

EyeTeractive: Sistema Multiplataforma de Rastreamento Ocular para Interação Assistiva de Baixo Custo

Jáder Louis de Souza Gonçalves¹, Lucas Marques de Cunha¹

¹ Departamento de Ciência da Computação – Universidade Federal de Rondônia (UNIR)
Código Postal 76801-059 – Porto Velho – RO – Brasil

{jaderlouis, lucas.marques}@unir.br

Abstract. Context: People with mild motor disabilities still face difficulties using digital technologies, especially due to the high cost and complexity of current assistive tools. **Objective:** To develop a low-cost, cross-platform eye-tracking system using only standard webcams, capable of interpreting eye movements as control commands. **Methods:** The system combines convolutional neural networks (ResNet-101) and MediaPipe for detecting the eye region. Paraconsistent logic is applied to interpret both the direction and intensity of eye movements under uncertainty. **Results:** The model achieved 98,8% accuracy in classifying nine distinct eye movement directions. **Conclusion:** This solution provides an accessible and affordable alternative for assistive interaction through webcam-based eye tracking.

Keywords Eye-tracking, assistive technology, paraconsistent logic, accessibility, webcam.

Resumo. Contexto: Pessoas com deficiências motoras leves ainda encontram dificuldades no uso de tecnologias digitais, devido ao custo e complexidade de ferramentas assistivas disponíveis. **Objetivo:** Desenvolver um sistema de rastreamento ocular multiplataforma e de baixo custo, que funcione com webcams convencionais e traduza movimentos oculares em comandos de controle. **Métodos:** A abordagem combina redes neurais convolucionais (ResNet-101) com o MediaPipe para detecção da região ocular. A lógica paraconsistente é utilizada para interpretar a direção e a intensidade do movimento ocular, mesmo sob incertezas. **Resultados:** O sistema obteve 98,8% de acurácia na classificação de nove direções distintas do olhar. **Conclusão:** A solução representa uma alternativa acessível e eficaz para interação assistiva baseada em rastreamento ocular com webcams.

Palavras-Chave Rastreamento ocular; tecnologia assistiva; lógica paraconsistente; acessibilidade; webcam.

1. Introdução

A interação com sistemas computacionais tornou-se essencial, mas pessoas com deficiência motora ainda enfrentam barreiras devido à dependência de mouse e teclado [Mozetic 2024]. O rastreamento ocular surge como alternativa promissora [Khan e Lee 2019], porém seu uso é limitado por custos elevados, hardware proprietário e complexidade de configuração [Housholder et al. 2021, Ehinger et al. 2019, Rakhmatulin 2020, Kumar 2006, Valliappan et al. 2020, Yang e Du 2024, Arbabi et al. 2017].

Apesar dos avanços em acessibilidade, soluções assistivas continuam voltadas principalmente à deficiência visual, deixando usuários com limitações motoras leves em segundo plano [Chemnad e Othman 2024, Fernández et al. 2024, Borges e Mendes 2024]. No Brasil, aplicativos móveis apresentam falhas recorrentes de acessibilidade [Neto et al. 2021, Chen et al. 2022, Zhang et al. 2021], e ferramentas existentes permanecem inacessíveis devido ao alto custo e à complexidade de uso [Gajos et al. 2007].

Este trabalho apresenta o *EyeTeractive*, um sistema multiplataforma de rastreamento ocular de baixo custo, projetado para funcionar em dispositivos móveis e *desktop*, utilizando apenas *webcams* convencionais. A proposta visa ampliar a acessibilidade digital por meio de uma solução leve, adaptável e com configuração mínima, focada em usuários com deficiências motoras leves ou situacionais.

2. Fundamentação teórica

O rastreamento ocular permite interações computacionais por meio do olhar, sendo útil para avaliar atenção e usabilidade [Tanenhaus e Spivey-Knowlton 1996, Štěpán Novák et al. 2024]. A detecção de direções discretas é mais acessível e robusta que o mapeamento contínuo [Garde et al. 2021, Fernández et al. 2016], com métodos como tempo de permanência e piscadas viabilizando o controle [MacKenzie 2012]. Para lidar com incertezas visuais, utiliza-se *lógica paraconsistente*, que representa direções com diferentes graus de certeza [Khaan 2024]. *Redes neurais convolucionais*, especialmente as *Residual Networks (ResNets)*, são eficazes na classificação de padrões oculares [He et al. 2016, Koonce 2021]. Técnicas como *Wavelet Transform–Support Vector Machine (WT-SVM)*, *Electrooculography (EOG)*, *Linear Discriminant Analysis (LDA)* e *Hidden Markov Models (HMM)* também têm sido aplicadas ao controle ocular [Wang et al. 2021, Heo et al. 2017, Mishra et al. 2017, Fujii et al. 2018]. Este trabalho utiliza um *dataset* externo anonimizado, mais estável que bibliotecas como *WebGazer.js* [Papoutsaki et al. 2016], permitindo controle das condições experimentais.

3. Metodologia

A pesquisa adotou uma abordagem iterativa, iniciando com revisão bibliográfica e testes em dados anonimizados. Para contornar limitações da lógica paraconsistente isolada, foram integradas redes neurais convolucionais, resultando em um modelo posteriormente validado com usuários. Utilizou-se um conjunto de vídeos focados na região ocular, coletados por terceiros em ambiente controlado. A extração da região de interesse foi realizada com a biblioteca *MediaPipe* [Lugaresi et al. 2019], escolhida por sua maior estabilidade em relação à *Dlib* [King 2009].

A direção do olhar foi estimada com base nas distâncias entre o centróide da íris e pontos faciais extraídos via *MediaPipe*, resultando em contribuições percentuais nos eixos horizontal e vertical. Valores positivos indicam olhar para cima/direita, e negativos para baixo/esquerda. A detecção de piscadas foi realizada com a razão de aspecto ocular (EAR), inspirada em [Xu et al. 2023], considerando um limiar inferior sustentado por múltiplos quadros.

A lógica paraconsistente foi aplicada para interpretar direções ambíguas do olhar, combinando graus de certeza e incerteza por eixo. Para lidar com limitações,

especialmente no eixo vertical, foram integradas redes neurais convolucionais (ResNet-50 e ResNet-101), que aumentaram a precisão em cenários ruidosos e com ambiguidades.

O sistema adota uma arquitetura *cliente-servidor em tempo real*, com o cliente responsável pela captura da imagem e o servidor pela inferência. A versão *mobile* foi reimplementada em *Kotlin* para garantir operação contínua em segundo plano. A estrutura apresentou bom desempenho mesmo em dispositivos com *hardware* modesto, mas requer conexão ativa para funcionar.

4. Resultados e Discussões

A interface do aplicativo adota um design minimalista, com dois botões principais: *Iniciar* e *Calibrar*, reduzindo a carga visual e cognitiva. A calibração ocorre automaticamente por meio do *MediaPipe*, que ajusta as coordenadas conforme a resolução da tela. Na versão *mobile*, o sistema funciona em segundo plano, transmitindo continuamente os dados de rastreamento ocular ao servidor. A direção do olhar é inferida em tempo real, e um *cursor virtual* reflete essa posição no dispositivo.

4.1. Classificação direcional e piscadas

O sistema obteve acurácia global de 98,8% na classificação de direções oculares, com melhor desempenho nos eixos horizontal e central, e falhas concentradas em diagonais e verticais. Testes simulados mostraram impacto mínimo de óculos comuns, redução de desempenho com lentes fotossensíveis e inviabilidade com óculos escuros; o estrabismo causou queda acentuada na acurácia e aumento no tempo de execução, enquanto variações na cor da pele não afetaram o resultado. Na detecção de piscadas, alcançou acurácia de 94,00%, precisão de 94,83%, *recall* de 99,10% e *F1-score* de 96,92%, mantendo boa separação entre eventos reais e ruídos.

4.2. Digitação

A digitação com o *EyeTeractive* foi avaliada utilizando *webcam* convencional, sem calibração prévia e com o teclado padrão do sistema operacional. Mesmo sem uma interface de entrada otimizada, o sistema demonstrou funcionamento estável, alcançando uma acurácia de caracteres de 94,7%, com velocidade média de 5,9 palavras por minuto e taxa de erro de 5,2%. Esses resultados refletem o potencial do sistema para uso acessível em contextos reais, embora melhorias na interface de entrada possam elevar sua eficiência e reduzir a taxa de erro em futuras versões.

4.3. Comparação entre modelos para classificação

Na classificação das direções do olhar, a *ResNet50* obteve 89,31% de acurácia, resultado insatisfatório. Com a *ResNet101*, o desempenho subiu para 97,15%. No sistema final, essa rede é combinada com lógica paraconsistente, que resolve ambiguidades ao analisar a coerência entre a predição e os vetores de tendência, aumentando a robustez em condições adversas.

4.4. Comparação com outros trabalhos

A Tabela 1 resume métodos relevantes de controle ocular. Soluções com *eye-trackers* [Ban et al. 2023, Pérez-Reynoso et al. 2021] ou sensores

EOG [Lopes et al. 2018, Chao et al. 2020] apresentam alta acurácia, mas requerem *hardware* dedicado ou contato direto. Já o *EyeTeractive* atinge 98,8% apenas com *webcam*, sem calibração, destacando-se pela acessibilidade e uso de lógica paraconsistente para lidar com incertezas visuais.

Tabela 1. Comparação compacta entre métodos de controle ocular.

Método	Acurácia	Sensor
<i>EyeTeractive</i>	98,8%	<i>Webcam</i>
[Ban et al. 2023]	99,99%	<i>Eye-tracker</i>
[Wang et al. 2021]	96,3%	<i>Webcam</i>
[Lopes et al. 2018]	97,0%	EOG
[Chao et al. 2020]	96,8%	EOG
[Pérez-Reynoso et al. 2021]	98,4%	<i>Eye-tracker</i>

Cuidados Éticos

A pesquisa utilizou dados anonimizados fornecidos por terceiros, coletados em contexto educacional sem finalidade científica. Conforme a Resolução CNS nº 510/2016 e a LGPD, não há identificação dos participantes, nem necessidade de submissão ao Comitê de Ética.

5. Conclusão geral

Os objetivos deste trabalho foram alcançados. O sistema *EyeTeractive* demonstrou ser funcional, responsivo e leve, com alta acurácia na classificação direcional do olhar a partir de imagens capturadas por *webcam* comum, mesmo sem validação direta com o público-alvo. Sua compatibilidade com diferentes ambientes e dispositivos reforça seu potencial como solução assistiva acessível.

Entre as contribuições, destaca-se a integração inédita entre lógica paraconsistente e redes neurais (*ResNet*), ampliando a robustez frente a incertezas visuais. O modelo obteve bom desempenho na classificação de direções discretas e serve como base replicável para sistemas assistivos. As limitações incluem a dependência de servidor e a abordagem categórica, que pode não refletir com precisão a natureza contínua dos movimentos oculares.

Como trabalhos futuros, propõe-se a transição para estimativas contínuas de posição do olhar na tela, utilizando a lógica paraconsistente como mecanismo de mapeamento, e o desenvolvimento de modelos próprios de segmentação ocular, substituindo o *MediaPipe*, com foco na melhoria da precisão vertical e na redução da sensibilidade a variações fisiológicas.

Agradecimentos

Este trabalho contou com apoio pontual de ferramentas de inteligência artificial generativa para revisão linguística e sugestões de estrutura textual, sempre sob supervisão e revisão crítica dos autores.

Referências

- Arbabi, E., Shabani, M., e Yarigholi, A. (2017). A low cost non-wearable gaze detection system based on infrared image processing.
- Ban, S., Lee, Y. J., Yu, K. J., Chang, J. W., Kim, J.-H., e Yeo, W.-H. (2023). Persistent human–machine interfaces for robotic arm control via gaze and eye direction tracking. *Advanced Intelligent Systems*, 5(7):2200408.
- Borges, W. F. e Mendes, E. G. (2024). Tecnologia assistiva e baixa visão: apps e recursos de acessibilidade em dispositivos móveis. *Cadernos Brasileiros de Terapia Ocupacional*, 32:e3746.
- Chao, C., Zhou, P., Belkacem, A. N., Lu, L., Xu, R., Wang, X., Tan, W., Qiao, Z., Li, P., Gao, Q., e SHIN, D. (2020). Quadcopter robot control based on hybrid brain–computer interface system. *Sensors and Materials*, 32:991.
- Chemnad, K. e Othman, A. (2024). Digital accessibility in the era of artificial intelligence—bibliometric analysis and systematic review. *Frontiers in Artificial Intelligence*, Volume 7 - 2024.
- Chen, S., Chen, C., Fan, L., Fan, M., Zhan, X., e Liu, Y. (2022). Accessible or not? an empirical investigation of android app accessibility. *IEEE Transactions on Software Engineering*, 48(10):3954–3968.
- Ehinger, B. V., Groß, K., Ibs, I., e König, P. (2019). A new comprehensive eye-tracking test battery concurrently evaluating the pupil labs glasses and the eyelink 1000. *PeerJ*, 7:e7086.
- Fernández, A., Usamentiaga, R., Carús, J. L., e Casado, R. (2016). Driver distraction using visual-based sensors and algorithms. *Sensors*, 16(11).
- Fernández, G., Eizaguirre, B., Gonzalez, C., Marinangeli, A., Ciufia, N., Bacigalupe, L., Silva, B., Cohen, L., Pita, C., Garcea, O., Casas, M., Lazaro, L., Pardo, G., e Alonso, R. (2024). Abnormal eye movements increase as motor disabilities and cognitive impairments become more evident in multiple sclerosis: A novel eye-tracking study. *Multiple Sclerosis Journal - Experimental, Translational and Clinical*, 10(2):20552173241255008. Epub ahead of print, 2024 Apr-Jun.
- Fujii, K., Gras, G., Salerno, A., e Yang, G.-Z. (2018). Gaze gesture based human robot interaction for laparoscopic surgery. *Medical Image Analysis*, 44:196–214.
- Gajos, K. Z., Wobbrock, J. O., e Weld, D. S. (2007). Automatically generating user interfaces adapted to users' motor and vision capabilities. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, page 231–240, New York, NY, USA. Association for Computing Machinery.
- Garde, G., Larumbe-Bergera, A., Bossavit, B., Porta, S., Cabeza, R., e Villanueva, A. (2021). Low-cost eye tracking calibration: A knowledge-based study. *Sensors*, 21(15).
- He, K., Zhang, X., Ren, S., e Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.
- Heo, J., Yoon, H., e Park, K. S. (2017). A novel wearable forehead eog measurement system for human computer interfaces. *Sensors*, 17(7).

- Housholder, A., Reaban, J., Peregrino, A., Votta, G., e Mohd, T. K. (2021). Evaluating accuracy of the tobii eye tracker 5. In *International Conference on Intelligent Human Computer Interaction*, pages 379–390. Springer.
- Khaan, H. (2024). Nothingness and paraconsistent logic. *Revista Latinoamericana de Filosofia*, 50:51–62.
- Khan, M. Q. e Lee, S. (2019). Gaze and eye tracking: Techniques and applications inadas. *Sensors*, 19(24).
- King, D. E. (2009). Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, 10:1755–1758.
- Koonce, B. (2021). *ResNet 50*, pages 63–72. Apress, Berkeley, CA.
- Kumar, M. (2006). Reducing the cost of eye tracking systems. Technical report, Stanford University, HCI Group, Gates Building, Room 382, 353 Serra Mall, Stanford, CA 94303-9035, USA. Stanford University Technical Report.
- Lopes, A., Fernández, M., Rodríguez, H., Ferrero, F., e Postolache, O. (2018). Development of an eog-based system to control a serious game. *Measurement*, 127:481–488. Scimago Journal Ranking: 0.72 (2018).
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Ubweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M., Lee, J., et al. (2019). Mediapipe: A framework for perceiving and processing reality. In *Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR)*, volume 2019.
- MacKenzie, I. S. (2012). Evaluating eye tracking systems for computer input. In Majaranta, P., Aoki, H., Donegan, M., Hansen, D. W., Hansen, J. P., Hyrskykari, A., e Räihä, K.-J., editors, *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies*, pages 205–225. IGI Global, Hershey, PA. [software].
- Mishra, S., Norton, J. J. S., Lee, Y., Lee, D. S., Agee, N., Chen, Y., Chun, Y., e Yeo, W.-H. (2017). Soft, conformal bioelectronics for a wireless human-wheelchair interface. *Biosensors and Bioelectronics*, 91:796–803. PMCID: PMC5323068.
- Mozetic, V. A. (2024). Transformação digital centrada no ser humano: redefinindo o espaço público digital. *Revista Internacional CONSINTER de Direito*, 10(19):291–306. ISSN 2183-6396 (impresso), 2183-9522 (online). Acesso em: 25 abr. 2025.
- Neto, N. G. d. S., Medeiros, F. P. A. d., Araújo, R. P., e Silva, A. M. d. (2021). Acessibilidade em dispositivos móveis: uma análise sob a perspectiva das pesquisas em interação humano computador no brasil/ accessibility on mobile devices: an analysis from the perspective of human computer interaction researches in brazil. *Brazilian Journal of Development*, 7(4):34137–34150.
- Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., e Hays, J. (2016). WebGazer: Scalable webcam eye tracking using user interactions. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*, pages 3839–3845. AAAI.
- Pérez-Reynoso, F. D., Rodríguez-Guerrero, L., Salgado-Ramírez, J. C., e Ortega-Palacios, R. (2021). Human–machine interface: Multiclass classification by machine learning on 1d eog signals for the control of an omnidirectional robot. *Sensors*, 21(17).

- Rakhmatulin, I. (2020). A review of the low-cost eye-tracking systems for 2010–2020. *SSRN Electronic Journal*, pages 1–11. Posted: 9 Mar. 2021.
- Tanenhaus, M. K. e Spivey-Knowlton, M. J. (1996). Eye-tracking. *Language and Cognitive Processes*, 11(6):583–588. PsycINFO Database Record (c) 2016 APA, all rights reserved.
- Valliappan, N., Dai, N., Steinberg, E., He, J., Rogers, K., Ramachandran, V., Xu, P., Shojaeizadeh, M., Guo, L., Kohlhoff, K., e Navalpakkam, V. (2020). Accelerating eye movement research via accurate and affordable smartphone eye tracking. *Nature Communications*, 11(1):4553.
- Wang, X., Xiao, Y., Deng, F., Chen, Y., e Zhang, H. (2021). Eye-movement-controlled wheelchair based on flexible hydrogel biosensor and wt-svm. *Biosensors*, 11(6):198. PMC8234407.
- Xu, J., Huang, Z., Liu, L., Li, X., e Wei, K. (2023). Eye-gaze controlled wheelchair based on deep learning. *Sensors*, 23(13).
- Yang, S. e Du, W. (2024). Tri-cam: Practical eye gaze tracking via camera network. <https://arxiv.org/abs/2409.19554>. Acesso em: 20 ago. 2025.
- Zhang, X., de Greef, L., Swearngin, A., White, S., Murray, K., Yu, L., Shan, Q., Nichols, J., Wu, J., Fleizach, C., Everitt, A., e Bigham, J. P. (2021). Screen recognition: Creating accessibility metadata for mobile applications from pixels.
- Štěpán Novák, J., Masner, J., Benda, P., Šimek, P., e and, V. M. (2024). Eye tracking, usability, and user experience: A systematic review. *International Journal of Human–Computer Interaction*, 40(17):4484–4500.