EV-Eye: Rethinking High-frequency Eye Tracking through the Lenses of Event Cameras

Guangrong Zhao[1],**Yurun Yang**[1], **Jingwei Liu**[1], **Ning Chen**[1],
Yiran Shen[1],[*] **Hongkai Wen**[2] **and Guohao Lan**[3]
[1]School of Software, Shandong University, China
[2]Department of Computer Science, University of Warwick, UK
[3]Department of Software Technology, Delft University of Technology, NL
{guangrong.zhao, yiran.shen}@sdu.edu.cn
hongkai.wen@warwick.ac.uk
g.lan@tudelft.nl

Abstract

In this paper, we present **EV-Eye**, a first-of-its-kind large-scale multimodal eye tracking dataset aimed at inspiring research on high-frequency eye/gaze tracking. EV-Eye utilizes the emerging bio-inspired event camera to capture independent pixel-level intensity changes induced by eye movements, achieving sub-microsecond latency. Our dataset was curated over two weeks and collected from 48 participants encompassing diverse genders and age groups. It comprises over **1.5 million** near-eye grayscale images and **2.7 billion** event samples generated by two DAVIS346 event cameras. Additionally, the dataset contains **675 thousand** scene images and **2.7 million** gaze references captured by a Tobii Pro Glasses 3 eye tracker for cross-modality validation. Compared with existing event-based high-frequency eye tracking datasets, our dataset is significantly larger in size, and the gaze references involve more natural and diverse eye movement patterns, i.e., fixation, saccade, and smooth pursuit. Alongside the event data, we also present a hybrid eye tracking method as a benchmark, which leverages both the near-eye grayscale images and event data for robust and high-frequency eye tracking. We show that our method achieves higher accuracy for both pupil and gaze estimation tasks compared to the existing solution.

1 Introduction

Eye tracking is a technique that continuously measures the movement of the eyes [1] and has shown great promise in a wide range of scientific fields [2] and everyday applications [3]. Traditional eye tracking systems in current mainstream leverage traditional CCD/CMOS cameras to capture the appearance of eyes for tracking [4, 5, 6, 7, 8]. Unfortunately, constrained by the frame rate and the limited bandwidth of CCD/CMOS camera, the tracking frequency of conventional eye tracking systems is often bounded by a few hundred hertz, e.g., the latest Tobii Pro Glasses 3 eye tracker [9] has a tracking frequency of 100Hz. Similarly, Pupil Labs [10] achieves up to 200Hz gaze estimation. Although such a modest frequency is enough for many common use cases, e.g., human-computer interaction [11, 12] and activity recognition [13, 14], it stands in the way of enabling *game-changing* applications that may require over kilohertz tracking frequency, such as the diagnosis of mental disorders [2, 15], and gaze-based user authentication [16, 17]. The peak angular speed of the human eye in saccade can reach up to $700^\circ/s$ [18] and the eye's acceleration is up to $24,000^\circ/s^2$ [19]. The high-frequency eye tracking can reveal latent information encoded in high-speed and irregular

[*] Corresponding Author

37th Conference on Neural Information Processing Systems (NeurIPS 2023) Track on Datasets and Benchmarks.

eye movements. For example, during the diagnosis of mental disorders, an increase in delays for visually guided saccades can help to characterize ADHD [20]. Unfortunately, achieving accurate eye tracking beyond kilohertz frequency requires a significant increase in camera bandwidth, which becomes a fundamental barrier for mainstream CCD/CMOS camera-based systems. Some high-end eye tracking systems (costing tens of thousands of US dollars), such as the EyeLink 1000 [21], can provide eye tracking at 1KHz by utilizing high-speed cameras. However, the high frame rate presents a considerable computational burden on downstream processing.

This challenge has inspired the use of the emerging bio-inspired dynamic vision sensor [22, 23], also known as the event camera, for eye tracking [24, 25]. In contrast to traditional CCD/CMOS cameras, which acquire information in a frame-based principle with a fixed frame rate, an event camera perceives the scene by capturing independent pixel-level light-intensity changes and producing asynchronous event streams that indicate the locations and polarity of the intensity changes. Due to its asynchronous nature and cost-saving readout, an event camera can achieve sub-microsecond latency [22]. Moreover, an event camera operates adaptively: the faster the targeted motion, the more events are generated per second, and vice versa. In a near-eye scenario, the light-intensity changes induced by eye movement are sparse in both time and space [25]. Thus, the event camera can adapt to the density of events based on the speed of eye movements and utilizes the camera bandwidth more efficiently than traditional cameras. These properties make event cameras ideal for high-frequency eye tracking.

While event-based eye tracking has demonstrated significant advantages over its frame-based counterpart, it is still in its early stage and requires substantial research efforts and public resources, such as datasets. First, similar to traditional camera-based eye tracking [26, 8, 27, 28], the success of event-based eye tracking heavily relies on the availability of large-scale event-based datasets. However, as detailed later in Section 3, the only existing dataset, EBVEYE [25], provides sparse gaze references with a limited range of eye movements (mainly fixation) collected from a limited number of subjects. This limitation poses a risk of high sensitivity to subject and tracking condition changes. Secondly, the state-of-the-art event-based eye tracking solution adopts the model-based approach [25], which is not robust to subject diversity and sensor noise. This work aims to shed light on the existing challenges and advances the field of event-based eye tracking by making the following contributions:

- We introduce the largest and most diverse multi-modal frame-event dataset for high-frequency eye tracking in the literature (over 170Gb in total).[2] To the best of our knowledge, ours is the largest multi-modal dataset aimed at high-frequency eye tracking in the literature.
- We propose a novel hybrid frame-event eye-tracking benchmarking approach tailored to the collected dataset, capable of tracking the pupil at a frequency up to 38.4KHz. As demonstrated using our dataset, the proposed approach significantly outperforms the existing solution [25] in both pupil and gaze estimation by a large margin.

2 Related Work

Applications of Event Camera. The methods of event-based applications normally start by converting the asynchronous event streams into formatted representations. For example, image-like representations convert event streams into event-based "images" by calculating the distributions of the events and their timestamps. Image-like representations have been widely used for object tracking [29], gait recognition [30], and optical flow estimation [31]. To better preserve the temporal information of event streams, graph-based representations [32, 33], point cloud-based representations [34] and voxel-grid-based representations [35] have been proposed. Convolutional neural networks, graph-based convolution, and PointNet [36] can be applied to the corresponding representations for various vision tasks, including object/ action classification [32], gait recognition [33], gesture recognition [34], frames interpolation [37] etc.

Conventional Eye Tracking. Current solutions for eye-tracking are either model-based [38, 39] or appearance-based [5, 26]. The model-based methods rely on corneal reflection or pupil shape to infer the Point of Gaze (PoG) from parametric features, such as pupil centers and iris. However,

[2] The EV-Eye dataset and the implementation of our benchmarking methodologies can be found at: https://github.com/Ningreka/EV-Eye.

these methods are sensitive to ambient light conditions and require high-resolution images to ensure good tracking accuracy. By contrast, appearance-based methods leverage neural networks to map images of eyes to the PoG, which have shown significant improvement in gaze estimation accuracy and system robustness [5, 26]. However, limited by the frame rate of conventional cameras, the eye tracking frequency of existing methods is usually constrained to around 200Hz. In addition, existing conventional eye-tracking datasets rarely contain videos, and when they do, they are only of participants gazing at fixed points [40]. While there are some datasets including natural eye movement patterns such as saccades, fixations, and smooth pursuits [41, 42, 43], they do not provide data in high temporal resolution and dense gaze references.

Event-based Eye Tracking. There are also works that leverage event streams for eye tracking. For instance, Feng et al. [44] propose an event-driven eye segmentation method for eye tracking. However, the event streams were generated by the event simulator rather than collected in real-world experiments. Ryan et al. [45] propose an event-based neural network for real-time face and eye detection. However, their focus is primarily on the detection and localization of eyes, rather than pupil tracking. More recently, Stoffregen et al. [24] introduce a coded differential lighting method to track eye movement at a frequency of 1KHz. Angelopoulos et al. [25] is the most relevant work to ours, in which the authors propose a hybrid frame-event approach that achieves up to 10KHz eye tracking. However, our method exhibits greater robustness across user diversity and achieves higher accuracy in eye tracking. Additionally, our dataset consists of 48 subjects, which is twice that of EBVEYE. Lastly, our dataset includes dense gaze references for multiple eye movements i.e., fixation, saccade, and smooth pursuit (the definition can be found in Section 3.2 Eye Tracking Model), while EBVEYE only provides sparse references during fixation status.

3 The EV-Eye Dataset

3.1 Dynamic Vision Sensors

To make the paper self-contained, we will brief the principal background of event cameras. Unlike traditional RGB cameras, event cameras do not produce synchronous video frames at a fixed rate, but asynchronous event streams. Specifically, pixels of the event camera work independently, to detect the change in the light intensity of the scene as,

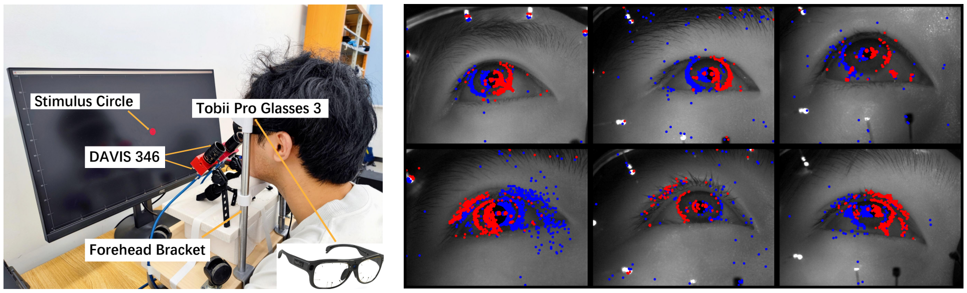
$$|log\, I(x, y, t_{now}) - log\, I(x, y, t_{previous})| < C \tag{1}$$

where $I(x, y, t)$ is the intensity value of pixel (x, y) at time t. When the change of intensity at the pixel is over the threshold C, an event will be released immediately. An event stream is a collection of events over time and is represented as a stream of quadruplet $\{x, y, t, p\}$. When the event corresponds to a positive change, the polarity p is $+1$ otherwise it is -1. Compared with traditional RGB cameras, event cameras possess several unique characteristics. As an event is launched as soon as a change is detected without global synchronization, the event streams are high in temporal resolution and low in response latency (in the order of microseconds). Event cameras save sensing energy and bandwidth as they produce events only when changes are detected. The high dynamic range (140 dB vs. 60 dB of traditional RGB cameras) enables them to work greatly under challenging lighting conditions. These characteristics make event cameras have great potential for high-speed motion capture and working on resource-constrained devices.

3.2 Preliminaries

Eye Tracking Model. In this paper, we consider a common eye-tracking model. It starts by localizing the centroids of the pupil area in the image domain. Subsequently, the corresponding Point of Gazes (PoGs) are determined, representing where the users are looking in realworld, through polynomial regression [46]. Eye movements are various and can be vastly categorized into three major types in natural viewing scenarios [41], i.e., fixation, saccade, and smooth pursuit. **fixation** is the status of eyes when users stare at a fixed point; **saccade** represents the fast eye movement towards a point of interest; **smooth pursuit** is the eye movement when the users follow a smoothly moving object in a predictable route.

Eye Tracking Modalities. As shown in Figure 1(a), EV-Eye adopts three different sensing modalities. These modalities include near-eye grayscale images and event streams captured by two sets of DAVIS346 event cameras [47], and gaze references provided by Tobii Pro Glasses 3 [9].



(a) Collection setup

(b) Collected samples

Figure 1: **Illustration of dataset collection setup (a) and collected samples (b)**. The near-eye grayscale images overlaid with 40ms of events.

Event Streams: The event streams are collected by two sets of DAVIS346 event cameras with a resolution of 346×240. They comprise positive and negative events triggered by intensity changes resulting from eye movements, winks, and other subtle motions. This setup provides high temporal resolution, enabling high-frequency eye tracking.

Grayscale Images: The DAVIS346 cameras also record near-eye grayscale image sequences at a frame rate of 25fps. They complement the event streams by providing rich semantic information about the eyes and facilitating robust pupil segmentation in the image domain.

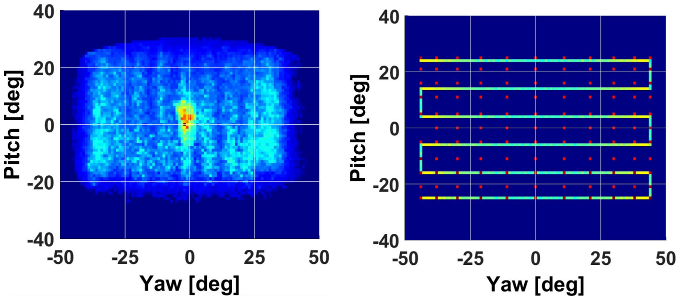
Gaze References: As obtaining ground truth for Points of Gaze (PoGs) during eye movement is not feasible, we adopt a commercialized device, Tobii Pro Glasses 3 [9], to acquire gaze references. It provides the PoGs and pupil diameters of the users at 100Hz. Tobii Glasses have a field of view (FoV) of 95°×63°, and it can achieve 0.6 angular error in gaze estimation in the central area of the virtual screen.

3.3 Dataset Curation

Data Collection Setup. The setup of the data collection is illustrated in Figure 1(a). During the collection, the subject's head is securely fixed on an ophthalmic headrest to prevent any unintended movement. Two DAVIS346 event cameras are positioned closely to record the movement of each eye. The aperture and exposure time of DAVIS346 are adjusted to prevent overexposure. We leverage a 32-inch monitor with a resolution of 1920×1080 to display the visual stimulus to guide the gaze movement of the subject. The distance between the monitor and the subject is about 33cm, which leads to a field of view (FoV) of 95°×63°. The visual stimulus is a solid red circle displayed on the monitor, with a diameter of 60px. In our setup, the subject wears Tobii Glasses to obtain reference PoGs. Additionally, the scene camera of Tobii records scene images throughout the experiments. All devices are synchronized before each data collection session.

Acquisition Protocol. We recruited 48 participants (28 male and 20 female)[3] aged between 21 and 35 years. They have various vision corrections, and each of them participates in four sessions of data collection. For the first two sessions, the screen is equally divided into 11 × 11 grid, resulting in 121 squared blocks. The stimulus appears in the center of one block in a random sequence with each appearance lasting for 1.5s. The subject quickly finds the stimulus and focuses on it when it is stationary. The random sequence for each subject is the same, and each session ends when all blocks are visited. These two sessions allow us to trigger and capture both saccade and fixation states of the eye movement. For the last two sessions, we ask the subject to fixate on a smoothly moving stimulus across the screen. The stimulus first moves horizontally from the lower-right to the upper-left corner of the screen then it moves vertically from the upper-left to the lower-right. The movement of the stimulus is a predictable squared wave trajectory. The space between the adjacent horizontal trajectories is 54px, which results in 20 horizontal paths. Similarly, the space between adjacent vertical trajectories is 96px and leads to 20 vertical paths. The moving pattern of the stimulus allows us to record eye movement in smooth pursuit.

[3] Our data collection was approved by our local IRB council.



(a) EV-Eye (ours) (b) EBVEYE [25]

Figure 2: **Distributions of the gaze references of datasets, the gaze references provided in our dataset are significantly more dense and involve all states.**

Data Annotation. We leverage the VGG Image Annotator [48] to label the pupil region of 9,011 near-eye images selected uniformly across the whole image dataset. Normally, the pupil region is regarded as an ellipse. Therefore, we label the region by adjusting the major axis, minor axis, tilt angle, and center of the ellipse [49]. Then inpolygon function [50] is applied to generate binarized masks G as the ground truth according to the region of the ellipse. Besides the data collected by DAVIS346, our dataset also includes 2.7 million PoGs estimated by the Tobii Pro eye tracker and 675 thousand images recorded by its scene camera. As the ground truth of gaze is impossible to obtain, data collected by Tobii is regarded as a reference for gaze estimation.

3.4 Data Characteristics

Our dataset includes multi-modal data collected from two DAVIS346 cameras and one Tobii Pro Glasses 3. The two DAVIS346 cameras produce 1.5 million near-eye grayscale images and more than 2.7 billion events. Figure 1(b) shows some samples of near-eye grayscale images from nine subjects. The images are overlaid with 40ms accumulated events between two consecutive frames.

Figure 2(a) demonstrates the distributions of the gaze references provided in our dataset. From the figure, we can observe PoGs of our dataset are densely distributed across a two-dimensional space with approx. 95° in yaw and 63° in pitch directions. However, as shown in Figure 2(b), EBVEYE [25] only provides very sparse gaze references. The red dots are the locations of the appearance of the stimulus for fixation state and the squared wave lines are the trajectory of the stimulus during smooth pursuit. EBVEYE assumes the human gaze can follow the guidance of the stimulus, though it is often not the case in practice. By comparing the two datasets, the gaze references provided in our dataset are significantly more dense and involve all states, i.e., fixation, saccade, and smooth pursuit. Then Tobii Glasses eye tracker can provide richer temporal information to enable researching on gaze estimation and eye movement dynamics [41]. It is worth noting that, in addition to the dynamic gaze references provided by Tobii, EV-Eye also provides the sparse coordinates of PoGs on the monitor screen during fixation states same as those "ground truth" provided by EBVEYE and static image-based datasets such as ETH-XGaze [51] and [52], the sparse coordinates are recorded by the scene camera of Tobii Pro Glasses 3 eye tracker synchronously[53].

4 Benchmarking Methodologies

In this section, we will describe our proposed benchmarking approach, which leverages both the near-eye grayscale images and the event streams generated by the event camera for accurate and high-frequency eye tracking. The overview of our method is shown in Figure 3.

4.1 Frame-based Pupil Segmentation

DL-based Pupil Segmentation. We employ U-Net [54] for pupil segmentation, which has been proven to achieve state-of-the-art accuracy and adopted as the backbone of several DL-based eye tracking works [55, 56, 57, 58]. The detailed settings of the model can be found in Appendix B. The pupil segmentation component outputs a binarized mask to extract the pupil area.

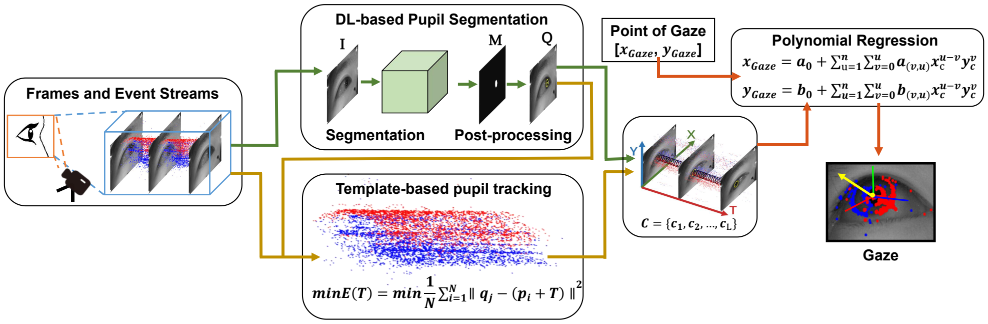

Frames and Event Streams
DL-based Pupil Segmentation
I
M
Q
Segmentation
Post-processing
Point of Gaze
$[x_{Gaze}, y_{Gaze}]$
Polynomial Regression
$x_{Gaze} = a_0 + \sum_{u=1}^{n}\sum_{v=0}^{u} a_{(v,u)} x_c^{u-v} y_c^{v}$
$y_{Gaze} = b_0 + \sum_{u=1}^{n}\sum_{v=0}^{u} b_{(v,u)} x_c^{u-v} y_c^{v}$
Template-based pupil tracking
$minE(T) = min\frac{1}{N}\sum_{i=1}^{N} \| q_j - (p_i + T) \|^2$
Y
X
T
$C = \{c_1, c_2, ..., c_L\}$
Gaze

Figure 3: **Overview of the proposed benchmarking approach of EV-Eye**.

Post-processing. After obtaining the binarized mask $\mathcal{M}$, we adopt morphological closings [59] to remove additional noise, such as glint, in the segmented pupil area. Then, we consider the centroid of the segmented mask as the pupil center c and employ an edge detector to find the pupil boundaries Q as the pupil template for the following high-frequency tracking.

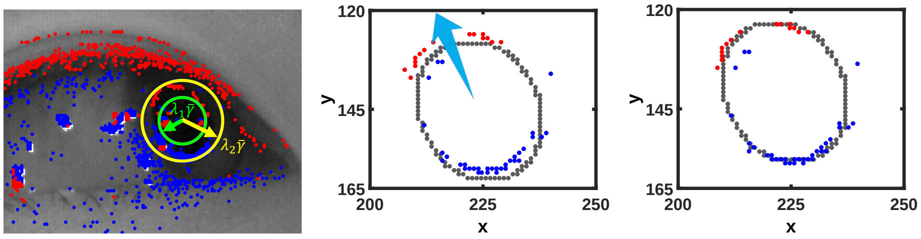
4.2 High-Frequency Event-based Pupil Tracking

Candidate Points Subset. After the pupil area is segmented out, we devise a method for selecting a subset of event points for high-frequency pupil tracking. This enables us to filter out noisy events caused by the movement of eyelashes and eyelids, which are unrelated to the actual pupil movement.

An example of the event points selection is shown in Figure 4(a). First, for each pixel on the boundary Q of the pupil template, we calculate its distance γ to the template center c. We denote the averaged distance for all pixels on Q as $\bar{\gamma}$. Then, we select a set of N event points to form a candidate point set P with $|P| = N$ by the following rule:

$$\lambda_1\bar{\gamma} < ||(x, y) - c|| < \lambda_2\bar{\gamma}, \tag{2}$$

where (x, y) is the coordinate of the current event point on the image; $||(x, y) - c||$ is the distance from the current event point to the template center; $0 < \lambda_1 < 1$ and $\lambda_2 > 1$ are two scale factors (we set $\lambda_1 = 0.8$, $\lambda_2 = 1.2$, respectively in our experiment). As shown in Figure 4(a), we accumulate N event points lying between the two concentric circles to form the event points subset generated by pupil movement.



(a) Candidate subset (b) Before matching (c) After matching

Figure 4: **An example of selected candidate point subset (a)**: the event points lying between the two concentric circles form the candidate point subset and **an example of points-to-edge matching(b)**, the candidate point subset guides pupil updating, the blue arrow is the moving direction of the pupil.

Points-to-edge Matching. Next, we propose a points-to-edge matching approach to update the template center based on the accumulated candidate points set P.

We first calculate the translation of the current pupil boundary Q using the candidate events set P. In our case, the events are generated due to the pupil movement in the horizontal and vertical directions in the camera space. As an example shown in Figure 4(b), when the pupil moves toward the upper-left corner, some of the iris pixels change to pupil pixels and generate red events with negative polarity (intensity decreases); similarly, some of the pupil pixels become iris pixels and produce blue events with positive polarity (intensity increases). Thus, the goal of the points-to-edge matching approach is to find the optimal translation that optimizes the ℓ_2-norm error E:

$$\min E(T) = \min \frac{1}{N} \sum_{i=1}^{N} \left\| q_i - (p_i + T) \right\|^2 , \tag{3}$$

where T is the optimal translation; q_i and p_i are samples in Q and P, respectively. More specifically, T can be obtained in three steps. First, for each event sample $p_i \in P$, we find its closest pixel q'_{i*} in Q by the nearest neighbor search and obtain $Q' = \{q'_1, ...q'_{i*}, ..., q'_N\}$, which is the set of nearest neighbors for all event points in P. Meanwhile, for each $p_i \in P$, we obtain its horizontal distance ΔT_{xi} and vertical distance ΔT_{y_i} to its nearest pixel q'_{i*} in Q. In the second step, we shift each $p_i \in P$ along the horizontal and vertical axes by $\Delta\bar{T}_x$ and $\Delta\bar{T}_y$, respectively, where $\Delta\bar{T}_x = \frac{1}{N}\sum_{i=1}^{N} \Delta T_{xi}$ and $\Delta\bar{T}_y = \frac{1}{N}\sum_{i=1}^{N} \Delta T_{y_i}$. Finally, we repeat the first two steps and update the translation T by:

$$T = T + \Delta\bar{T}, \tag{4}$$

where $\Delta\bar{T} = \{\Delta\bar{T}_x, \Delta\bar{T}_y\}$. The iteration terminates when $\Delta\bar{T}$ diminishes, i.e., $\Delta\bar{T}/T < 0.01$.

Template Center Updating. After obtaining the optimal translation T, we use the N candidate events in P to update the template center by:

$$c^{t+1} = c^t - T, \tag{5}$$

where c^t is the last template center. By updating the center of the pupil template with a few events, the eye movement can be tracked in high frequency.

4.3 Gaze Estimation

After the pupil centers are obtained, we leverage the polynomial regression [46] to estimate the PoGs on the screen. Given the coordinate of the pupil center $c = (x_c, y_c)$, the corresponding PoG is obtained by n order polynomial transformations:

$$x_{Gaze} = a_0 + \sum_{u=1}^{n} \sum_{v=0}^{u} a_{(v,u)} x_c^{u-v} y_c^{v}, \tag{6}$$

where x_{Gaze} is the estimated horizontal coordinate of the PoG on the screen; n is the polynomial order; $a_{(v,u)}$ and $b_{(v,u)}$ are the coefficients. The vertical coordinate y_{Gaze} can be obtained following the same protocol. Due to the subject heterogeneity, e.g., different kappa angles and cornea radius of human eyes, the coefficients are obtained through a subject-dependent calibration.

5 Evaluation on EV-Eye Dataset

In this section, we implement two benchmarking approaches for evaluating the EV-Eye dataset: our proposed method and the model-based method introduced in EVBEYE [25].

5.1 Implementation and Evaluation Metrics

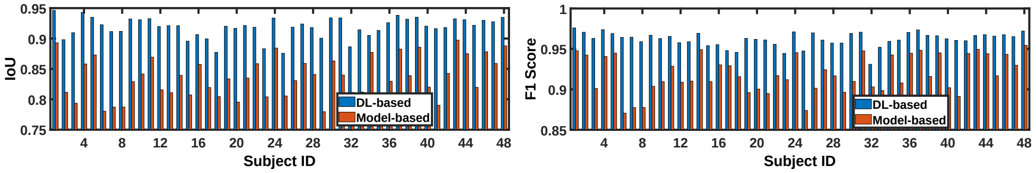
We implement the benchmarking methods with pytroch 3.8. When training the DL-based segmentation network, Adam [60] optimizer and ReduceLROnPlateau [61] scheduler are adopted. The initial learning rate is set to $1e^{-3}$, then it decays by multiplying a scale factor of 0.1 when the dice coefficient is not improved after two consecutive epochs. The batch size is set to 8. Both training and testing are implemented on an NVIDIA GeForce RTX 3090 GPU with 24GB VRAM.

Four metrics are adopted for the dataset evaluation (the detailed description can be found in Appendix Section C.1): **Intersection over union (IoU)** is a widely used metric for pupil region segmentation [62], which corresponds to the overlap between the estimated and groudtruth pupil region. **Dice coefficient (F1 Score)** is another commonly used metric for eye segmentation tasks [55, 63], which measures the similarity between the estimated and groundtruth pupil region. **Pixel error (PE) of eye tracking [8, 64, 65]** is the localization accuracy of eye tracking demonstrated as Euclidean distance in pixels between the estimated and groundtruth pupil centers. **Difference of direction (DoD) in gaze tracking** is the difference between the estimated and reference gaze directions to show the performance of gaze tracking. This metric is similar to the “gaze estimation error” on existing eye tracking literature [41, 52, 51].

5.2 Evaluation on frame-based pupil segmentation

Both our method and EVBEYE [25] contain the frame-based pupil segmentation component. The difference is that we adopt DL-based pupil segmentation with post-processing (DL-based method) rather than model-based method. The 9,011 manually-segmented images are used for the user-independent evaluation: in each round, the manually labeled images from one subject are selected for testing and those from the rest 47 subjects are used for training. The IoU, F1 score, and PE of each subject for two different methods i.e., DL-based method and model-based method are reported.

IoU and F1 score. The IoUs and F1 scores of different subjects obtained by the two methods are presented in Figure 5(a) and Figure 5(b), respectively. The DL-based method achieves significantly higher IoU and FI score for all the subjects compared with the model-based method. The average of IoU value is 0.9187 and 0.8360 for DL-based method and model-based method, respectively; while the average F1 score of these two methods are 0.9560 and 0.9075, respectively. In average, the IoU and F1 score of our DL-based method are 8.27% and 4.85% higher than the model-based method in pupil segmentation task.

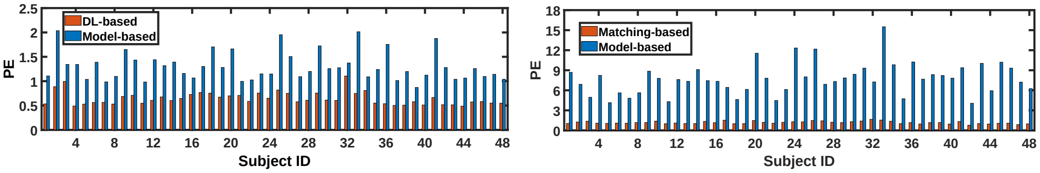


(a) IoUs

(b) F1 scores

Figure 5: **IoUs (a) and F1 scores (b) on frame-based pupil segmentation.**

PE of frame-based pupil segmentation. The PEs of the two methods are shown in Figure 6(a). Our DL-based method achieves significantly lower PE than model-based method for all subjects. The average PE of DL-based method and model-based method is 0.64px and 1.3px, respectively. Therefore, DL-based method achieves a significant 50.7% improvement in frame-based pupil segmentation over model-based method.



(a) Frame-based Pe

(b) Event-based Pe

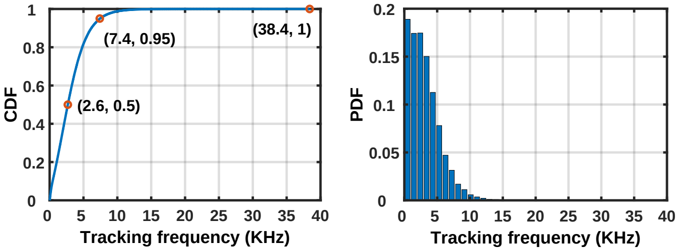
Figure 6: **The pixel error of frame-based (a) and event-based (b) pupil tracking**.

5.3 Evaluation on event-based pupil tracking

Below, we compare and evaluate the accuracy of our template matching-based method and EVBEYE [25]'s model-based method on event-based pupil tracking.

PE of event-based pupil tracking. As we are not able to obtain the event-wise groundtruth, we consider the 9,011 labeled frames as reference to assess the accuracy of event-based pupil tracking. Specifically, the two methods start with obtaining the pupil region of the last image before the labeled one. Then, event-based pupil tracking module runs through the events between the two images. The last pupil center obtained by the event-based module is compared with the groundtruth of the labeled image to obtain the tracking accuracy. The number of events for each event-based update is set to be 20 (same to that used in EVBEYE [25]). The PE of the two methods for each subject is shown in Figure 6(b). Our matching-based method achieves significantly lower PE for all subjects compared with model-based method. The average PE over all subjects is reduced by about 6.5×, i.e., from 7.7px to 1.2px.

5.4 Evaluation on update frequency



(a) CDF

(b) PDF

Figure 7: **CDF and PDF of instantaneous tracking frequency achieved by our method**.

We evaluate the temporal resolution of pupil tracking using EV-Eye by calculating the cumulative distribution function (CDF) and probability density function (PDF) of the tracking frequency. We update the pupil location as soon as a candidate points subset with 20 events is accumulated. The time difference between the first and last event is $T_{interval}$ and the instantaneous tracking frequency is defined as $F = 1/T_{interval}$. The CDF and PDF of tracking frequencies are presented in Figure 7. The tracking frequency of our method is dynamic because the time of generating 20 events are adaptive to pupil movement. Moreover, the peak tracking frequency F_{peak} is up to 38.4KHz. This indicates our method is capable of capturing ultra-high-speed saccadic movement.

Adaptive update strategy: the tracking frequency of the benchmarking approaches is not at a fixed rate: it is determined by how long it takes to accumulate 20 events caused by the movement of the pupil region. Therefore, the tracking frequency is adaptive and proportional to the movement speed. 38.4KHz is the maximum tracking frequency calculated in our dataset (in saccade state). In fact, the extremely-high tracking frequency is an inherent property of an event camera which is high in temporal resolution (around tens of microseconds). Another benefit of the adaptive update strategy is its energy and computation efficiency: it is not necessary to work at very high frequency all the time. For example, when eyes are in a fixed or slow-moving state, the generated events are relatively few, leading to low updating frequency. This property can better utilize the limited resources on computation, energy, and bandwidth.

5.5 Evaluation on gaze tracking

The polynomial regression model (discussed in Section 4.3) is used to map the pupil centers in image domain to the PoGs for gaze tracking. As groundtruth is not available, we utilize the gaze references provided by Tobii for calibration following the similar protocol in [25]. We show the DoDs of the two methods for different subjects in Figure 8, since it's a general indicator on gaze estimation task [52, 41, 5, 66] . From the results we can observe, our method (Ours) achieves significantly lower DoDs for all subjects than model-based method from EVBEYE. The average DoD of Ours and model-based method for the whole field of view is $4.71°$ and $9.72°$ respectively, which indicates the gaze estimation results of our method are significantly closer to the commercialized eye tracker than model-based method proposed in EVBEYE.

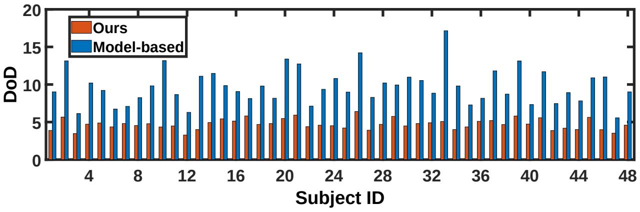


Figure 8: **DoDs of model-based method vs. ours with respect to the gaze references.** The average DoD of our method and model-based method for the whole field of view is $4.71°$ and $9.72°$ respectively.

6 Limitation

The collected multi-modal dataset, EV-Eye, and the proposed benchmark, as presented and evaluated above, offer accessible resources for research in developing new eye tracking algorithms with high frequency. These resources support applications that require the analysis of dynamic eye movement in high temporal resolution, such as the diagnosis of mental disorders. However, further efforts are required in the future to enhance the usability of the dataset. Firstly, additional labeling work is necessary to improve its usability. Currently, approximately 9,000 sparsely selected and labeled images are available to evaluate the accuracy of pupil region segmentation. To support a wider range of applications, a significantly larger number of images with diverse types of labels is required. For instance, we plan to label specific periods of different eye movement statuses, including saccade, fixation, smooth pursuit, blinks, etc. Furthermore, more texture details of the eye, such as regions of the pupil, iris, eyelids, and eyebrows, need to be labeled. Secondly, the dataset lacks ground truth for gaze. Obtaining ground truth for gaze tracking is challenging in real-world settings. Therefore, we utilized the commercial gaze tracker, Tobii Glasses3, to provide gaze references. However, a number of studies[67, 68] have shown that eye-trackers generally are less accurate for large view angles and this seems to be a common issue in many other eye-tracking systems such as ETH-XGaze, MPIIGaze and Angelopoulos et al [51, 52, 25]. Hence, methods to acquire more trustworthy and accurate gaze references need further exploration. Thirdly, regarding subject diversity, we are aware that at the moment the dataset is collected solely from an academic institution, which may introduce slight bias in some sense (e.g. no very diverse ethnicities). We plan to make every effort to expand our dataset to include more diverse participants. Lastly, in the context of the current work, we only select three major types of gaze behavior in natural viewing scenarios, i.e., fixation, saccade and smooth pursuit. However, we realize that considering more diverse scenarios in the real-world could enrich the gaze behavior in the dataset and thus broaden community use. In the next stage, we plan to use video stimuli such as CRCNS, DIEM [41], and Gaze360 [69]in the wild environment to capture more diverse types of eye movements.

7 Conclusion

In this paper, we introduce the most diverse and largest event-based multimodal dataset EV-Eye for high-frequency eye tracking, collected from 48 subjects with different devices. Frames and events from two DAVIS346 can describe the eye movement in extremely high temporal resolution and a commercialized eye tracker can provide densely distributed gaze references for cross modality comparison. Then we propose a novel hybrid frame-event eye-tracking approach to uncover the potential of the multi-modal dataset to achieve a tracking frequency of up to 38.4KHz. The extensive evaluations on EV-Eye demonstrate our method achieves significantly higher accuracy and is more robust to the diverse dataset than the state-of-the-art hybrid frame-event eye tracking method.

Acknowledgments and Disclosure of Funding

This work is supported by Shandong Provincial Natural Science Foundation, China, Grant No.2022HWYQ-040.

References

[1] Eye tracking. `https://en.wikipedia.org/wiki/Eye_tracking`.

[2] Maria K Eckstein, Belén Guerra-Carrillo, Alison T Miller Singley, and Silvia A Bunge. Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*, 25:69–91, 2017.

[3] Andrew T Duchowski. A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4):455–470, 2002.

[4] Yihua Cheng, Haofei Wang, Yiwei Bao, and Feng Lu. Appearance-based gaze estimation with deep learning: A review and benchmark. *arXiv preprint arXiv:2104.12668*, 2021.

[5] Carlos H. Morimoto and Marcio R. M. Mimica. Eye gaze tracking techniques for interactive applications. *Comput. Vis. Image Underst.*, 98(1):4–24, apr 2005.

[6] Nachiappan Valliappan, Na Dai, Ethan Steinberg, Junfeng He, Kantwon Rogers, Venky Ramachandran, Pingmei Xu, Mina Shojaeizadeh, Li Guo, Kai Kohlhoff, et al. Accelerating eye movement research via accurate and affordable smartphone eye tracking. *Nature Communications*, 11(1):1–12, 2020.

[7] Xinming Wang, Jianhua Zhang, Hanlin Zhang, Shuwen Zhao, and Honghai Liu. Vision-based gaze estimation: a review. *IEEE Transactions on Cognitive and Developmental Systems*, 2021.

[8] Joohwan Kim, Michael Stengel, Alexander Majercik, Shalini De Mello, David Dunn, Samuli Laine, Morgan McGuire, and David Luebke. Nvgaze: An anatomically-informed dataset for low-latency, near-eye gaze estimation. In Proceedings of the 2019 CHI Conference, New York, NY, USA, 2019. Association for Computing Machinery.

[9] Tobii pro 3 glasses. `https://www.tobii.com/products/eye-trackers/wearables/tobii-pro-glasses-3`.

[10] Pupil Labs eye tracker. `https://pupil-labs.com/`.

[11] Tiffany CK Kwok, Peter Kiefer, Victor R Schinazi, Benjamin Adams, and Martin Raubal. Gaze-guided narratives: Adapting audio guide content to gaze in virtual and real environments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

[12] Alexander Mariakakis, Mayank Goel, Md Tanvir Islam Aumi, Shwetak N Patel, and Jacob O Wobbrock. SwitchBack: Using focus and saccade tracking to guide users' attention for mobile task resumption. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2953–2962, 2015.

[13] Guohao Lan, Bailey Heit, Tim Scargill, and Maria Gorlatova. GazeGraph: Graph-based few-shot cognitive context sensing from human visual behavior. In *Proceedings of the ACM Conference on Embedded Networked Sensor Systems (SenSys)*, pages 422–435, 2020.

[14] Namrata Srivastava, Joshua Newn, and Eduardo Velloso. Combining low and mid-level gaze features for desktop activity recognition. *In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4):1–27, 2018.

[15] Zillah Boraston and Sarah-Jayne Blakemore. The application of eye-tracking technology in the study of autism. *The Journal of Physiology*, 581(3):893–898, 2007.

[16] Dillon Lohr, Henry Griffith, and Oleg V Komogortsev. Eye know you: Metric learning for end-to-end biometric authentication using eye movements from a longitudinal dataset. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 2022.

[17] Dillon Lohr and Oleg V Komogortsev. Eye know you too: Toward viable end-to-end eye movement biometrics for user authentication. *IEEE Transactions on Information Forensics and Security*, 17:3151–3164, 2022.

[18] Wikipedia saccade. `https://en.wikipedia.org/wiki/Saccade`.

[19] Richard A Abrams, David E Meyer, and Sylvan Kornblum. Speed and accuracy of saccadic eye movements: characteristics of impulse variability in the oculomotor system. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3):529, 1989.

[20] Nanda N. J. Rommelse, Stefan Van der Stigchel, and Joseph A. Sergeant. A review on eye movement studies in childhood and adolescent psychiatry. *Brain and Cognition*, 68:391–414, 2008.

[21] A highly accurate, precise, and versatile eye tracker. `https://www.sr-research.com/eyelink-1000-plus/`.

[22] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):154–180, 2022.

[23] Christoph Posch, Daniel Matolin, and Rainer Wohlgenannt. A qvga 143 db dynamic range frame-free pwm image sensor with lossless pixel-level video compression and time-domain cds. *IEEE Journal of Solid-State Circuits*, 46(1):259–275, 2011.

[24] Timo Stoffregen, Hossein Daraei, Clare Robinson, and Alexander Fix. Event-based kilohertz eye tracking using coded differential lighting. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2515–2523, 2022.

[25] Anastasios N Angelopoulos, Julien NP Martel, Amit P Kohli, Jörg Conradt, and Gordon Wetzstein. Event-based near-eye gaze tracking beyond 10,000 Hz. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2577–2586, 2021.

[26] Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. Appearance-based gaze estimation in the wild. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4511–4520, 2015.

[27] Xucong Zhang, Seonwook Park, Thabo Beeler, Derek Bradley, Siyu Tang, and Otmar Hilliges. ETH-XGaze: A large scale dataset for gaze estimation under extreme head pose and gaze variation. In *Proceedings of European Conference on Computer Vision (ECCV)*, pages 365–381. Springer, 2020.

[28] Kyle Krafka, Aditya Khosla, Petr Kellnhofer, Harini Kannan, Suchendra Bhandarkar, Wojciech Matusik, and Antonio Torralba. Eye tracking for everyone. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2176–2184, 2016.

[29] Jiang Zhao, Shilong Ji, Zhihao Cai, Yiwen Zeng, and Yingxun Wang. Moving object detection and tracking by event frame from neuromorphic vision sensors. *Biomimetics*, 7(1):31, 2022.

[30] Yanxiang Wang, Bowen Du, Yiran Shen, Kai Wu, Guangrong Zhao, Jianguo Sun, and Hongkai Wen. Ev-gait: Event-based robust gait recognition using dynamic vision sensors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6358–6367, 2019.

[31] Patrick Bardow, Andrew J. Davison, and Stefan Leutenegger. Simultaneous optical flow and intensity estimation from an event camera. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 884–892, 2016.

[32] Yin Bi, Aaron Chadha, Alhabib Abbas, Eirina Bourtsoulatze, and Yiannis Andreopoulos. Graph-based spatio-temporal feature learning for neuromorphic vision sensing. *IEEE Transactions on Image Processing*, 29:9084–9098, 2020.

[33] Yanxiang Wang, Xian Zhang, Yiran Shen, Bowen Du, Guangrong Zhao, Lizhen Cui, and Hongkai Wen. Event-stream representation for human gaits identification using deep neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[34] Qinyi Wang, Yexin Zhang, Junsong Yuan, and Yilong Lu. Space-time event clouds for gesture recognition: From rgb cameras to event cameras. In *Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1826–1835. IEEE, 2019.

[35] Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Unsupervised event-based optical flow using motion compensation. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, pages 0–0, 2018.

[36] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 652–660, 2017.

[37] S. Tulyakov, D. Gehrig, S. Georgoulis, J. Erbach, M. Gehrig, Y. Li, and D. Scaramuzza. Time lens: Event-based video frame interpolation. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16150–16159, 2021.

[38] Kang Wang and Qiang Ji. Real time eye gaze tracking with 3d deformable eye-face model. In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, pages 1003–1011, 2017.

[39] E.D. Guestrin and M. Eizenman. General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on Biomedical Engineering*, 53(6):1124–1133, 2006.

[40] Qiong Huang, Ashok Veeraraghavan, and Ashutosh Sabharwal. Tabletgaze: Dataset and analysis for unconstrained appearance-based gaze estimation in mobile tablets. *Mach. Vision Appl.*, 28(5–6):445–461, aug 2017.

[41] Kang Wang, Hui Su, and Qiang Ji. Neuro-inspired eye tracking with eye movement dynamics. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9823–9832, 2019.

[42] Kenneth Alberto Funes Mora, Florent Monay, and Jean-Marc Odobez. Eyediap: A database for the development and evaluation of gaze estimation algorithms from rgb and rgb-d cameras. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '14, page 255–258, New York, NY, USA, 2014. Association for Computing Machinery.

[43] Seonwook Park, Emre Aksan, Xucong Zhang, and Otmar Hilliges. Towards end-to-end video-based eye-tracking. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII*, page 747–763, Berlin, Heidelberg, 2020. Springer-Verlag.

[44] Y. Feng, N. Goulding-Hotta, A. Khan, H. Reyserhove, and Y. Zhu. Real-time gaze tracking with event-driven eye segmentation. In *2022 IEEE on Conference Virtual Reality and 3D User Interfaces (VR)*, pages 399–408, Los Alamitos, CA, USA, mar 2022. IEEE Computer Society.

[45] Cian Ryan, Brian O'Sullivan, Amr Elrasad, Aisling Cahill, Joe Lemley, Paul Kielty, Christoph Posch, and Etienne Perot. Real-time face & eye tracking and blink detection using event cameras. *Neural Networks*, 141:87–97, 2021.

[46] Z.R. Cherif, A. Nait-Ali, J.F. Motsch, and M.O. Krebs. An adaptive calibration of an infrared light device used for gaze tracking. In *Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference*, volume 2, pages 1029–1033 vol.2, 2002.

[47] Davis346 event camera. `https://inivation.com/wp-content/uploads/2019/08/DAVIS346.pdf`,.

[48] Vgg image annotator. `http://www.robots.ox.ac.uk/vgg/software/via`.

[49] Parameters of ellipse traced out by tip of a polarized field vector. `https://www.mathworks.com/help/phased/ref/polellip.html`,.

[50] Points located inside or on edge of polygonal region. `https://ch.mathworks.com/help/matlab/ref/inpolygon.html`,.

[51] Xucong Zhang, Seonwook Park, Thabo Beeler, Derek Bradley, Siyu Tang, and Otmar Hilliges. Eth-xgaze: A large scale dataset for gaze estimation under extreme head pose and gaze variation. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 365–381, Cham, 2020. Springer International Publishing.

[52] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling. Mpiigaze: Real-world dataset and deep appearance-based gaze estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(01):162–175, jan 2019.

[53] tobiipro. Pro glasses 3 developer guide. `https://www.tobii.com/products/eye-trackers/wearables/tobii-proglasses-3#form`, 2022.

[54] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, 2015.

[55] Aayush K. Chaudhary, Rakshit Kothari, Manoj Acharya, Shusil Dangi, Nitinraj Nair, Reynold Bailey, Christopher Kanan, Gabriel Diaz, and Jeff B. Pelz. RITnet: Real-time semantic segmentation of the eye for gaze tracking. In *Proceedings of IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE, oct 2019.

[56] Chengdong Lin, Xinlin Li, Zhenjiang Li, and Junhui Hou. Finding stars from fireworks: Improving non-cooperative iris tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(9):6137–6147, 2022.

[57] Caiyong Wang, Jawad Muhammad, Yunlong Wang, Zhaofeng He, and Zhenan Sun. Towards complete and accurate iris segmentation using deep multi-task attention network for non-cooperative iris recognition. *IEEE Transactions on Information Forensics and Security*, 15:2944–2959, 2020.

[58] Sheng Lian, Zhiming Luo, Zhun Zhong, Xiang Lin, Songzhi Su, and Shaozi Li. Attention guided u-net for accurate iris segmentation. *Journal of Visual Communication and Image Representation*, 56:296–304, 2018.

[59] Luc Vincent. Morphological area openings and closings for grey-scale images. In Ying-Lie O, Alexander Toet, David Foster, Henk J. A. M. Heijmans, and Peter Meer, editors, *Shape in Picture*, pages 197–208, 1994.

[60] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. 2015.

[61] Manzil Zaheer, Sashank Reddi, Devendra Sachan, Satyen Kale, and Sanjiv Kumar. Adaptive methods for nonconvex optimization. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[62] Hamid Rezatofighi, Nathan Tsoi, JunYoung Gwak, Amir Sadeghian, Ian Reid, and Silvio Savarese. Generalized intersection over union: A metric and a loss for bounding box regression. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 658–666, 2019.

[63] Stephan J. Garbin, Yiru Shen, Immo Schuetz, Robert Cavin, Gregory Hughes, and Sachin S. Talathi. Openeds: Open eye dataset, 2019.

[64] Wolfgang Fuhl, Thiago C Santini, Thomas Kübler, and Enkelejda Kasneci. Else: Ellipse selection for robust pupil detection in real-world environments. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*, pages 123–130, 2016.

[65] Wolfgang Fuhl, Thiago Santini, Gjergji Kasneci, Wolfgang Rosenstiel, and Enkelejda Kasneci. Pupilnet v2.0: Convolutional neural networks for cpu based real time robust pupil detection. *ArXiv*, abs/1711.00112, 2017.

[66] Murthy L R D and Pradipta Biswas. Appearance-based gaze estimation using attention and difference mechanism. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 3137–3146, 2021.

[67] Dodou D. de Winter J.C.F. Onkhar, V. Evaluating the tobii pro glasses 2 and 3 in static and dynamic conditions. *Behav Res*, 2023.

[68] Iqbal S. Pearson J. Johnson E. N. MacInnes, J. J. Wearable eye-tracking for research: Automated dynamic gaze mapping and accuracy/precision comparisons across devices. *BioRxiv*, 2018.

[69] Petr Kellnhofer, Adria Recasens, Simon Stent, Wojciech Matusik, and Antonio Torralba. Gaze360: Physically unconstrained gaze estimation in the wild. In *IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[70] Shruti Jadon. A survey of loss functions for semantic segmentation. In *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 1–7, 2020.

[71] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, 2021.