

Uma ferramenta baseada em aprendizado profundo para o suporte à decisão de diagnóstico de distúrbios vestibulares periféricos

A deep learning-based tool for the diagnostic decision support of peripheral vestibular disorders

Una herramienta basada en el aprendizaje profundo para apoyo a la toma de decisiones en el diagnóstico de desórdenes vestibulares periféricos

Recebido: 07/03/2022 | Revisado: 15/03/2022 | Aceito: 19/03/2022 | Publicado: 26/03/2022

Antônia de Maria Rodrigues de Sousa Castro

ORCID: <https://orcid.org/0000-0001-6951-0242>

Universidade Federal do Piauí, Brasil

E-mail: antoniademariacastro123@gmail.com

Ariel Soares Teles

ORCID: <https://orcid.org/0000-0002-0840-3870>

Instituto Federal do Maranhão, Brasil

E-mail: ariel.teles@ifma.edu.br

Lucas Daniel Batista Lima

ORCID: <https://orcid.org/0000-0002-4278-6975>

Universidade Federal do Delta do Parnaíba, Brasil

E-mail: lucas.daniel.bp@gmail.com

José Everton da Silva Fontenele

ORCID: <https://orcid.org/0000-0001-5982-4707>

Universidade Federal do Delta do Parnaíba, Brasil

E-mail: vetophp@gmail.com

Victor Hugo do Vale Bastos

ORCID: <https://orcid.org/0000-0001-7425-4730>

Universidade Federal do Delta do Parnaíba, Brasil

E-mail: victorhugobastos@ufpi.edu.br

Silmar Silva Teixeira

ORCID: <https://orcid.org/0000-0002-9240-1228>

Universidade Federal do Delta do Parnaíba, Brasil

E-mail: silmarteixeira@ufpi.edu.br

Resumo

O nistagmo é o movimento involuntário dos olhos, caracterizado pelo movimento suave, chamado de fase lenta do nistagmo, interrompido pela fixação rápida na direção contrária. Ele é um dos fatores preponderantes no diagnóstico de desordens vestibulares. Este estudo apresenta o *Smart Nystagmography*, uma proposta de ferramenta baseada em visão computacional para o suporte ao diagnóstico de disfunções vestibulares periféricas, que engloba todo o processo, desde o dispositivo para coleta do movimento ocular até o classificador do distúrbio. A solução proposta é baseada em vetores de características que apresentam padrões de movimento ocular, os quais são identificados com o uso de aprendizado de máquina, em particular, Aprendizado Profundo (AP). A técnica de videonistagmografia e suas diferentes provas foram realizadas pelos indivíduos a fim de gerar um conjunto de dados representativo para indivíduos tanto saudáveis quanto com disfunção vestibular. Os métodos de pré-processamento de dados, assim como uma técnica de otimização de hiperparâmetros dos algoritmos de AP foram empregados com o propósito de melhorar o desempenho dos modelos do estado da arte. Os resultados de desempenho para a identificação da presença de disfunção vestibular periférica chegaram a uma acurácia de 96,7% para o melhor modelo, depois de passar pelo processo de otimização. Os resultados mostram a eficiência do *Smart Nystagmography*, o qual possui uma solução que envolve desde o dispositivo de coleta de vídeos até o sistema com as técnicas de preparação dos dados e o modelo de AP implantado. Estudos clínicos adicionais são necessários para validar a solução.

Palavras-chave: Nistagmo; Videonistagmografia; Vestibulopatia periférica; Diagnóstico; Inteligência Artificial; Visão computacional; Aprendizado profundo.

Abstract

Nystagmus is involuntary eye movement characterized by smooth movement, called the slow phase of nystagmus, interrupted by rapid fixation in the opposite direction. It is one of the preponderant factors in the diagnosis of vestibular disorders. This study presents *Smart Nystagmography*, a proposal for a computer vision-based tool to support the diagnosis of peripheral vestibular disorders, which encompasses the entire process, from the eye movement collection

device to the disorder classifier. The proposed solution is based on feature vectors that present eye movement patterns, which are identified using machine learning, in particular, Deep Learning (DL). The videonystagmography technique and its different tests were performed by the subjects in order to generate a representative dataset for both healthy subjects and those with vestibular dysfunction. Data pre-processing methods, as well as a hyperparameter optimization technique of the DL algorithms were employed with the purpose of improving the performance of state-of-the-art models. The performance results for identifying the presence of peripheral vestibular dysfunction reached an accuracy of 96.7% for the best model, after going through the optimization process. The results show the efficiency of *Smart Nystagmography*, which has a solution that involves from the video collection device to the system with data preparation techniques and the DL model deployed. Additional clinical studies are needed to validate the solution

Keywords: Nystagmus; Videonystagmography; Peripheral vestibulopathy; Diagnosis; Artificial intelligence; Computer vision; Deep learning.

Resumen

El nistagmo es un movimiento ocular involuntario caracterizado por un movimiento suave, llamado fase lenta del nistagmo, interrumpido por una fijación rápida en la dirección opuesta. Es uno de los factores preponderantes en el diagnóstico de los trastornos vestibulares. Este estudio presenta el *Smart Nystagmography*, una propuesta de herramienta basada en visión artificial para apoyar el diagnóstico de trastornos vestibulares periféricos, que abarca todo el proceso, desde el dispositivo de recolección de movimientos oculares hasta el clasificador de trastornos. La solución propuesta se basa en vectores de características que presentan patrones de movimiento ocular, que se identifican mediante el aprendizaje automático, en particular, el aprendizaje profundo (AP). La técnica de videonistagmografía y sus diferentes pruebas fueron realizadas por los individuos para generar un conjunto de datos representativo tanto para individuos sanos como para aquellos con disfunción vestibular. Se emplearon métodos de preprocesamiento de datos, así como una técnica de optimización de hiperparámetros de los algoritmos AP con el fin de mejorar el rendimiento de los modelos de última generación. Los resultados de rendimiento para identificar la presencia de disfunción vestibular periférica alcanzaron una precisión del 96,7% para el mejor modelo, luego de pasar por el proceso de optimización. Los resultados muestran la eficiencia de lo *Smart Nystagmography*, que tiene una solución que involucra desde el dispositivo de captura de video hasta el sistema con técnicas de preparación de datos y el modelo AP desplegado. Se necesitan estudios clínicos adicionales para validar la solución.

Palabras clave: Nistagmo; Videonistagmografía; Vestibulopatía periférica; Diagnóstico; Inteligencia artificial; Visión por computador; Aprendizaje profundo.

1. Introdução

A disfunção vestibular comumente se manifesta com sintomas como tontura e vertigem com alta prevalência (entre 15% e 20%) e 1,4% de incidência em pessoas da população geral (Karatatou, 2008; Neuhauser, 2016). Além disso, pode manifestar-se em diferentes etiologias, tais como: a Vertigem Posicional Paroxística Benigna (VPPB) e Ménière (distúrbios periféricos) (Staton e Freeman, 2022), Esclerose Múltipla e fatores de risco vasculares centrais (distúrbios centrais) (Mostafa et al., 2014).

A avaliação de rotina de pacientes com disfunção de tontura e vertigem inclui registro e análise dos tipos de nistagmo e outros movimentos oculares (Papageorgiou et al., 2014). O nistagmo é um movimento involuntário, rápido, rítmico e oscilatório de um ou ambos os olhos com ao menos uma fase lenta (Eggers, 2019), seguido de movimento compensatório central rápido para a posição primária dos olhos (Strupp et al., 2021).

Os sistemas mais comuns, que têm permitido investigar a integridade das vias vestibulares, são a eletronistagmografia, a vectoeletronistagmografia e a videonistagmografia (VNG). Em particular, o VNG registra a partir de vídeos os movimentos oculares e quantifica com alta precisão a amplitude, latência, velocidade e frequência (Savage, 2011). A VNG usando câmera de vídeo em óculos ou máscaras (Pietkiewicz et al., 2012; West, Sheppard & King, 2012) permite detectar a existência de nistagmo relacionado a muitos sintomas vestibulares (Slama & Maghraoui, 2014). Por outro lado, os sistemas de VNG possuem diferentes abordagens no processamento de vídeo, o que pode comprometer os resultados das avaliações, principalmente para remover artefatos e também possui custo elevado (Mantokoudis et al., 2022). Neste contexto, novos sistemas que utilizam Inteligência Artificial (IA), em particular, o Aprendizado de Máquina (AM), podem melhorar a qualidade da identificação das disfunções vestibulares.

Vários algoritmos de AM têm sido propostos na literatura para a detecção automática de artefatos e para categorizar as disfunções vestibulares (Groezinger et al., 2010; Visscher et al., 2019 e Kamogashira et al., 2020). Exemplos de algoritmos de Aprendizado Profundo (AP), o qual é considerado um ramo do AM, são as Redes Neurais Recorrentes e as Redes Neurais Convolucionais, do inglês: *Recurrent Neural Network* (RNN) e *Convolutional Neural Network* (CNN), respectivamente (SLAMA et al., 2019). Estes modelos são utilizados em abordagens que podem ser utilizadas na solução de problemas variados, incluindo o propósito deste trabalho, na realização de tarefas de classificação (Alpaydin, 2020).

No trabalho de Lim et al. (2019), os autores propuseram um algoritmo baseado em AP para identificar canais semicirculares e diagnosticar a VPPB usando dados de VNG. Embora o modelo tenha apresentado alto valor de recall e especificidade, ele não consegue detectar a VPPB associada a múltiplos canais. Outras abordagens têm usado o rastreamento do movimento ocular para detectar automaticamente a disfunção vestibular. Slama et al. (2017) usaram dados de VNG de 60 pacientes e propuseram um método automático para melhorar o diagnóstico de neurite vestibular. Para tal, eles utilizaram a Análise de Componentes Principais (ACP), combinada com um algoritmo de Rede Neural Multicamadas (RNM). Os autores ressaltam que a RNM apenas classifica satisfatoriamente a neurite vestibular, ao realizar a redução dos parâmetros originais ao pré-processar com a ACP. Em outro estudo, Slama et al. (2020) extraíram e reduziram as características temporais e de frequência dos movimentos da pupila e, em seguida, aplicaram uma CNN a fim de classificar a Neurite Vestibular e doença de Ménière. Em Mao et al. (2020) foi desenvolvido um rastreador ocular, e aplicado às informações coletadas um tipo de RNN capaz de guardar memória durante a análise de uma série temporal, nomeada, em inglês: *Long Short-term Memory* (LSTM). Este método de classificação configura uma abordagem de AP especial para estabelecer um modelo de discriminação de doenças de acordo com o movimento dos globos oculares extraídos das provas optocinéticas e busca ocular, para classificação de paciente saudável, com lesão cerebral e com vertigem.

Este estudo objetivou desenvolver o *Smart Nystagmography*, uma proposta de ferramenta apoiada em visão computacional e algoritmos de AP para suportar o diagnóstico de disfunções vestibulares periféricas. A solução proposta visa identificar objetivamente a presença ou ausência de disfunção vestibular periférica a partir de vídeos que registram o movimento ocular quando o indivíduo realiza diferentes provas vestibulares.

Os sistemas de vectoeletronistagmografia e videonistagmografia sempre demandam por uma interpretação subjetiva dos relatórios emitidos. As soluções do estado da arte que aplicam algoritmos de AP ao diagnóstico de disfunções vestibulares visam fornecer um resultado objetivo. Portanto, as soluções que usam AP têm a principal vantagem de não demandar essa análise especializada realizada pelos profissionais de saúde examinadores. O *Smart Nystagmography* possui essa característica, por emitir um relatório objetivo que informa a presença ou ausência de disfunção vestibular periférica. Adicionalmente às soluções do estado da arte que usam AP, este trabalho apresenta uma ferramenta completa que realiza todo o processo de exame, fornecendo desde o dispositivo para a coleta e gravação por vídeo dos movimentos do globo ocular até a sua análise e classificação em dois grupos: presença ou ausência de disfunção vestibular periférica.

O restante desse artigo está organizado da seguinte forma. A Seção 2 descreve a metodologia utilizada para o desenvolvimento da solução. A Seção 3 apresenta uma avaliação de desempenho com o sistema desenvolvido e seus resultados. A Seção 4 discute a solução proposta, os resultados de desempenho alcançados, bem como as limitações do estudo. Finalmente, na Seção 5, as considerações finais são feitas.

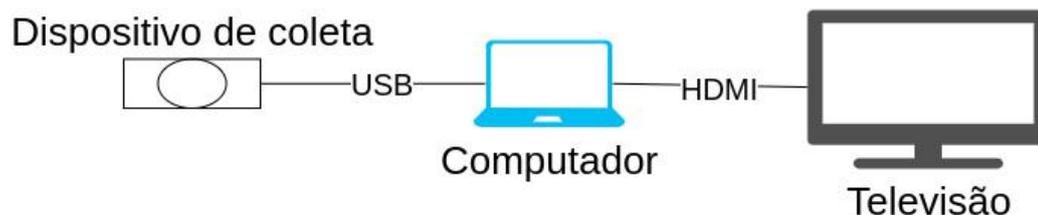
2. Metodologia

Esta seção apresenta a ferramenta proposta para suportar a decisão de diagnóstico de distúrbios vestibulares periféricos e, mais especificamente, o seu processo de desenvolvimento. Este um estudo adotou os procedimentos metodológicos para pesquisa científica com inovação tecnológica, descritos por Wazlawick (2014).

2.1 Visão geral da ferramenta

A ferramenta proposta foi idealizada e desenvolvida no Laboratório de Neuroinovação Tecnológica e Mapeamento Cerebral (NitLab) da Universidade Federal do Delta do Parnaíba (UFDPAr). Ela possui três componentes principais ilustrados no diagrama da Figura 1, com fotos ilustrativas na Figura 2, e descritos a seguir.

Figura 1 - Diagrama de componentes da ferramenta proposta.



Fonte: Autores.

Figura 2 - Ferramenta proposta em funcionamento (esquerda) e visão frontal do dispositivo de coleta (direita).



Fonte: Autores.

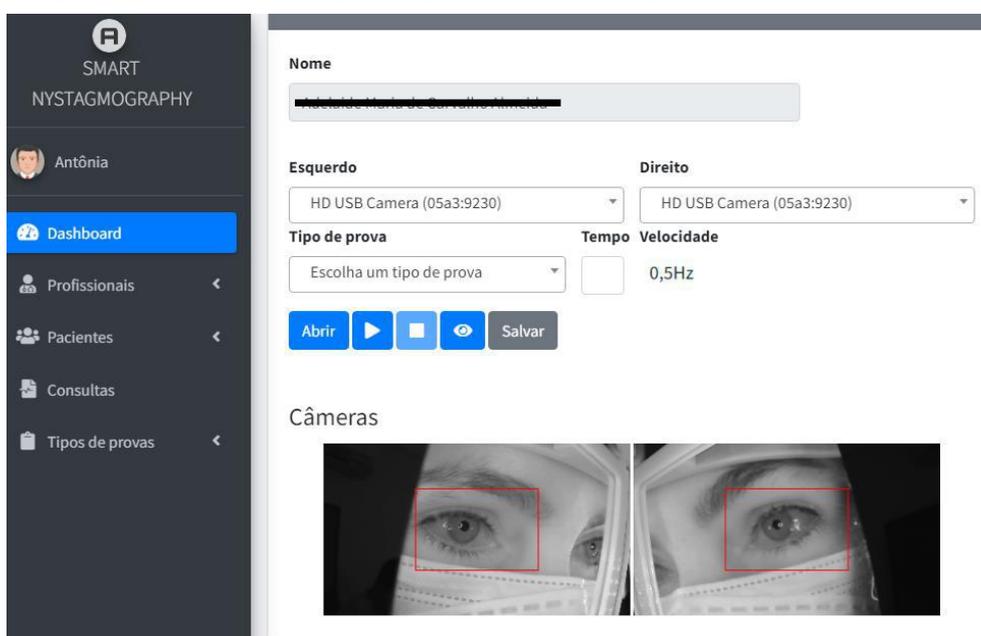
O primeiro componente é um dispositivo de coleta de vídeos que adequa-se à cabeça do indivíduo, o qual tem um formato de óculos. Os vídeos coletados objetivam capturar os movimentos do globo ocular. A estrutura desse dispositivo foi modelada com a ferramenta *Fusion 360* e impressa na impressora 3D *Stella* usando filamento de políácido láctico (PLA). A estrutura é ajustável, o que possibilita a utilização aos demais biotipos faciais. O projeto do dispositivo de coleta possibilitou otimizar o posicionamento de duas câmeras e dois espelhos dicrômico ajustáveis, o que permitiu obter um campo de acuidade visual livre para registro do movimento do globo ocular, como visto na Figura 2 à direita. A estrutura contém faixas de náilon reguláveis para a fixação segura. As câmeras captam imagens com resolução de 1080x720 e taxa de 30 quadros por segundo.

O segundo componente é o sistema computacional responsável pela gerência do dispositivo de coleta dos vídeos, via conexões USB com as duas câmeras, e manipulação das tarefas exibidas no terceiro componente, a televisão conectada por um cabo de interface multimídia de alta definição (HDMI). O sistema de gerenciamento da ferramenta foi desenvolvido com tecnologias *web*, tais como HTML, CSS e JavaScript, para o front-end, além do *framework Laravel* e o Sistema de Gerenciamento de Banco de Dados (SGBD) MariaDB, para o back-end. Esse sistema é responsável pela gestão das coletas dos vídeos, ou seja, ele obtém os vídeos através das câmeras e os armazena para serem analisados. Em um primeiro momento, essa análise objetivou o treinamento dos algoritmos de AP. Desta forma, após o treinamento, o objetivo é de que o sistema passe a utilizar o modelo

de melhor desempenho para realizar a classificação de padrões de movimento do globo ocular identificados a partir dos vídeos, para reconhecer a presença de disfunção vestibular.

Com a utilização do sistema, o profissional examinador pode realizar o cadastro dos pacientes e a anamnese. Após estas duas etapas iniciais, o sistema permite a realização da coleta de vídeos. Neste momento, como visto na Figura 3, o profissional escolhe a câmera do olho direito e esquerdo, e o tipo de prova que será realizada. Ao escolher as câmeras, é exibido um vídeo de teste para confirmar o funcionamento (Figura 3). Dois quadros de bordas vermelhas aparecem para a centralização dos olhos, os quais são usados como calibração para determinar a área de interesse para o algoritmo de rastreamento ocular. Isso é necessário para obter imagens com enquadramento correto durante toda a execução das provas, diminuindo artefatos. Em seguida, o profissional define o “Tipo de prova” e clica em “Abrir”, o que irá permitir a janela relativa à prova ser compartilhada no terceiro componente, a Televisão. Por fim, a prova é iniciada com o estímulo luminoso quando o profissional examinador, avisando o indivíduo, indicar no botão de “Play”.

Figura 3 - Sistema de configuração para a realização das provas e obtenção dos vídeos.



Fonte: Autores.

As provas em que a ferramenta permite serem realizadas já são conhecidas na literatura e estão presentes na vectonistagmografia e videonistagmografia (Wu et al., 2022). Para o propósito deste estudo, as seguintes provas foram consideradas: Optocinética Horizontal Esquerda (OHE), Optocinética Horizontal Direita (OHD), Calórica Quente - Orelha Direita (CQOD), Calórica Quente - Orelha Esquerda (CQOE), Calórica Fria - Orelha Direita (CFOD), Calórica Fria - Orelha Esquerda (CFOE) e Pré-Calórica (PC).

2.2 Coleta de dados

Para realizar o treinamento de algoritmos de AP, foi necessário primeiro realizar a coleta de dados para serem utilizados como exemplos de ambas as classes: positivo (classe 1) e negativo (classe 0) para distúrbio vestibular periférico. Para isso, foram recrutados 44 participantes, de ambos os sexos, na faixa etária de 20 a 60 anos, divididos em dois grupos: 22 saudáveis (classe 0) e 22 com distúrbio vestibular periférico (classe 1). A coleta dos dados ocorreu no período de dezembro de 2020 a setembro

de 2021. Antes da realização das provas, os participantes foram esclarecidos sobre o propósito da pesquisa, seus objetivos e procedimentos, e orientados a assinar o Termo de Consentimento Livre e Esclarecido (TCLE), garantindo-lhes o anonimato, bem como assegurado o direito de desistir de participar do estudo em qualquer momento. O estudo foi submetido ao Comitê de Ética da Universidade Federal do Piauí- UFPI e aprovado com o parecer de número 3.662.523.

Os participantes foram orientados a evitar o uso de medicamentos como tranquilizantes, narcóticos, anti-histamínicos, anti vertiginosos, antialérgicos por um período de 72 horas, bem como não ingerir bebida alcoólica nas 48 horas e não usar produtos com cafeína e cigarro nas 24 horas que antecederam o exame. Todavia, não recomendamos a suspensão de medicamentos essenciais, como aqueles usados para o tratamento de hipertensão arterial, cardiopatia, diabetes mellitus, dentre outras patologias. Além disso, fez-se necessário um jejum de 3 horas, pois a estimulação realizada durante o exame poderia provocar sintomas neurovegetativos (náuseas, vômitos, dentre outros).

As provas foram realizadas em uma sala preparada exclusivamente para essa finalidade. Durante a realização, as luzes foram apagadas, o indivíduo foi sentado confortavelmente em uma cadeira com suporte para os braços e cabeça, ficando a uma distância de um metro dos estímulos luminosos, expostos em uma televisão de LED de 45 polegadas. Nesse momento, o dispositivo de coleta foi posicionado na região orbitária da face e fixada em torno da cabeça, e então os vídeos foram coletados individualmente para todas as provas.

A prova optocinética é registrada durante o acompanhamento visual de vários estímulos luminosos em movimento, com direção fixa no sentido horizontal para a direita e depois para a esquerda, à velocidade de 20 graus por segundo. A prova calórica é considerada o padrão ouro na avaliação do reflexo vestibulo ocular e tem sido amplamente utilizada desde a sua introdução por Róbert Bárány no início do século XX (Vallim et al., 2021). Logo após a realização da prova optocinética, o indivíduo é reposicionado e colocado na posição de I de Brunnings (ou seja, decúbito dorsal em inclinação de 30° em um plano horizontal), para dar início a prova calórica com a busca do nistagmo pré-calórico sem estimulação auditiva e, em seguida, realizar a estimulação com ar nas temperaturas de 42°C e depois 21°C em cada orelha separadamente, durante 60 segundos (Felipe, Staggs e Hunnicutt, 2021).

2.3 Preparação e extração de características

Com a evolução das tecnologias de captura e processamento de vídeo, a VNG tornou-se o principal método para obter informações sobre o movimento dos olhos (Mao et al., 2020). Desta forma, o sistema proposto possui várias etapas, e uma das mais cruciais é o rastreamento ocular a partir dos vídeos coletados. Após realizar as provas clínicas e registrar todo o processo, os vídeos resultantes foram submetidos à análise usando técnicas de visão computacional. Para o rastreador proposto, foram utilizadas 7 provas, as quais são descritas no Quadro 1. Para cada prova, são produzidos dois vídeos, um para o olho esquerdo e outro para o olho direito. Portanto, a partir do movimento ocular são extraídas as coordenadas x e y do centro geométrico, bem como a área da pupila. As coordenadas e valores de área da pupila são extraídos de cada quadro, gerando assim sequências ou vetores de valores com tamanho correspondente ao número de quadros. Esses vetores são as características dos parâmetros utilizados neste estudo (Quadro 1).

Quadro 1 - Vetores de características: coordenada x, coordenada y e área da pupila, representados por **x**, **y** e **a**, respectivamente.

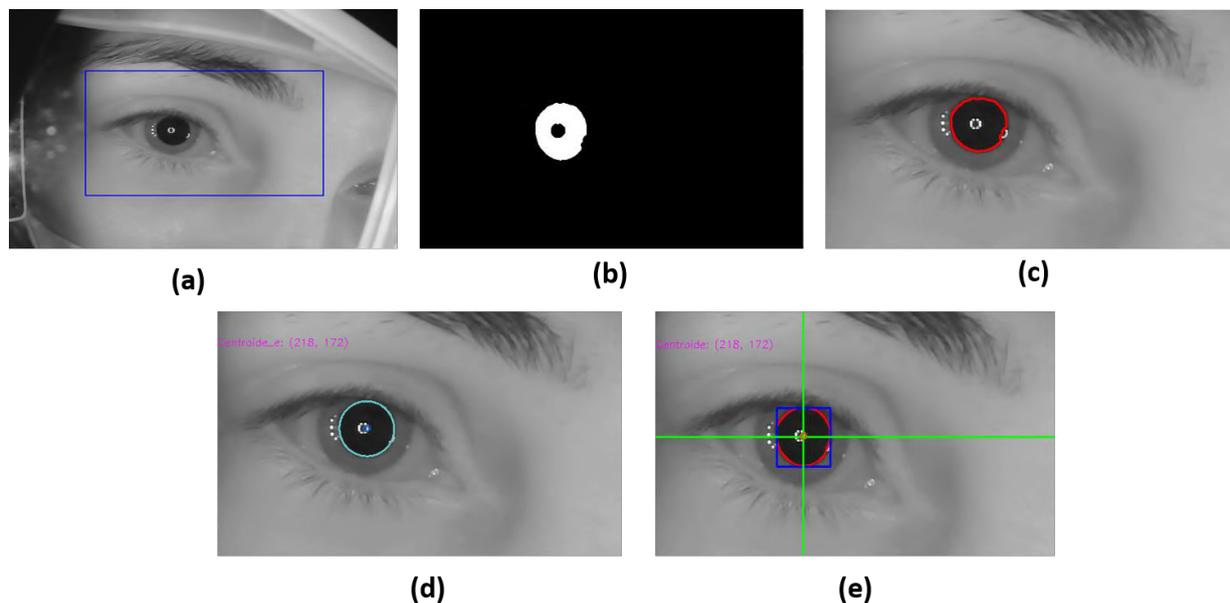
Provas	Vídeos - Coletas	x	y	a
OHE	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6
OHD	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6
CQOD	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6
CQOE	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6
CFOD	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6
CFOE	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6
PC	Olho direito	V1	V2	V3
	Olho esquerdo	V4	V5	V6

Fonte: Autores.

O conjunto de vetores de características ($V1$, $V2$, $V3$, $V4$, $V5$ e $V6$) listados no Quadro 1 são parâmetros temporais, ou seja, que variam com o tempo, considerados como entrada para o treinamento dos modelos de classificação. Juntos, esses vetores formam uma série temporal multivariada e mantêm dependências entre si. Cada conjunto de 3 vetores ($V1$ ao $V3$ e $V4$ ao $V6$) representam os valores relativos às coordenadas x, y e à área da pupila, respectivamente. Os vetores $V1$, $V2$ e $V3$ são resultados da análise das coletas de vídeo das provas aplicadas ao olho direito, enquanto os vetores $V4$, $V5$ e $V6$ são resultados correspondentes ao olho esquerdo.

O rastreamento ocular é iniciado na calibração, realizada no momento em que as provas são aplicadas e os vídeos são gravados. Nesse momento, há um ajuste das câmeras para que a área de captura encaixe o olho ao centro para permitir mais facilmente o rastreamento do globo ocular, reduzindo o espaço analisado e evitando confundidores, como sombras ou reflexos (Figura 4a). Depois disso, a biblioteca de VC *OpenCV* é usada para processar todos os vídeos, quadro a quadro. Cada quadro passa pelo processo de limiarização (do inglês, *thresholding*) para destacar a pupila e assim detectar seu contorno (Figura 4b).

Figura 4 - Processo de identificação do globo ocular: (a) Área de interesse; (b) Aplicando a limiarização; (c) Aplicando a envoltória convexa; (d) Encaixe da elipse circunscrita; e (e) Cálculo do centróide.



Fonte: Autores.

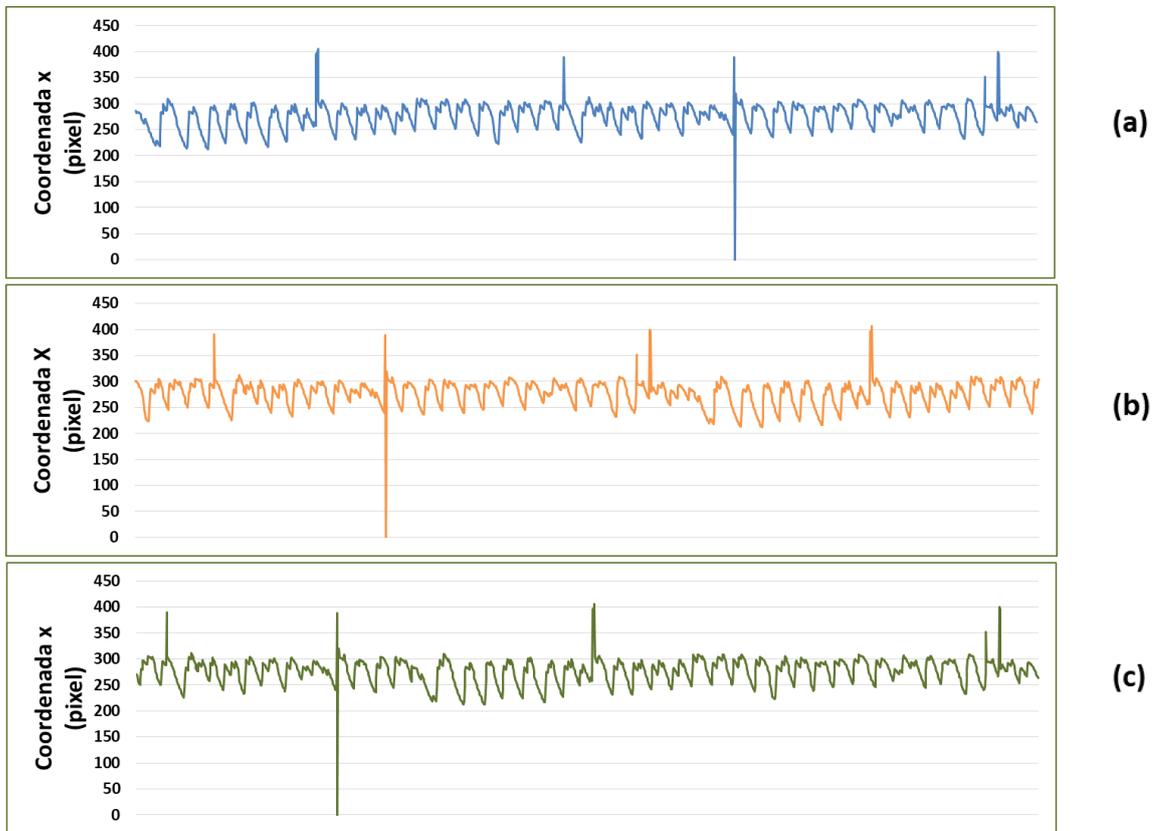
A envoltória convexa (do inglês, *convex hull* ou *convex envelope*), a qual se encaixa em um problema de geometria computacional, consiste em: dado vários pontos, somente os pontos mais externos e que formam o menor polígono que “fechem” todos os outros pontos, são considerados. Esses pontos são ligados, formando assim um polígono ou invólucro que abarque todo o conteúdo interno. No caso da pupila, é usado para expandir o contorno encontrado através da limiarização, que em muitos momentos pode apresentar formas geométricas côncavas (Figura 4c). Em seguida, o encaixe da elipse circunscrita é aplicado para aproximar o invólucro geométrico de uma forma circular ou elíptica, o que se aproxima mais da pupila (Figura 4d). A determinação dessa forma aproximada permitirá calcular o centro geométrico, ou centróide, da pupila (coordenadas x e y) e sua área (Figura 4e).

2.4 Aumento de dados e normalização vetorial

Existem métodos para gerar dados sintéticos que preservam as características fundamentais dos dados originais. O aumento de dados tende a aumentar também a capacidade de generalização dos modelos treinados, reduzindo o sobreajuste (do inglês, *overfitting*) (Shorten & Khoshgoftaar, 2019). No mundo real, em que a escassez de dados pode afetar a decisão de um modelo, o aumento deste é especialmente importante para um conjunto desequilibrado (Olson et al., 2018) ou limitado (Hasibi et al., 2019). Nesse sentido, o conjunto de dados construído neste estudo consiste em séries temporais multivariadas, que foram submetidas a técnicas de aumento de dados específicas para dados unidimensionais.

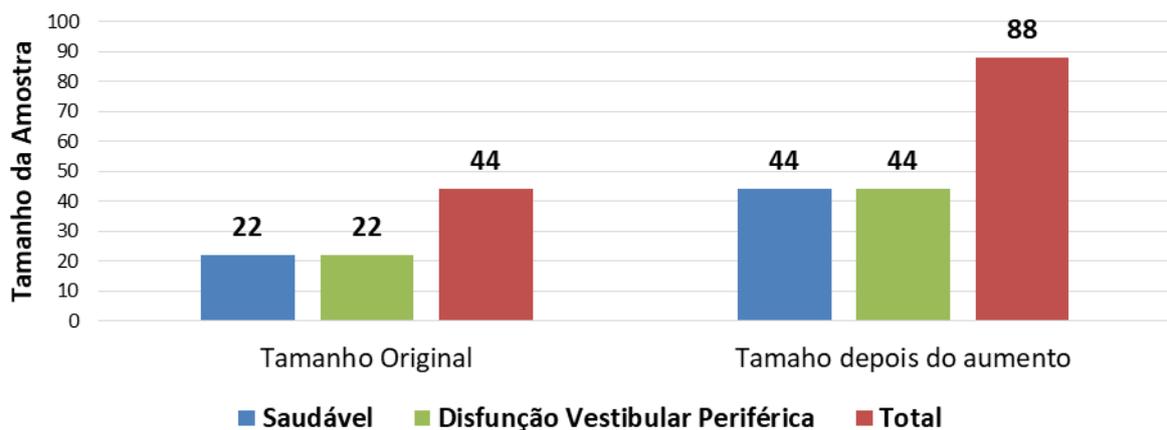
Cada vetor possui 1500 elementos, nos quais a técnica de permutação é aplicada para gerar dados sintéticos. A permutação consiste em dividir o vetor em N segmentos, os quais são realocados temporal e aleatoriamente. Onde N varia entre os valores 2 e 7, e o número total de elementos é dividido igualmente pelo número de segmentos ou tendo uma distribuição também aleatória, cuidando para que o vetor resultante seja um dado sintético válido. A Figura 5 mostra a permutação aplicada duas vezes ao vetor original, resultando em um novo vetor sintético sempre que aplicado. A técnica foi aplicada aos dados originais até que eles duplicassem de tamanho, ampliando assim a quantidade de amostras de 44 para 88 (Figura 6).

Figura 5 - Permutação aplicada 2 vezes (b e c) ao vetor de coordenada X original (a) obtido do vídeo durante a prova OHE, do olho direito.



Fonte: Autores.

Figura 6 - Tamanho da amostra antes e depois do aumento de dados.



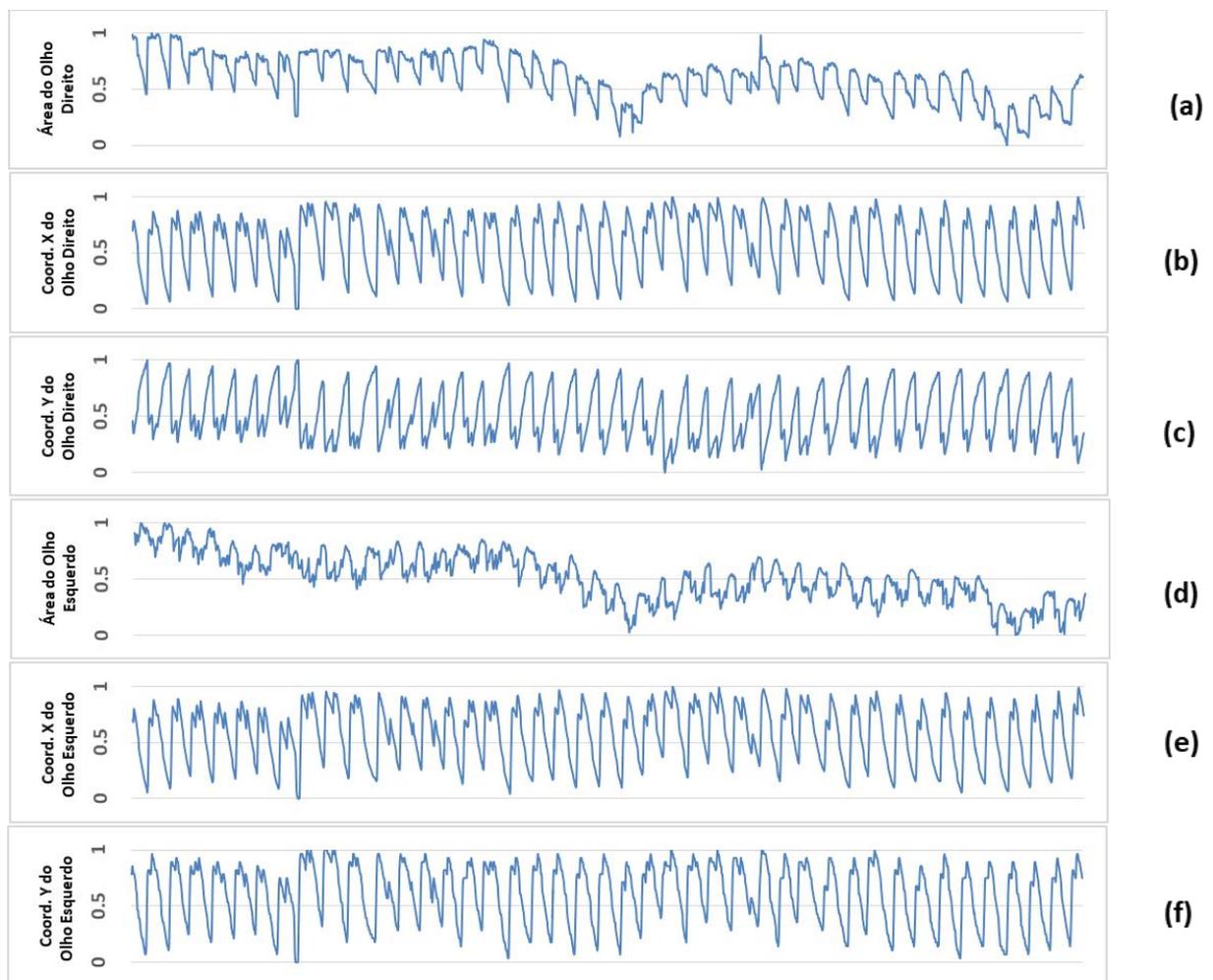
Fonte: Autores.

A normalização dos dados é considerada uma etapa necessária no treinamento de algoritmos de AP, tendo seu desempenho afetado caso não tenha sido feito apropriadamente (Passalis, 2019). Dada a variação de tamanho da pupila para cada paciente e também as diferentes proporções do vídeo em relação ao espaço ocupado pelo olho, foi utilizada a abordagem *Mim-Max Scalling* para normalizar os dados de entradas, de acordo com a Equação 1 abaixo.

$$X_{normalizado} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (\text{Equação 1})$$

A Equação 1 é aplicada a cada vetor da Tabela 1. Considerando os valores de cada vetor, X representa o valor de cada elemento do vetor, X_{min} é o menor valor dentre os elementos, e X_{max} o maior, resultando assim no valor normalizado $X_{normalizado}$ para cada um deles. Essa equação é conhecida na prática como Normalização Min-Max. Um exemplo dos resultados da aplicação da Equação 1 nos vetores coletados (com valores que variam de 0 a 1) de uma prova específica são apresentados na Figura 7.

Figura 7 - Vetores de características normalizados: (a) Área da pupila do Olho Direito; (b) Coordenada X do Olho Direito; (c) Coordenadas Y do Olho Direito; (d) Área da pupila do Olho Esquerdo; (e) Coordenada X do Olho Esquerdo; (f) Coordenada Y do Olho Esquerdo.



Fonte: Autores.

2.5 Treinamento de algoritmos de aprendizado profundo

Dado que a preparação dos dados foi realizada, a próxima etapa consistiu em criar os modelos de AP, visando encontrar um classificador que melhor desempenho para o problema abordado e conjunto de dados criado. Partindo do conhecimento obtido a partir referencial teórico e das soluções do estado da arte, o presente estudo escolheu trabalhos com duas metodologias a serem utilizados como base, são eles: a arquitetura CNN com uma dimensão proposta por Ben Slama et al. (2019) e o método

que utiliza classificadores fracos baseados em LSTM, proposto por Mao (2020). Para o propósito de rápido endereçamento dos modelos no restante do artigo, estes serão referenciados por 1D-CNN SLAMA e LSTM MAO, respectivamente.

Modelos baseados em CNN são geralmente usados na classificação de dados com duas dimensões, o que inclui imagens. Esse tipo de rede pode ser usada de forma adaptada para conjuntos de uma dimensão (1D-CNN), dessa forma se torna possível aplicá-la em tarefas de classificação em séries temporais. O processo de coleta de dados realizado neste trabalho se diferencia do método proposto para 1D-CNN SLAMA. A quantidade de provas utilizadas se diferencia em número, o que leva a uma quantidade de características também diferentes. A formatação aplicada na camada de *input* da 1D-CNN SLAMA requer entradas configuradas como matrizes de tamanho 100x3, em que a dimensão de tamanho 3 representa as coordenadas x, y e z. Neste estudo, as entradas são configuradas como matrizes de tamanho 1500x42, ou seja, vetores de 1500 elementos, em que a dimensão de tamanho 42 representa os valores x, y e área da pupila para o olho esquerdo e direito, o que necessário em termos de não-simetria, em cada prova realizada pelo indivíduo, assim como visto na Figura 7.

O estudo realizado por Mao et al. (2020) propõe a classificação de disfunções vestibulares utilizando o algoritmo LSTM, a qual é adequada para predições em séries temporais. Para isso, cada valor de entrada tal como x ou y variam com o tempo, formando vetores que podem apresentar padrões de interesse que caracterizam o nistagmo e determinam a classificação. Os eventos ou padrões de interesse ocorrem em espaços de tempo desconhecidos e, nesse caso, a LSTM possui uma memória de longo prazo para lidar com situações em que esses espaços de tempo são longos. Desta forma, as camadas em uma LSTM ficam responsáveis por analisar vetores de características e retornar uma classificação. No método de classificação LSTM MAO, uma estrutura algoritmos LSTM é utilizada para cada vetor de características, os quais são considerados classificadores fracos. Eles juntos formam um classificador forte que, por fim, retorna uma classificação definitiva. Nesse caso, algumas adaptações em nosso estudo tiveram que ser feitas também.

O estudo desenvolvido por Mao et al. (2020) utilizou somente duas provas e coletou 3 vetores de características para cada uma delas (x, y e área da pupila), sendo necessários 6 classificadores fracos no total, um para cada vetor de característica. Deste modo, no trabalho proposto neste artigo, são realizadas 7 provas para ambos os olhos, totalizando 14 vídeos. De cada vídeo, são coletados os vetores de características x, y e área da pupila, sugerindo a necessidade de 42 classificadores fracos. Para a implementação dos modelos de AP, foi utilizada a biblioteca *TensorFlow 2* (TensorFlow, 2022) e, como parte dela, a biblioteca *Keras* (Chollet., 2015). Nessas bibliotecas, foi possível implementar tanto uma 1D-CNN quanto uma LSTM para a análise de séries temporais multivariadas. Nesse sentido, o método de classificação LSTM MAO, utilizado para classificar o conjunto de dados especificado neste estudo, usou 7 classificadores fracos, um para cada prova. Portanto, cada classificador fraco recebeu como entrada uma matriz de tamanho 1500x6, em que o valor 6 representa os vetores de características x, y e área da pupila de cada olho (esquerdo e direito). Para a classificação final ou forte, o somatório das classificações fracas são computadas levando em consideração o risco de erro de cada classificador.

A função dos modelos analisados consiste em realizar uma classificação binária para cada amostra, diferenciando-a entre indivíduos saudáveis (classe negativa) ou com disfunção vestibular periférica (classe positiva). Para alcançar esse propósito, o modelo precisa ser treinado, de forma a conseguir generalizar as características aprendidas nessa fase e reconhecer amostras desconhecidas como pertencentes à classe correta. Portanto, para cada um dos 44 pacientes (22 de cada classe) as provas foram aplicadas e os vetores de características foram coletados, os quais foram normalizados e aumentados de forma a completar 88 amostras compostas por vetores de elementos com valores entre 0 e 1. Dessa forma, esse conjunto de dados com características específicas, decorrente do modo como foi construído, e levando em consideração que também é desconhecido para os autores deste estudo os detalhes dos conjuntos de dados dos modelos desenvolvidos pelos outros estudos (Slama et al., 2019; Mao et al.

2020), chega-se à conclusão de que é necessário realizar um ajuste fino dos hiperparâmetros no processo de treinamento dos algoritmos.

O método de ajuste fino *Hyperband* proposto por Li et al. (2017) foi utilizado como maneira de encontrar um conjunto de hiperparâmetros otimizados para cada um dos algoritmos, o que permite identificar o modelo com melhor desempenho, a ser implantado no sistema. No caso da arquitetura 1D-CNN SLAMA foram refinados os hiperparâmetros de quantidade de filtros, tamanho do kernel e tamanho do passo (*strides*) para a camada de convolução. Seguindo para a camada de *MaxPooling*, o tamanho do *pool* e do passo (*strides*) foram os objetos da otimização, e para a camada de *Dropout*, o valor da taxa de *dropout*. Por fim, a quantidade de unidades da camada totalmente conectada é refinada. Quando o mesmo processo é aplicado no modelo LSTM MAO, separadamente para cada um dos classificadores fracos, a otimização é reduzida à quantidade de unidades na camada de recorrência. Para todos os modelos há também a procura pelo número apropriado de épocas usado.

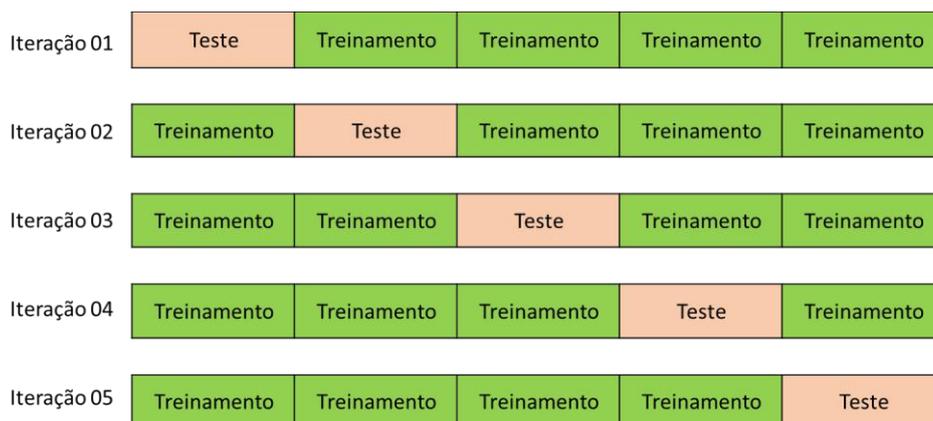
Este método de refinamento propõe acelerar a busca randômica através da alocação adaptativa de recursos e *early-stopping*, em que, dado algumas condições, o treinamento utilizando um conjunto de hiperparâmetros não segue em frente se for considerado longe do ótimo. Essa fase de treinamento dos algoritmos foi implementada utilizando a biblioteca *Keras-Tuner* (Keras, 2022).

3. Avaliação de desempenho

3.1 Metodologia da avaliação

A fim de avaliar a efetiva classificação dos modelos criados, tanto com os mesmos hiperparâmetros utilizados pelos estudos originais e também com eles refinados, foi realizada a validação cruzada com 5 *folds* (Figura 5). A validação cruzada foi aplicada nos modelos 1D-CNN SLAMA e 1D-CNN SLAMA Tuned, assim como em todos os classificadores fracos do método de classificação LSTM MAO e sua versão refinada LSTM MAO Tuned.

Figura 8 - Validação cruzada com 5 *folds*.



Fonte: Autores.

Os dados foram divididos no formato 80-20%, ou seja, 80% do conjunto de dados foi utilizado para realizar o treinamento/validação e 20% para o teste. Ressalta-se que, para a validação realizada para os modelos baseados no método descrito em Mao et al. (2020), há a necessidade de que os vetores de características analisados separadamente pelos classificadores fracos sejam da mesma amostra. Desta forma, o método *train_test_split* do módulo *scikit-learn* foi utilizado com o valor do parâmetro *random_state* igual quando aplicado aos dados, distribuindo o conjunto de vetores diferentes da mesma amostra para os classificadores fracos. Isso garante que em cada *fold* de cada iteração da validação cruzada, as mesmas amostras

sejam distribuídas entre os classificadores fracos, mesmo que estes estejam analisando características diferentes desta amostra, ou seja estejam analisando vetores diferentes, de provas diferentes.

3.2 Métricas

As métricas utilizadas são baseadas na matriz de confusão apresentada no Quadro 2, com os valores descritos abaixo:

- Verdadeiro Positivo (VP): quando há previsão correta de valor positivo;
- Verdadeiro Negativo (VN): quando há uma previsão correta de valor negativo;
- Falso Positivo (FP): quando há uma previsão errada (ou seja, uma contradição) de valor negativo;
- Falso Negativo (FN): quando há uma previsão errada de valor positivo.

Quadro 2: Matriz de Confusão.

		Valores Reais		Total
		Positivo	Negativo	
Valores previstos	Positivo	<i>Verdadeiro Positivo</i>	<i>Falso Positivo</i>	VP + FP
	Negativo	<i>Falso Negativo</i>	<i>Verdadeiro Negativo</i>	FN + VN
Total		VP + FN	FP + VN	

Fonte: Autores.

Nós analisamos as seguintes métricas de desempenho bem conhecidas (Ojala e Garriga, 2010; Sokolova e Lapalme, 2009): acurácia (Equação 2), Precisão (Equação 3), *Recall* (Equação 4) e F-score (Equação 5).

$$Acurácia = \frac{(VP + VN)}{(VP + VN + FP + FN)} \quad \text{(Equação 2)}$$

$$Precisão = \frac{VP}{VP + FP} \quad \text{(Equação 3)}$$

$$Recall = \frac{VP}{VP + FN} \quad \text{(Equação 4)}$$

$$F\text{-score} = \frac{2 * (Precisão * Recall)}{Precisão + Recall} \quad \text{(Equação 5)}$$

Visto que o objetivo dos classificadores envolve propor um diagnóstico com a maior probabilidade de acerto possível, a métrica principal escolhida para a avaliação dos modelos foi a acurácia. Ela define a taxa de acerto, ou seja, é um valor relacionado às classificações corretas dentre todas as amostras apresentadas ao classificador. A acurácia é a métrica mais utilizada entre os trabalhos relatados na literatura como parâmetro de comparação de desempenho. As outras métricas (*Recall*, Precisão e F-score) também são analisadas nesse estudo de forma secundária.

Visando verificar a eficácia do uso da área da pupila como vetor de entrada e da capacidade do aumento de dados em melhorar o desempenho dos modelos de classificação, realizamos diferentes experimentos para identificar a configuração que consegue alcançar o melhor desempenho. Além disso, também consideramos como variável de análise a aplicação da normalização dos valores de entrada. O Quadro 3 apresenta todas as configurações da preparação dos dados de entrada considerados na análise.

Quadro 3 - Diferentes configurações da preparação dos dados de entrada.

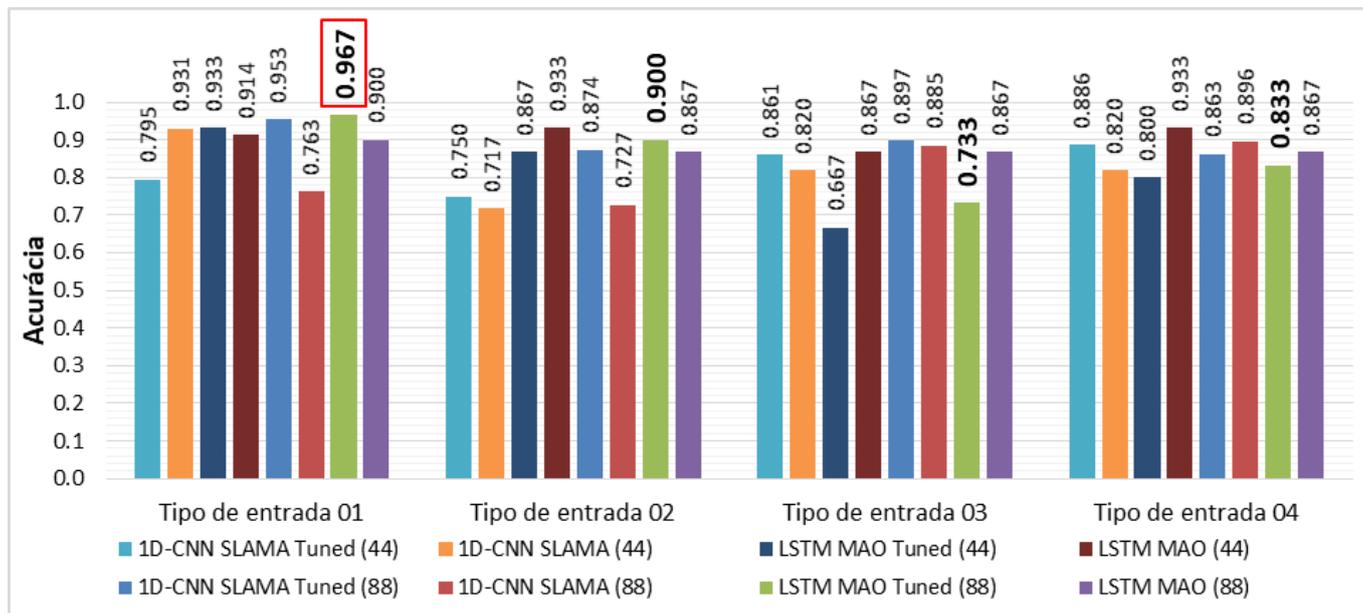
Vetores de características		
Tamanho da amostra	Tipo de entrada	Identificador
44 amostras	Coordenadas X e Y	Entrada 01
	Coordenadas X, Y e área da pupila	Entrada 02
	Coordenadas X e Y (valores normalizados)	Entrada 03
	Coordenadas X, Y e área da pupila (valores normalizados)	Entrada 04
88 amostras	Coordenadas X e Y	Entrada 01
	Coordenadas X, Y e área da pupila	Entrada 02
	Coordenadas X e Y (valores normalizados)	Entrada 03
	Coordenadas X, Y e área da pupila (valores normalizados)	Entrada 04

Fonte: Autores.

3.3 Resultados

Para comprovar a eficiência de predição dos modelos desenvolvidos, várias métricas para tarefas de classificação foram calculadas e comparadas para os tamanhos de amostra (44 e 88) e as oito variações de entrada (Quadro 3), para cada modelo com hiperparâmetros igualmente aos usados nos estudos e também ajustados. O modelo que obteve melhor acurácia nos testes, como visto na Figura 9, foi o LSTM MAO Tuned, ou seja, o método de classificação proposto em Mao et al. (2020) com a otimização dos hiperparâmetros através do ajuste fino, chegando a 96,7%. O ajuste no número de neurônios usados nas camadas da LSTM para cada classificador fraco sugere que os hiperparâmetros adaptados resultam em um melhor desempenho.

Figura 9 - Acurácia para todos os modelos utilizados, considerando os diferentes tipos de entrada e tamanhos de amostra.



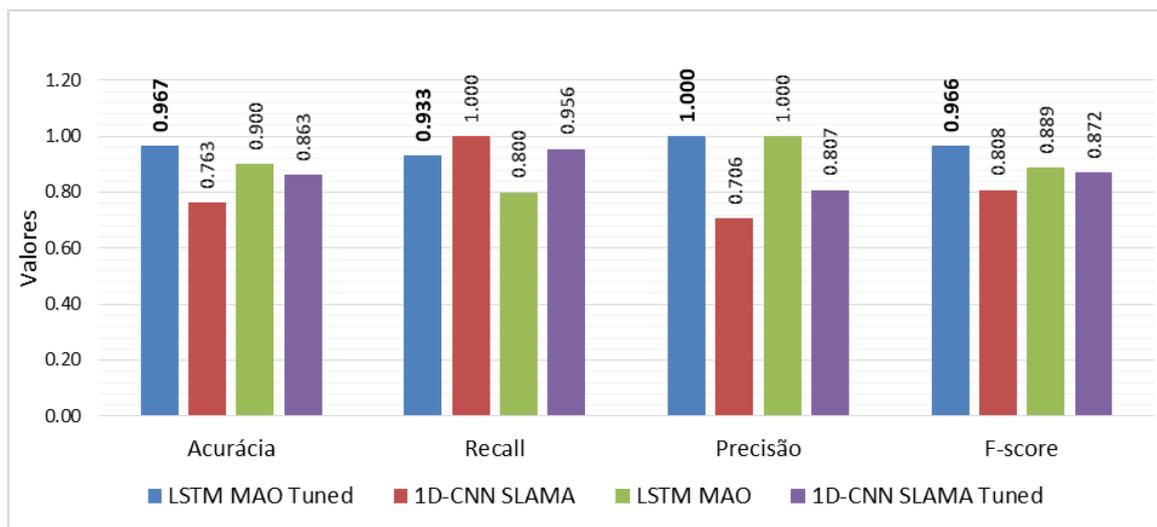
Fonte: Autores.

O impacto positivo do ajuste fino perde efeito nas entradas do tipo 03 e 04, as quais são normalizadas previamente. Essa tendência pode ser endereçada às condições de parada ou *early-stopping* aplicadas na abordagem de ajuste de hiperparâmetros. Ou seja, o algoritmo de ajuste fino *hyperband* somente prossegue com o treinamento de algoritmos que se mostram precocemente bons. Isso significa que os modelos que atingem uma alta acurácia nas primeiras épocas são escolhidos como candidatos ótimos, enquanto os outros são descartados. Dessa forma, o algoritmo consegue ser mais rápido do que um algoritmo que não utiliza nenhuma heurística, por exemplo, o algoritmo *GridSearch*. O algoritmo utilizado, apesar de mais rápido e na sua maioria, sugerir conjuntos de hiperparâmetros próximos do ótimo, pode acabar sugerindo um conjunto não muito eficiente para modelos que precisam de mais tempo de treinamento para convergir. O *HyperbandSearch* é um algoritmo robusto e eficiente quanto a otimização de hiperparâmetros, mas, para os modelos empregados nesse conjunto de dados específico, com valores normalizados, uma abordagem que considere mais tempo de treinamento para descarte de uma configuração candidata poderia ser mais eficiente. No entanto, deve ser pontuado o bom desempenho do refinamento nos primeiros tipos de entradas (entradas 01 e 02).

No gráfico da Figura 9 também pode ser observado o ganho de desempenho quando o aumento dos dados é implementado. Em 75% dos casos, quando o mesmo modelo classifica um conjunto de dados com 44 amostras e com 88 amostras, onde se incluem os dados sintéticos, é possível observar um aumento na acurácia.

Na Figura 10 são apresentados os resultados na avaliação dos algoritmos de AP, trazendo as métricas analisadas neste estudo na situação em que os algoritmos tiveram o melhor desempenho: para a amostra de tamanho 88 e entradas do tipo “Entrada 04”. O LSTM MAO Tuned mantém o melhor desempenho para a maioria das métricas, tendo uma acurácia de 96,7%, a melhor entre os modelos criados.

Figura 10 - Comparação dos resultados. Métricas: Acurácia, Recall, Precisão e F-score. Considerando o tamanho da amostra: 88 amostras, e o tipo de entrada: “Tipo 04” (Coordenadas X, Y e área da pupila - com valores normalizados).



Fonte: Autores.

4. Discussão

A área de VC vem recebendo grande atenção da comunidade de pesquisadores de saúde, com intuito de automatizar e aperfeiçoar processos clínicos. O uso de IA e, mais precisamente, de abordagens de AP, ajudam na análise dessas informações, evitando vieses inerentes ao ser humano ou mesmo suprindo a falta de um profissional treinado. Entende-se que a versatilidade de métodos de diagnósticos automatizados é uma tendência inclusive no ramo da otorrinolaringologia, onde este trabalho vem propor uma solução.

Pesquisas médicas comprovam que o movimento dos olhos está relacionado a várias doenças psicológicas e físicas. O sistema proposto é principalmente uma ferramenta de suporte ao diagnóstico que, a partir de informações obtidas do movimento dos olhos, informa se o respectivo paciente possui disfunção periférica, com acurácia superior a 96%. O processo de desenvolvimento deste sistema passou por várias etapas:

1. Impressão 3D do dispositivo de coleta utilizado para a obtenção do movimento dos olhos;
2. Desenvolvimento de uma interface web intuitiva para o profissional de saúde;
3. Desenvolvimento de um rastreador ocular baseado em VC;
4. Análise e seleção de um algoritmo de AP para automatizar o suporte diagnóstico de disfunção vestibular periférica.

Dessa forma, busca-se ocupar um espaço ainda pouco explorado de ferramentas ou sistemas de diagnósticos de doenças vestibulares, visando o baixo custo, acesso fácil e o manuseio intuitivo. Pode-se verificar que, com um rastreo preciso do globo ocular é possível treinar algoritmos de AP que possam fornecer de forma confiável uma ajuda ao processo diagnóstico realizado pelo profissional de saúde.

O trabalho apresentado neste artigo constitui uma versão inicial do sistema *Smart Nystagmography*. No entanto, mesmo sendo uma versão inicial, muitas dificuldades foram enfrentadas para o seu desenvolvimento. Primeiramente, o conjunto de procedimentos aplicados para a coleta dos dados foi um desafio, mostrando a complexidade de criar um conjunto de dados clínicos. Outra dificuldade foi a adaptação com o dispositivo de coleta e com a interface do sistema, os quais são componentes da ferramenta que demandam ajustes após o primeiro uso pelo examinador, até a geração de ruídos. Por fim, o algoritmo de

reconhecimento da pupila precisou ser ajustado para desconsiderar artefatos, tais como reflexos e sombras indesejadas no processo de coleta.

Algumas limitações deste estudo precisam ser reconhecidas. Como visto em outros trabalhos da literatura, técnicas para lidar com artefatos, como a piscada dos olhos, por exemplo, poderiam trazer mais robustez e confiança para o classificador usado. Apesar do sistema apresentar uma taxa de acerto de 96,7%, entende-se que poderia haver bastante benefício em um conjunto de dados mais amplo e otimizado com técnicas para lidar com artefatos. Outra limitação a ser reconhecida é a versatilidade das técnicas escolhidas para o aumento de dados. Este trabalho limitou a permutação de vetores para implementar o aumento, dado o seu caráter simples, eficiente e a grande capacidade de multiplicar o número de amostras. Como citado em Iwana e Uchida (2021), atualmente existe uma variedade de técnicas de aumento de dados para séries temporais, as quais poderiam ser aplicadas aos vetores de características criadas neste trabalho. Uma última limitação é o tamanho da amostra, a qual poderia ser maior. Muitos indivíduos se recusaram a participar do estudo devido a atual situação em que o mundo se encontra com a pandemia da COVID-19.

5. Considerações Finais

Este estudo objetivou desenvolver o *Smart Nystagmography*, uma ferramenta para suportar o diagnóstico de disfunções vestibulares periféricas a partir de vídeos que registraram o movimento ocular quando indivíduos realizavam diferentes provas optocinéticas e calóricas. A solução proposta conseguiu identificar objetivamente a presença ou ausência de disfunção vestibular periférica com uma acurácia de 96,7%. Dessa forma, o sistema tem sua eficiência comprovada perante os experimentos realizados. Estudos clínicos adicionais são necessários para validar a solução.

Os trabalhos futuros estão sendo pensados para o estudo de novas arquiteturas de redes neurais profundas, bem como na otimização de hiperparâmetros dos algoritmos. Isso poderá melhorar o desempenho dos modelos para a tarefa de classificação de disfunções vestibulares periféricas. Ainda como trabalhos futuros também fica o estudo de outros métodos de aumento de dados, diferentes opções de algoritmos para a classificação de séries temporais, e também aumentar continuamente o número de amostras para, assim, incrementar o desempenho do processo de classificação de disfunções vestibulares periféricas.

Referências

- Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.
- Chollet, F. (2015). Keras documentation. *keras.io*, 33.
- Eggers, S. D. (2019). Approach to the Examination and Classification of Nystagmus. *Journal of Neurologic Physical Therapy*, 43, S20-S26.
- Felipe, L., Staggs, A., & Hunnicutt, S. (2021). Can Type of Dizziness Influence the Vestibular Caloric Test Result?. *Journal of Primary Care & Community Health*, 12, 21501327211030120.
- Ganância, M. M., Caovilla, H. H., & Ganância, F. F. (2010). Electronystagmography versus videonystagmography. *Brazilian Journal of Otorhinolaryngology*, 76(3), 399-403.
- Mantokoudis, G., Otero-Millan, J., & Gold, D. R. (2022). Current concepts in acute vestibular syndrome and video-oculography. *Current opinion in neurology*, 35(1), 75-83.
- Groeziinger, M., Huppert, D., Strobl, R., & Grill, E. (2020). Development and validation of a classification algorithm to diagnose and differentiate spontaneous episodic vertigo syndromes: results from the DizzyReg patient registry. *Journal of neurology*, 267(1), 160-167.
- Hasibi, R., Shokri, M., & Dehghan, M. (2019). Augmentation scheme for dealing with imbalanced network traffic classification using deep learning. *arXiv preprint arXiv:1901.00204*.
- Iwana, B. K., & Uchida, S. (2021). An empirical survey of data augmentation for time series classification with neural networks. *Plos one*, 16(7), e0254841.
- Kamogashira, T., Fujimoto, C., Kinoshita, M., Kikkawa, Y., Yamasoba, T., & Iwasaki, S. (2020). Prediction of vestibular dysfunction by applying machine learning algorithms to postural instability. *Frontiers in neurology*, 11, 7.

- Karatas, M. (2008). Central vertigo and dizziness: epidemiology, differential diagnosis, and common causes. *The neurologist*, 14(6), 355-364.
- Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2017). Hyperband: A novel bandit-based approach to hyperparameter optimization. *The Journal of Machine Learning Research*, 18(1), 6765-6816.
- Lim, E. C., Park, J. H., Jeon, H. J., Kim, H. J., Lee, H. J., Song, C. G., & Hong, S. K. (2019). Developing a diagnostic decision support system for benign paroxysmal positional vertigo using a deep-learning model. *Journal of clinical medicine*, 8(5), 633.
- Mao, Y., He, Y., Liu, L., & Chen, X. (2020). Disease Classification Based on Synthesis of Multiple Long Short-Term Memory Classifiers Corresponding to Eye Movement Features. *IEEE Access*, 8, 151624-151633.
- Mostafa, B. E., Kahky, A. O. E., Kader, H. M. A., & Rizk, M. (2014). Central vestibular dysfunction in an otorhinolaryngological vestibular unit: incidence and diagnostic strategy. *International archives of otorhinolaryngology*, 18, 235-238.
- Neuhauser, H. K. (2016). The epidemiology of dizziness and vertigo. *Handbook of clinical neurology*, 137, 67-82.
- Olson, M., Wyner, A., & Berk, R. (2018). Modern neural networks generalize on small data sets. *Advances in Neural Information Processing Systems*, 31.
- Papageorgiou, E., McLean, R. J., & Gottlob, I. (2014). Nystagmus in childhood. *Pediatrics & Neonatology*, 55(5), 341-351.
- Passalis, N., Tefas, A., Kannianen, J., Gabbouj, M., & Iosifidis, A. (2019). Deep adaptive input normalization for time series forecasting. *IEEE transactions on neural networks and learning systems*, 31(9), 3760-3765.
- Pietkiewicz, P., Pepaś, R., Sułkowski, W. J., Zielińska-Bliźniewska, H., & Olszewski, J. (2012). Electronystagmography versus videonystagmography in diagnosis of vertigo. *International journal of occupational medicine and environmental health*, 25(1), 59-65.
- Sahli, H., Ben Slama, A., Bouzaiane, S., Marrakchi, J., Boukriba, S., & Sayadi, M. (2020). VNG technique for a convenient vestibular neuritis rating. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 8(5), 571-580.
- Savage, C. O. (2011, January). Eye position prediction in the case of nystagmus and refixations. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 7924-7927). IEEE.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- Slama, A. B., Machraoui, A. N., & Sayadi, M. (2014, November). Pupil tracking using active contour model for videonystagmography applications. In *International Image Processing, Applications and Systems Conference* (pp. 1-5). IEEE.
- Slama, A. B., Mouelhi, A., Sahli, H., Manoubi, S., Mbarek, C., Trabelsi, H., ... & Sayadi, M. (2017). A new preprocessing parameter estimation based on geodesic active contour model for automatic vestibular neuritis diagnosis. *Artificial intelligence in medicine*, 80, 48-62.
- Slama, A. B., Sahli, H., Mouelhi, A., Marrakchi, J., Sayadi, M., & Trabelsi, H. (2020). DBN-DNN: discrimination and classification of VNG sequence using deep neural network framework in the EMD domain. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 8(6), 681-690.
- Slama, A. B., Mouelhi, A., Sahli, H., Manoubi, S., Lahiani, R., Salah, M. B., ... & Sayadi, M. (2018). A new neural network method for peripheral vestibular disorder recognition using VNG parameter optimisation. *International Journal of Biomedical Engineering and Technology*, 27(4), 321-336.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4), 427-437.
- Stanton, M., & Freeman, A. M. (2021). Vertigo. In *StatPearls [Internet]*. StatPearls Publishing.
- Strupp, M. L., Straumann, D., & Helmchen, C. (2021). Nystagmus: Diagnosis, Topographic Anatomical Localization and Therapy. *Klinische Monatsblätter für Augenheilkunde*, 238(11), 1186-1195.
- TensorFlow. Uma plataforma completa de código aberto para machine learning. Recuperado em 16 de março, 2022, de <https://www.tensorflow.org/>
- Vallim, M. G. B., Gabriel, G. P., Mezzalira, R., Stoler, G., & Chone, C. T. (2021). Does the video head impulse test replace caloric testing in the assessment of patients with chronic dizziness? A systematic review and meta-analysis. *Brazilian journal of otorhinolaryngology*, 87, 733-741.
- Visscher, R. M., Feddermann-Demont, N., Romano, F., Straumann, D., & Bertolini, G. (2019). Artificial intelligence for understanding concussion: Retrospective cluster analysis on the balance and vestibular diagnostic data of concussion patients. *PloS one*, 14(4), e0214525.
- Wazlawick, R. S. (2014). Metodologia de Pesquisa Para Ciência Da Computação (2nd ed.). Elsevier.
- West, P. D. B., Sheppard, Z. A., & King, E. V. (2012). Comparison of techniques for identification of peripheral vestibular nystagmus. *The Journal of Laryngology & Otology*, 126(12), 1209-1215.
- Wipperman, J. (2014). Dizziness and vertigo. Primary Care: Clinics in Office Practice, 41(1), 115-131.
- Wu, C. N., Luo, S. D., Chen, S. F., Huang, C. W., Chiang, P. L., Hwang, C. F., ... & Li, Y. L. (2022). Applicability of Oculomotor Tests for Predicting Central Vestibular Disorder Using Principal Component Analysis. *Journal of Personalized Medicine*, 12(2), 203.
- Yiu, Y. H., Aboulatta, M., Raiser, T., Ophay, L., Flanagan, V. L., Zu Eulenburg, P., & Ahmadi, S. A. (2019). DeepVOG: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning. *Journal of neuroscience methods*, 324, 108307.