e Alternativa de pessoas que vivem com a ELA, o Ecossistema Autonomus atende ao que foi estabelecido no Art. 74º do Estatuto da Pessoa com

Deficiência, onde é dada a garantia à pessoa com deficiência acesso a produtos, recursos, estratégias, práticas, processos, métodos e serviços de tecnologia assistiva que maximizem sua autonomia, mobilidade pessoal e qualidade de vida (Brasil 2015). Especialmente, por ser uma tecnologia de código fonte aberto e livre, com licença creative commons 4.0¹, no qual é possível adaptar e compartilhar, a Plataforma Autonomus do Projeto RevELA contribui diretamente para reduzir as desigualdades no acesso às tecnologias assistivas, aspecto fundamental para que o Brasil cumpra com a implantação das Metas 3, 9, 10, 11, 16 e 17 (Figura 7.1) no sentido de melhorar a qualidade de vida das pessoas que vivem com a ELA, não apenas no Brasil, mas em qualquer parte do mundo.

Figura 7.1: Visão geral dos Objetivos de Desenvolvimento Sustentável aos quais o Ecosistema Autonomus alinha-se.



No Brasil, aproximadamente 18,6 milhões de pessoas, incluindo aquelas que vivem com a ELA, enfrentam dificuldades funcionais. Dentro deste grupo, além dos desafios relacionados às limitações funcionais, também se observam limitações orçamentárias e financeiras. Cerca de 11,8 milhões (64%) destas pessoas vivem sem renda ou com rendimento mensal domiciliar per capita de até um salário mínimo (1.320 reais na data atual). Esse valor está abaixo dos 1.625 reais, que é o rendimento domiciliar per capita para o Brasil (Brasil 2023a, Brasil 2023b). Portanto, diante da pobreza e desigualdade de renda no Brasil, bem como dos custos associados ao acesso a tecnologias assistivas comerciais

¹Acesse: https://creativecommons.org/licenses/by/4.0/deed.pt_BR

ofertadas pelo setor privado e à manutenção do bem-estar e qualidade de vida, o Ecossistema Autonomus revela-se como uma alternativa com potencial contribuição na vida das pessoas, especialmente, na vida de pessoas com Esclerose Lateral Amiotrófica.

No contexto das pessoas que vivem com a Esclerose Lateral Amiotrófica, o impacto das tecnologias assistivas é ainda mais significativo. Isso ocorre porque a doença compromete todo o sistema motor e, com a sua progressão inevitável, leva o indivíduo a perder sua autonomia, independência, capacidade de realizar movimentos naturais, de se comunicar, de interagir com o ambiente e de ter participação social (Fernandes et al. 2023, Goutman et al. 2022, Saadeh et al. 2017, Hardiman et al. 2017, van Es et al. 2017). Com essa premissa, a solução de saúde digital para a Comunicação Aumentativa e Alternativa de pessoas com Esclerose Lateral Amiotrófica, apresentada neste estudo, emerge como um recurso de tecnologia assistiva acessível (gratuito), viável e essencial para a manutenção da capacidade de se comunicar, da autonomia, da interação e participação social e, sobretudo, para promover o bem-estar e a qualidade de vida.

Para construir o instrumento tecnológico de Comunicação Aumentativa e Alternativa, dois componentes elementares, presentes na arquitetura do Autonomus, foram desenvolvidos e apresentados como contribuições desta tese: os módulos de Interação Humano-Computador e de interface, destacados nos capítulos 5 e 6, respectivamente. Especialmente para o módulo de Interação Humano-Computador, um fluxo de trabalho (*pipeline*) foi desenvolvido, validado e testado em 10 pessoas saudáveis. As tecnologias subjacentes a este módulo envolvem técnicas de visão computacional, para a detecção da face do usuário e pontos relacionados a região de interesse (Liu et al. 2015, Sagonas et al. 2016), e aprendizado de máquina, com destaque para o modelo baseado em uma rede neural convolucional (Li et al. 2018, Xiong et al. 2017, Anas et al. 2017). Para a Interação Humano-Computador baseada em imagens dos olhos, os resultados mostraram-se promissores (92,63% de F1-score) e funcionais para o Ecossistema Autonomus (de Lima Medeiros et al. 2022).

Além do recurso tecnológico, uma importante contribuição foi gerada para a comunidade científica. Como parte dos resultados do módulo de Interação Humano-Computador, foi disponibilizado o primeiro conjunto de dados de acesso público que dispõe de imagens de 10 pessoas saudáveis para a detecção voluntária do piscar dos olhos (de Lima Medeiros et al. 2022). Com base nesses resultados alcançados, é possível afirmar, como resposta à questão de pesquisa Q2, que é possível utilizar técnicas de visão computacional e aprendizado de máquina para desenvolver tecnologias que possam prover Comunicação Aumentativa e Alternativa para pessoas que vivem com a Esclerose Lateral Amiotrófica, considerando os dispositivos móveis neste contexto.

O segundo componente refere-se à interface do sistema computacional, que foi desenvolvida centrada no usuário, isto é, pensando na pessoa com ELA. A interface desenvolvida inclui elementos iconográficos e contextos, associados ao cotidiano da pessoa com ELA, que podem ser customizáveis. Além disso, a interface conta com um teclado virtual que permite a escrita livre usando os olhos. A integração desses dois elementos da Arquitetura Autonomus constituem a solução de saúde digital acessível, e de código aberto, para a Comunicação Aumentativa e Alternativa de pessoas que vivem com a Esclerose Lateral Amiotrófica. Neste sentido, como resposta à questão de pesquisa Q1, esta tese in-

dica que é possível desenvolver tecnologias assistivas de baixo custo e com código fonte livre e aberto, diferentemente dos que estão disponíveis no mercado, para pessoas que vivem com a Esclerose Lateral Amiotrófica, com o objetivo de lhes garantir a comunicação e lhes devolver a autonomia.

Capítulo 8

Considerações Finais

Nesta tese foi apresentada uma solução de saúde digital para a Comunicação Aumentativa e Alternativa de pessoas que vivem com Esclerose Lateral Amiotrófica (ELA). Trata-se de um recurso de tecnologia assistiva de acesso gratuito¹ que busca intermediar o processo comunicativo e promover a autonomia, a inclusão e participação social, a interação com o ambiente e, sobretudo, a qualidade de vida de pessoas com ELA ou com dificuldades funcionais relacionadas à comunicação. A implementação de um modelo computacional para a Interação Humano-Computador, baseado em uma Rede Neural Convolucional e técnicas de visão computacional, aliada ao desenvolvimento de uma interface de usuário inclusiva confirma a hipótese de que é possível desenvolver tecnologias assistivas baseadas em métodos computacionais para garantir formas de Comunicação Aumentativa e Alternativa para pessoas que vivem com Esclerose Lateral Amiotrófica (ELA).

8.1 Limitações

Esta tese apresenta limitações, principalmente relacionadas a testes de usabilidade com pessoas que vivem com Esclerose Lateral Amiotrófica (ELA). Entretanto, testes estão sendo conduzidos com especialistas em comunicação alternativa, de forma antecipada, para aprimorar o *software* Autonomus e preparar o cenário adequado para os experimentos com o grupo de pessoas com ELA. Quanto ao módulo de Interação Humano-Computador, o modelo proposto foi treinado apenas com o grupo de controle (pessoas saudáveis), o que pressupõe um modelo com viés, se considerar usuários com ELA ou com outras dificuldades funcionais relacionadas ao processo comunicativo.

Como o trabalho partiu da concepção da arquitetura do Autonomus, da estratégia de interação e da interface do usuário, não foi possível desenvolver outros módulos de Interação humano-Computador baseados em imagens dos olhos. É importante ressaltar, no contexto do objeto de estudo e de pessoas com Esclerose Lateral Amiotrófica, a ausência de dados na literatura para treinar, avaliar e testar modelos algorítmicos de *machine learning*. Além disso, há uma limitação natural relacionada à qualidade das imagens provenientes de câmeras integradas aos *notebooks* ou *smartphones* (câmera frontal). Isso

¹Faça o *download*: <https://revela.lais.ufrn.br/projetos/autonomus/>

umenta ainda mais o desafio de desenvolver um recurso de Comunicação Aumentativa e Alternativa baseado no rastreamento ocular e direciona a necessidade de esforços nos campos de visão computacional e *machine learning*.

8.2 Perspectivas

Esta tese aponta possibilidades de contribuições significativas para o Ecossistema Autonomus. O foco principal deste trabalho foi o desenvolvimento do módulo de Interação Humano-Computador baseado em *eye-blink* e a concepção da interface do usuário, assim como da Arquitetura que engloba todo o sistema computacional de Comunicação Aumentativa e Alternativa. A Figura 6.1 destaca outros componentes com potenciais áreas de pesquisa ainda em aberto e relevantes para o Ecossistema Autonomus, tanto do ponto de vista de *software* quanto de *hardware*.

No contexto da Interação Humano-computador no Ecossistema Autonomus, os módulos de *Eye-Gaze* e *Eye-Tracking* têm potencial para melhorar o processo de seleção de itens ou caracteres na interface do usuário e no Teclado Virtual. Isso, conseqüentemente, deve minimizar o esforço cognitivo da pessoa que vive com Esclerose Lateral Amiotrófica. Ainda observando a Arquitetura do Autonomus, representada através da Figura 6.1, nota-se o módulo do Teclado Virtual conectado ao sistema de recomendação. Essa ilustração representa a possibilidade do desenvolvimento de técnicas baseadas em inteligência artificial para otimizar o tempo de seleção de caracteres.

Outro segmento a ser explorado, que traz grandes benefícios para a pessoa que vive com Esclerose Lateral Amiotrófica, é o módulo de Automação. Neste módulo, há a possibilidade de desenvolver diversos recursos de tecnologias assistivas, tais como órteses e cadeiras de rodas automatizadas. O estudo realizado é apenas uma amostra de sua potencialidade no campo de tecnologias assistivas, com foco em Comunicação Aumentativa e Alternativa, e enseja a continuidade do projeto Autonomus através de trabalhos complementares nos campos de pesquisa acima mencionados.

Referências Bibliográficas

- Abe, Kiyohiko, Shoichi Ohi e Minoru Ohyama (2011), Eye-gaze detection by image analysis under natural light, *em* J. A. Jacko, ed., ‘Human-Computer Interaction. Interaction Techniques and Environments’, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 176–184.
- Ahangaran, Meysam e Adriano Chiò (2020), *AIM in Amyotrophic Lateral Sclerosis*, Springer International Publishing, Cham, pp. 1–13.
- Ahangaran, Meysam, Adriano Chiò, Fabrizio D’Ovidio, Umberto Manera, Rosario Vasta, Antonio Canosa, Cristina Moglia, Andrea Calvo, Behrouz Minaei-Bidgoli e Mohammad-Reza Jahed-Motlagh (2022), ‘Causal associations of genetic factors with clinical progression in amyotrophic lateral sclerosis’, *Computer Methods and Programs in Biomedicine* **216**, 106681.
- Aharonson, V., V. Y. Coopoo, K. L. Govender e M. Postema (2020), ‘Automatic pupil detection and gaze estimation using the vestibulo-ocular reflex in a low-cost eye-tracking setup’, *SAIEE Africa Research Journal* **111**(3), 120–124.
- Al-gawwam, Sarmad e Mohammed Benaissa (2018), ‘Robust eye blink detection based on eye landmarks and savitzky–golay filtering’, *Information* **9**(4).
URL: <https://www.mdpi.com/2078-2489/9/4/93>
- Ambikapathy, Bakiya, Kamalanand Kirshnamurthy e Rajinikanth Venkatesan (2018), ‘Assessment of electromyograms using genetic algorithm and artificial neural networks’, *Evolutionary Intelligence* pp. 1–11.
- An, Kwanghoon, Myungjong Kim, Kristin Teplansky, Jordan Green, Thomas Campbell, Yana Yunusova, Daragh Heitzman e Jun Wang (2018), Automatic early detection of amyotrophic lateral sclerosis from intelligible speech using convolutional neural networks, *em* ‘Proc. Interspeech 2018’, pp. 1913–1917.
URL: <http://dx.doi.org/10.21437/Interspeech.2018-2496>
- Anas, Essa R., P. Henríquez e B. Matuszewski (2017), Online eye status detection in the wild with convolutional neural networks, *em* ‘VISIGRAPP’.
- Anumanchipalli, Gopala K., Josh Chartier e Edward F. Chang (2019), ‘Speech synthesis from neural decoding of spoken sentences’, *Nature* **568**(7753), 493–498.
URL: <https://doi.org/10.1038/s41586-019-1119-1>

- Aslam, Zaryab, Aisha Zahid Junejo, Aarish Memon, Ali Raza, Juveria Aslam e Liaquat Ali Thebo (2019), Optical assistance for motor neuron disease (mnd) patients using real-time eye tracking, *em* '2019 8th International Conference on Information and Communication Technologies (ICICT)', pp. 61–65.
- Atasoy, N. A., A. Çavuşoğlu e F. Atasoy (2016), 'Real-time motorized electrical hospital bed control with eye-gaze tracking', *Turkish Journal of Electrical Engineering & Computer Sciences* **24**(6), 5162–5172.
- Barbalho, Ingridy M. P., Aleika L. Fonseca, Felipe Fernandes, Jorge Henriques, Paulo Gil, Danilo Nagem, Raquel Lindquist, Thaísa Santos-Lima, João Paulo Q. Santos, Jailton C. Paiva, Antônio H. F. Morais e Ricardo A. M. Valentim (2023), 'Digital health solution for monitoring and surveillance of amyotrophic lateral sclerosis in brazil', *Frontiers in Public Health* **11**, 1209633.
URL: <https://doi.org/10.3389/fpubh.2023.1209633>
- Barbalho, Ingridy M. P., Felipe Fernandes, Daniele M. S. Barros, Jailton C. Paiva, Jorge Henriques, Antônio H. F. Morais, Karilany D. Coutinho, Giliate C. Coelho Neto, Arthur Chioro e Ricardo A. M. Valentim (2022), 'Electronic health records in brazil: Prospects and technological challenges', *Frontiers in Public Health* **10**, 963841.
URL: <https://doi.org/10.3389/fpubh.2022.963841>
- Barbalho, Ingridy, Ricardo Valentim, Mário Dourado Júnior, Daniele Barros, Hércules Pedrosa Júnior, Felipe Fernandes, César Teixeira, Thaísa Lima, Jailton Paiva e Danilo Nagem (2021), 'National registry for amyotrophic lateral sclerosis: a systematic review for structuring population registries of motor neuron diseases', *BMC Neurology* **21**(1), 269.
- Bede, Peter, Aizuri Murad e Orla Hardiman (2022), 'Pathological neural networks and artificial neural networks in als: diagnostic classification based on pathognomonic neuroimaging features', *Journal of Neurology* **269**(5), 2440–2452.
URL: <https://doi.org/10.1007/s00415-021-10801-5>
- Bekios-Calfa, Juan, José M Buenaposada e Luis Baumela (2014), 'Robust gender recognition by exploiting facial attributes dependencies', *Pattern recognition letters* **36**, 228–234.
- Belhumeur, Peter N, David W Jacobs, David J Kriegman e Neeraj Kumar (2013), 'Localizing parts of faces using a consensus of exemplars', *IEEE transactions on pattern analysis and machine intelligence* **35**(12), 2930–2940.
- Blignaut, Pieter (2017), 'Development of a gaze-controlled support system for a person in an advanced stage of multiple sclerosis: a case study', *Universal Access in the Information Society* **16**(4), 1003–1016.
URL: <https://doi.org/10.1007/s10209-016-0493-9>
- Bradski, Gary R (1998), 'Computer vision face tracking for use in a perceptual user interface'.

- Brasil (2015), Lei nº 13.146, de 6 de julho de 2015. lei brasileira de inclusão da pessoa com deficiência (estatuto da pessoa com deficiência). Presidência da República. Secretaria-Geral. Subchefia para Assuntos Jurídicos. Available from: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/113146.htm. [Acesso 10 Ago 2023].
- Brasil (2018), Estratégia nacional de ciência, tecnologia e inovação 2016/2022: Sumário executivo. Ministério da Ciência, Tecnologia, Inovações e Comunicações. Available from: https://antigo.mctic.gov.br/mctic/export/sites/institucional/ciencia/SEPED/Arquivos/PlanosDeAcao/PACTI_Sumario_executivo_Web.pdf. [Acesso 10 Ago 2023].
- Brasil (2019), Projeto de lei nº 4691, de 2019. Atividade Legislativa. Senado Federal. Brasília, DF. Available from: <https://www25.senado.leg.br/web/atividade/materias/-/materia/138326>. [Accessed 10 July 2023].
- Brasil (2021a), Decreto nº 10.645, de 11 de março de 2021. regulamenta o art. 75 da lei nº 13.146, de 6 julho de 2015, para dispor sobre as diretrizes, os objetivos e os eixos do plano nacional de tecnologia assistiva. Presidência da República. Secretaria-Geral. Subchefia para Assuntos Jurídicos. Available from: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2021/Decreto/D10645.htm#art1. [Acesso 10 Jul 2023].
- Brasil (2021b), Lei nº 10924, de 10 de junho de 2021. Diário Oficial do Rio Grande do Norte. Available from: http://diariooficial.rn.gov.br/dei/dorn3/docview.aspx?id_jor=00000001&data=20210611&id_doc=726286. [Acesso 10 Jul 2023].
- Brasil (2021c), Portaria mcti nº 5.366, de 02.12.2021. plano nacional de tecnologia assistiva (pnta). Ministério da Ciência, Tecnologia e Inovações. Comitê Interministerial de Tecnologia Assistiva. Available from: https://www.gov.br/mcti/pt-br/centrais-de-conteudo/publicacoes-mcti/plano-nacional-de-tecnologia-assistiva/pnta_documento_web.pdf. [Acesso 16 Ago 2023].
- Brasil (2022), Portaria mcti nº 6.033, de 24.06.2022. sistema nacional de laboratórios de tecnologia assistiva (sisassistiva-mcti). Ministério da Ciência, Tecnologia e Inovações. Available from: https://antigo.mctic.gov.br/mctic/opencms/legislacao/portarias/Portaria_MCTI_n_6033_de_24062022.html. [Acesso 10 Ago 2023].
- Brasil (2023a), IBGE - Instituto Brasileiro de Geografia e Estatística. Censo Demográfico 2022. População e domicílios. Available from: <https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-demografico-2022.html?edicao=37225&t=resultados>. [Acesso em 7 Jul 2023].

- Brasil (2023b), IBGE - Pesquisa Nacional por Amostra de Domicílios Contínua - PNAD Contínua. Pessoas com deficiência 2022. Available from: https://biblioteca.ibge.gov.br/visualizacao/livros/liv102013_informativo.pdf. [Acesso em 7 Jul 2023].
- Brereton, Pearl, Barbara A. Kitchenham, David Budgen, Mark Turner e Mohamed Khalil (2007), 'Lessons from applying the systematic literature review process within the software engineering domain', *Journal of Systems and Software* **80**(4), 571 – 583. Software Performance.
URL: <http://www.sciencedirect.com/science/article/pii/S016412120600197X>
- Cedarbaum, Jesse M., Nancy Stambler, Errol Malta, Cynthia Fuller, Dana Hilt, Barbara Thurmond e Arline Nakanishi (1999), 'The alsfrs-r: a revised als functional rating scale that incorporates assessments of respiratory function', *Journal of the Neurological Sciences* **169**(1), 13–21.
URL: <https://www.sciencedirect.com/science/article/pii/S0022510X99002105>
- Chang, Won-Du, Ho-Seung Cha, Do Yeon Kim, Seung Hyun Kim e Chang-Hwan Im (2017), 'Development of an electrooculogram-based eye-computer interface for communication of individuals with amyotrophic lateral sclerosis', *Journal of NeuroEngineering and Rehabilitation* **14**(1), 89.
URL: <https://doi.org/10.1186/s12984-017-0303-5>
- Chareonsuk, W., S. Kanhaun, K. Khawkam e D. Wongsawang (2016), Face and eyes mouse for als patients, em '2016 Fifth ICT International Student Project Conference (ICT-ISPC)', pp. 77–80.
- Chatterjee, S., K. Samanta, N. R. Choudhury e R. Bose (2019), 'Detection of myopathy and als electromyograms employing modified window stockwell transform', *IEEE Sensors Letters* **3**(7), 1–4.
- Chau, Michael e Margrit Betke (2005), 'Real time eye tracking and blink detection with usb cameras'.
URL: <https://open.bu.edu/handle/2144/1839>
- Chaudhary, Ujwal, Ioannis Vlachos, Jonas B. Zimmermann, Arnau Espinosa, Alessandro Tonin, Andres Jaramillo-Gonzalez, Majid Khalili-Ardali, Helge Topka, Jens Lehmberg, Gerhard M. Friehs, Alain Woodtli, John P. Donoghue e Niels Birbaumer (2022), 'Spelling interface using intracortical signals in a completely locked-in patient enabled via auditory neurofeedback training', *Nature Communications* **13**(1), 1236.
URL: <https://doi.org/10.1038/s41467-022-28859-8>
- Chaudhary, Ujwal, Niels Birbaumer e Ander Ramos-Murguialday (2016), 'Brain-computer interfaces for communication and rehabilitation', *Nature Reviews Neurology* **12**(9), 513–525.
URL: <https://doi.org/10.1038/nrneurol.2016.113>

- Chen, Jiachen e W. Kenneth Jenkins (2017), Facial recognition with pca and machine learning methods, em ‘2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)’, pp. 973–976.
- Chen, Qiu-Feng, Xiao-Hong Zhang, Nao-Xin Huang e Hua-Jun Chen (2020), ‘Identification of amyotrophic lateral sclerosis based on diffusion tensor imaging and support vector machine’, *Frontiers in Neurology* **11**.
URL: <https://www.frontiersin.org/articles/10.3389/fneur.2020.00275>
- Cooney, C., R. Folli e D. Coyle (2019), Optimizing layers improves cnn generalization and transfer learning for imagined speech decoding from eeg, em ‘2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)’, pp. 1311–1316.
- Corcia, Philippe, Christian Lunetta, Patrick Vourc’h, Pierre-François Pradat e H  l  ne Blasco (2023), ‘Time for optimism in amyotrophic lateral sclerosis’, *European Journal of Neurology* **30**(5), 1459–1464.
- Dash, Debadatta, Paul Ferrari, Angel Hernandez, Daragh Heitzman, Sara G. Austin e Jun Wang (2020), Neural speech decoding for amyotrophic lateral sclerosis, em ‘Proc. Interspeech 2020’, pp. 2782–2786.
URL: <http://dx.doi.org/10.21437/Interspeech.2020-3071>
- Dash, Debadatta, Paul Ferrari e Jun Wang (2020), ‘Decoding imagined and spoken phrases from non-invasive neural (meg) signals’, *Frontiers in Neuroscience* **14**, 290.
URL: <https://www.frontiersin.org/article/10.3389/fnins.2020.00290>
- de Lima Medeiros, Paulo Augusto, Gabriel Vin  cius Souza da Silva, Felipe Ricardo dos Santos Fernandes, Ignacio S  nchez-Gendr  z, Hertz Wilton Castro Lins, Daniele Montenegro da Silva Barros, Danilo Alves Pinto Nagem e Ricardo Alexsandro de Medeiros Valentim (2022), ‘Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam’, *Expert Systems with Applications* **188**, 116073.
URL: <https://doi.org/10.1016/j.eswa.2021.116073>
- de Souza, Aline Alves, Stephano Tomaz da Silva, Karen de Medeiros Pondofe, Vanessa Regiane Resqueti, Luciana Prot  sio de Melo, Ricardo Alexsandro de Medeiros Valentim e Tatiana Souza Ribeiro (2022), ‘Remote versus face-to-face home-based exercise programme in people with amyotrophic lateral sclerosis: protocol for a randomised clinical trial’, *BMJ Open* **12**(5).
URL: <https://bmjopen.bmj.com/content/12/5/e056323>
- Dietrich-Neto, Fl  via, Dagoberto Callegaro, Elza Dias-Tosta, Helga Almeida Silva, Maria Elizabeth Ferraz, Jos   Mauro Braz De Lima e Acary Souza Bulle Oliveira (2000), ‘Amyotrophic lateral sclerosis in brazil: 1998 national survey’, *Arquivos de Neuro-Psiquiatria* **58**, 607–615.

- Doulah, A. B. M. S. U., S. A. Fattah, W. . Zhu e M. O. Ahmad (2014), ‘Wavelet domain feature extraction scheme based on dominant motor unit action potential of emg signal for neuromuscular disease classification’, *IEEE Transactions on Biomedical Circuits and Systems* **8**(2), 155–164.
- Drutarovsky, Tomas e Andrej Fogelton (2014), Eye blink detection using variance of motion vectors, *em ‘European Conference on Computer Vision’*, Springer, pp. 436–448.
- Ebert, Achim, Nahum D. Gershon e Gerrit C. van der Veer (2012), ‘Human-computer interaction’, *KI - Künstliche Intelligenz* **26**(2), 121–126.
URL: <https://doi.org/10.1007/s13218-012-0174-7>
- Eom, Yeonkyung, Shenglin Mu, Shibata Satoru e Tungkuan Liu (2019), A method to estimate eye gaze direction when wearing glasses, *em ‘2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)’*, pp. 1–6.
- Ephrat, A., I. Mosseri, O. Lang, T. Dekel, K Wilson, A. Hassidim, W. T. Freeman e M. Rubinstein (2018), ‘Looking to listen at the cocktail party: A speaker-independent audio-visual model for speech separation’, *arXiv preprint arXiv:1804.03619* .
- Erazo, Daniells, Jaime Luna, Pierre-Marie Preux, Farid Boumediene e Philippe Couratier (2022), ‘Epidemiological and genetic features of amyotrophic lateral sclerosis in latin america and the caribbean: a systematic review’, *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **23**(1-2), 4–15. PMID: 33871294.
URL: <https://doi.org/10.1080/21678421.2021.1909066>
- Everingham, Mark, Luc Gool, Christopher K. Williams, John Winn e Andrew Zisserman (2010), ‘The pascal visual object classes (voc) challenge’, *Int. J. Comput. Vision* **88**(2), 303–338.
URL: <https://doi.org/10.1007/s11263-009-0275-4>
- Faceli, Katti, Ana Carolina Lorena, João Gama, Tiago Agostinho de Almeida e Andre Carlos Ponce de Leon Ferreira de Carvalho (2021), *Inteligência artificial: uma abordagem de aprendizado de máquina*, 2ª edição, LTC, Rio de Janeiro, RJ, Brasil.
- Fathi, Abdolhossein e Fardin Abdali-Mohammadi (2015), ‘Camera-based eye blinks pattern detection for intelligent mouse’, *Signal, Image and Video Processing* **9**(8), 1907–1916.
URL: <https://doi.org/10.1007/s11760-014-0680-1>
- Fernandes, Felipe e Ingridy Barbalho (2021), ‘Camera-based eye interaction techniques for amyotrophic lateral sclerosis individuals: A systematic review’. PROSPERO 2021 CRD42021230721.
- Fernandes, Felipe, Ingridy Barbalho, Arnaldo Bispo Júnior, Luca Alves, Danilo Nagem, Hertz Lins, Ernano Arrais Júnior, Karilany D. Coutinho, Antônio H. F. Morais, João Paulo Q. Santos, Guilherme Medeiros Machado, Jorge Henriques, César Teixeira,

- Mário E. T. Dourado Júnior, Ana R. R. Lindquist e Ricardo A. M. Valentim (2023), 'Digital alternative communication for individuals with amyotrophic lateral sclerosis: What we have', *Journal of Clinical Medicine* **12**(16), 5235.
- Fernandes, Felipe, Ingridy Barbalho, Daniele Barros, Ricardo Valentim, César Teixeira, Jorge Henriques, Paulo Gil e Mário Dourado Júnior (2021), 'Biomedical signals and machine learning in amyotrophic lateral sclerosis: a systematic review', *BioMedical Engineering OnLine* **20**(1), 61.
URL: <https://doi.org/10.1186/s12938-021-00896-2>
- Ferraro, Pilar M, Federica Agosta, Nilo Riva, Massimiliano Copetti, Edoardo Gioele Spinelli, Yuri Falzone, Gianni Sorarù, Giancarlo Comi, Adriano Chiò e Massimo Filippi (2017), 'Multimodal structural mri in the diagnosis of motor neuron diseases', *NeuroImage: Clinical* **16**, 240–247.
- Fidelix, Emanuela Coriolano, Glauciane Costa Santana, Daniele Montenegro da Silva Barros e Mário Emílio Teixeira Dourado Junior (2023), 'Telehealth for amyotrophic lateral sclerosis in a multidisciplinary service in a brazilian reference center', *Arq Neuropsiquiatr* **81**(05), 469–474.
- Fogelton, Andrej e Wanda Benesova (2016), 'Eye blink detection based on motion vectors analysis', *Computer Vision and Image Understanding* **148**, 23–33.
- Fogelton, Andrej e Wanda Benesova (2018), 'Eye blink completeness detection', *Computer Vision and Image Understanding* **176**, 78–85.
- Forsyth, David A e Jean Ponce (2011), *Computer vision: a modern approach*, 2ª edição, Prentice Hall, Upper Saddle River, New Jersey.
- Genge, Angela e Adriano Chio (2023), 'The future of als diagnosis and staging: where do we go from here?', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **24**(3-4), 165–174. PMID: 36503299.
URL: <https://doi.org/10.1080/21678421.2022.2150555>
- Géron, Aurélien (2022), *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*, 3ª edição, "O'Reilly Media, Inc.", Sebastopol, CA.
- Gillespie, Jacqueline, Antoine Przybylak-Brouillard e Christine L Watt (2021), 'The palliative care information needs of patients with amyotrophic lateral sclerosis and their informal caregivers: A scoping review', *Journal of Pain and Symptom Management* .
URL: <https://www.sciencedirect.com/science/article/pii/S0885392421002323>
- Goetz, Christopher G. (2000), 'Amyotrophic lateral sclerosis: Early contributions of jean-martin charcot', *Muscle & Nerve* **23**(3), 336–343.
- Gokgoz, Ercan e Abdulhamit Subasi (2014), 'Effect of multiscale pca de-noising on emg signal classification for diagnosis of neuromuscular disorders', *Journal of medical systems* **38**(4), 31.

- Gokgoz, Ercan e Abdulhamit Subasi (2015), ‘Comparison of decision tree algorithms for emg signal classification using dwt’, *Biomedical Signal Processing and Control* **18**, 138–144.
- Goodfellow, Ian, Yoshua Bengio e Aaron Courville (2016), *Deep Learning*, MIT Press.
<http://www.deeplearningbook.org>.
- Gordon, Jonathan e Boaz Lerner (2019), ‘Insights into amyotrophic lateral sclerosis from a machine learning perspective’, *Journal of clinical medicine* **8**(10), 1578.
- Goutman, Stephen A, Orla Hardiman, Ammar Al-Chalabi, Adriano Chió, Masha G Savelieff, Matthew C Kiernan e Eva L Feldman (2022), ‘Emerging insights into the complex genetics and pathophysiology of amyotrophic lateral sclerosis’, *The Lancet Neurology* **21**(5), 465–479.
URL: <https://www.sciencedirect.com/science/article/pii/S1474442221004142>
- Graimann, Bernhard, Brendan Allison e Gert Pfurtscheller (2009), Brain–computer interfaces: A gentle introduction, *em* ‘Brain-computer interfaces’, Springer, pp. 1–27.
- Grauman, K., M. Betke, J. Gips e G.R. Bradski (2001), Communication via eye blinks - detection and duration analysis in real time, *em* ‘Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001’, Vol. 1, pp. I–I.
- Greco, Alberto, Maria Rosa Chiesa, Ilaria Da Prato, Anna Maria Romanelli, Cristina Dolciotti, Gabriella Cavallini, Silvia Maria Masciandaro, Enzo Pasquale Scilingo, Renata Del Carratore e Paolo Bongioanni (2021), ‘Using blood data for the differential diagnosis and prognosis of motor neuron diseases: a new dataset for machine learning applications’, *Scientific Reports* **11**(1), 3371.
URL: <https://doi.org/10.1038/s41598-021-82940-8>
- Grollemund, Vincent, Gaétan Le Chat, Marie-Sonia Secchi-Buhour, François Delbot, Jean-François Pradat-Peyre, Peter Bede e Pierre-François Pradat (2020), ‘Development and validation of a 1-year survival prognosis estimation model for amyotrophic lateral sclerosis using manifold learning algorithm umap’, *Scientific Reports* **10**(1), 13378.
URL: <https://doi.org/10.1038/s41598-020-70125-8>
- Grollemund, Vincent, Gaétan Le Chat, Marie-Sonia Secchi-Buhour, François Delbot, Jean-François Pradat-Peyre, Peter Bede e Pierre-François Pradat (2021), ‘Manifold learning for amyotrophic lateral sclerosis functional loss assessment’, *Journal of Neurology* **268**(3), 825–850.
URL: <https://doi.org/10.1007/s00415-020-10181-2>
- Grollemund, Vincent, Pierre-François Pradat, Giorgia Querin, François Delbot, Gaétan Le Chat, Jean-François Pradat-Peyre e Peter Bede (2019), ‘Machine learning in

- amyotrophic lateral sclerosis: Achievements, pitfalls, and future directions’, *Frontiers in Neuroscience* **13**, 135.
URL: <https://www.frontiersin.org/article/10.3389/fnins.2019.00135>
- Gromicho, Marta, Tiago Leão, Miguel Oliveira Santos, Susana Pinto, Alexandra M. Carvalho, Sara C. Madeira e Mamede De Carvalho (2022), ‘Dynamic bayesian networks for stratification of disease progression in amyotrophic lateral sclerosis’, *European Journal of Neurology* **29**(8), 2201–2210.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/ene.15357>
- Gu, Zaiwang, Jun Cheng, Huazhu Fu, Kang Zhou, Huaying Hao, Yitian Zhao, Tianyang Zhang, Shenghua Gao e Jiang Liu (2019), ‘Ce-net: Context encoder network for 2d medical image segmentation’, *IEEE transactions on medical imaging* **38**(10), 2281–2292.
- Güell, Maria Rosa, Antonio Antón, Ricardo Rojas-García, Carmen Puy, Jesus Pradas et al. (2013), ‘Comprehensive care of amyotrophic lateral sclerosis patients: a care model’, *Archivos de Bronconeumología (English Edition)* **49**(12), 529–533.
- Gutmann, Michelle L e Rajinder Koul (2023), Aac intervention for persons with acquired disorders, em D. R.Fuller e L. L.Lloyd, eds., ‘Principles and Practices in Augmentative and Alternative Communication’, 1edª edição, Vol. 1, SLACK Incorporated, Thorofare, NJ 08086 USA, pp. 415–437.
- Hardiman, Orla, Ammar Al-Chalabi, Adriano Chio, Emma M. Corr, Giancarlo Logroscino, Wim Robberecht, Pamela J. Shaw, Zachary Simmons e Leonard H. van den Berg (2017), ‘Amyotrophic lateral sclerosis’, *Nature Reviews Disease Primers* **3**(1), 17071.
URL: <https://doi.org/10.1038/nrdp.2017.71>
- Harezlak, Katarzyna e Pawel Kasproski (2018), ‘Application of eye tracking in medicine: A survey, research issues and challenges’, *Computerized Medical Imaging and Graphics* **65**, 176 – 190. Advances in Biomedical Image Processing.
URL: <http://www.sciencedirect.com/science/article/pii/S0895611117300435>
- Hazarika, A., L. Dutta, M. Barthakur e M. Bhuyan (2019), ‘A multiview discriminant feature fusion-based nonlinear process assessment and diagnosis: Application to medical diagnosis’, *IEEE Transactions on Instrumentation and Measurement* **68**(7), 2498–2506.
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren e Jian Sun (2015), ‘Delving deep into rectifiers: Surpassing human-level performance on imagenet classification’, *CoRR* **abs/1502.01852**.
- Herff, Christian, Dominic Heger, Adriana de Pestors, Dominic Telaar, Peter Brunner, Gerwin Schalk e Tanja Schultz (2015), ‘Brain-to-text: decoding spoken phrases from phone representations in the brain’, *Frontiers in Neuroscience* **9**, 217.
URL: <https://www.frontiersin.org/article/10.3389/fnins.2015.00217>

- Holmqvist, Kenneth, Saga Lee Örbom, Ignace T. C. Hooge, Diederick C. Niehorster, Robert G. Alexander, Richard Andersson, Jeroen S. Benjamins, Pieter Blignaut, Anne-Marie Brouwer, Lewis L. Chuang, Kirsten A. Dalrymple, Denis Drieghe, Matt J. Dunn, Ulrich Ettinger, Susann Fiedler, Tom Foulsham, Jos N. van der Gest, Dan Witzner Hansen, Samuel B. Hutton, Enkelejda Kasneci, Alan Kingstone, Paul C. Knox, Ellen M. Kok, Helena Lee, Joy Yeonjoo Lee, Jukka M. Leppänen, Stephen Macknik, Päivi Majaranta, Susana Martinez-Conde, Antje Nuthmann, Marcus Nyström, Jacob L. Orquin, Jorge Otero-Millan, Soon Young Park, Stanislav Popelka, Frank Proudlock, Frank Renkewitz, Austin Roorda, Michael Schulte-Mecklenbeck, Bonita Sharif, Frederick Shic, Mark Shovman, Mervyn G. Thomas, Ward Venrooij, Raimondas Zemblys e Roy S. Hessels (2022), ‘Eye tracking: empirical foundations for a minimal reporting guideline’, *Behavior Research Methods*.
URL: <https://doi.org/10.3758/s13428-021-01762-8>
- Hori, J., K. Sakano e Y. Saitoh (2004), Development of communication supporting device controlled by eye movements and voluntary eye blink, *em* ‘The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society’, Vol. 2, pp. 4302–4305.
- Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto e Hartwig Adam (2017), ‘Mobilenets: Efficient convolutional neural networks for mobile vision applications’, *CoRR* **abs/1704.04861**.
URL: <http://arxiv.org/abs/1704.04861>
- Howard, Ileana M. e Kathleen Burgess (2021), ‘Telehealth for amyotrophic lateral sclerosis and multiple sclerosis’, *Physical Medicine and Rehabilitation Clinics of North America* **32**(2), 239–251. Telerehabilitation.
URL: <https://www.sciencedirect.com/science/article/pii/S1047965120301078>
- Hwang, Ing-Shiou, Yi-Ying Tsai, Bo-Han Zeng, Chien-Ming Lin, Huei-Sheng Shiue e Gwo-Ching Chang (2020), ‘Integration of eye tracking and lip motion for hands-free computer access’, *Universal Access in the Information Society*.
URL: <https://doi.org/10.1007/s10209-020-00723-w>
- Iadanza, Ernesto, Rachele Fabbri, Francesco Goretti, Giovanni Nardo, Elena Niccolai, Caterina Bendotti e Amedeo Amedei (2022), ‘Machine learning for analysis of gene expression data in fast- and slow-progressing amyotrophic lateral sclerosis murine models’, *Biocybernetics and Biomedical Engineering* **42**(1), 273–284.
URL: <https://www.sciencedirect.com/science/article/pii/S0208521622000043>
- ISO (1998), Iso 9241-11:1998 - ergonomic requirements for office work with visual display terminals (vdts) – part 11: Guidance on usability. Available from: <https://www.iso.org/standard/16883.html>. [Acesso em 15 Ago 2023].
- Jaramillo-Yáñez, Andrés, Marco E. Benalcázar e Elisa Mena-Maldonado (2020), ‘Real-time hand gesture recognition using surface electromyography and machine learning’.

- ning: A systematic literature review’, *Sensors* **20**(9).
URL: <https://www.mdpi.com/1424-8220/20/9/2467>
- Karam, Chafic Y, Sabrina Paganoni, Nanette Joyce, Gregory T Carter e Richard Bedlack (2016), ‘Palliative care issues in amyotrophic lateral sclerosis: an evidenced-based review’, *American Journal of Hospice and Palliative Medicine*® **33**(1), 84–92.
- Kaushik, Rishabh, Tushar Arora, Sukanya e Rohit Tripathi (2018), Design of eyewriter for als patients through eyecan, em ‘2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)’, pp. 991–995.
- Kavale, Kunal, Kiran Kokambe e Sunita Jadhav (2018), taskeye: "a novel approach to help people interact with their surrounding through their eyes", em ‘2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)’, pp. 311–313.
- Kayadibi, Ismail, Gür Emre Güraksın, Uçman Ergün e Nurgül Özmen Süzme (2022), ‘An eye state recognition system using transfer learning: Alexnet-based deep convolutional neural network’, *International Journal of Computational Intelligence Systems* **15**(1), 49.
URL: <https://doi.org/10.1007/s44196-022-00108-2>
- Kazemi, Vahid e Josephine Sullivan (2014), One millisecond face alignment with an ensemble of regression trees, em ‘Proceedings of the IEEE conference on computer vision and pattern recognition’, pp. 1867–1874.
- Keele, Staffs (2007), Guidelines for performing systematic literature reviews in software engineering, Relatório técnico, Keele University and University of Durham, UK. Version 2.3, EBSE Technical Report, EBSE-2007-01.
- Khorasani, Abed, Mohammad Reza Daliri e Mohammad Pooyan (2016), ‘Recognition of amyotrophic lateral sclerosis disease using factorial hidden markov model’, *Biomedical Engineering/Biomedizinische Technik* **61**(1), 119–126.
- King, Davis E. (2009), ‘Dlib-ml: A machine learning toolkit’, *Journal of Machine Learning Research* **10**, 1755–1758.
- Kitchenham, Barbara (2004), Procedures for performing systematic reviews, Relatório técnico, Keele University, Department of Computer Science, Software Engineering Group and Empirical Software Engineering National ICT Australia Ltd., Keele, Staffs, ST5 5BG, UK.
- Kitchenham, Barbara Ann, David Budgen e Pearl Brereton (2016), *Evidence-Based Software Engineering and Systematic Reviews*, 1ª edição, Chapman and Hall/CRC, New York.
- Kocar, Thomas D., Anna Behler, Albert C. Ludolph, Hans-Peter Müller e Jan Kassubek (2021), ‘Multiparametric microstructural mri and machine learning classification

- yields high diagnostic accuracy in amyotrophic lateral sclerosis: Proof of concept', *Frontiers in Neurology* **12**, 745475.
URL: <https://www.frontiersin.org/articles/10.3389/fneur.2021.745475>
- Kocar, Thomas D., Hans-Peter Müller, Albert C. Ludolph e Jan Kassubek (2021), 'Feature selection from magnetic resonance imaging data in als: a systematic review', *Therapeutic Advances in Chronic Disease* **12**, 20406223211051002. PMID: 34729157.
URL: <https://doi.org/10.1177/20406223211051002>
- Krapic, Luka, Kristijan Lenac e Sandi Ljubic (2015), 'Integrating blink click interaction into a head tracking system: implementation and usability issues', *Universal Access in the Information Society* **14**(2), 247–264.
URL: <https://doi.org/10.1007/s10209-013-0343-y>
- Krizhevsky, Alex, Ilya Sutskever e Geoffrey E Hinton (2012), Imagenet classification with deep convolutional neural networks, em 'Advances in neural information processing systems', pp. 1097–1105.
- Królak, Aleksandra e Paweł Strumiłło (2012), 'Eye-blink detection system for human–computer interaction', *Universal Access in the Information Society* **11**(4), 409–419.
URL: <https://doi.org/10.1007/s10209-011-0256-6>
- Kumar, L Ashok e D Karthika Renuka (2023), *Deep Learning Approach for Natural Language Processing, Speech, and Computer Vision: Techniques and Use Cases*, 1ª edição, CRC Press, Boca Raton, FL.
- Kwon, Kyung-Ah, Rebecca Shipley, Mohan Edirisinghe, Daniel Ezra, Geoff Rose, Serena Best e Ruth Cameron (2013), 'High-speed camera characterization of voluntary eye blinking kinematics', *Journal of the Royal Society, Interface / the Royal Society* **10**, 20130227.
- Larson, A., J. Herrera, K. George e A. Matthews (2017), Electrooculography based electronic communication device for individuals with als, em '2017 IEEE Sensors Applications Symposium (SAS)', pp. 1–5.
- Le, Vuong, Jonathan Brandt, Zhe Lin, Lubomir Bourdev e Thomas S Huang (2012), Interactive facial feature localization, em 'European conference on computer vision', Springer, pp. 679–692.
- LeCun, Yann, Léon Bottou, Yoshua Bengio e Patrick Haffner (1998), 'Gradient-based learning applied to document recognition', *Proceedings of the IEEE* **86**(11), 2278–2324.
- Leão, Tiago, Sara C. Madeira, Marta Gromicho, Mamede de Carvalho e Alexandra M. Carvalho (2021), 'Learning dynamic bayesian networks from time-dependent and time-independent data: Unraveling disease progression in amyotrophic lateral sclerosis', *Journal of Biomedical Informatics* **117**, 103730.
URL: <https://www.sciencedirect.com/science/article/pii/S1532046421000599>

- Li, Y., M. Chang e S. Lyu (2018), In icu oculi: Exposing ai created fake videos by detecting eye blinking, *em* '2018 IEEE International Workshop on Information Forensics and Security (WIFS)', pp. 1–7.
- Likert, Rensis (1932), 'A technique for the measurement of attitudes', *Archives of psychology* .
- Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár e C. Lawrence Zitnick (2014), Microsoft coco: Common objects in context, *em* D.Fleet, T.Pajdla, B.Schiele e T.Tuytelaars, eds., 'Computer Vision – ECCV 2014', Springer International Publishing, Cham, pp. 740–755.
- Lingegowda, D. R., K. Amrutesh e S. Ramanujam (2017), Electrooculography based assistive technology for als patients, *em* '2017 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)', pp. 36–40.
- Linse, Katharina, Elisa Aust, Markus Joos e Andreas Hermann (2018), 'Communication matters—pitfalls and promise of hightech communication devices in palliative care of severely physically disabled patients with amyotrophic lateral sclerosis', *Frontiers in Neurology* **9**.
- Linse, Katharina, Wolfgang Rüter, Markus Joos, Henning Schmitz-Peiffer, Alexander Storch e Andreas Hermann (2018), 'Usability of eyetracking computer systems and impact on psychological wellbeing in patients with advanced amyotrophic lateral sclerosis', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **19**(3-4), 212–219. PMID: 29092645.
URL: <https://doi.org/10.1080/21678421.2017.1392576>
- Liu, Shuo Samuel, Andrew Rawicz, Teng Ma, Cheng Zhang, Kyle Lin, Siavash Rezaei e Eion Wu (2012), 'An eye-gaze tracking and human computer interface system for people with als and other locked-in diseases', *Journal of Medical and Biological Engineering* **32**(2).
- Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu e Alexander C. Berg (2015), 'SSD: single shot multibox detector', *CoRR* **abs/1512.02325**.
URL: <http://arxiv.org/abs/1512.02325>
- Liu, Yi, Bu-Sung Lee, Deepu Rajan, Andrzej Sluzek e Martin J. McKeown (2019), 'Camtype: assistive text entry using gaze with an off-the-shelf webcam', *Machine Vision and Applications* **30**(3), 407–421.
URL: <https://doi.org/10.1007/s00138-018-00997-4>
- Liu, Yi-Hung, Shiuan Huang e Yi-De Huang (2017), 'Motor imagery eeg classification for patients with amyotrophic lateral sclerosis using fractal dimension and fisher's criterion-based channel selection', *Sensors* **17**(7), 1557.

- Londral, Ana, Anabela Pinto, Susana Pinto, Luis Azevedo e Mamede De Carvalho (2015), 'Quality of life in amyotrophic lateral sclerosis patients and caregivers: Impact of assistive communication from early stages', *Muscle & Nerve* **52**(6), 933–941.
- Mainsah, B O, L M Collins, K A Colwell, E W Sellers, D B Ryan, K Caves e C S Throckmorton (2015), 'Increasing BCI communication rates with dynamic stopping towards more practical use: an ALS study', *Journal of Neural Engineering* **12**(1), 016013.
- Martin, Robert C (2019), *Arquitetura limpa: o guia do artesão para estrutura e design de software*, Alta Books Editora.
- Martinez, Aleix e Shichuan Du (2012), 'A model of the perception of facial expressions of emotion by humans: Research overview and perspectives', *J. Mach. Learn. Res.* **13**(null), 1589–1608.
- McCane, Lynn M., Eric W. Sellers, Dennis J. McFarland, Joseph N. Mak, C. Steve Carmack, Debra Zeitlin, Jonathan R. Wolpaw e Theresa M. Vaughan (2014), 'Brain-computer interface (bci) evaluation in people with amyotrophic lateral sclerosis', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **15**(3-4), 207–215. PMID: 24555843.
URL: <https://doi.org/10.3109/21678421.2013.865750>
- Melo, Adriana e Ricardo Abelheira (2015), *Design Thinking & Thinking Design: Metodologia, ferramentas e uma reflexão sobre o tema*, Novatec Editora.
- Mercadante, Sebastiano e Lou'i Al-Husinat (2023), 'Palliative care in amyotrophic lateral sclerosis', *Journal of Pain and Symptom Management* .
- Miao, Yangyang, Erwei Yin, Brendan Z Allison, Yu Zhang, Yan Chen, Yi Dong, Xingyu Wang, Dewen Hu, Andrzej Chchocki e Jing Jin (2020), 'An erp-based bci with peripheral stimuli: validation with als patients', *Cognitive Neurodynamics* **14**(1), 21–33.
- Missimer, Eric e Margrit Betke (2010), Blink and wink detection for mouse pointer control, em 'Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments', PETRA '10, Association for Computing Machinery, New York, NY, USA.
- Mitchell, Tom M (1997), *Machine learning*, Vol. 1, 1ª edição, McGraw-Hill, New York, NY, USA.
- Moura, Mirian Conceicao, Luiz Augusto Casulari e Maria Rita Carvalho Garbi Novaes (2016), 'Ethnic and demographic incidence of amyotrophic lateral sclerosis (als) in brazil: A population based study', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **17**, 275–281.

- Mu, Shenglin, Satoru Shibata, Kuo chun Chiu, Tomonori Yamamoto e Tung kuan Liu (2022), ‘Study on eye-gaze input interface based on deep learning using images obtained by multiple cameras’, *Computers and Electrical Engineering* **101**, 108040.
URL: <https://www.sciencedirect.com/science/article/pii/S0045790622003019>
- Myszczyńska, Monika A., Poojitha N. Ojames, Alix M. B. Lacoste, Daniel Neil, Amir Saffari, Richard Mead, Guillaume M. Hautbergue, Joanna D. Holbrook e Laura Ferraiuolo (2020), ‘Applications of machine learning to diagnosis and treatment of neurodegenerative diseases’, *Nature Reviews Neurology* **16**(8), 440–456.
URL: <https://doi.org/10.1038/s41582-020-0377-8>
- Nagem, Danilo, Ricardo Valentim, Milena Duarte e Karilany Coutinho (2023), Projeto revela: desenvolvimento científico e tecnológico aplicado à esclerose lateral amiotrófica - relatório parcial. Available from: https://repositorio.ufrn.br/bitstream/123456789/52691/1/ProjetoRevELA_Nagem_2023. [Acesso em 14 Jul 2023].
- Nakazawa, Nobuaki, Shinnosuke Aikawa e Toshikazu Matsui (2018), Development of communication aid device for disabled persons using corneal surface reflection image, em ‘Proceedings of the 2nd International Conference on Graphics and Signal Processing’, ICGSP’18, Association for Computing Machinery, New York, NY, USA, p. 16–20.
URL: <https://doi-org.ez18.periodicos.capes.gov.br/10.1145/3282286.3282298>
- Nussbaumer, H. J. (1981), *Two-Dimensional Convolution and DFT Computation*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 37–88.
URL: <https://doi.org/10.1007/BFb0057595>
- Oliveira, Roberto Teodoro Gurgel De, Glauciane Costa Santana, Maria De Jesus Gonçalves, Guilherme Augusto De Freitas Fregonezi, Sancha Helena De Lima Vale, Lucia Leite-Lais e Mário Emílio Teixeira Dourado Júnior (2023), ‘A geographical study on amyotrophic lateral sclerosis in rio grande do norte, brazil, from 2005 to 2018’, *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **24**(1-2), 117–124. PMID: 35916197.
URL: <https://doi.org/10.1080/21678421.2022.2102429>
- Ouzzani, Mourad, Hossam Hammady, Zbys Fedorowicz e Ahmed Elmagarmid (2016), ‘Rayyan—a web and mobile app for systematic reviews’, *Systematic Reviews* **5**(1), 210.
URL: <http://dx.doi.org/10.1186/s13643-016-0384-4>
- Oyabu, Yuki, Hironobu Takano e Kiyomi Nakamura (2012), Development of the eye input device using eye movement obtained by measuring the center position of the pupil, em ‘2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)’, pp. 2948–2952.

- Page, Matthew J, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, Roger Chou, Julie Glanville, Jeremy M Grimshaw, Asbjørn Hróbjartsson, Manoj M Lalu, Tianjing Li, Elizabeth W Loder, Evan Mayo-Wilson, Steve McDonald, Luke A McGuinness, Lesley A Stewart, James Thomas, Andrea C Tricco, Vivian A Welch, Penny Whiting e David Moher (2021), ‘The prisma 2020 statement: an updated guideline for reporting systematic reviews’, *BMJ* **372**.
URL: <https://www.bmj.com/content/372/bmj.n71>
- Paguirigan, Raegan Faith F., Mikelene Beron B. Camero, Mark Angerlo Equias, Mideth Abisado e Gabriel Avelino Sampedro (2023), Machine learning approaches to facial recognition: A survey, *em* ‘2023 International Conference on Electronics, Information, and Communication (ICEIC)’, pp. 1–6.
- Pan, G., L. Sun, Z. Wu e S. Lao (2007), Eyeblink-based anti-spoofing in face recognition from a generic webcam, *em* ‘2007 IEEE 11th International Conference on Computer Vision’, pp. 1–8.
- Papaiz, Fabiano, Mario Emílio Teixeira Dourado, Ricardo Alexsandro de Medeiros Valentim, Antonio Higor Freire de Moraes e Joel Perdiz Arrais (2022), ‘Machine learning solutions applied to amyotrophic lateral sclerosis prognosis: A review’, *Frontiers in Computer Science* **4**, 869140.
URL: <https://www.frontiersin.org/articles/10.3389/fcomp.2022.869140>
- Park, Jung Hoon e Jin Bae Park (2016), ‘A novel approach to the low cost real time eye mouse’, *Computer Standards & Interfaces* **44**, 169–176.
URL: <https://www.sciencedirect.com/science/article/pii/S0920548915000458>
- Paszke, Adam, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga e Adam Lerer (2017), ‘Automatic differentiation in pytorch’.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot e E. Duchesnay (2011), ‘Scikit-learn: Machine learning in Python’, *Journal of Machine Learning Research* **12**, 2825–2830.
- Phinyomark, Angkoon, Pornchai Phukpattaranont e Chusak Limsakul (2012), ‘Feature reduction and selection for emg signal classification’, *Expert Systems with Applications* **39**(8), 7420 – 7431.
URL: <http://www.sciencedirect.com/science/article/pii/S0957417412001200>
- Pinheiro, Carlos G., Eduardo LM Naves, Pierre Pino, Etienne Losson, Adriano O. Andrade e Guy Bourhis (2011), ‘Alternative communication systems for people with severe motor disabilities: a survey’, *BioMedical Engineering OnLine* **10**(1), 31.
URL: <https://doi.org/10.1186/1475-925X-10-31>

- Pinto, Susana, Stefano Quintarelli e Vincenzo Silani (2020), 'New technologies and amyotrophic lateral sclerosis – which step forward rushed by the covid-19 pandemic?', *Journal of the Neurological Sciences* **418**, 117081.
URL: <https://www.sciencedirect.com/science/article/pii/S0022510X20304184>
- Pondofe, Karen, Ana Aline Marcelino, Tatiana Souza Ribeiro, Rodrigo Torres-Castro, Roberto Vera-Urbe, Guilherme AF Fregonezi e Vanessa R Resqueti (2022), 'Effects of respiratory physiotherapy in patients with amyotrophic lateral sclerosis: protocol for a systematic review of randomised controlled trials', *BMJ Open* **12**(5).
URL: <https://bmjopen.bmj.com/content/12/5/e061624>
- Radlak, Krystian e Bogdan Smolka (2013), Blink detection based on the weighted gradient descriptor, em R.Burduk, K.Jackowski, M.Kurzynski, M.Wozniak e A.Zolnierek, eds., 'Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013', Springer International Publishing, Heidelberg, pp. 691–700.
- Rahnama-ye Moqaddam, Reza e Hamed Vahdat-Nejad (2015), Designing a pervasive eye movement-based system for als and paralyzed patients, em '2015 5th International Conference on Computer and Knowledge Engineering (ICCKE)', pp. 218–221.
- Rakhmatulin, Ildar e Andrew T. Duchowski (2020), 'Deep neural networks for low-cost eye tracking', *Procedia Computer Science* **176**, 685–694. Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 24th International Conference KES2020.
URL: <https://www.sciencedirect.com/science/article/pii/S1877050920319360>
- Rakshita, R (2018), Communication through real-time video oculography using face landmark detection, em '2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)', pp. 1094–1098.
- Ramakrishnan, Jayabrabu, Dinesh Mavaluru, Ramkumar Siva Sakthivel, Abdulrahman Saad Alqahtani, Azath Mubarakali e Mervin Retnadhas (2020), 'Brain-computer interface for amyotrophic lateral sclerosis patients using deep learning network', *Neural Computing and Applications* .
URL: <https://doi.org/10.1007/s00521-020-05026-y>
- Rangayyan, Rangaraj M (2015), *Biomedical signal analysis*, John Wiley & Sons.
- Redmon, Joseph, Santosh Divvala, Ross Girshick e Ali Farhadi (2016), You only look once: Unified, real-time object detection, em 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 779–788.
- Ren, P., S. Tang, F. Fang, L. Luo, L. Xu, M. L. Bringas-Vega, D. Yao, K. M. Kendrick e P. A. Valdes-Sosa (2017), 'Gait rhythm fluctuation analysis for neurodegenerative diseases by empirical mode decomposition', *IEEE Transactions on Biomedical Engineering* **64**(1), 52–60.

- Ren, Shaoqing, Kaiming He, Ross Girshick e Jian Sun (2015), Faster r-cnn: Towards real-time object detection with region proposal networks, *em* 'Advances in neural information processing systems', pp. 91–99.
- Rooney, James, Susan Byrne, Mark Heverin, Katy Tobin, Alison Dick, Colette Donaghy e Orla Hardiman (2015), 'A multidisciplinary clinic approach improves survival in als: a comparative study of als in ireland and northern ireland', *Journal of Neurology, Neurosurgery & Psychiatry* **86**(5), 496–501.
- Rosa Silva, Jéssica Paloma, José Bomfim Santiago Júnior, Elizabete Lima dos Santos, Fernanda Oliveira de Carvalho, Iandra Maria Pinheiro de França Costa e Deise Maria Furtado de Mendonça (2020), 'Quality of life and functional independence in amyotrophic lateral sclerosis: A systematic review', *Neuroscience & Biobehavioral Reviews* **111**, 1 – 11.
URL: <http://www.sciencedirect.com/science/article/pii/S0149763418309448>
- Rozado, David, Francisco B. Rodriguez e Pablo Varona (2012), 'Low cost remote gaze gesture recognition in real time', *Applied Soft Computing* **12**(8), 2072–2084.
URL: <https://www.sciencedirect.com/science/article/pii/S1568494612001159>
- Rozado, David, Javier S. Agustin, Francisco B. Rodriguez e Pablo Varona (2012), 'Gliding and saccadic gaze gesture recognition in real time', *ACM Trans. Interact. Intell. Syst.* **1**(2).
URL: <https://doi-org.ez18.periodicos.capes.gov.br/10.1145/2070719.2070723>
- Rupanagudi, Sudhir Rao, Varsha G. Bhat, B. S. Ranjani, Abhiram Srisai, Shreya K. Gurikar, M. R. Pranay e S Chandana (2018), A simplified approach to assist motor neuron disease patients to communicate through video oculography, *em* '2018 International Conference on Communication information and Computing Technology (ICCICT)', pp. 1–6.
- Russell, Stuart e Peter Norving (2022), *Artificial intelligence: A Modern Approach*, 4ª edição, Pearson Education Limited, United Kingdom.
- Saadeh, W., M. A. B. Altaf e S. A. Butt (2017), A wearable neuro-degenerative diseases detection system based on gait dynamics, *em* '2017 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)', pp. 1–6.
- Sagonas, Christos, Epameinondas Antonakos, Georgios Tzimiropoulos, Stefanos Zafeiriou e Maja Pantic (2016), '300 faces in-the-wild challenge: Database and results', *Image and vision computing* **47**, 3–18.
- Saleh, Neven e Aya Tarek (2021), Vision-based communication system for patients with amyotrophic lateral sclerosis, *em* '2021 3rd Novel Intelligent and Leading Emerging Sciences Conference (NILES)', pp. 19–22.

- Schuster, Christina, Orla Hardiman e Peter Bede (2016), ‘Development of an automated mri-based diagnostic protocol for amyotrophic lateral sclerosis using disease-specific pathognomonic features: a quantitative disease-state classification study’, *PLoS One* **11**(12).
- Shorten, Connor e T. Khoshgoftaar (2019), ‘A survey on image data augmentation for deep learning’, *Journal of Big Data* **6**, 1–48.
- Silva, Emília M. G. S., Ledycnarf J. Holanda, Gustavo K. B. Coutinho, Fernanda S. Andrade, Gabriel I. S. Nascimento, Danilo A. P. Nagem, Ricardo A. de M. Valentim e Ana Raquel Lindquist (2021), ‘Effects of active upper limb orthoses using brain–machine interfaces for rehabilitation of patients with neurological disorders: Protocol for a systematic review and meta-analysis’, *Frontiers in Neuroscience* **15**, 661494.
URL: <https://www.frontiersin.org/articles/10.3389/fnins.2021.661494>
- Silva, Gustavo Henrique Souto da, Antônio Higor Freire de Morais, Ricardo Alexsandro de Medeiros Valentim, Rummenigge Rudson Dantas, Leila Raulino Câmara Cavalcanti e Mário Emílio Teixeira Dourado Júnior (2013), ‘An angel for als: architecture based on computer vision applied for on-line monitoring respiratory parameters of patients with amyotrophic lateral sclerosis (als) in hospital environment’, *Brazilian Journal of Biomedical Engineering* **29**(2), 206–210.
URL: <https://doi.org/10.4322/rbeb.2013.018>
- Silva, Stephano Tomaz, Aline Alves de Souza, Karen Pondofe, Luciana Protásio de Melo, Vanessa R Resqueti, Ricardo Alexsandro de Medeiros Valentim e Tatiana Souza Ribeiro (2022), ‘Physical therapy for the management of motor symptoms in amyotrophic lateral sclerosis: protocol for a systematic review’, *BMJ Open* **12**(11).
URL: <https://bmjopen.bmj.com/content/12/11/e063689>
- Singh, Hari e Jaswinder Singh (2018), ‘Real-time eye blink and wink detection for object selection in hci systems’, *Journal on Multimodal User Interfaces* **12**(1), 55–65.
URL: <https://doi.org/10.1007/s12193-018-0261-7>
- Singh, Hari e Jaswinder Singh (2019a), ‘Object acquisition and selection in human computer interaction systems: A review’, *International Journal of Intelligent Systems and Applications in Engineering* **7**(1), 19–29.
URL: <https://ijisae.org/index.php/IJISAE/article/view/805>
- Singh, Hari e Jaswinder Singh (2019b), ‘Object acquisition and selection using automatic scanning and eye blinks in an hci system’, *Journal on Multimodal User Interfaces* **13**(4), 405–417.
URL: <https://doi.org/10.1007/s12193-019-00303-0>
- Singh, Himanshu (2019), *Practical machine learning and image processing: for facial recognition, object detection, and pattern recognition using Python*, 1ª edição, Springer, Apress Berkeley, CA.

- Snyder, Hannah (2019), 'Literature review as a research methodology: An overview and guidelines', *Journal of Business Research* **104**, 333 – 339.
URL: <http://www.sciencedirect.com/science/article/pii/S0148296319304564>
- Soltanpour, Sima, Boubakeur Boufama e QM Jonathan Wu (2017), 'A survey of local feature methods for 3d face recognition', *Pattern Recognition* **72**, 391–406.
- Song, Fengyi, Xiaoyang Tan, Xue Liu e Songcan Chen (2014), 'Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients', *Pattern Recognition* **47**(9), 2825 – 2838.
URL: <http://www.sciencedirect.com/science/article/pii/S0031320314001228>
- Sorbello, R., S. Tramonte, M. E. Giardina, V. La Bella, R. Spataro, B. Allison, C. Guger e A. Chella (2018), 'A human–humanoid interaction through the use of bci for locked-in als patients using neuro-biological feedback fusion', *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **26**(2), 487–497.
- Soriani, M-H e C Desnuelle (2017), 'Care management in amyotrophic lateral sclerosis', *Revue neurologique* **173**(5), 288–299.
- Soukupová, Tereza e Jan Cech (2016), Eye blink detection using facial landmarks, *em '21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia'*, pp. 1–8.
- Stegmann, Gabriela M., Shira Hahn, Julie Liss, Jeremy Shefner, Seward Rutkove, Kerisa Shelton, Cayla Jessica Duncan e Visar Berisha (2020), 'Early detection and tracking of bulbar changes in als via frequent and remote speech analysis', *npj Digital Medicine* **3**(1), 132.
URL: <https://doi.org/10.1038/s41746-020-00335-x>
- Stigberg, Susanne Koch (2020), Human computer interfaces reconsidered: A conceptual model for understanding user interfaces, *em M.Kurosu, ed., 'Human-Computer Interaction. Design and User Experience'*, Springer International Publishing, Cham, pp. 160–171.
- Stone, Debbie, Caroline Jarrett, Mark Woodroffe e Shailey Minocha (2005), *User interface design and evaluation*, Elsevier.
- Suhas, B., J. Mallela, A. Illa, B. Yamini, N. Atchayaram, R. Yadav, D. Gope e P. K. Ghosh (2020), Speech task based automatic classification of als and parkinson's disease and their severity using log mel spectrograms, *em '2020 International Conference on Signal Processing and Communications (SPCOM)'*, pp. 1–5.
- Tamura, Hiroki, Mingmin Yan, Keiko Sakurai e Koichi Tanno (2016), 'Eog-semg human interface for communication', *Computational intelligence and neuroscience* **2016**.
- Tan, Mingxing e Quoc V. Le (2019), 'Efficientnet: Rethinking model scaling for convolutional neural networks', *CoRR* **abs/1905.11946**.
URL: <http://arxiv.org/abs/1905.11946>

- Tavazzi, Erica, Sebastian Daberdaku, Alessandro Zandonà, Rosario Vasta, Beatrice Nefussy, Christian Lunetta, Gabriele Mora, Jessica Mandrioli, Enrico Grisan, Claudia Tarlarini, Andrea Calvo, Cristina Moglia, Vivian Drory, Marc Gotkine, Adriano Chiò, Barbara Di Camillo, A. Chiò, Rita Levi Montalcini, A. Calvo, C. Moglia, A. Canosa, U. Manera, R. Vasta, F. Palumbo, A. Bombaci, M. Grassano, M. Brunetti, F. Casale, G. Fuda, P. Salomone, B. Iazzolino, L. Peotta, P. Cugnasco, G. De Marco, M. C. Torrieri, S. Gallone, M. Barberis, L. Sbaiz, S. Gentile, A. Mauro, L. Mazzini, F. Marchi, L. Corrado, S. D'Alfonso, A. Bertolotto, M. Gionco, D. Leotta, E. Oddenino, R. Cavallo, M. De Mattei, G. Gusmaroli, C. Comi, C. Labate, F. Poglio, L. Ruiz, D. Ferrandi, L. Testa, E. Rota, M. Aguggia, N. Di Vito, P. Meineri, P. Ghiglione, N. Launaro, M. Dotta, A. Sapio, M. Giovanni, J. Mandrioli, N. Fini, I. Martinelli, E. Zucchi, G. Gianferrari, C. Simonini, M. Vinceti, S. Meletti, V. Vacchiano, R. Liguori, Fabrizio Salvi, Ilaria Bartolomei, Roberto Michelucci, P. Cortelli, A. M. Borghi, A. Zini, R. Rinaldi, E. Sette, V. Tugnoli, M. Pugliatti, E. Canali, L. Codeluppi, F. Valzania, L. Zinno, G. Pavesi, D. Medici, G. Pilurzi, E. Terlizzi, D. Guidetti, S. Pasqua, M. Santangelo, M. Bracaglia, P. DeMassis, M. Casmiro, P. Querzani, S. Morresi, M. Longoni, A. Patuelli, S. Malagù, M. Currò Dossi, S. Vidale e for the Emilia Romagna Registry for ALS (ERRALS) For the Piemonte, Valle d'Aosta Register for ALS (PARALS) (2022), 'Predicting functional impairment trajectories in amyotrophic lateral sclerosis: a probabilistic, multifactorial model of disease progression', *Journal of Neurology* **269**(7), 3858–3878.
URL: <https://doi.org/10.1007/s00415-022-11022-0>
- Thome, Janine, Robert Steinbach, Julian Grosskreutz, Daniel Durstewitz e Georgia Koppe (2022), 'Classification of amyotrophic lateral sclerosis by brain volume, connectivity, and network dynamics', *Human Brain Mapping* **43**(2), 681–699.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.25679>
- Tonin, Alessandro, Andres Jaramillo-Gonzalez, Aygul Rana, Majid Khalili-Ardali, Niels Birbaumer e Ujwal Chaudhary (2020), 'Auditory electrooculogram-based communication system for als patients in transition from locked-in to complete locked-in state', *Scientific Reports* **10**(1), 8452.
URL: <https://doi.org/10.1038/s41598-020-65333-1>
- Umeyama, Shinji (1991), 'Least-squares estimation of transformation parameters between two point patterns', *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(4), 376–380.
URL: <https://doi.org/10.1109/34.88573>
- United Nations (2015), Take action for the sustainable development goals. New York, NY, USA. Available from: <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>. [Accessed 10 Apr 2023].
- United Nations (2022), World health organization (who) and the united nations children's fund (unicef). global report on assistive technology. New York, NY, USA. Available

from: <https://www.who.int/publications/i/item/9789240049451>. [Acesso 10 Abr 2023].

Valentim, Ricardo, Carlos A.P. Oliveira, Eloiza Oliveira, Eduardo Ribeiro, Soneide Costa, Ione Morais, Felipe Fernandes, Alexandre Caitano, Cristine Gusmão, Aliete Cunha-Oliveira, Maria C. Rêgo, Karilany Coutinho, Daniele Barros e Ricardo Ceccim (2022), ‘Virtual learning environment of the brazilian health system (avasus): Efficiency of results, impacts, and contributions’, *Frontiers in Medicine*. 9:896208. 9, 896208.

URL: <https://doi.org/10.3389/fmed.2022.896208>

Vallejo, Marcela, Carlos J Gallego, Leonardo Duque-Muñoz e Edilson Delgado-Trejos (2018), ‘Neuromuscular disease detection by neural networks and fuzzy entropy on time-frequency analysis of electromyography signals’, *Expert Systems* 35(4), e12274.

van der Burgh, Hannelore K, Ruben Schmidt, Henk-Jan Westeneng, Marcel A de Reus, Leonard H van den Berg e Martijn P van den Heuvel (2017), ‘Deep learning predictions of survival based on mri in amyotrophic lateral sclerosis’, *NeuroImage: Clinical* 13, 361–369.

van Es, Michael A, Orla Hardiman, Adriano Chio, Ammar Al-Chalabi, R Jeroen Pasterkamp, Jan H Veldink e Leonard H van den Berg (2017), ‘Amyotrophic lateral sclerosis’, *The Lancet* 390(10107), 2084 – 2098.

URL: <http://www.sciencedirect.com/science/article/pii/S0140673617312874>

Vansteensel, Mariska J., Elmar G.M. Pels, Martin G. Bleichner, Mariana P. Branco, Timothy Denison, Zachary V. Freudenburg, Peter Gosselaar, Sacha Leinders, Thomas H. Ottens, Max A. Van Den Boom, Peter C. Van Rijen, Erik J. Aarnoutse e Nick F. Ramsey (2016), ‘Fully implanted brain–computer interface in a locked-in patient with als’, *New England Journal of Medicine* 375(21), 2060–2066. PMID: 27959736.

URL: <https://doi.org/10.1056/NEJMoa1608085>

Vieira, H., N. Costa, T. Sousa, S. Reis e L. Coelho (2019), ‘Voice-based classification of amyotrophic lateral sclerosis: Where are we and where are we going? a systematic review’, *Neurodegenerative Diseases* 19(5-6), 163–170.

URL: <https://doi.org/10.1159/000506259>

Villanueva, Arantxa, Gintautas Daunys, Dan Witzner Hansen, Martin Böhme, Rafael Cabeza, André Meyer e Erhardt Barth (2009), ‘A geometric approach to remote eye tracking’, *Universal Access in the Information Society* 8(4), 241.

URL: <https://doi.org/10.1007/s10209-009-0149-0>

Wang, Jun, Prasanna V. Kothalkar, Beiming Cao e Daragh Heitzman (2016), Towards automatic detection of amyotrophic lateral sclerosis from speech acoustic and articulatory samples, em ‘Interspeech 2016’, pp. 1195–1199.

URL: <http://dx.doi.org/10.21437/Interspeech.2016-1542>

- Wang, Jun, Prasanna V. Kothalkar, Myungjong Kim, Andrea Bandini, Beiming Cao, Yana Yunusova, Thomas F. Campbell, Daragh Heitzman e Jordan R. Green (2018), 'Automatic prediction of intelligible speaking rate for individuals with als from speech acoustic and articulatory samples', *International journal of speech-language pathology* **20**(6), 669–679. PMC6506394[pmcid].
URL: <https://doi.org/10.1080/17549507.2018.1508499>
- Wang, Xiaoxing (2017), Eye-blink detection based on svm, Relatório técnico, Shanghai Jiao Tong University.
- Wang, Ziqin, Jun Xu, Li Liu, Fan Zhu e Ling Shao (2019), Ranet: Ranking attention network for fast video object segmentation, *em* 'Proceedings of the IEEE international conference on computer vision', pp. 3978–3987.
- Welsh, Robert, Laura Jelsone-Swain e Bradley Foerster (2013), 'The utility of independent component analysis and machine learning in the identification of the amyotrophic lateral sclerosis diseased brain', *Frontiers in Human Neuroscience* **7**, 251.
URL: <https://www.frontiersin.org/article/10.3389/fnhum.2013.00251>
- Wisler, Alan, Kristin Teplansky, Daragh Heitzman e Jun Wang (n.d.), 'The effects of symptom onset location on automatic amyotrophic lateral sclerosis detection using the correlation structure of articulatory movements', *Journal of Speech, Language, and Hearing Research* pp. 1–11.
- Wisler, Alan, Kristin Teplansky, Jordan Green, Yana Yunusova, Thomas Campbell, Daragh Heitzman e Jun Wang (2019), Speech-based estimation of bulbar regression in amyotrophic lateral sclerosis, *em* 'Proceedings of the Eighth Workshop on Speech and Language Processing for Assistive Technologies', Association for Computational Linguistics, Minneapolis, Minnesota, pp. 24–31.
URL: <https://www.aclweb.org/anthology/W19-1704>
- Xia, Yi, Qingwei Gao e Qiang Ye (2015), 'Classification of gait rhythm signals between patients with neuro-degenerative diseases and normal subjects: Experiments with statistical features and different classification models', *Biomedical Signal Processing and Control* **18**, 254 – 262.
URL: <http://www.sciencedirect.com/science/article/pii/S1746809415000142>
- Xiong, Sijie, Sujie Zhu, Yisheng Ji, Binyao Jiang, Xiaohua Tian, Xuesheng Zheng e Xinbing Wang (2017), Iblink: Smart glasses for facial paralysis patients, *em* 'Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services', MobiSys '17, Association for Computing Machinery, New York, NY, USA, p. 359–370.
URL: <https://doi.org/10.1145/3081333.3081343>
- Xu, Chu-Lian e Chyi-Yeu Lin (2017), Eye-motion detection system for mnd patients, *em* '2017 IEEE 4th International Conference on Soft Computing & Machine Intelligence (ISCMi)', pp. 99–103.

- Yildiz, Metin e Muhammet Yorulmaz (2019), Gaze-controlled turkish virtual keyboard application with webcam, *em* ‘2019 Medical Technologies Congress (TIPTEKNO)’, pp. 1–4.
- Zeiler, Matthew D (2012), ‘Adadelta: an adaptive learning rate method’, *arXiv preprint arXiv:1212.5701* .
- Zhang, R., S. He, X. Yang, X. Wang, K. Li, Q. Huang, Z. Yu, X. Zhang, D. Tang e Y. Li (2019), ‘An eog-based human–machine interface to control a smart home environment for patients with severe spinal cord injuries’, *IEEE Transactions on Biomedical Engineering* **66**(1), 89–100.
- Zhang, X., P. E. Barkhaus, W. Z. Rymer e P. Zhou (2014), ‘Machine learning for supporting diagnosis of amyotrophic lateral sclerosis using surface electromyogram’, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **22**(1), 96–103.
- Zhang, Xiaoyi, Harish Kulkarni e Meredith Ringel Morris (2017), Smartphone-based gaze gesture communication for people with motor disabilities, *em* ‘Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems’, CHI ’17, Association for Computing Machinery, New York, NY, USA, p. 2878–2889.
URL: <https://doi-org.ez18.periodicos.capes.gov.br/10.1145/3025453.3025790>
- Zhao, Qijie, Xinming Yuan, Dawei Tu e Jianxia Lu (2015), ‘Eye moving behaviors identification for gaze tracking interaction’, *Journal on Multimodal User Interfaces* **9**(2), 89–104.
URL: <https://doi.org/10.1007/s12193-014-0171-2>
- Zhu, Yanjia, Hongxiang Cai, Shuhan Zhang, Chenhao Wang e Yichao Xiong (2020), ‘Tinaface: Strong but simple baseline for face detection’, *arXiv preprint arXiv:2011.13183* .
- Zizzi, Christine, Jamison Seabury, Spencer Rosero, Danae Alexandrou, Ellen Wagner, Jennifer S. Weinstein, Anika Varma, Nuran Dilek, John Heatwole, Joanne Wu, James Caress, Richard Bedlack, Volkan Granit, Jeffrey M. Statland, Paul Mehta, Michael Benatar e Chad Heatwole (2023), ‘Patient reported impact of symptoms in amyotrophic lateral sclerosis (prism-als): a national, cross-sectional study’, *eClinicalMedicine* **55**, 101768.
URL: <https://www.sciencedirect.com/science/article/pii/S2589537022004977>
- Zuiderveld, Karel (1994), ‘Contrast limited adaptive histogram equalization’, *Graphics gems* pp. 474–485.
- Özden O. Dalgiç, Haoran Wu, F. Safa Erenay, Mustafa Y. Sir, Osman Y. Özaltın, Brian A. Crum e Kalyan S. Pasupathy (2021), ‘Mapping of critical events in disease progression through binary classification: Application to amyotrophic lateral sclerosis’, *Journal of Biomedical Informatics* **123**, 103895.
URL: <https://www.sciencedirect.com/science/article/pii/S1532046421002240>