INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 3, PP. 1-6
Manuscript received May 14,2026; revised June, 2026. doi: 10.24425/ijet.2026.1734

Real Time Face-to-Screen Distance Estimation on Smartphones Using Iris Diameter and the Pinhole Camera Model

Daniel Al-Burgan*, and Jerzy Weremczuk

***Abstract*—This paper presents a proof-of-concept architecture that investigates if a standard smartphone front camera can estimate face-to-screen distance in real time by using the average human iris diameter as a biological scale reference within a pinhole camera model. The proposed Android implementation uses automatic eye region localization, camera intrinsic calibration, grayscale preprocessing, and random sample consensus circle fitting to measure iris diameter in pixels. To support natural device use, a gaze rectification step compensates for perspective distortion when the user looks at the screen instead of directly into the camera. The method was evaluated in a pilot study with 48 recordings from one participant across eight distances from 200 to 550 millimeters, captured under both screen-directed and camera-directed gaze conditions. The system remained stable over the 250 to 500 millimeter range and produced comparable performance for screen and camera gaze. While accuracy degraded markedly at 550 millimeters. The main limitation was a systematic underestimation of roughly 20 to 35 millimeters, affected by the ground truth measurement setup and the static biological assumption of iris diameter. The findings of this paper show that an on-device, real-time pipeline for iris based monocular distance estimation is computationally feasible for mobile applications.**

***Keywords*—face-to-screen distance estimation; iris based scaling; digital biomarkers; smartphone front camera; pinhole camera model**

I. Introduction

HOLDING a smartphone in a close distance to the face during prolonged use is associated with increased visual discomfort and eye strain. A reliable method for estimating this distance can enable smartphones to analyze user behavior and give feedback on viewing distance habits. Smartphones such as iPhones already incorporate face distance awareness, by relying on specialized hardware. These capabilities are not available universally across smartphones. Direct face-to-screen distance measurement typically requires dedicated hardware, such as infrared or time-of-flight sensors. Most Android smartphones rely solely on conventional front facing cameras, which do not provide explicit depth information, and limits distance estimation to front camera based methods operating without additional sensors.

From a metrological point of view, monocular distance estimation using a single camera cannot provide absolute distance without additional reference information. One approach to obtain this information is to use anatomical features with approximately consistent dimensions in most humans as reference points for distance estimation. In face based estimation, ocular features are particularly suitable, for example, pupillary distance which ranges approximately between 54mm to 74mm, and iris diameter, as used in this study, which ranges 11.64 ± 0.49 mm [1]. The practical suitability of such references for real-time distance estimation on consumer smartphones still requires systematic evaluation.

The aim of this work is to assess the architectural and computational feasibility of estimating face-to-screen distance on Android devices using an assumed average human iris diameter of 11.64 mm as a biological scaling reference. Rather than targeting a specific clinical application, as done in [2] and [3], this study focuses on the implementation and analysis of a real-time, on-device measurement pipeline. This pipeline combines ML Kit eye-region localization, camera intrinsic parameters extracted via the Camera2 API, and a native C++ RANSAC circle-fitting algorithm. Additionally, the system introduces a gaze rectification step to compensate for perspective distortion, for maintaining computational stability even when the user looks naturally at the screen rather than the camera lens. By evaluating the system's computational performance and mathematical stability, this paper seeks to demonstrate a viable software architecture for future mobile eye health and digital biomarker applications.

II. Literature Review

Face-to-screen distance can be accurately measured using specialized hardware, that is commonly used in Apple’s devices. A recent study evaluating iPad eye tracking, for example, highlighted how Apple’s TrueDepth infrared system which uses facial depth maps, is precise enough to be used to track eye movements [4]. Dedicated depth sensors are not universally available on Android devices. Because of that, alternative biological references, such as the human iris, can be used to estimate distance using a standard front camera.

Monocular camera based distance estimation methods commonly use a geometric object with known size as a reference. This object, in regards to the human face, must be a feature that is commonly close in size in humans, such as interpapillary distance (IPD) or iris diameter.

Literature on using IPD for distance estimation includes [5] where similarly to our study, a smartphone method was developed, to estimate face-to-screen distance using front camera. The method used accelerometer to trigger photo capture only when specific motion is detected, rather than

Authors are with Institute of Electronic Systems, Faculty of Electronics and Information Technology, Warsaw University of Technology, Poland (e-mails: daniel.al-burgan.dokt@pw.edu.pl, jerzy.weremczuk@pw.edu.pl).

processing a continuous camera feed to reduce power consumption. In [6], an IPD based method using a webcam mounted on a precision controlled uniaxial displacement system, with physical ground truth established by a laser distance meter. As mentioned in [6], a fundamental limitation of using IPD is its natural demographic variability, making population-averaged measurements less reliable for individual users.

The visible iris diameter can be used as a reference scale to calibrate and convert linear facial measurements from images of unknown scale into real world dimensions in millimeters [7]. The study demonstrated that the iris is a suitable reference point for monocular scaling, but focused on static image analysis.

Iris based distance estimation relies on the limited variability of the human iris. Anthropometric studies validate its use as a reference, showing that horizontal iris diameter has a remarkably narrow spread between 11.6mm and 11.7mm, with no statistically significant dependence on age or gender [8]. This aligns with earlier literature which established the mean diameter as 11.64 ± 0.49 mm [1]. Based on these findings, an assumed biological reference of 11.64 mm is used in this study.

Literature on using the iris diameter for distance estimation using smartphone front camera, similar in methodology to this study, include a recent study [2], in which the authors developed a smartphone application to diagnose eye convergence insufficiency by continuously estimating the smartphone to eye distance using the using a pinhole camera model. Their system relied on an assumed iris diameter of 11.7 mm for reference, which is close to ours 11.64 mm. The author’s ground truth was the RulerT, which is a test where a doctor moves a target toward the patient until the patient reports blurry vision. Their method achieved a clinical accuracy margin of ±30 mm against a human doctor. Additionally, another similar study [3] developed a smartphone application for myopia screening that estimated eye to screen distance operating on similar geometric principles, but instead of directly using a pinhole camera model with native focal lengths, their system calculated distance by applying an empirical curve fitting coefficient, by manually plotting the physical distance against the pixel output for a specific phone model (the iPhone X) to find the exact mathematical ratio that linked them together, with the assumption of iris diameter of 12 mm. To validate their distance estimation accuracy, the authors compared their system's scaling against a clinical pupilometer as the ground truth, achieving a minimal mean bias of 0.17 mm with 95% limits of agreement of ±4.08 mm.

During natural device usage, users focus their eyes on the screen rather than staring directly into the camera lens, which creates perspective distortion on the captured facial features. A recent study [9] explored this phenomenon in the context of mobile gaze estimation, which is a to predict exactly where a user is looking on a device's screen using the front camera images. While the authors were focused on tracking eye movements rather than measuring distance, they evaluated how natural user movement impacts visual model accuracy. They found that standard camera models are vulnerable to changes in head pose and device orientation. Because users typically hold devices below head level and look down at the screen, the resulting off axis perspective causes significant accuracy degradation in methods that lack geometric normalization. The authors concluded that explicitly accounting for these orientation and head pose shifts remains an unmet need in mobile tracking systems.

The necessity of correcting this geometric distortion is also mentioned in a recent study on iris based eye tracking focused on XR and VR headsets [10], where the authors highlighted the that camera perspective projection phenomenon, where viewing a circular object, like the iris, from an angle causes it to appear squished into an oval in the 2D image. To accurately measure the eye's true physical parameters, the authors mathematically adjusted the raw image data to simulate what the iris would look like if the camera were viewing it perfectly straight on. In our study, this exact off axis viewing behavior directly affects the shape and apparent horizontal diameter of the iris, which would otherwise skew distance calculations. To address this and maintain measurement stability, our proposed distance estimation pipeline introduces a gaze rectification step that mathematically unwarps the horizontal iris gradients using the eye's yaw angle before fitting the circle.

III. Materials and Methods

The proposed method estimates face-to-screen distance using a monocular camera by exploiting a known anatomical reference dimension. The front facing camera captures a feed of images, from which, the two eye regions are extracted, and the horizontal diameter of user’s irises are expressed in pixels. Iris diameter is assumed to be 11.64mm and is used as a biological reference. Distance estimation is performed using a pinhole camera model, in which the face-to-screen distance is computed as the ratio between the focal length of the smartphone in pixels and the measured iris diameter in pixels, scaled by the assumed physical iris diameter, as given by

$$D = \frac{(f_{px} * W_{iris})}{P_{iris}} \tag{1}$$

where: D - estimated face-to-screen distance in mm, f_{px} - focal length of the front-facing camera expressed in pixels, obtained from the Camera2 API, W_{iris} - assumed physical iris diameter (11.64mm used in this study) , P_{iris} - measured horizontal iris diameter from the image in pixels.

The system was implemented as an Android application consisting of a Flutter user interface, native Android code written in Kotlin, and native C++ components accessed via the Android NDK. Keeping biometric computations local is a deliberate choice for data privacy. Unauthorized data gathering and privacy violations are increasingly common in mobile and IoT frameworks [10]. Flutter provides a responsive user interface, but the computations required for continuous frame-by-frame image analysis need high execution speeds, because of that, the biometric algorithms, such as histogram equalization, spatial bounding, and RANSAC (Random Sample Consensus) for circle fitting, were implemented on a native C++ environment. The system processes Camera frames in real time. The continuous real-time processing pipeline, detailing the step-by-step data flow from the raw camera frame capture to the final pinhole camera distance calculation, is illustrated in Fig. 1.

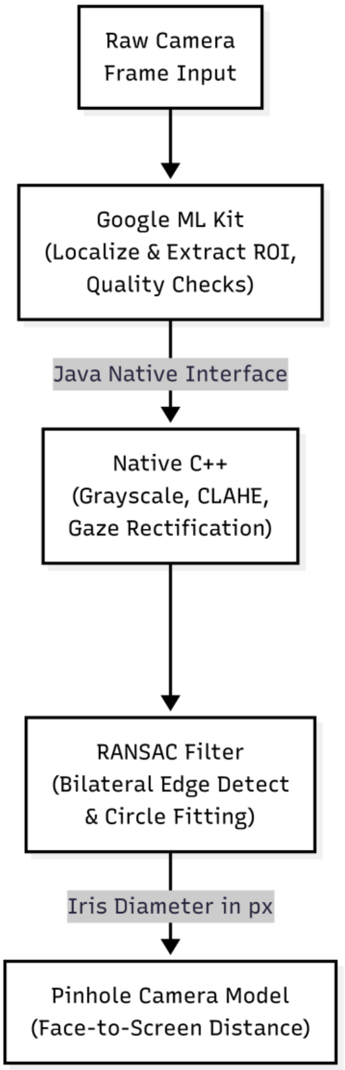

Fig. 1. System architecture and data flow of the proposed real-time face-to-screen distance estimation pipeline

Face and eye localization are performed using the Google ML Kit Face Detection framework, in fast performance mode with contour, landmark, and classification outputs enabled. For each camera frame, the detector returns the face bounding box together with eye contour points, eye landmark positions, eye open probability estimates, and head pose information expressed as Euler angles. The eye contour points are used as region-of-interest sent to the C++ layer.

Frames affected by closed eyes, where the user's head yaw or roll exceeded ±12° (this threshold was determined to prevent perspective distortion from skewing the horizontal iris geometry, while retaining a sufficiently high frame rate during natural device usage), or insufficient image quality (frames with an ML Kit eye detection confidence score below 0.10) are excluded. Additional temporal consistency constraints are applied to suppress abrupt, implausible changes in measured iris diameter, and the last valid measurement is retained for short intervals when valid detections are momentarily unavailable.

Iris diameter estimation is implemented in native C++ using OpenCV and invoked via the Java Native Interface. The algorithm first improves the cropped grayscale image of the eye. It applies Contrast Limited Adaptive Histogram Equalization to make the iris boundaries more visually distinct and suppresses bright light reflections (specular glints) that might obscure the true shape of the eye, this is done to isolate the circular edge of the iris.

Iris radial edge detection is restricted to the left and right sides of the iris. The algorithm only scans a ±30° area along the horizontal axis. This ±30° limit was chosen for two reasons: it provides enough data points for the RANSAC algorithm to accurately calculate a circle. Second, it is narrow enough to avoid the upper and lower eyelids, which can cover the top and bottom of the iris when a user looks down at a smartphone. Looking at a screen from a close distance causes a perspective distortion of the iris (iris appears as a ellipse rather not a circle). To correct this, the measured horizontal gradients are mathematically unwrapped using the eye's yaw angle, in other words, the algorithm applies a gaze rectification step, that stretches the horizontal width of the detected iris to restore its true circular shape. These rectified points are then processed by RANSAC circle-fitting algorithm. RANSAC filter enforces strict bilateral geometric support. It requires valid edge points on both the left and right sides of the iris, at least 10% of the supporting data points come from the opposite side of the iris center, if this condition isn't met, then the incomplete detection fails gracefully to avoid producing a false measurement.

Finally, the face-to-screen distance is computed using a pinhole camera model. The system uses a focal length obtained from the Android Camera2 API. This focal length is adjusted by calculating the ratio between the full physical sensor array and the active cropped region to account for digital zoom and aspect ratio changes. The distance is then calculated using the stabilized iris diameter, the adjusted focal length in pixels, and the assumed physical iris diameter of 11.64 mm. The iris diameter is based on measurements from [1], as mentioned in the Related Work section.

To assess the feasibility of the proposed method, a system validation pilot study was conducted producing 48 datasets from one person. 8 distances were measured ranging from 200mm to 550mm in 50mm increments. Each of the eight distances was measured six times: three times anchoring the tape measure to the left upper eyelid, and three times to the right upper eyelid. This range is within natural human smartphone interaction according to literature, as highlighted by [12], where the authors highlighted that mean working distance for text messages to be 362mm with range between 175mm to 580mm, and for internet viewing to be 322mm with range between 190mm to 600mm. Additionally, in [12], the authors presented mean viewing distance while using a smartphone over 60 minutes to be from 219 to 365mm (29.2 ± 7.3 cm). Testing cutoffs of 200mm and 550mm were determined by the system's limits observed during testing, as distances of 200mm and below captured less total valid frames, while distances of 550 mm and beyond resulted in unacceptably high Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

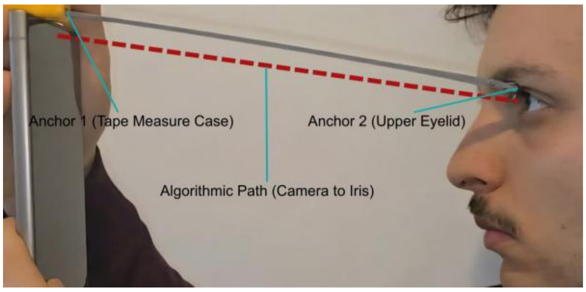

Fig. 2. Experimental setup showing the source of systematic measurement offset. The ground truth was measured using a tape measure anchored from the top edge of the device, more precisely, from tape measure case (Anchor 1) to the subject's upper eyelid region (Anchor 2). The depth difference between the front camera and the tape measure case, combined with the gap between the eyelid anchor and the iris, accounts for a physical measurement error between the two paths.

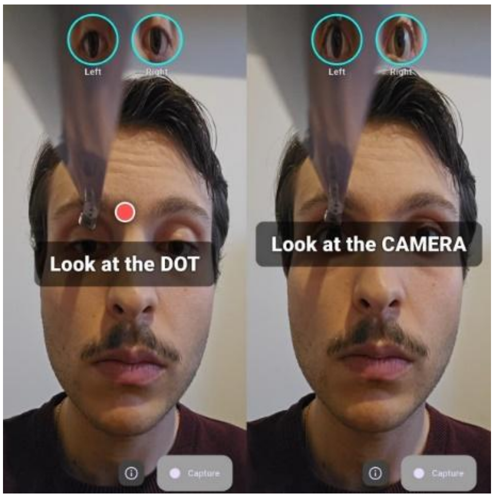

Fig. 3. Application User Interface and Data Collection Workflow. The application overlay during the two phase capture process. Left: The subject focuses on red dot at the center of the screen for 25 seconds to capture perspective distortion on the iris. Right: The subject focuses directly on the camera lens. Real-time grayscale eye crops are visible at the top, demonstrating successful region-of-interest isolation before C++ processing.

As illustrated in Fig. 2, the physical ground truth was established using a tape measure. However, the physical distance between the two tape measure anchors is slightly shorter than the actual optical path from the front camera lens to the iris. This depth error was not compensated for in the measurements. Data collection was done in standard indoor office lighting, and followed a strict two phase workflow as shown in Fig. 3. Each of the 48 datasets consisted of a 50 second capture: 25 seconds where the subject focused on a red dot at the center of the screen, followed by 25 seconds focusing directly on the front camera lens. This protocol was used to evaluate the algorithm's ability to handle perspective distortion using horizontal gradient scanning when the iris shifts position or is partially occluded by eyelids. The data collection application used a timed workflow, to prevent user reaction time from corrupting the measurements, each 25 second active recording phase was preceded by a 3 second instruction reading period and a 3 second countdown. A 4 second transition pause was put between the screen directed and camera directed phases, to make sure that the subject's gaze had shifted and stabilized on the new target before the second recording started.

IV. Result and Discussion

The performance of the iris based distance estimation pipeline was evaluated across the 48 captured datasets. Absolute measurement accuracy was quantified using MAE, and RMSE, while algorithmic precision (frame-to-frame mathematical stability) was measured via Standard Deviation. The results across varying distances and both screen and camera gaze targets are presented in Table I.

Data analysis reveals three core findings regarding the algorithm's performance, along with specific system limitations:

1) *Systematic Measurement Offset*

At distances from 250 mm to 500 mm, the system exhibited a consistent systematic underestimation of approximately 20 mm to 35 mm. Rather than indicating an algorithmic failure, this offset is attributed to two factors: the uncompensated physical depth difference between the tape measure anchor (the device casing) and the recessed camera sensor, and the use of population average iris diameter as a constant (11.64 mm). While the absolute distance was shifted by this constant offset, the low standard deviations indicate that the algorithm successfully and stably tracked relative user movement. At 550 mm, the algorithm's performance drastically decreased, most probably due to the limitation of pixels displaying the iris.

2) *The Efficacy of Gaze Rectification*

The objective of the proposed architecture was to maintain measurement stability during natural smartphone usage when the user is looking at the screen rather than the camera lens. The data confirms that the C++ gaze rectification pipeline successfully mitigated perspective distortion. At distances from 300 mm to 450 mm, the MAE and RMSE for "screen" targets were comparable to the "camera" targets, demonstrating that the horizontal iris gradient is effective.

3) *Algorithmic Precision and Focal Limits*

The standard deviation column highlights the optical limitations of the fixed focus methodology. To prevent autofocus lens breathing from distorting intrinsic parameters, the camera focus was locked to 3.0 diopters (approximately 333 mm). At distances near this focal plane (350 mm to 500 mm), the RANSAC algorithm demonstrated high precision and stability. However, at closer proximities (200 mm to 250 mm), the iris falls outside the focal plane. At 200 mm and below, the optical blur and extreme gaze convergence angles caused the system's graceful failure checks to trigger frequently. While other distance tests captured roughly 375 valid frames, the 200 mm "screen" target dropped to almost half of its frames, yielding only 197 valid frames.

4) *Algorithmic Precision and Focal Limits*

As a proof-of-concept architecture, this study is limited by its single subject pilot design. The primary goal was to validate the computational performance and local processing viability of the pipeline, not to establish clinical accuracy. Consequently, the current RANSAC circle-fitting implementation has not been evaluated against diverse eye morphologies (e.g., epicanthic folds), varying iris coloration, or different skin-tone contrasts. Relying on an assumed population-average iris diameter (11.64 mm) introduces an unavoidable systematic offset for any individual user whose iris deviates from this mean.

TABLE I
SUMMARY OF ALGORITHM MEASUREMENTS AND ERROR METRICS ACROSS ALL TESTED DISTANCES AND GAZE TARGETS

Ground Truth (mm)	Gaze Target	Mean Calc. (mm)		MAE (mm)		RMSE (mm)		Std Dev (mm)		Valid Frames	
		Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
200	Camera	174.83	188.80	25.17	22.17	28.25	34.66	12.85	32.85	288	322
	Screen	180.52	202.59	21.55	29.67	25.62	51.61	16.69	51.62	197	320
250	Camera	226.55	217.17	23.45	32.83	27.58	36.32	14.52	15.56	375	377
	Screen	236.62	227.65	25.86	22.86	38.86	24.56	36.53	10.19	376	376
300	Camera	271.71	279.01	29.14	21.09	35.17	23.50	20.91	10.58	376	377
	Screen	278.35	280.22	22.45	21.36	23.35	23.28	8.75	12.28	377	376
350	Camera	325.48	326.60	24.68	23.42	26.52	25.00	10.11	8.79	377	376
	Screen	328.55	329.10	21.86	20.92	22.97	22.13	8.21	7.28	376	375
400	Camera	372.71	374.24	27.29	25.76	28.39	26.19	7.83	4.74	375	375
	Screen	375.57	370.81	24.43	29.19	24.98	29.83	5.21	6.18	375	377
450	Camera	415.31	424.61	34.69	25.39	40.65	26.23	21.23	6.60	358	375
	Screen	421.85	423.81	28.23	26.19	28.82	27.00	6.18	6.56	376	377
500	Camera	468.56	466.89	31.44	33.11	31.59	33.37	3.12	4.22	375	375
	Screen	468.07	464.44	31.93	35.56	32.17	36.15	3.91	6.55	376	376
550	Camera	406.90	457.05	143.10	92.95	143.37	94.30	8.78	15.91	376	376
	Screen	415.81	437.50	134.19	112.50	135.87	116.22	21.31	29.23	376	376

V. CONCLUSION

This paper presented a proof-of-concept architecture for monocular face-to-screen distance estimation using a smartphone's front camera, by utilizing the average human iris diameter as a scaling reference within a pinhole camera model.

The real-time C++ pipeline successfully integrated eye region localization, gaze rectification, and RANSAC based circle fitting while keeping all biometric computations local to the device. The methodology maintained mathematical stability across a 250 mm to 500 mm range. Accuracy for "screen" gaze targets was comparable to "camera" targets, demonstrating that the system can effectively mitigate perspective distortion and eyelid interference during natural smartphone usage.

The study identified a consistent systematic distance underestimation (20 mm to 35 mm), affected by the physical ground truth measurement setup and the static biological assumption of the iris diameter. Furthermore, precision degraded beyond 550 mm. Despite these constraints, the results support the computational feasibility of deploying real time, iris based scaling pipelines for future mobile eye health applications. Future research will expand the dataset to a diverse demographic to evaluate algorithmic performance across varying eye morphologies. It will also include incorporating this architecture into an active digital biomarker test for smartphone viewing distance analysis.

REFERENCES

[1] D. K. Martin and B. A. Holden, "A New Method for Measuring the Diameter of the In Vivo Human Cornea," American Journal of Optometry and Physiological Optics, vol. 59, no. 5, pp. 436–441, May 1982.
[2] A. Khatib et al., "AI-Powered Smartphone Diagnostics for Convergence Insufficiency," Journal of Clinical & Translational Ophthalmology, vol. 3, no. 2, p. 8, 2025.
[3] Q. Su et al., "Innovative Myopic Screening Platform Based on Smartphones," Frontiers in Bioengineering and Biotechnology, vol. 13, p. 1678800, 2025.
[4] A. Taore, M. Tiang, and S. C. Dakin, "(The Limits of) Eye-Tracking with iPads," Journal of Vision, vol. 24, no. 7, p. 1, 2024.
[5] Z. Li et al., "Look into My Eyes: Fine-Grained Detection of Face–Screen Distance on Smartphones," in Proc. 12th Int. Conf. Mobile Ad-Hoc and Sensor Networks (MSN), 2016, pp. 258–265.
[6] M. C. Temneanu, C. Donciu, and E. Serea, "Distance Measurement Between a Camera and a Human Subject Using Statistically Determined Interpupillary Distance," AppliedMath, vol. 5, no. 3, p. 118, 2025.
[7] Z. Pirayesh et al., "A Deep Learning Framework to Scale Linear Facial Measurements to Actual Size Using Horizontal Visible Iris Diameter: A Study on an Iranian Population," Scientific Reports, vol. 13, p. 13755, 2023.
[8] J. P. Driessen, H. Vuyk, and J. Borgstein, "New Insights into Facial Anthropometry in Digital Photographs Using Iris Dependent Calibration," International Journal of Pediatric Otorhinolaryngology, vol. 75, no. 4, pp. 579–584, 2011.
[9] Y. Lei et al., "Quantifying the Impact of Motion on 2D Gaze Estimation in Real-World Mobile Interactions," arXiv preprint, 2025. [Online]. Available: https://arxiv.org/abs/2502.10570. Accessed: May 11, 2026.
[10] D. Shmunk, "Eye Pose Estimation and Tracking Using Iris as a Base Feature," IEEE Open Journal on Immersive Displays, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11189046. Accessed: May 11, 2026.

[11] Y. Sharrab et al., "Analysis and Evaluation of ToN IoT Windows and Garage Door Datasets," in Proc. 2025 5th Intelligent Cybersecurity Conf. (ICSC), 2025.
[12] Y. Bababekova et al., "Font Size and Viewing Distance of Handheld Smart Phones," Optometry and Vision Science, vol. 88, no. 7, pp. 795–797, 2011.
[13] J. Long et al., "Viewing Distance and Eyestrain Symptoms with Prolonged Viewing of Smartphones," Clinical and Experimental Optometry, vol. 100, no. 2, pp. 133–137, 2017.