GlobeMetrics: A Healthcare Framework for Video Based Saccade Characterization

Debanga Raj Neog †
Mehta Family School of DS & AI
Indian Institute of Technology Guwahati
Guwahati-781039, Assam, India
dneog@iitg.ac.in

Pharvesh Salman Choudhary *
Indian Institute of Technology Guwahati
Guwahati-781039, Assam, India
pharvesh@iitg.ac.in

Mayurdeep Pathak *
Indian Institute of Science
Bangalore-560012, Karnataka, India
mayurdeepp@iisc.ac.in

***Abstract*—Eye movement analysis is extensively utilized in understanding mechanisms governing perception, cognition, and action and proves valuable in exploring neurological and neurodegenerative diseases. This paper introduces GlobeMetrics, a healthcare application designed for ocular analysis from video data. Our emphasis in this study is on estimating the saccadic profile of the subjects utilizing the GlobeMetrics framework. The framework includes a setup for data recording, an appearance-based gaze estimation system, a module for analyzing gaze data specifically for saccade analysis, and an interactive GUI application that interfaces with each aspect of the framework. The proposed gaze estimation network consists of a convolutional neural network (CNN) based segmentation and regression network that maps input frames to gaze points. The proposed gaze estimation architecture achieves a prediction error of 0.467± 0.133 cm on our database. Additionally, the segmentation network attains mean IOUs of 95.19 and 97.39 for sclera and iris, respectively. Our proposed framework, GlobeMetrics, offers an interactive platform for conducting ocular analysis in clinical settings. This application seamlessly integrates data recording, stimulus generation, database management, and data analysis within a unified framework. The overall framework is accurate, robust, and generalizes well to new subjects.**

***Index Terms*—eye movement, saccade analysis, healthcare application, convolutional neural network**

I. Introduction

Gaze tracking finds extensive application in the examination of perception, cognition, and action mechanisms [1] and is useful for investigating neurological and neurodegenerative diseases [2]. Analysis of eye movement patterns, especially rapid shifts in gaze, also known as saccadic eye movements, is often crucial for non-invasive diagnosis of neurocognitive diseases [3]. In clinical practice, saccade analysis often involves assessment of eye movements to a visual stimulus. This stimulus may appear randomly or in a predetermined sequence, requiring subjects to quickly and accurately shift their gaze. By examining patterns, velocities, and accuracies of these eye movements, clinicians can identify markers of cognitive health.

Earlier gaze analysis techniques relied on simple visual observation, often necessitating the use of specks or eye-cups placed over the eyeball. As eye movement research has advanced, non-intrusive methods like video-based eye-gaze tracking, which relies on video cameras, have become increasingly popular. In particular, the appearance-based gaze estimation method directly estimates gaze direction through eye images. This approach eliminates the necessity for specialized hardware that can be restrictive and expensive. Recently, many works have incorporated appearance-based methods to a broad range of gaze-related tasks, including emotion analysis [4], ASD diagnosis [5], and visual attention detection [6].

In this work, we developed a gaze recording setup along with an interactive application designed specifically for clinicians. The application produces visual stimuli on a screen, and the associated eye movements are captured using a camera placed beneath the display. Furthermore, we developed an automated processing pipeline to directly analyse saccades from video data. Initially, a segmentation map is obtained for both the sclera and pupil. This segmentation map is then utilised to create a mapping for acquiring the 2D gaze point, which, in turn, is employed for saccade analysis. The overall system allows for the recording and analysing of subjects' eye movements, providing an interactive graphical interface for evaluating ocular movement patterns.

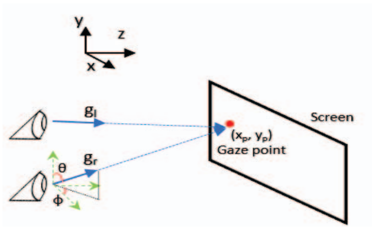


Fig. 1. Gaze direction corresponding to a gaze point on screen.

Eye movements allow for the selection and processing of visual information from our environment [7]. Broadly, eye movements can be classified into fixations, smooth pursuit, and saccades. Fixations refers to a state in which the eyes are relatively still and remain focused on a particular point. Smooth pursuit involves a fluent eye motion smoothly following a visual stimulus in motion, and saccades refer to rapid transitions from one fixation point to another [8]. These eye movements can be estimated from gaze data, which is generally recorded

*These authors contributed equally to this work
† Corresponding author

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00125

using some eye-tracking method. The vector representing gaze data can take two forms: a 2D Cartesian point for spatial location on the screen or a 2D polar angle for direction and magnitude. The analysis of saccades has proven crucial for distinguishing and monitoring neurocognitive diseases [3]. Saccadic eye movements are assessed using spatiotemporal and kinematic variables like velocity, acceleration, frequency, timing, and duration. For the analysis of saccades, the first step involves eye detection in the video frame, which is then utilized to estimate the gaze point at that instant of time. Gaze estimation, especially appearance-based methods [9], has drawn considerable interest in a variety of research fields and applications. Such an approach directly predicts gaze direction or gaze point (see Fig. 1) from eye images [10]. In general, these techniques leverage CNNs to acquire hierarchical features from eye images, employing them to estimate gaze direction [11], [12]. In this paper, we initially acquire segmentation masks for the sclera and iris of both eyes utilizing a CNN-based architecture, specifically UNet, in light of the successes of UNet architecture in many medical imaging applications [13], [14] and computer vision tasks [15], [16]. Subsequently, we estimate the gaze point using a lightweight CNN network.

Many software packages such as SMI BeGaze [17], Tobii Pro Lab [18], and imotions [19] have been utilised in literature to study and analyse eye movements. These provide solutions to eye tracking, offering in-depth insights into visual attention patterns across a broad spectrum of application domains. While such solutions provide a holistic approach to eye analysis, they can be complex and expensive, with a steep learning curve for less experienced professionals. In this paper, we propose to provide a lightweight and open-source platform for ocular studies, emphasising its simplicity to cater to the needs of healthcare professionals.

In summary, the main contributions of this paper are listed as follows:

- We propose GlobeMetrics, a healthcare framework for ocular analysis.
- We introduce the characterization of saccades from gaze data as a module within the framework. To accomplish this, we utilize an appearance-based deep learning method.
- We have compiled a database "GlobeMetrics Dataset" consisting of data from eight subjects. Subsequently, we conducted thorough experiments on this dataset for the purpose of saccade characterization.

II. Methodology

GlobeMetrics Framework is a software package for ocular research with a focus on studying patient responses to visual stimuli. It comprises of: (i) Recording setup and calibration, (ii) Visual stimuli apps, (iii) Synchronization, (iv) Preprocessing, (v) Experimental modules, and (vi) GUI interface. Our software package provides an interactive user interface, which allows for patient-wise data collection, including grouping patients based on health records. We introduce the characterization of saccades from gaze data as a module within the framework, comprising three primary stages: Eye segmentation, Gaze prediction, and Saccade analysis. Our objective is to assess an individual's saccadic profile through video data analysis while also developing a user-friendly graphical interface for clinicians. The overview has been presented in Fig. 2. Other ocular experiments can be added as experiment modules. GlobeMetrics allows the combination of different experiment modules and the export of results to data visualization software.

A. *GlobeMetrics Framework*

1) Recording Setup and Calibration: The experimental setup comprises a desk-mounted chin rest for head stability, with a 27-inch LCD screen positioned 60 cm in front. A high fps camera, capable of recording 120 fps or higher, is centrally located just below the screen to capture each session (see Fig. 2). The camera must also have audio recording capabilities, as the audio track will be utilized to synchronize the footage with the data from the visual stimuli app. It is ensured that the recording is done with minimal background noise. Measurements, such as the height from the desk to the observer's eye, the height from the desk to the screen's bottom edge, and the height and width of the screen, are manually recorded. For determining the perceived centre of the screen, the visual target apps include a calibration step; therein, the observer looks straight at the screen, and the perceived centre is adjusted using the mouse pointer and is recorded. The patient is instructed to place their head on the chin rest in a relaxed position. To mitigate eye fatigue during extended experiment sessions, the visual stimuli apps incorporate interval breaks.

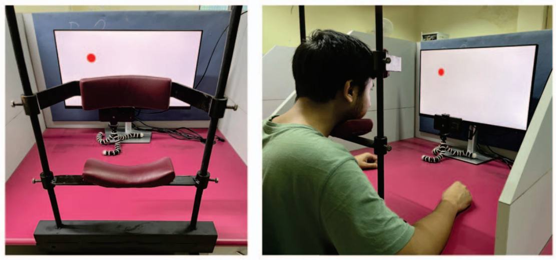


Fig. 2. Experimental Setup: A participant seated comfortably at the desk-mounted chin rest, engaged in the eye-tracking experiment. The 27-inch LCD screen, positioned 60 cm in front, and the high-speed camera below the screen capture precise eye movements. Calibration ensures accurate data synchronization with visual stimuli, providing a controlled environment for the study.

2) Generating Visual Stimuli: For saccade characterization, the visual targets app contains a red circular target on a white background, displaying these targets sequentially after a pause of a few seconds (see Fig. 3). Additionally, the app provides options for generating various sequential patterns, such as moving the target serially across the full grid, following a diamond pattern on the grid, employing a completely random sequence, or adopting a random sequence with a constraint within a specified radius from the last displayed position. The app includes calibration tools to record measurements, including the perceived centre of the screen, and an option to include interval breaks. The target coordinates are converted to polar coordinates, theta and phi, using the calibration data. At the start of each session, the app produces a buzzer sound, which is later used to synchronize the recorded footage. At the end of each session, the app outputs a JSON file, which contains a dictionary of timestamps and corresponding target polar coordinates, as well as the measurements and calibration data. Together, this JSON file and the recorded footage are the inputs to the preprocessing pipeline.

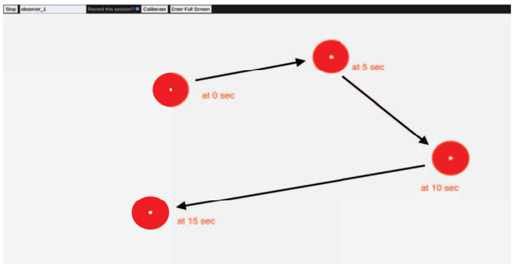


Fig. 3. Illustration of the red circular target's movement on the app screen, showcasing key features for precise saccade characterization in eye-tracking experiments.

3) Synchronization: Due to a disparity in the system clocks between the recording camera and the system running GlobeMetrics, there arises a necessity for synchronization between the two clocks. Saccade analysis requires precise timing at the millisecond level, with the average duration of a saccade being 20–200 ms, depending on its amplitude. Therefore, the GlobeMetrics package is equipped with a synchronization tool to address this requirement. The synchronization tool can be found under the Capture tab (see Fig. 4). Users are required to upload recorded footage and the corresponding JSON file from the visual target app for preprocessing. As mentioned before, the visual targets app generates three consecutive beep sounds, and the camera captures the video footage along with the audio. The synchronization tool extracts the audio channel from the captured footage and displays an audio waveform and an interactive slider that controls a cue point on the audio track. The three beeps in the recording align with three clear peaks in the early part of the waveform, easily identifiable through visual inspection. The user moves the slider to accurately position the cue point at the conclusion of the third peak. The tool computes the time difference between the system clock and the camera clock and designates the frame at the instant of the 3rd beep with the zeroth timestamp. Subsequently, all frames are assigned timestamps accordingly. This synchronization step between the video footage and the system clock is necessary to give meaningful utility to the timestamps in the JSON file.

4) Preprocessing: After the footage has been synchronized, the preprocessing pipeline then extracts all frames and per-

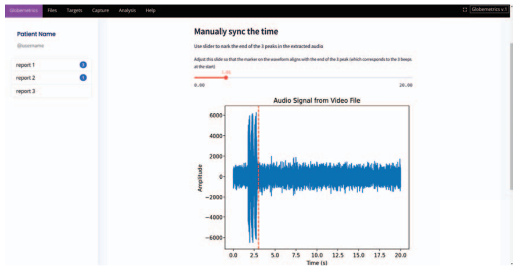
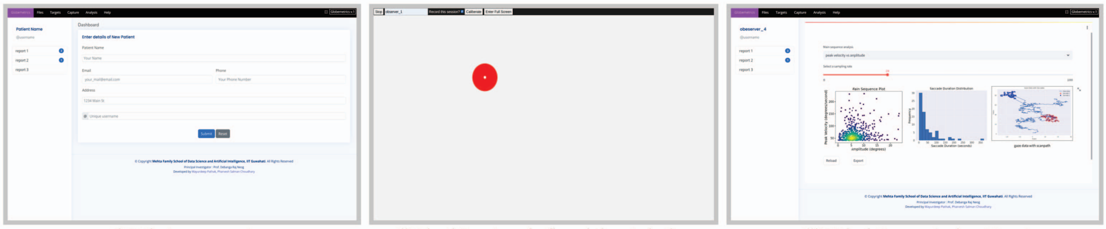


Fig. 4. Interactive slider aligns with audio peaks for clock synchronization in GlobeMetrics, ensuring precise timestamps for saccade analysis.

forms facial landmark detection using the Dlib 68-pt landmark detector [20] to obtain a bounding box around the eyes. It then crops in on the bounding box for both the Left and Right eye. All images are named descriptively. The frames that match with the timestamps in the JSON file represent the key moments when the eye fixates on the target, while the frames in between timestamps capture the natural movement of the eye before fixating on the next target. These frames between targets are of great interest as they correspond to saccadic movements or smooth pursuits. The resulting dataset is a sequence of eye images with the target coordinates serving as labels.

5) Experiment Modules: GlobeMetrics is designed to accommodate a range of experiment modules, each featuring a unique pipeline encompassing preprocessing stages, models, and post-processing with visualization steps. The outcomes of these experiments are exportable to other data visualization software. We explore a novel eye saccade analysis module in detail.

6) GUI Interface: The GlobeMetrics software package comes with an interactive user interface. The GUI comprises different tabs. Patient Management tab allows managing patient and health records (see Fig. 5). It opens up a section that allows viewing, creation, deletion, and modification of patients and the systematic data collection of patient health records. A section within the Patient Management tab displays a list of experiments that have been performed, its associated processed data, and outputs of the experiment modules. Next, the Visual Targets tab contains different visual target apps. Each app contains its own set of options that are specific to the task. For saccade characterization, these apps include tasks such as moving the target serially across the full grid, following a diamond pattern on the grid, and employing a completely random sequence, among others. All visual targets app includes a calibration tool to record measurements, including the perceived centre of the screen. The Capture tab contains tools related to capturing and preprocessing. The upload sub-menu under the Capture tab allows the user to upload the recorded footage and the JSON output file from the visual target app. The synchronize sub-menu under the Capture tab brings up the interactive audio waveform tool, enabling users to set the cue point. Various parameters that affect the



(i) Patient management (ii) Visual Target app for Saccad characterization (iii) Upload, Process, Analyze, Export

Fig. 5. Patient management in the GlobeMetrics Framework involves an interactive user interface for systematic data collection and grouping based on health records. Several types of visual target apps exist within the Targets tab, such as moving the target serially across the entire grid, following a diamond pattern on the grid, and employing a completely random sequence. These apps include a calibration tool to record measurements, including the perceived centre of the screen. The Capture tab contains tools that allow users to upload the recorded footage and the corresponding JSON file from the visual target app. The Analyze tab contains experiment modules such as the eye saccade characterization module.

preprocessing stage can be adjusted within the preprocessing section. The Analyze tab contains various experiment modules, including a built-in eye saccade characterization module. Custom ocular experiments can be added as experiment modules to GlobeMetrics. The outputs from the experiment modules are organized patient-wise, and the GUI allows this data to be exported to data visualization software.

B. *Saccade Analysis Module*

The module provides a comprehensive examination of the saccadic profile of individuals through a detailed analysis of video data. Fig. 6 gives a schematic overview of the saccade analysis pipeline. The process of saccade determination begins with the estimation of gaze, involving the extraction of eye position from each frame. Subsequently, the gaze data is analyzed to identify abrupt eye movements within the video sequence. For the estimation of gaze, the objective is to regress the 2D gaze point by learning features from a series of eye images. The overall structure consists of two high-level building blocks: Eye segmentation network and a gaze prediction network. The segmentation network delineates the sclera and iris regions, and the gaze prediction network produces the gaze direction from the sequence of segmentation masks.

a) Eye Segmentation: The segmentation network is based on U-Net architecture [21]. It employs a symmetrical encoder-decoder structure to learn contextual information with high-level semantics. The encoder consists of convolutional blocks, with each block containing two 3×3 convolution layers activated with ReLU. A Max pooling layer with a kernel size of 2 is applied after each block, which essentially downsamples the feature representations in successive convolutional blocks while the number of feature channels is gradually increased. The Decoder, having a similar architecture to the encoder, takes the encoded feature representation and reconstructs the spatial information through upsampling and skip connections.

The model receives cropped eye patches as input and generates segmentation masks with three distinct classes: background, sclera, and iris. Given an input frame image, $I \in \mathcal{R}^{H\times W\times 3}$ with corresponding labels representing the three classes. For the task of eye segmentation, the network is trained to minimize cross-entropy loss, which is given by Eq. 1.

$$\mathcal{L}_s = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{n,c}\log P_s \tag{1}$$

Where $\mathcal{L}_s$ represents the segmentation loss, N represents the total number of input images in a given batch, C is the total segmentation classes, and P_s represents the probability scores of segmentation by the network. Both the left and right eyes utilize the same segmentation network. The resulting segmentation mask is then input into the gaze regression network for further processing.

b) Gaze Regression: The gaze regression network comprises of two branches of CNN corresponding to the left and right eye, one for the left eye and one for the right eye. Each branch incorporates two stacked convolutional layers with ReLU activation and a stride of 2, followed by a subsequent max pooling layer. The features extracted by both branches for the eyes are subsequently fused and input into a fully connected network for regression. It utilizes the segmentation mask of the eyes to regress to the gaze point, employing the Huber loss defined in Eq. 2.

$$\mathcal{L}_r(\hat{g}_p, g_p) = \begin{cases} 0.5(\hat{g}_p - g_p)^2, & \text{if } |\hat{g}_p - g_p| < \beta. \\ \beta * (|\hat{g}_p - g_p| - 0.5 * \beta), & \text{otherwise} \end{cases} \tag{2}$$

Where $\mathcal{L}_r$ represents the regression loss, $\hat{g}_p$ refers to the predicted gaze point for the corresponding ground truth of $\hat{g}_p$. β represents the threshold governing the transition between the two components of the loss; for this work, β was set to 1.

III. Saccade Analysis: Results and Discussion

In this section, we first evaluate the performance of our segmentation network.

A. *Database*

1) UBIPr Dataset: UBIPr dataset [22] comprises 10,254 images taken from 261 subjects, capturing diverse lighting conditions, varying distances, and common occlusions encountered in everyday environments. The dataset includes

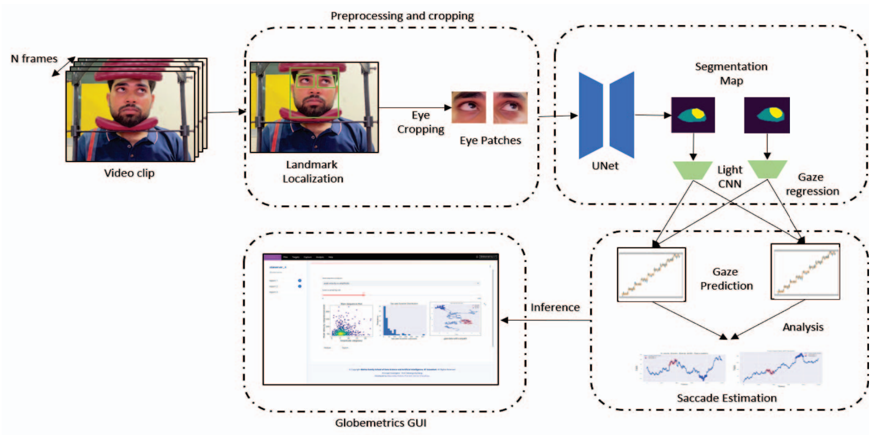


Fig. 6. Schematic overview of the proposed approach.

segmentation masks that outline the skin, eyebrows, sclera, and iris. We employ this database exclusively to pre-train our segmentation model.

2) GlobeMetrics Dataset: A database contains a set of video clips of 8 healthy subjects. The videos are recorded with a fixed head motion under laboratory conditions. Participants were instructed to maintain their gaze on these red dots as they appeared sequentially at ten random positions. Following this sequence, a 15-second rest period was recorded. This entire process was repeated five times for each subject during a single session. For the purpose of calibration, an additional recording involving four subjects was conducted. In this configuration, the generation of dots was systematically adjusted in a linear manner across the screen. This was undertaken to capture all potential gaze movements achievable within the specified setup.

B. Preprocessing

Given the limited number of images in our calibration dataset, utilizing them directly for training the segmentation network could lead to overfitting the network. To overcome this, we employed data augmentation in the form of horizontal and vertical flips along with Gaussian blur.

C. Evaluation Metrics

To evaluate the performance of the segmentation model, we employ four common quantitative metrics, including Dice similarity coefficient (DSC), positive predictive value (PPV), Intersection over union (IOU) and Hausdorff distance (HD). For gaze estimation, we utilize the Euclidean distance in centimetres relative to the screen as the performance metric.

D. Eye Segmentation

For training, we first pre-train the network using images from the UBIPr Dataset consisting of 10,254 images. For subsequent downstream training, we employ the calibration set from the GlobeMetrics dataset, which comprises a total of 1152 images. The images from our calibration dataset are preprocessed using the preprocessing pipeline described in Section III-B. Table 1 provides the segmentation performance across the testing database. For sclera segmentation, the model achieved average DSC, HD, PPV and IOU for 95.12%, 3.910 mm, 95.71% and 95.19, respectively and similarly, for iris the model achieved an IOU of 97.39 and HD of 1.347 mm. These metrics collectively indicate a high level of accuracy and preciseness in delineating the sclera and iris region within the eye patches. Additionally, Fig. 7 visually presents segmentation maps for four subjects, providing a qualitative representation of the model's performance.

TABLE I
SEGMENTATION PERFORMANCE

Region	DSC(↑)	HD(↓)	PPV(↑)	mIOU(↑)
Sclera	0.951	3.910	95.71	0.952
Iris	0.971	1.347	97.46	0.974

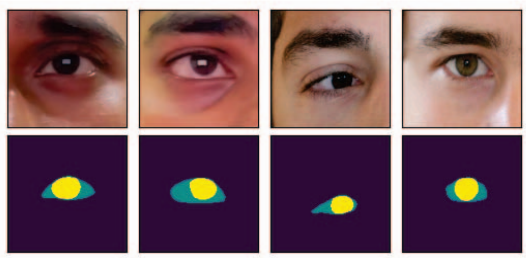


Fig. 7. Visualization of segmentation mask (bottom row) obtained by the segmentation network for the input eye patches (top row).

Table 2 presents the performance of gaze estimation. As can be observed that the proposed approach, incorporating segmentation maps of both sclera and iris, achieved the lowest estimation error with a mean Euclidean error of 0.46 ± 0.133 cm. The setup utilizing direct eye patches resulted in a slightly larger prediction error of 0.798 ± 0.410 cm. Configurations using segmentation maps of either sclera or iris exhibited the highest errors of 1.099 ± 1.164 cm and 0.956 ± 0.536 cm, indicating the importance of incorporating complementary information from both sclera and iris for accurate gaze estimation.

TABLE II
GAZE ESTIMATION ERRORS ON GLOBEMETRICS DATASET FOR DIFFERENT CONFIGURATIONS

Configuration	Euclidean error (cm)
Without segmentation	0.798 ± 0.410
Only sclera	1.099 ± 1.164
Only iris	0.956 ± 0.536
Both sclera andd iris	0.467 ± 0.133

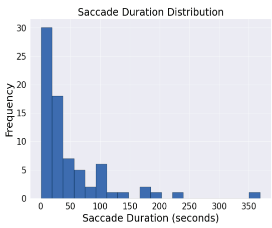


Fig. 8. Histogram of saccade distribution for a subject.

In Fig. 8, we showcase a saccade distribution plot, which represents the occurrence and duration of saccadic eye movements. It illustrates the frequency of saccades across a range of durations. Each bar on the plot represents a specific duration range, and the corresponding height or value indicates the number of saccades falling within that duration. It helps understand the temporal dynamics of eye movement behaviour and may reveal underlying conditions influencing saccadic patterns.

IV. Conclusion

In this study, we presented GlobeMetrics, a healthcare framework for ocular analysis. GlobeMetrics is developed specifically to aid clinicians and hospital administrators in conducting and evaluating ocular studies. GlobeMetrics features an easily accessible GUI via the web server. We demonstrated saccade characterization as an application within the framework. The framework serves as an integrated platform encompassing various functionalities, including data recording, health record management, analysis, and visualization. In the future, we aim to expand the GlobeMetrics framework by integrating additional application modules for ocular studies and as well as behavioural studies. In addition, we want to deploy our method as an adjunct to clinical decision support systems.

Acknowledgment

The project is supported by Science and Engineering Research Board Startup Research Grant (SRG/2022/001775).

References

[1] J. M. Henderson, "Human gaze control during real-world scene perception," *Trends in cognitive sciences*, vol. 7, no. 11, pp. 498–504, 2003.
[2] R. Andersson, L. Larsson, K. Holmqvist, M. Stridh, and M. Nyström, "One algorithm to rule them all? an evaluation and discussion of ten eye movement event-detection algorithms," *Behavior research methods*, vol. 49, pp. 616–637, 2017.
[3] H.-Y. Lai, G. Saavedra-Pena, C. G. Sodini, V. Sze, and T. Heldt, "Measuring saccade latency using smartphone cameras," *IEEE journal of biomedical and health informatics*, vol. 24, no. 3, pp. 885–897, 2019.
[4] S. D'Mello, A. Olney, C. Williams, and P. Hays, "Gaze tutor: A gaze-reactive intelligent tutoring system," *International Journal of human-computer studies*, vol. 70, no. 5, pp. 377–398, 2012.
[5] J. Li, Z. Chen, Y. Zhong, H.-K. Lam, J. Han, G. Ouyang, X. Li, and H. Liu, "Appearance-based gaze estimation for asd diagnosis," *IEEE Transactions on Cybernetics*, vol. 52, no. 7, pp. 6504–6517, 2022.
[6] H. Sattar, S. Muller, M. Fritz, and A. Bulling, "Prediction of search targets from fixations in open-world settings," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 981–990.
[7] J. M. Henderson, "Eye movements," *The Oxford handbook of cognitive psychology*, pp. 69–82, 2013.
[8] T. Santini, W. Fuhl, T. Kübler, and E. Kasneci, "Bayesian identification of fixations, saccades, and smooth pursuits," in *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, 2016, pp. 163–170.
[9] X. Wang, J. Zhang, H. Zhang, S. Zhao, and H. Liu, "Vision-based gaze estimation: a review," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 316–332, 2021.
[10] Y. Sugano, Y. Matsushita, and Y. Sato, "Appearance-based gaze estimation using visual saliency," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 2, pp. 329–341, 2012.
[11] K. Krafka, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, "Eye tracking for everyone," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2176–2184.
[12] D. Lian, L. Hu, W. Luo, Y. Xu, L. Duan, J. Yu, and S. Gao, "Multiview multitask gaze estimation with deep convolutional neural networks," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 10, pp. 3010–3023, 2018.
[13] K. He, C. Lian, B. Zhang, X. Zhang, X. Cao, D. Nie, Y. Gao, J. Zhang, and D. Shen, "Hf-unet: learning hierarchically inter-task relevance in multi-task u-net for accurate prostate segmentation in ct images," *IEEE transactions on medical imaging*, vol. 40, no. 8, pp. 2118–2128, 2021.
[14] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
[15] Z. Wang, J. Chai, and S. Xia, "Realtime and accurate 3d eye gaze capture with dcnn-based iris and pupil segmentation," *IEEE transactions on visualization and computer graphics*, vol. 27, no. 1, pp. 190–203, 2019.
[16] Y. Wang, J. Wang, and P. Guo, "Eye-unet: A unet-based network with attention mechanism for low-quality human eye image segmentation," *Signal, Image and Video Processing*, vol. 17, no. 4, pp. 1097–1103, 2023.
[17] "SMI BeGaze," [Accessed 24-12-2023]. [Online]. Available: https://gazeintelligence.com/
[18] "Tobii Pro Lab," [Accessed 24-12-2023]. [Online]. Available: https://www.tobii.com/products/software/behavior-research-software/tobii-pro-lab
[19] "iMotions," [Accessed 24-12-2023]. [Online]. Available: https://imotions.com/eye-tracking/
[20] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.
[21] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.
[22] C. N. Padole and H. Proenca, "Periocular recognition: Analysis of performance degradation factors," in *2012 5th IAPR international conference on biometrics (ICB)*. IEEE, 2012, pp. 439–445.