Improving biomedical diagnosis through light-based technologies and machine learning

Doctoral Networks (MSCA-DN)

HORIZON - MSCA - 2022 - DN

Grant Agreement Nº 101119924

WP1: *Eye diseases and vision*

Deliverable D1.1
Eye tracking records

Lorenzo Fratini, IMO

Clara Mestre, UPC

Date: 16 10 2025

Type: DATA - data sets, microdata, etc.

Dissemination Level: PU-Public

Funded by the European Union

DELIVERABLE IDENTIFICATION SHEET

Document version	Date	Editor	Change(s)
1	16-10-2025	Lorenzo Fratini	First version

Grant Agreement Nº	101119924
Call Identifier	HORIZON - MSCA - 2022 – DN
Acronym	BE-LIGHT
Full title	Improving biomedical diagnosis through light-based technologies and machine learning
Project URL	http://www.belightproject.eu/

Deliverable Nº and Name	D1.1 Eye tracking records
Work Package	WP1 Eye diseases and vision
Lead Beneficiary	IMO
Type	DATA - data sets, microdata, etc.
Dissemination Level	PU-Public
Due date	Month 25

Authors (Partner)	Lorenzo Fratini, IMO Clara Mestre, UPC			
Responsible Author	Clara Mestre		**Email**	clara.mestre@upc.edu
	Partner	UPC	**Phone**	+34 937398314

Abstract (for dissemination)	This deliverable presents two studies conducted within the Be-Light project at CD6-UPC and the Instituto de Microcirugía Ocular (IMO). The first study validated a webcam-based eye tracking system using appearance methods to record saccadic eye movements, comparing its performance with a research-grade eye tracker (Tobii Pro Spectrum). Participants fixated sequential visual stimuli to perform saccades of known amplitudes, allowing synchronous comparison between both systems and evaluation of low-cost eye-tracking alternatives for research and clinical applications. The second study was a preliminary investigation of paediatric amblyopia, employing a head-mounted eye tracker (Pupil Core, Pupil Labs) during a naturalistic visual search task to record eye movements in realistic viewing conditions. The collected data provide initial insights into the oculomotor behaviour of amblyopic patients. Future work will extend these experiments to assess vergence and fixation stability, supporting the development of accessible diagnostic tools for early detection of visual impairments.
Keywords	Eye tracking, Saccades, Webcam, Mediapipe, Amblyopia

Table of Contents

DELIVERABLE IDENTIFICATION SHEET **2**

EXECUTIVE SUMMARY **4**

1 INTRODUCTION **4**

2 MATERIAL AND METHODS **5**

2.1 Experimental Setup and Data Collection 5

2.2 Data processing 6

3 RESULTS **7**

4 SUMMARY **8**

REFERENCES **9**

EXECUTIVE SUMMARY

In this deliverable, a technical description of two studies performed within the framework of the Be-Light project at the CD6-UPC and the Instituto de Microcirugía Ocular (IMO) is presented.

The first study was performed to validate a webcam-based eye tracking system based on appearance methods to record saccadic eye movements and compare it with a research-grade commercial infrared eye tracker (Tobii Pro Spectrum). Participants fixated a visual stimulus that sequentially appeared at predefined screen positions to elicit horizontal saccades of known amplitudes. Both systems operated synchronously, and this configuration enabled direct performance comparison between the research-grade infrared tracker and a conventional RGB camera, allowing validation of low-cost eye tracking alternatives for large-scale or clinical use.

The second experiment represented a preliminary study to collect eye movements data from paediatric patients with amblyopia with a head-mounted eye tracker (Pupil Core, from Pupil Labs). A head-mounted eye tracker was used because it allowed us to record eye movements during naturalistic tasks that resemble everyday activities. For this preliminary study, a visual search task was designed. Further experiments are planned to deepen the analysis of the capabilities of the webcam-based eye tracking system to measure vergences and fixation stability in healthy and pathological individuals in research and clinical settings, contributing to the development of accessible diagnostic tools for early detection of oculomotor impairments.

1 INTRODUCTION

In recent decades, video oculography has become the most popular eye tracking technique due to its performance, versatility, and minimal intrusiveness. Currently, most commercial video-based eye trackers utilise the pupil-corneal reflection technique, which is based on the detection of the pupil and one or more corneal reflections from an IR light source [1]. Video oculography has become essential in many fields such as psychology, marketing, or vision science and clinical diagnostics, enabling analysis of eye movements such as saccades, fixations, vergence and smooth pursuit. High-performance commercial eye trackers offer precise measurements; however, their high-cost limits widespread accessibility, particularly for clinical and home-based applications, and they require tightly controlled laboratory setups [1]. Recently, deep learning has revolutionised conventional eye-tracking methods based on regressing gaze from the human eye appearance [2]. Deep learning-based eye tracking offers several advantages such as low hardware requirements, increased robustness to head movements, and reduced need for individual calibration procedures.

In this context, part of the work presented in this deliverable is related to the use of an eye tracking system based on a deep learning framework designed for real-time face landmark detection (MediaPipe Face Landmarker, Google). The system uses a conventional laptop webcam to capture face images. In the second set of work, a commercial head-mounted eye tracker is used to track eye movements while participants, in this case children with amblyopia, do naturalistic tasks. This represents an ecologically valid approach to study the impact of this visual disorder on patients' performance in daily-live activities.

2 MATERIAL AND METHODS

2.1 Experimental Setup and Data Collection

Binocular eye movements from 7 healthy adults with normal binocular vision (28.12±3.04 years) were recorded using the Tobii Pro Spectrum eye tracker and a laptop webcam (Lenovo Thinkbook 14 G7 IML) simultaneously at frame rates of 1200 Hz and 30 Hz, respectively.

Stimuli were presented on the native Tobii Pro Spectrum screen with a resolution of 1920 x 1080 pixels (52.8 x 29.7 cm), and both systems were aligned along the same vertical axis. The Tobii Pro Spectrum was positioned centrally below the screen, while a laptop webcam was mounted below it (Fig 1). A chinrest was used to minimize head movement and maintain a constant viewing geometry across trials. The viewing distance from participants to the screen was 70 cm.

Visual stimuli consisted of a red fixation target sequentially presented at seven horizontal positions, corresponding to physical displacements of ±5 cm, ±15 cm, and ±20 cm from the screen center (0 cm). This produced the following sequence of visual target's positions: 0 → +15 → –15 → +5 → –5 → +20 → –20 → 0 cm, giving seven horizontal saccades per trial.

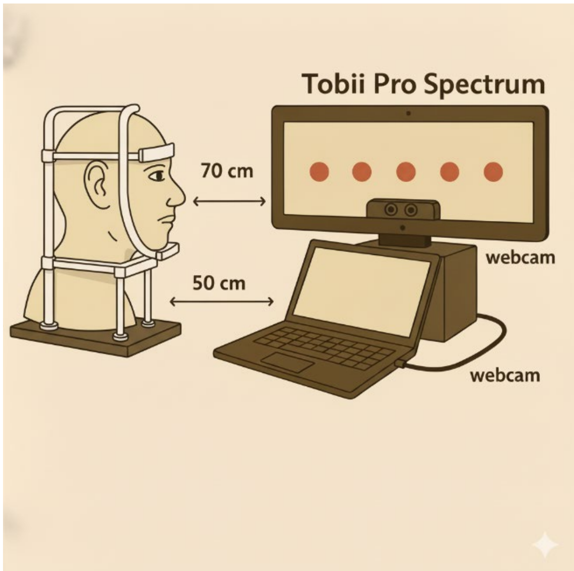

Figure 1: Schematic representation of the experimental set-up.

The two systems were software-synchronized through a custom multithreaded Python interface, ensuring parallel acquisition, timestamp alignment and avoiding bottlenecks. Each thread handled one device independently, and synchronization markers were logged using a common clock reference at the onset of each stimulus change. This approach minimized temporal offsets and ensured that corresponding gaze position data from both systems could be analyzed on a frame-by-frame basis.

Access to the Tobii data stream required the Tobii Pro SDK, providing binocular gaze coordinates, gaze direction vectors, and 3D eye position data at 1200 Hz. Data from left and right_gaze_point_on_display_area columns were used for this analysis. The webcam-based tracker captured RGB frames at 30 Hz, processed in real time using MediaPipe Face Mesh for facial landmark and iris detection [3][4]. From these features, left and right x-y positions on the screen were estimated. All data were normalized to the screen coordinate space (0–1). An amount of about 50k gaze points for each eye and each participant were collected by the Tobii, and an amount of about 1200 eye positions by the webcam.

In the second study, eye movements from 5 patients with amblyopia (6.40 ± 0.90 years) were recorded with a Pupil Core (Pupil Labs) eye tracker at 200 Hz. Four of them had anisometropic amblyopia and one had mixed amblyopia. All of them wore their habitual refractive correction. Participants did a visual search task in which stimuli were printed on a DIN A4 paper. Their task was to find the crocodile that wears the same sock as the one shown in a template (Fig 2). Whereas the stimulus was initially presented at a distance of 40 cm approximately, participants were free to move their head. Data were collected using the Pupil Capture software, and exported from the recorded videos using Pupil Player.

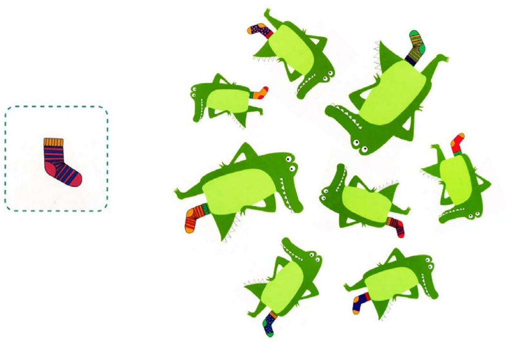

Figure 2. Example of the stimulus shown for the visual search task.

2.2 Data processing

Fig 3 shows an example of horizontal gaze position measured with the Tobii Pro Spectrum (blue) and the webcam system (red) for a representative participant. The instantaneous gaze velocity was obtained by numerical differentiation of the gaze-position signal (frame-to-frame differences) and smoothed with a Savitzky–Golay filter. Saccades were automatically detected using a velocity threshold-based algorithm, with onset and offset defined when the velocity exceeded or fell below adaptive thresholds [5].

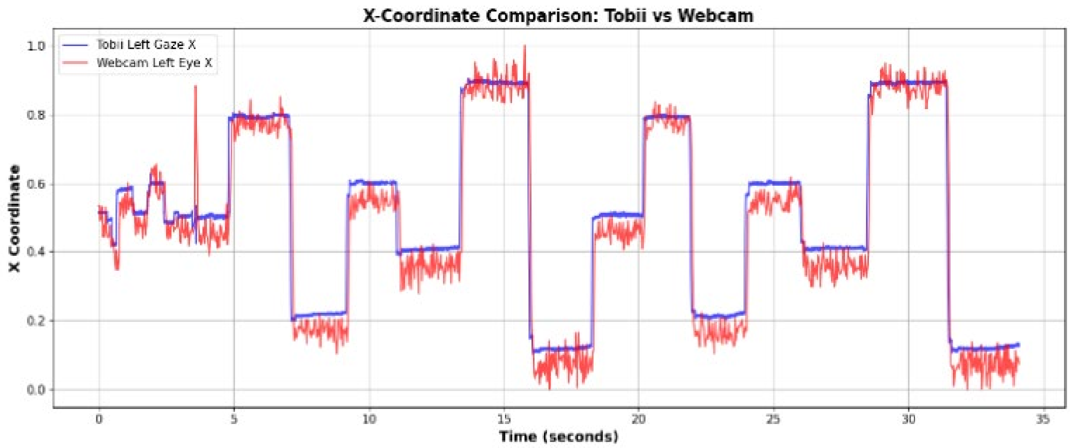

Figure 3: Raw horizontal left eye position data in FSU obtained with the webcam (red) and the Tobii Pro Spectrum (blue).

From each detected event, amplitude, peak velocity, and duration were extracted, allowing comparison between systems. Data was collected in Fixed screen units (FSU) which are dimensionless screen coordinates scaled to the interval [0, 1], with 0 representing the left edge of the display, 0.5 the centre, and 1 the right edge; units were converted to degrees for better clarification.

In the visual search task of the second study, saccades were detected using a similar velocity threshold-based algorithm [6]. Their amplitude, peak velocity and duration were extracted.

3 Results

A total of 91 saccades were detected and their amplitude were analysed. A Bland–Altman analysis revealed a mean bias of 1.00° and a standard deviation of 2.15° in the measured amplitudes between the webcam-based system and the Tobii Pro Spectrum. The mean absolute error (MAE) in saccade amplitude between systems ranged from 0.9° to 2.7° with the standard deviation (SD) of the differences varying between 0.5° and 1.9° depending on saccade amplitude (Table 1). The lowest errors were observed for the smallest saccades (≈8°), while larger movements (≈30°) exhibited slightly higher variability.

Whereas the raw data collected from the 5 participants with amblyopia, together with data from the first study, are available online [7], more analyses are required for the second study.

Amplitude (°)	MAE (°)	SD (°)
8.1	0.9	0.5
12.1	0.9	0.8
15.9	1.8	1.6
19.7	2.6	1.8
23.2	2.0	1.7
29.7	2.3	1.4

Table 1. For each visual target displacement (Amplitude), mean absolute error (MAE) between the amplitude measured with the webcam and the Tobii Pro Spectrum systems, and standard deviation (SD) of the differences.

4 Summary

This deliverable reports the outcomes of two eye-tracking studies conducted within the Be-Light project at CD6-UPC and the Instituto de Microcirugía Ocular (IMO).

The first study evaluated the performance of a webcam-based eye-tracking system employing appearance-based gaze estimation to record saccadic eye movements, and compared it with the Tobii Pro Spectrum. Synchronous recordings demonstrated a close correspondence between systems, with a mean bias of 1.0° and standard deviation of 2.15° in saccade amplitude. The mean absolute error ranged between 0.9° and 2.3°, with the lowest errors observed for the smallest saccades, while larger movements exhibited slightly higher variability.

The second study was a preliminary investigation in paediatric amblyopia, using a head-mounted eye tracker (Pupil Core, Pupil Labs) to record eye movements during a naturalistic visual search task. The approach proved feasible for collecting data in ecological conditions and provided initial insights into the oculomotor behaviour of amblyopic children.

Together, these studies validate the feasibility of affordable eye-tracking technologies for both laboratory and clinical contexts. Future work will extend the analysis to vergence movements and fixation stability, supporting the development of accessible diagnostic tools for early detection of visual and oculomotor disorders.

REFERENCES

[1] Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

[2] Cheng, Y., Wang, H., Bao, Y., Lu, F. (2024). Appearance-Based Gaze Estimation With Deep Learning: A Review and Benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46, 7509–7528.

[3] Al-Nuimi, A. M., & Mohammed, G. J. (2021, August). Face direction estimation based on mediapipe landmarks. In *2021 7th International Conference on Contemporary Information Technology and Mathematics (ICCITM)* (pp. 185-190). IEEE.

[4] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M. (2019, June). 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)* (Vol. 2019).

[5] Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods*, *42*(1), 188-204.

[6] Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*(9), 1035–1045.

[7] Instituto de Microcirugía Ocular. (2025). Eye-tracking records of patients with visual disorders [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17046553