PopEYE - Infrared Ocular Image Dataset for Eye State and Gaze-Direction Classification

Giovanni Gibertoni*[1,2], Guido Borghi[3], and Luigi Rovati[1]

[1] *Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy*
[2] *Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, Modena, Italy*
[3] *Department of Computer Science and Engineering, University of Bologna, Bologna, Italy*

February 2026

Abstract

The PopEYE dataset comprises 14,976 near-infrared (NIR) images of the human eye, designed to support the development and benchmarking of computer vision algorithms for eye-state detection and coarse gaze-direction classification. All images are provided in 8-bit grayscale PNG format with a fixed resolution of 772 × 520 pixels. Data were acquired using the OptoCheck system, a modular ophthalmic imaging architecture comprising a board-level CMOS camera equipped with a 16 mm S-mount lens. The setup ensures stable 850 nm infrared illumination and high-contrast visualization of the anterior eye segment, including the pupil, iris, and eyelids. The dataset is manually annotated into six mutually exclusive classes: *correct* (8,160 images), *closed* (1,790 images), *up* (1,379 images), *down* (1,015 images), *left* (1,296 images), and *right* (1,336 images). This resource supports the training of machine learning models for real-time eye monitoring and positioning assessment in ophthalmic measurement systems.

Keywords: infrared eye imaging, ocular image dataset, gaze direction classification, eye state detection, near-infrared imaging, computer vision, machine learning, biomedical imaging

1 PopEYE Dataset

This document provides a technical description of the OptoCheck instrument setup and the acquisition procedure used to generate the PopEYE dataset. Section 1.2 details the hardware configuration of the acquisition device, while Section 1.3 illustrates the experimental methodology and the manual annotation process.

1.1 Dataset release and access

The PopEYE dataset is publicly available on Zenodo. The Concept DOI (latest version) is:

`10.5281/zenodo.18430186`.

A version-specific DOI is also provided by Zenodo for reproducibility (see the Zenodo record).

1.2 Acquisition Setup

The images included in the PopEYE dataset were acquired using the *OptoCheck* system [1], a compact near-infrared (NIR) ocular imaging module designed for real-time eye-state monitoring and positioning assessment. The hardware architecture, detailed in Figure 1, employs a modular design that ensures compatibility with various ophthalmic instruments by separating the NIR acquisition path from the visible stimulation path.

The illumination is provided by eight infrared LEDs (850 nm SFH4250, ams-OSRAM AG, Premstätten, Austria) arranged circularly around the optical axis to provide uniform lighting of the anterior eye

*Corresponding author: giovanni.gibertoni@unimore.it

segment. The acquisition core consists of a high-resolution board-level CMOS camera (mvBlueFOX3-3M, Matrix Vision GmbH, Germany) a 16 mm S-mount wide-angle lens. To ensure stable image contrast and suppress visible light interference during examinations, a 12 mm infrared long-pass filter ($\lambda_{\mathrm{cutoff}} = 750\,\mathrm{nm}$) is mounted directly in front of the camera sensor.

The camera is positioned orthogonally to the main optical axis through a 45° dichroic mirror (DMSP750B, Thorlabs, Newton, NJ, USA). This component is essential as it reflects the NIR pupil image toward the sensor while allowing the transmission of visible light (yellow dashed line in Figure 1), thus enabling simultaneous eye stimulation or clinical inspection without compromising the original functionality of the host ophthalmic device.

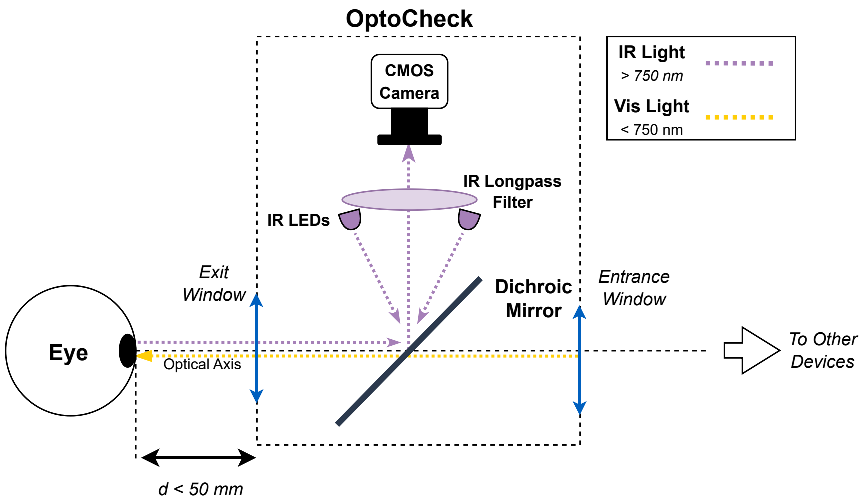


Figure 1: Optical diagram of the OptoCheck system. The infrared (IR) path (dashed purple line) facilitates eye monitoring, while the visible path (dashed yellow line) allows for direct stimulation or inspection. The dichroic mirror separates the two paths, ensuring modularity and clinical compatibility.

The system is housed in a modular optomechanical enclosure with a compact footprint of approximately 50 mm × 50 mm × 50 mm. The design features standard C-MOUNT internal threading and is compatible with SM1A10 adapters (Thorlabs), allowing for straightforward integration into the optical paths of existing devices such as slit lamps or OCT scanners. The mechanical layout, shown in Figure 2, supports rapid prototyping via FDM 3D printing, making the system highly adaptable to different professional optical interfaces. This modular design was successfully employed and validated in the development of a simple Maxwellian-view optical system designed to investigate photoreceptor contributions to the pupillary light reflex [2].

1.3 Acquisition Procedure

The images used for this dataset were collected from both the left and right eyes of 22 subjects (6 female and 16 male), with ages between 20 and 38 years old. The system was set to acquire continuous monochromatic 8-bit images at 50 Hz. Output frames recorded from the sensor (SONY IMX178 CMOS) were resized and stored as uncompressed video files with output resolution 772 × 520 pixels (width and height, respectively).

During acquisition, subjects were instructed to observe a circular light target of fixed intensity and, under voice command, move the gaze in the four cardinal directions: right, down, left, and up. Natural blinks were recorded, as well as a controlled closure task (closing the eye completely and slowly opening the lids). Individual frames were extracted from the saved video files with temporal sub-sampling at 10 Hz, reducing nearly identical frames.

Each frame was manually labeled into one of six classes: `correct`, `closed`, `right`, `down`, `left`, and `up`. Example images for each class are shown in Figure 3. As shown, every image contains the eyeball as seen through lens **L2** (Figure 1). The region outside the lens area is dark because it is not illuminated by the infrared LEDs.

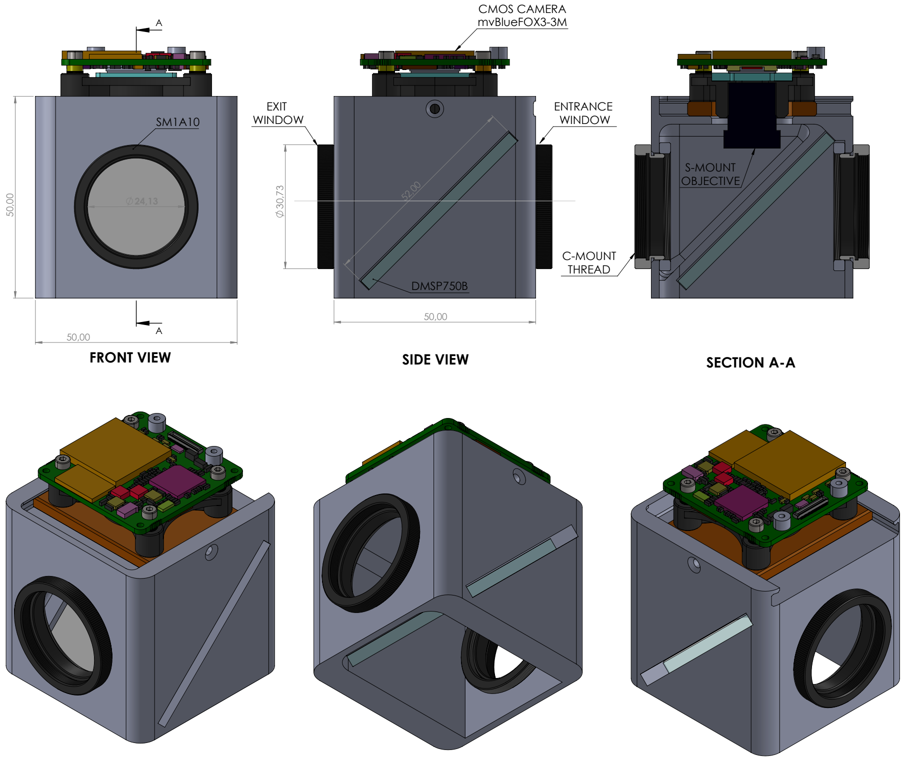
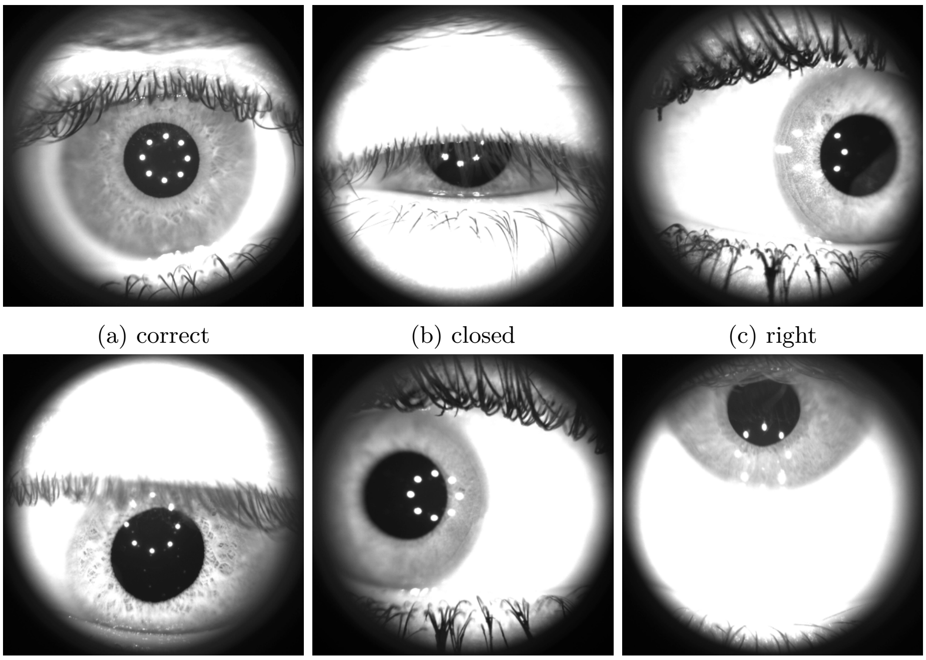


Figure 2: Modular integration design for the mvBlueFOX3-3M camera. The enclosure includes C-mount threads for easy attachment, entrance/exit windows, and internal housing for the DMSP750B dichroic mirror and S-mount objective.



(a) correct (b) closed (c) right

(d) down (e) left (f) up

Figure 3: Example images from the PopEYE dataset. The images are labeled into six categories: correct, closed, right, down, left, and up. Typical features include iris texture, eyelashes/eyelids, and specular highlights due to infrared illumination.

Manual classification was performed by the same operator, supported by fixed visual aids overlaid on the images. In particular, each frame was cropped to a square-shaped image (520×520 pixels) and divided into four zones obtained by two diagonals. A circular area (radius = 69 pixels) defines the *central region* used for `correct` images. The central region size was chosen to avoid strong aberrations and non-uniform illumination in the analyzed area.

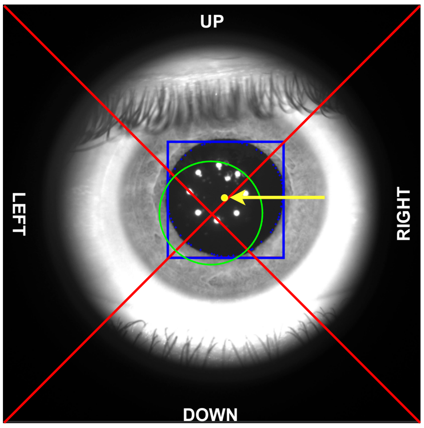


Figure 4: Visual aids used for manual annotation. The diagonals define the four cardinal regions (up, right, down, left). The green circle indicates the central region used for `correct` images.

The operator assigns one of the six labels based on the pupil center location and eyelid occlusion:

1. **Correct**: the pupil center is inside the central region (edge included), and the pupil area is uncovered (>90% visible);
2. **Closed**: the pupil center is inside the central region (edge included) and the pupil is significantly covered (>10% covered), or the pupil is not visible/detectable by the operator;
3. **Up, right, down, or left**: the pupil center is outside the central region, and the pupil is visible or partially occluded (>15% visible); the class depends on which triangular region contains the pupil center.

The final dataset contains 14,976 images and presents an unbalanced class distribution, which can affect model training and evaluation. The dataset is also used in previous works focusing on eye-image classification and real-time eye position monitoring [3, 1].

Table 1: PopEYE dataset composition across the six annotated classes.

Class	**N° Images**
Correct	8160
Closed	1790
Right	1336
Left	1296
Up	1379
Down	1015
Total	14976

References

[1] G. Gibertoni and L. Rovati, "Enhancing ophthalmic examinations with real-time eye position monitoring using SVM," in *Ophthalmic Technologies XXXV*, Y. K. Tao, D. X. Hammer, and D. Nankivil, Eds. San Francisco, United States: SPIE, Mar. 2025, p. 48.

[2] G. Gibertoni, V. D. Pinto, S. Cattini, F. Tramarin, M. Geiser, and L. Rovati, "A simple Maxwellian optical system to investigate the photoreceptors contribution to pupillary light reflex," in *Ophthalmic Technologies XXXII*, vol. 11941. SPIE, Mar. 2022, pp. 52–60.

[3] G. Gibertoni, G. Borghi, and L. Rovati, "Vision-Based Eye Image Classification for Ophthalmic Measurement Systems," *Sensors*, vol. 23, no. 1, p. 386, Jan. 2023.