Improving biomedical diagnosis through light–based technologies and machine learning

VALIDATION OF A LAPTOP-WEBCAM-BASED EYE TRACKING SYSTEM FOR SACCADE MEASUREMENT

Lorenzo Fratini[1], Jaume Pujol[2], José Luis Güel[1], Clara Mestre[2]

1. Instituto de Microcirugía Ocular (IMO) Barcelona Grupo Miranza, Carrer de Josep Maria Lladó, 3, Barcelona, Spain,
2. Centre for Sensors, Instruments and Systems Development, Universitat Politècnica de Catalunya, Rambla Sant Nebridi 10, Terrassa (Barcelona), Spain.

Contact: Lorenzo.fratini@imo.es

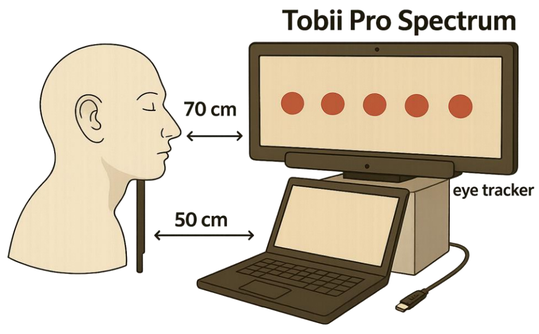
INTRODUCTION

- Deep learning (DL) is transforming eye-tracking technology by estimating gaze from eye appearance, enabling low-cost setups, greater head-movement tolerance, and minimal calibration.[1 2]

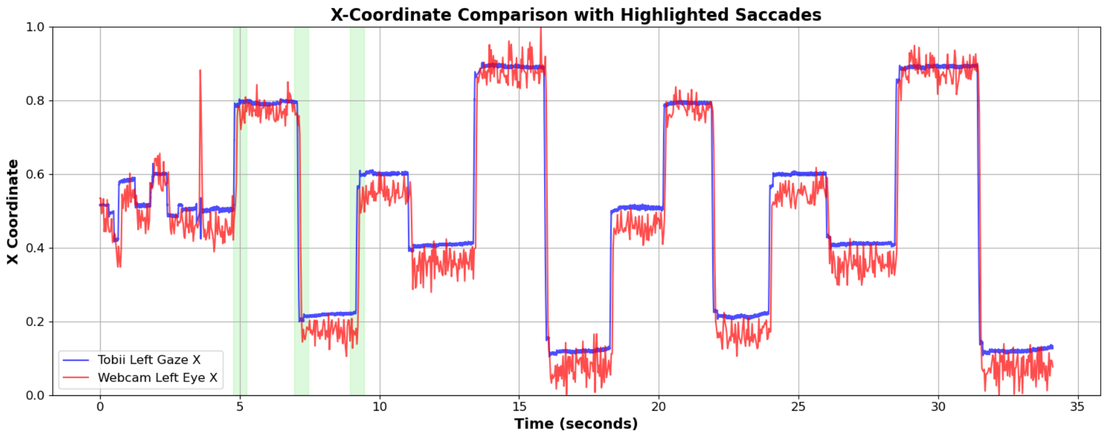
- **Goal:** To compare the performance of a DL laptop-webcam-based eye tracking system to detect and distinguish between saccades and fixations with a research-grade eye tracker (Tobii Pro Spectrum).

METHODS

- Binocular eye movements recorded from 8 healthy adults (28.12 ± 3.04 years) using a Tobii Pro Spectrum eye tracker (70 cm, @ 1200 Hz) and a laptop webcam (50 cm, @ 30 Hz). Head movements were restricted with a chinrest. Stimuli were presented on the Tobii screen (1920 x 1080 pixels; 52.8 x 29.7 cm).

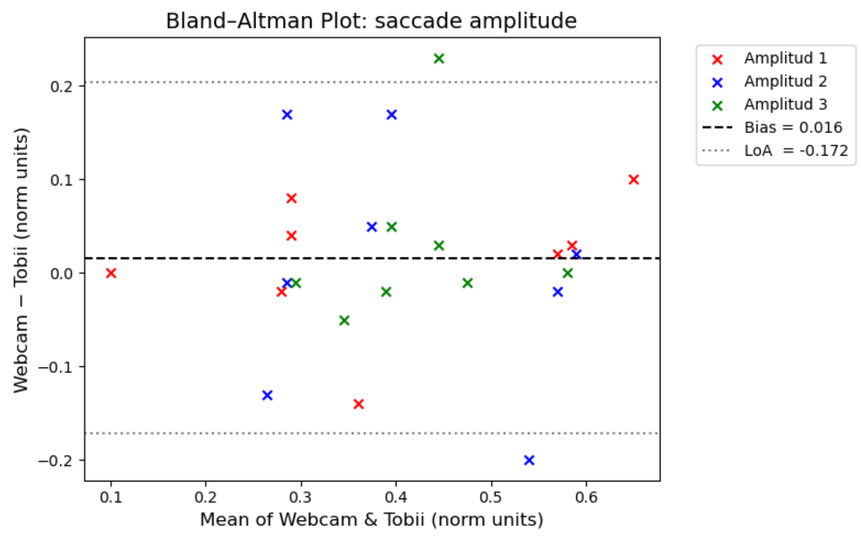


- Data were normalized to [0,1] relative to screen pixels and smoothed with a Savitzky-Golay filter. Eye velocities were derived from position differentiation, and saccades were detected using an adaptive velocity threshold algorithm.[3]



RESULTS

- The analysis included the amplitudes of saccades made in response to the first 3 stimulus' position changes [from the center of the screen (0.5 in normalized units, norm-units) to +15cm (0.78 norm-units), -15cm (0.22 norm-units) and +5cm (0.59 norm-units)].



- Bland-Altman analysis showed a bias of +0.016 norm-units with 95% limits of agreement from –0.172 to +0.208 norm-units.
- The webcam system agreed with the Tobii Pro Spectrum within ±0.19 norm-units in amplitude measurement for 95% of individual saccades, but the mean of webcam and Tobii Pro Spectrum measured amplitudes did not always match the expected ones.

Amplitudes (Norm-Units)	**Mean ±SD of differences in measured amplitude (Norm-Units)**
0.28	0.05 ± 0.04
0.56	0.09 ± 0.07
0.37	0.07 ± 0.04

CONCLUSIONS

- On average, amplitudes obtained with the webcam system matched those of a research-grade tracker, although we observed high variability among measured amplitudes. Future work will investigate its causes, expand the dataset, and assess spatial accuracy, precision, and resolution.

REFERENCES

1 Cheng, Y., Wang, H., Bao, Y., & Lu, F.; IEEE Transactions on Pattern Analysis and Machine Intelligence (2024).
2 Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M.; In Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR) (Vol. 2019).
3 Nyström, M., & Holmqvist, K.; Behavior research methods, 42(1), 188-204 (2010).

Acknowledgements: HORIZON–MSCA–2022–DN, *Improving BiomEdical diagnosis through LIGHT-based technologies and machine learning* "BE-LIGHT" (GA nº 101119924 – BE-LIGHT). Funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union