Towards Affordable Smartphone Eye Tracking for Nystagmus Analysis and Monitoring

Benjamin Duvieusart[1,2], Miguel Xochicale[3], Diego Kaski[2], and Terence S. Leung[1]

***Abstract*— Nystagmus, an involuntary oscillatory eye movement, is an important diagnostic marker for various neurological and vestibular disorders. However, current clinical eye tracking systems are expensive, require specialized equipment and training, thus limiting their usability at scale. This study proposes a smartphone-based eye tracking pipeline for nystagmus analysis, utilizing a deep learning-based segmentation model and a circle-fitting algorithm to estimate pupil position. Two variations of U-Net segmentation models — a 3-class (sclera, iris, background) and a 4-class (sclera, iris, pupil, background) U-Net — were trained on publicly available MOBIUS dataset. Then applied and evaluated on 12 optokinetic nystagmus (OKN) videos with varying gaze positions. On the test set the 3-class model performed best across cross entropy loss, intersection over union loss, and DICE score metrics (0.051, 0.081, and 0.956 respectively), compared to the 4-class model's 0.068, 0.162, and 0.903 respectively. On the tracking task, when evaluated against ground truth tracking measured using the ICS Impulse, both methods were able to capture the nystagmus movements reasonably well (avg R^2 of 0.755 and 0.768 for 3 and 4-class methods across 12 OKN tests). No statistical difference in the performance of the 3-class model compared to the 4-class model on R^2 (p-value 0.80). These findings highlight the potential of smartphone-based eye tracking as a cost-effective tool for objective nystagmus assessment, with further refinements needed to improve tracking performance. This approach could enhance diagnostic accuracy and accessibility in clinical settings.**

I. INTRODUCTION

Eye movements have been the subject of much research, from gaming to medicine they can carry significant information [1]. This is particularly true in medical contexts where eye movements have been used to evaluate conditions such as social disabilities in toddlers [2], poor sleep quality [3], and schizophrenia [4]. A particularly medically relevant eye movement is nystagmus — a small involuntary oscillatory movement of the eye typically consisting of two saccades, a slow pathological drift followed by a fast corrective jerk towards the desired direction of gaze, giving a sawtooth-like trace [5], [6]. Nystagmus is associated with a large range of underlying conditions, from benign paroxysmal positional vertigo [7] to life-threatening strokes [8] leading to significant prevalence in the general population (24 per 10,000 [9]).

Although not sufficient when considered alone, a detailed analysis of the variations between the types of nystagmus is instrumental in the correct diagnosis of the underlying cause [10]. This is particularly important in EDs where 3-5% of incoming patients require eye movement interpretation for correct diagnosis [11]. Accurate analysis and efficient triaging at this stage would allow patients to receive appropriate care immediately without delay. However, while nystagmus is an effective diagnostic differentiator in the hands of specialists, studies have shown that non-specialist physicians' low confidence and inadequate training frequently lead to incorrect interpretation of nystagmus and other eye movements [12]. This could be largely avoided by providing clinicians with adequate eye tracking tools for quantified and objective analysis of nystagmus [6].

The aim of this study is to put forth a novel eye tracking pipeline which uses only a smartphone camera to accurately track the small rapid eye movements associated with nystagmus. Two variations of this pipeline are considered: a 3-class segmentation U-Net classifying the periocular area (i.e. background), the sclera, and the iris including the pupil; this is then fed into a circle fitting which uses the iris/sclera boundary as input. The second variation uses a 4-class U-Net where the iris and pupil are classified separately; here, the pupil/iris boundary is used as input to the circle fitting algorithm. The center of the fitted circle is taken to be the center of the pupil hence allowing tracking of eye movements.

We aim to evaluate the performance of this technique on both neutral (eye looking forward) and eccentric (eyes looking left/right) positions using OKN stimulus with varying stimulus speeds, participant positions, and lighting conditions. As this is an exploratory study, the scope is limited to tracking horizontal saw-tooth jerk nystagmus in one individual. Further work, such as application to different nystagmus types or analysis of the output waveform, is outside the scope of this study but is relevant to the larger work concerning the development of an end-to-end nystagmus analysis application to support clinicians' diagnoses.

A. Nystagmus Tracking

The fundamental technology needed for nystagmus analysis is well-established. Specialist eye trackers (e.g. ICS Impulse [13], EyeSeeCam [14]) using infra-red cameras are considered to be the clinical gold standard. However, their

[1]B. Duvieusart and T. S. Leung are with the Smartphone Imaging Group, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, UK (email: benjamin.duvieusart.22@ucl.ac.uk)

[2]B. Duvieusart and D. Kaski are with the SENSE Research Unit, Queen Square Institute of Neurology, University College London, London WC1N 3BG, UK

[3]M. Xochicale is with the Research Software Development Group, Advanced Research Computing Center, University College London, London WC1V 6LJ, UK

This work is supported by the EPSRC-funded UCL center for Doctoral Training in Intelligent, Integrated Imaging in Healthcare (i4health) [EP/S021930/1].

widespread use is severely hampered by expensive costs — up to US $40,000 — and the need for a trained expert [11]. Cheaper versions of these devices have been developed [15], [16], but still require custom hardware or are not portable, making them inaccessible at scale. Commercial eye tracking hardware made for non-medical purposes — such as gaming (e.g. Tobii [17]) — can be cheaper (around US $350) but typically are designed for desktop computers with limited tracking range, lack the necessary accuracy, and/or require custom software which does not support medical tests.

In recent years, various types of machine learning models using smartphone imaging have emerged as potential competitors to the established technology [11]. Challenges have already been encountered, particularly regarding the low frame rates of 30-60 Hz of typical smartphone videos, which are significantly lower than the 200 Hz minimum recommended frame rate for accurate detection of peak saccade velocities and subtle movements [6]. The other key challenge is the significant reduction in eye tracking accuracy when the eye is in eccentric positions, when the amount of the iris and pupil visible to the camera is reduced [18]. This is particularly relevant to nystagmus, as eye oscillations are expected to change behavior depending on the gaze direction [8]. Other salient challenges include different lighting conditions, performance discrepancies between individuals, camera resolutions, and variations in phone software [18]. Examining the current literature reveals that approaches can generally be split between infrared and visible light tracking.

B. Infrared Tracking

Most commercial eye trackers use infrared (IR) or near-IR light as the strong retinal reflection under IR light gives a clear pupil outline [19], [20]. When combined with the corneal reflections, this method enables highly accurate gaze estimation [21], [20]. Although not standard, some high-end phones (e.g. iPhones) have integrated IR emitters and sensors, prompting research which leverages these tools [11], [22], [23], [24]. The larger contingent of IR-based smartphone tracking rely on Apple's integrated ARKit [25] which directly provides eye and head position output from the IR sensor [23]. Greinacher et al. [23] assessed ARKit's accuracy in a smooth pursuit task, reporting a gaze bias of 3.18° — an order of magnitude larger than dedicated trackers, but still usable for phone-based tracking.

Bastani et al. [11] and Parker et al. evaluated its performance on the head impulse tests (HIT) — a clinical test that involves a small abrupt rotation of the head while tracking the relative motion of the head and eyes to assess the vestibulo-ocular reflex [26]. Both studies detected pathological saccades in HIT, claiming similar accuracy to dedicated medical oculography goggles. A followup study by Parker et al. [10] evaluated their system for optokinetic nystagmus (OKN) — normal physiological movement induced by visual stimulus. Their algorithm performed well in the horizontal nystagmus but struggled tracking the vertical eye movements.

Despite its convenience, ARKit suffers from black box processing and OS limitations [23]; preventing custom tracking algorithms. For instance, Brousseau et al. [22] use a Huawei phone to access raw IR sensor data,used as input to a machine-learning feature extractor and 3D model to achieve a superior gaze bias of 0.72°. By far surpassing ARKit-based algorithms, but implementable only on a handful of devices. Even overlooking the other significant challenge specific to IR-based tracking (e.g. sunlight affecting the sensor due to the IR light it contains [20]) the sparsity of IR emitters and sensors in smartphones fundamentally prevents these algorithms from running on a large scale [22] and incentivises the switch to visible light [1], [23].

C. Visible Light Tracking

While more accessible, visible light tracking faces more immediate challenges such as the pupil being hard to differentiate from the iris in dim lighting conditions or interference from specular reflections [27]. Nevertheless, studies have shown very promising results.

Valliappan et al [1] used a front-facing video and eye corner positions as input to a deep learning model to track gaze on a smartphone screen, to achieve a 0.6-1° error with 100 frames of calibration - matching the performance of dedicated IR eye trackers. Reinhardt et al. [28] and Friedrich et al. [29], also used a deep learning approach to track the pupil in a nystagmus task. Reinhardt et al. reported matching an ENT clinician's performance in simple detection, with Friedrich et al. surpassing this by also extracting the speed of the slow phase of the detected nystagmus waveforms. Non-deep learning methods have also been developed. Yang et al. [4] used a circle search algorithm to find the iris boundary and center of the pupil, and successfully tracked reading patterns of schizophrenic patients. van Bonn et al. [30] opted to use gradient vector analysis to find the center of the eye; they were able to detect nystagmus and general characteristics such as direction and shape. "ScreenGlint", by Huang et al. [27], used the reflections off the cornea (i.e. glint) - typically considered interference - as the fundamental eye tracking feature. Essentially mimicking the use of corneal reflection typical of IR based tracking, and achieved an accuracy of 2.44° when tracking the eye gaze on the phone screen.

The limitation common to all tracking studies presented above is the restricted gaze direction, generally to a smartphone screen or neutral gaze position nystagmus. The one study which explicitly compared eccentric performance of smartphone imaging for gaze tracking [18] noted an important drop in performance in eccentric positions. This raises a serious question on the practicality of these algorithms and their applicability to the nystagmus task given how fundamental eccentric tracking is to an accurate nystagmus characterization.

II. Methods and Materials

This paper presents a two-step eye tracking method with two variations, iris and pupil based, using segmentation and circle fitting. The pipelines were evaluated on optokinetic nystagmus (OKN) videos of the left eye in the neutral,

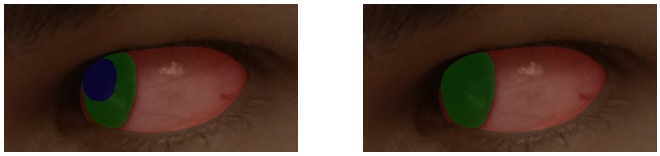
left eccentric, and right eccentric positions in a healthy individual; corresponding tracking of the right eye was measured using the ICS Impulse device [13] as a ground truth comparable.

A. U-Net model

The U-Net model architecture used for the segmentation is the unmodified model presented by Ronneberger et al. [31], who developed the architecture to specifically leverage data augmentation on a smaller training dataset. U-Nets have been shown to largely outperform conventional convolutional neural networks in image segmentation tasks due to improved localization ability [32].

For both the 3 and 4-class variations of the U-Net, the model is trained, validated, and tested on the MOBIUS dataset [33], [34], [35]. This dataset contains 16,717 eye images from 100 subjects. For each subject, 168 images were collected: for each eye, 2 sets of pictures were taken on 3 phones in 3 lighting conditions (natural sunlight, indoor lights, poor lighting) and in 4 gaze directions (neutral, left eccentric, right eccentric, upwards), with one additional 'bad' photo per gaze direction with randomized lighting. Of the total dataset, only 3,559 — belonging to 35 subjects — have matching annotations for the sclera, iris, pupil, and periocular region/background (see Fig 1a). Only these annotated images are used to train the models; they are randomly split into 75% (2,671) training, 15% (533) validation, and 10% (355) testing sets.

To fully leverage the U-Net horizontal and vertical flip, as well as rotations up to $\pm 45^\circ$ are applied to augment data. Afterwards the data is reshaped to 256×256 and color values standardized; for the 3-class model, the iris and pupil segments in the ground truth masks are combined to match the desired output (see Fig 1b). The models' $\sim$23.4M parameters are trained on a single GPU (24GiB memory) for 100 epochs using Python PyTorch libraries, with a learning rate of 0.01, 5 sample training batch size, Adam optimizer, and the epoch checkpoints with the lowest intersection over union (IoU) loss on the validation set are saved.



(a) 4-class: pupil (blue) , iris (green), sclera (red)

(b) 3-class: iris + pupil (green), sclera (red)

Fig. 1: 4 and 3-class ground truth masks for image `1_1i_Lr_2` from MOBIUS [33], [34], [35], periocular region (i.e. background) not shaded

B. Circle estimation

Circle estimation is not a trivial problem; classical least squares algorithms have no closed form solution and require computationally expensive and unreliable iterative schemes [36]. Bullock [37] proposes an alternative minimization scheme which allows for a closed form solution which is comparatively computationally cheap and stable.

In Bullock's method, N input points are first shifted to a zero mean such that $u_i = x_i - \bar{x}$ and $v_i = y_i - \bar{y}$, and the target circle has center (u_c, v_c) with radius $R = \sqrt{\alpha}$. The resulting minimization criterion S,

$$S(\alpha, u_c, v_c) = \sum_i ((u_i - u_c)^2 + (v_i - v_c)^2 - \alpha) \qquad (1)$$

can be differentiated to impose optimal solution gradient boundary conditions ($\delta S/\delta\alpha = \delta S/\delta u_c = \delta S/\delta v_c = 0$). Letting $S_{uu} = \sum_i u_i^2$, $S_{uv} = \sum_i u_i v_i$, *etc.* the following closed form solution is produced

$$\begin{bmatrix} S_{uu} & S_{uv} \\ S_{uv} & S_{vv} \end{bmatrix} \begin{bmatrix} u_c \\ v_c \end{bmatrix} = \begin{bmatrix} S_{uuu} + S_{uvv} \\ S_{vvv} + S_{uuv} \end{bmatrix} \qquad (2)$$

$$\alpha = u_c^2 + v_c^2 + \frac{S_{uu} + S_{vv}}{N} \qquad (3)$$

Bullock's method is applied for circle estimations using the iris/sclera and pupil/iris boundaries as input points for the 3 and 4-class segmentation methods respectively. The center of the fitted circle is taken as the estimated pupil center position.

C. Optokinetic nystagmus (OKN) video testing

In order to evaluate the performance of the eye tracking, horizontal OKN was induced in a healthy volunteer (no prior history of dizziness; no visual impairments; congruent eye movements) in the neutral, eccentric left, and eccentric right positions. The OKN stimulus consisted of vertical black bars moving rightward on a white background; generated using the Cyclops BalanceEye VOG software [38]. The eye movements of the right eye were recorded by the ICS Impulse device mounted on the head of the participant (see Fig 2). The movements of the left eye were recorded at 30 frames per second (fps) using a Samsung S22 positioned face on to the patients at 1m with a 3x optical zoom, such that the full face was in view of the camera (See Fig. 2). Each gaze position (left eccentric, neutral, and right eccentric) was repeated with OKN stimulus at 20, 40, and 60° per second, to test the range of typical nystagmus slow phase speeds in patients [39]. Additionally, the neutral gaze OKN was repeated under dim lighting conditions, for a total of 12 tests. For the eccentric gaze tests the participant was rotated by 25° using the Difra Minitorque Rotary chair [40].

The stimulus was run for no longer than 25 seconds to avoid patient discomfort. From this a 5-second clip with no blinking artifacts was extracted for pipeline evaluations. The subject's head was not restrained, but they were instructed to keep still. This study received ethical approval from the London-Central Research Ethics Committee (IRAS 321000).

III. Results

A. Segmentation

The DICE score, IoU, and cross-entropy loss of the segmentation model when applied to the MOBIUS test set can be found in Table I. Across all tracked metrics, the 3-class

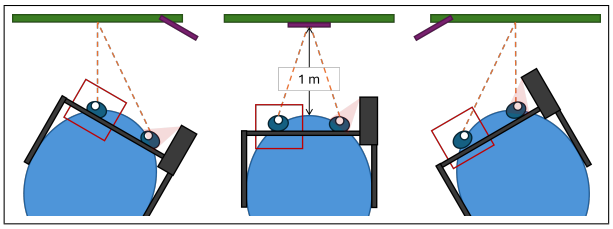


Fig. 2: From left, experimental set up for left eccentric, neutral, and right eccentric OKN stimulation. Larger screen (in green) displayed stimulus, the phone (in purple) recorded left eye (boxed) from a distance of 1m with 3x optical zoom, ICS Impulse [13] headset (in black) record the right eye using infrared tracking.

TABLE I: Cross-Entropy loss, intersection over union (IoU) and DICE score of 3-class and 4-class segmentation models on test set (epoch with lowest validation loss saved).

Model	Cross-Entropy Loss	IoU Loss	DICE Score
3-class (97th epoch)	0.051	0.081	0.956
4-class (82nd epoch)	0.068	0.162	0.903

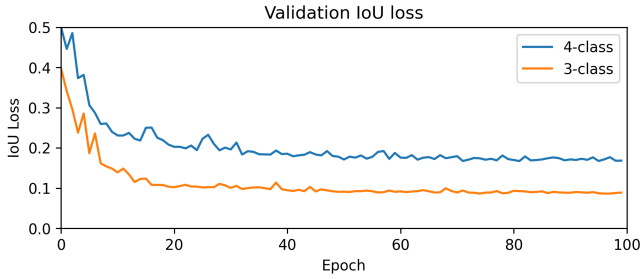
model outperforms the 4-class model, with a particularly large margin in the segmentation-specific IoU loss metric.

This trend is also seen in the evolution of IoU loss and DICE score during the training (see Fig. 3) where the 3-class model systematically outperforms the 4-class segmentation model. A class-wise break down of the IoU loss during training is shown in Fig. 4. These plots show that the drop in performance can be primarily attributed to the segmentation of the pupil, as it has the largest IoU class specific loss in the 4-class pipeline (see Fig.4b). This makes sense given the pupil is known to be significantly more complex to segment, particularly in eccentric positions and poor lighting images found in MOBIUS. This has a knock-on effect on the iris segmentation performance, causing its IoU loss to double from $\sim$0.1 to $\sim$0.2, while sclera and periocular segmentation performance remain constant.

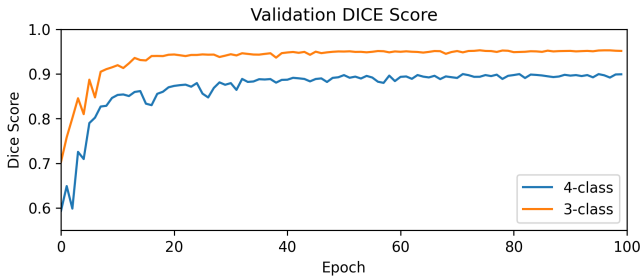
B. Tracking on video of OKN stimulus

After implementing the two models into a pipeline with the circle fitting to get a tracking of the pupil (see sample outputs in Fig 5), the pipelines were tested on twelve 5-second videos (150 frames per video) of OKN movements. The trace of the x-coordinate of the pupil and the ICS tracking data for eccentric right at 40°/sec (i.e. ER40) are shown in Fig. 6. Since the chosen OKN stimulus generated horizontal eye movements, vertical motions are not relevant and hence its tracking was not evaluated.

To better evaluate the accuracy of the segmentation tracking, each timeseries was passed through a linear regression (see Fig 7 for models on ER40 data). The results are time series which match well, all models having a p-value<0.001. However, there are significant variations in the R^2 values of the models ranging from 0.441 (3-class ER60) to 0.936 (3-class ER40), see Table II for all values. Applying the models

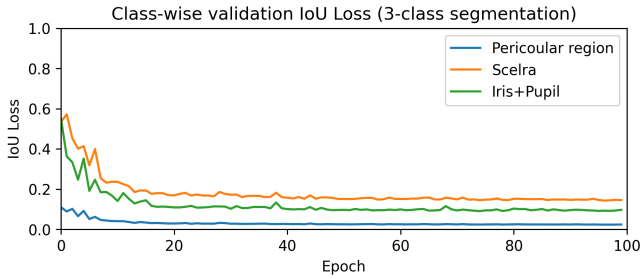


(a) Intersection over union (IoU) loss

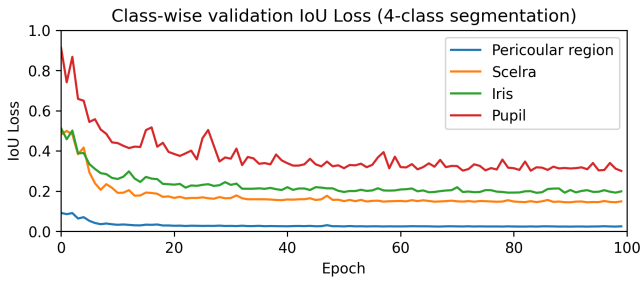


(b) DICE score

Fig. 3: Performance metrics on the validation set for the 3 and 4-class segmentations model during training.



(a) 3-class segmentation



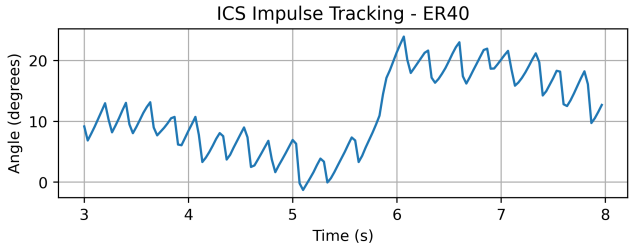
(b) 4-class segmentation

Fig. 4: Class-wise IoU loss on the validation set for the 3 and 4-class segmentation models during training.

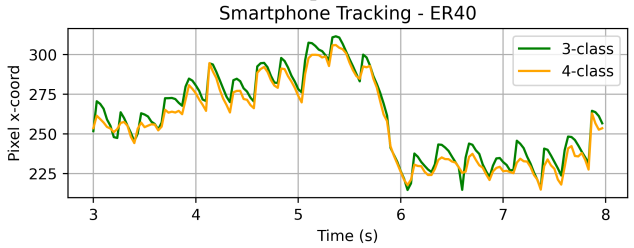
(a) 3-class, iris circled

(b) 4-class, pupil circled

Fig. 5: Segmentation and pupil position estimation for pupil and iris based tracking (frame 20 from OKN test ER40)



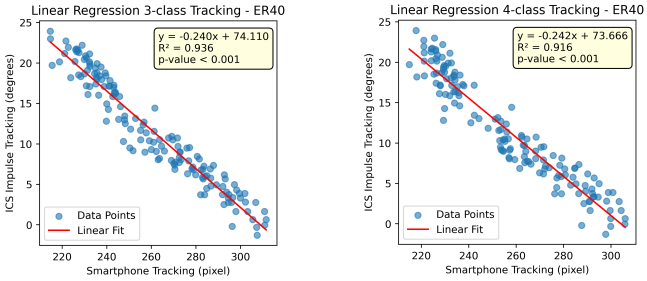
(a) ICS Impulse device



(b) 3 and 4 class segmentation pipeline

Fig. 6: Tracking data from ICS Impulse (in degrees) and both pipeline variations (pupil x-corrdinate) from optokinetic nystagmus test: right eccentric gaze 40°/sec stimulus.

to pixel data transforms the 2D coordinates into an estimated horizontal angle of rotation - see Fig 8 for the transformed timeseries plots for ER40 OKN test



(a) 3-class model (b) 4-class model

Fig. 7: Linear regressions from pupil x-coordinate to rotation angle for both pipeline variations from optokinetic nystagmus test: right eccentric gaze 40°/sec stimulus (ER40).

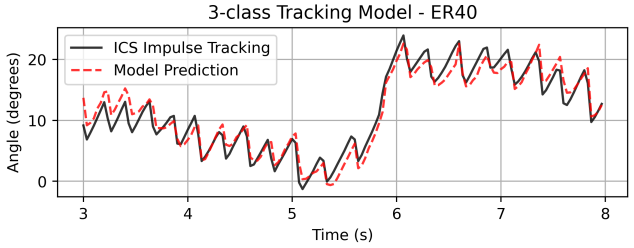
Across over all tests, the 4-class and 3-class perform very similarly. With the former having a slightly better averaged R^2 value (0.77) compared to 3-class (0.75), with no statistical difference found (p-value 0.80). There is some important variability in the accuracy of the tracking between tests, but again no statistical difference was found between the results of matched groups tested - neutral gaze vs eccentric (p-value 0.44), good vs dim lighting conditions (p-value 0.50), leftward gaze vs rightward gaze (p-value 0.99).

IV. Discussion

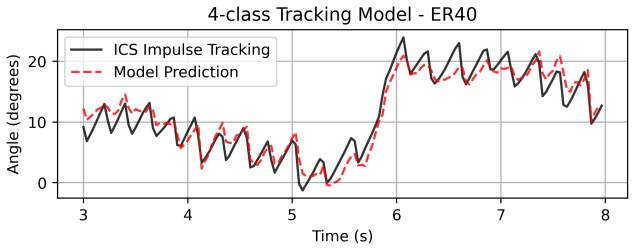
The pipelines presented in this work show promising results when tasked with tracking nystagmus movements — with both pipelines performing remarkably well in certain tests. Despite the variability, no statistically significant difference was found between any subset of tracking results.

TABLE II: R^2 for all linear regressions between smartphone tracking and ground truth (by method and OKN tests), all models have p<0.001. Test acronyms: EL - eccentric left, ER - eccentric right, NG - neutral gaze, dim - low light conditions; 20/40/60 - velocity of the OKN stimulus in °/sec

Test	R2 (3-class)	R2 (4-class)
PG20	0.738	0.782
PG40	0.628	0.678
PG60	0.787	0.680
PG20 (dim)	0.604	0.500
PG40 (dim)	0.888	0.889
PG60 (dim)	0.881	0.842
EL20	0.850	0.783
EL40	0.720	0.742
EL60	0.759	0.831
ER20	0.824	0.802
ER40	0.936	0.916
ER60	0.441	0.770
Mean	*0.755*	*0.786*



(a) 3 class



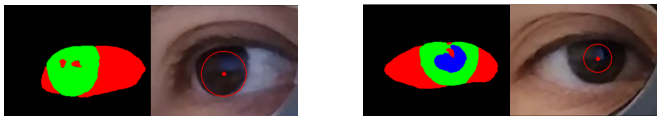
(b) 4 class

Fig. 8: Overlay of ground truth ICS Impulse tracking data and transformed smartphone tracking data for both pipeline variations from optokinetic nystagmus test: right eccentric gaze 40°/sec stimulus.

This goes against the literature, which states that the accuracy of the tracking tends to diminish as eccentricity of the eye position increases.

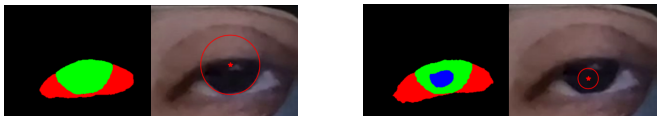
Similarly, no statistically significant difference was found between the 3-class and 4-class methods. This is counter to the initial hypothesis which claimed that the iris-based tracking would out-perform the pupil-based tracking as the lower accuracy of the pupil segmentation would directly translate to poorer tracking. However, the circle estimation algorithm limits these errors; as seen in Fig 9b, where a specular reflection causes poor segmentation of the pupil, but the circle estimation largely compensates for it. Similarly, the specular reflections complicate the segmentation of the iris in Fig 9a, but the circle estimation mitigates its impact. It is

important to emphasize that the study is limited to a single participant, and hence a trend may emerge with a larger dataset. Furthermore, this study only focuses on sections of the OKN test without blinking artifacts, but there is evidence that the iris-based tracking may be better suited to handle blinking interference (see Fig. 10).



(a) 3-class, iris circled (b) 4-class, pupil circled

Fig. 9: Interference due to specular reflections on the iris for both methods.



(a) 3-class, iris circled (b) 4-class, pupil circled

Fig. 10: Interference due to blinking, 3-class better able to handle the partial occulsion of the eye.

While segmentation inaccuracies account for some of the errors between both the smartphone and ICS Impulse tracking, there are several other potential sources of error. Firstly, while the participant was asked to limit head movement, their head was not restrained and inevitably some movement was seen. Given the smartphone is static and the ICS impulse device is attached to the head, these movements will only be visible to the smartphone camera, hence preventing the signals from being perfectly aligned. Furthermore, the linear interpolations necessary to down-sample the ground truth data to 30fps can also inject errors into the ground truth signal. Adding to this, since the smartphone and ICS Impulse are completely independent devices, the frames may not be in phase, even if down sampled to the same frame rate - hence introducing further error. Lastly, the relationship between the 2D linear motion of the pupil on a video and the 3D rotation in space is non-linear. Hence, there are approximation errors when attempting to use the linear regression model to map the pixel position data to the rotation data from the ICS Impulse. However, given in this experiment the gaze is in a limited range for any given test and at small angles the rotations can be approximated linearly, the error can be assumed to be minimal.

As this is exploratory work, the opportunity for future work is extensive. To address some of the limitations stated above, improvements can include tracking the movement of the head to isolate purely eye movement data, as well as increasing to 60fps smartphone video. Also developing a more robust conversion from 2D positional data to 3D rotational data, which captures the non-linear relationship between the two. Alternatively expanding the circle estimation to compensate for the distortion of the circle shape of the iris in eccentric positions, or even including hybrid 3D model-based tracking as implemented by Brousseau et al [22]. Furthermore, increasing the size of the dataset to include more participants with increased diversity (ethnicity/age/eye color/medical conditions) is key to the validation of these methods. Augmenting the training dataset to contain data with prominent specular reflections is also fundamental to the long-term success of these methods.

V. Conclusions

The pipelines presented in this exploratory study are largely able to capture the nystagmus movements, though notable improvements need to be made. The iris-based tracking method benefits from a simple segmentation problem with clear high contrast boundaries compared to the 4-class segmentation. Furthermore, the circle estimation step provides a robustness which compensates for many segmentation errors. Despite the current shortcomings of the work, the preliminary results suggest that the two-step method addresses some challenges facing visible eye light tracking (e.g. limitation of specific devices, sensitivity to light conditions and gaze direction). Efficient application of this method would provide affordable eye tracking software to virtually anybody with a smartphone, eliminating the need for specialized hardware. Embedding it into a clinical diagnostic system for nystagmus would aid non-specialist doctors when triaging and diagnosing patients, lowering the interpretative burden placed on them. By providing clinicians with the appropriate tool, there is potential to greatly reduce the number of nystagmus-related misdiagnoses. Looking beyond nystagmus, at other clinically relevant eye movements, low-cost and widely accessible eye tracking technology has the possibility of providing doctors with more information for improved diagnoses.

Acknowledgment

The images from the SBVPI DATASET (i.e MOBIUS) used in this work have been provided by the University of Ljubljana, Slovenia [33], [34], [35].

References

[1] N. Valliappan, N. Dai, E. Steinberg, J. He, K. Rogers, V. Ramachandran *et al.*, "Accelerating eye movement research via accurate and affordable smartphone eye tracking," *Nature Communications*, vol. 11, no. 1, Sep. 2020. [Online]. Available: http://dx.doi.org/10.1038/s41467-020-18360-5

[2] W. Jones, K. Carr, and A. Klin, "Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder," *Archives of General Psychiatry*, vol. 65, no. 8, p. 946, Aug. 2008. [Online]. Available: http://dx.doi.org/10.1001/archpsyc.65.8.946

[3] L. S. Stone, T. L. Tyson, P. F. Cravalho, N. H. Feick, and E. E. Flynn-Evans, "Distinct pattern of oculomotor impairment associated with acute sleep loss and circadian misalignment," *The Journal of Physiology*, vol. 597, no. 17, p. 4643–4660, Aug. 2019. [Online]. Available: http://dx.doi.org/10.1113/JP277779

[4] H. Yang, L. He, W. Li, Q. Zheng, Y. Li, X. Zheng, and J. Zhang, "An automatic detection method for schizophrenia based on abnormal eye movements in reading tasks," *Expert Systems with Applications*, vol. 238, p. 121850, Mar. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2023.121850

[5] R. J. Leigh and D. S. Zee, *The Neurology of Eye Movements.*, 4th ed. Oxford University Press, 2006.

[6] J.-R. Rizzo, M. Beheshti, W. Dai, and J. C. Rucker, "Eye movement recordings: Practical applications in neurology," *Seminars in Neurology*, vol. 39, no. 06, p. 775–784, Dec. 2019. [Online]. Available: http://dx.doi.org/10.1055/s-0039-1698742

[7] R. Palmeri and A. Kumar, "Benign paroxysmal positional vertigo," in *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing, Dec. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK470308/

[8] K. Ioannides, P. Tadi, and I. A. Naqvi, "Cerebellar infarct," in *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing, 2022. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29261863/

[9] N. Sarvananthan, M. Surendran, E. O. Roberts, S. Jain, S. Thomas, N. Shah *et al.*, "The prevalence of nystagmus: The leicestershire nystagmus survey," *Investigative Opthalmology &; Visual Science*, vol. 50, no. 11, p. 5201, Nov. 2009. [Online]. Available: http://dx.doi.org/10.1167/iovs.09-3486

[10] P. B. Bastani, H. Rieiro, S. Badihian, J. Otero-Millan, N. Farrell, M. Parker *et al.*, "Quantifying induced nystagmus using a smartphone eye tracking application (eyephone)," *Journal of the American Heart Association*, vol. 13, no. 2, Jan. 2024. [Online]. Available: http://dx.doi.org/10.1161/JAHA.123.030927

[11] P. B. Bastani, S. Badihian, V. Phillips, H. Rieiro, J. Otero-Millan, N. Farrell *et al.*, "Smartphones versus goggles for video-oculography: current status and future direction," *Research in Vestibular Science*, vol. 23, no. 3, p. 63–70, Sep. 2024. [Online]. Available: http://dx.doi.org/10.21790/rvs.2024.009

[12] A. E. Quimby, E. S. H. Kwok, D. Lelli, P. Johns, and D. Tse, "Usage of the HINTS exam and neuroimaging in the assessment of peripheral vertigo in the emergency department," *Journal of Otolaryngology - Head & Neck Surgery*, vol. 47, no. 1, Sep. 2018. [Online]. Available: https://doi.org/10.1186/s40463-018-0305-8

[13] Natus. Ics impulse - vestibular testing. [Online]. Available: https://natus.com/sensory/ics-impulse/

[14] Interacoustics. Eyeseecam vhit. [Online]. Available: https://www.interacoustics.com/balance-testing-equipment/eyeseecam-vhit

[15] J. Zimmermann, Y. Vazquez, P. W. Glimcher, B. Pesaran, and K. Louie, "*Oculomatic*: High speed, reliable, and accurate open-source eye tracking for humans and non-human primates," *Journal of Neuroscience Methods*, vol. 270, p. 138–146, Sep. 2016. [Online]. Available: http://dx.doi.org/10.1016/j.jneumeth.2016.06.016

[16] J. Casas and C. Chandrasekaran, "openeyetrack - a high speed multi-threaded eye tracker for head-fixed applications," *Journal of Open Source Software*, vol. 4, no. 42, p. 1631, Oct. 2019. [Online]. Available: http://dx.doi.org/10.21105/joss.01631

[17] Tobii. Tobii eye tracker 5. [Online]. Available: https://gaming.tobii.com/product/eye-tracker-5/?srsltid=AfmBOoo6p2hrPmiBQZbDytFj1j8vTr9Y-p4Nv71oV4yWFndse7XdGU0F

[18] T. M. Parker, S. Badihian, A. Hassoon, A. S. Saber Tehrani, N. Farrell, D. E. Newman-Toker, and J. Otero-Millan, "Eye and head movement recordings using smartphones for telemedicine applications: Measurements of accuracy and precision," *Frontiers in Neurology*, vol. 13, Mar. 2022. [Online]. Available: http://dx.doi.org/10.3389/fneur.2022.789581

[19] F. Hermens and D. İren, "Eye tracking eyes on the prize: Eye tracking for business value," *Actionable Data Science Whitepaper Series*, vol. 2, no. 1, 02 2023.

[20] K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, England: Oxford University Press, Sep. 2011, ch. 2.

[21] E. Guestrin and M. Eizenman, "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1124–1133, 2006.

[22] B. Brousseau, J. Rose, and M. Eizenman, "Hybrid eye-tracking on a smartphone with cnn feature extraction and an infrared 3d model," *Sensors*, vol. 20, no. 2, p. 543, Jan. 2020. [Online]. Available: http://dx.doi.org/10.3390/s20020543

[23] R. Greinacher and J.-N. Voigt-Antons, *Accuracy Assessment of ARKit 2 Based Gaze Estimation*. Springer International Publishing, 2020, p. 439–449. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-49059-1_32

[24] T. M. Parker, N. Farrell, J. Otero-Millan, A. Kheradmand, A. McClenney, and D. E. Newman-Toker, "Proof of concept for an "eyephone" app to measure video head impulses," *Digital Biomarkers*, vol. 5, no. 1, p. 1–8, Dec. 2020. [Online]. Available: http://dx.doi.org/10.1159/000511287

[25] Apple. Framework: Arkit. [Online]. Available: https://developer.apple.com/documentation/arkit

[26] M. Gottlieb, G. D. Peksa, and J. N. Carlson, "Head impulse, nystagmus, and test of skew examination for diagnosing central causes of acute vestibular syndrome," *Cochrane Database of Systematic Reviews*, vol. 2022, no. 8, Aug. 2022. [Online]. Available: http://dx.doi.org/10.1002/14651858.CD015089

[27] M. X. Huang, J. Li, G. Ngai, and H. V. Leong, "Screenglint: Practical, in-situ gaze estimation on smartphones," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. ACM, May 2017. [Online]. Available: http://dx.doi.org/10.1145/3025453.3025794

[28] S. Reinhardt, J. Schmidt, J. Schneider, E. Schulte, C. Schüle, M. Leuschel, and J. Schipper, "Smartphone-based videonystagmography using artificial intelligence," *Current Directions in Biomedical Engineering*, vol. 9, no. 1, p. 528–531, Sep. 2023. [Online]. Available: http://dx.doi.org/10.1515/cdbme-2023-1132

[29] M. U. Friedrich, E. Schneider, M. Buerklein, J. Taeger, J. Hartig, J. Volkmann, R. Peach, and D. Zeller, "Smartphone video nystagmography using convolutional neural networks: Convng," *Journal of Neurology*, vol. 270, no. 5, p. 2518–2530, Nov. 2022. [Online]. Available: http://dx.doi.org/10.1007/s00415-022-11493-1

[30] S. M. van Bonn, S. P. Behrendt, B. L. Pawar, S. P. Schraven, R. Mlynski, and T. Schuldt, "Smartphone-based nystagmus diagnostics: development of an innovative app for the targeted detection of vertigo," *European Archives of Oto-Rhino-Laryngology*, vol. 279, no. 12, p. 5565–5571, Apr. 2022. [Online]. Available: http://dx.doi.org/10.1007/s00405-022-07385-9

[31] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015. [Online]. Available: https://arxiv.org/abs/1505.04597

[32] A. Oğuz and O. F. Ertuğrul, *Introduction to deep learning and diagnosis in medicine*. Elsevier, 2023, p. 1–40. [Online]. Available: http://dx.doi.org/10.1016/B978-0-323-96129-5.00003-2

[33] M. Vitek, M. Bizjak, P. Peer, and V. Struc, "Ipad: Iterative pruning with activation deviation for sclera biometrics," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 8, p. 101630, September 2023. [Online]. Available: https://doi.org/10.1016/j.jksuci.2023.101630

[34] M. Vitek, A. Das, D. R. Lucio, L. A. Zanlorensi, D. Menotti, J. N. Khiarak *et al.*, "Exploring bias in sclera segmentation models: A group evaluation approach," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 190–205, 2023.

[35] M. Vitek, A. Das, Y. Pourcenoux, A. Missler, C. Paumier, S. Das *et al.*, "Ssbc 2020: Sclera segmentation benchmarking competition in the mobile environment," in *2020 IEEE International Joint Conference on Biometrics (IJCB)*, 2020, pp. 1–10.

[36] N. Chernov and C. Lesort, "Least squares fitting of circles and lines," 2003. [Online]. Available: https://arxiv.org/abs/cs/0301001

[37] R. Bullock, "Least-squares circle fit," Oct. 2006. [Online]. Available: https://www.scribd.com/document/224546035/Least-Squares-Circle-Fit-Bullock-2006

[38] C. Medtech. Balanceeye vog - videonystagmography. [Online]. Available: https://www.cyclopsmedtech.com/videonystagmography-balanceeye-vng

[39] G. Wei, S. H. Lafortune, D. J. Ireland, and R. M. Jell, "Stimulus velocity dependence of human vertical optokinetic nystagmus and afternystagmus," *Journal of vestibular research : equilibrium & orientation*, vol. 2, no. 2, p. 99–106, 1992.

[40] I. Difra. Minitorque rotary chair. [Online]. Available: https://difra.be/shop/00005580-minitorque-rotary-chair-programmed-with-reclining-backrest-to-the-horizontal-4101?category=66#attr=