Preprints.org

Article

Validation of an Eye Tracking Algorithm Based on Smartphone Videos: A Pilot Study

Wanzi Su , Damon Hoad , Leandro Pecchia , Davide Piaggio [*]

Article

Validation of an Eye Tracking Algorithm Based on Smartphone Videos: A Pilot Study

Wanzi Su [1], Damon Hoad [2], Leandro Pecchia [3] and Davide Piaggio [1,*]

[1] School of Engineering, University of Warwick, Library Road, CV4 7AL; wanzi.su@warwick.ac.uk

[2] Warwick Medical School, University of Warwick, Gibbet Hill Road, CV4 7AL; damon.hoad@swft.nhs.uk

[3] Department of Engineering, University Campus Bio-Medico of Rome, Via Alvaro del Portillo 21, 00128 Roma, Italy; l.pecchia@warwick.ac.uk

[*] Correspondence: davide.piaggio@warwick.ac.uk.

Abstract: This project aimed to develop and validate an efficient eye tracking algorithm suitable for the analysis of images captured in the visible light spectrum using a smartphone camera. In particular, the investigation primarily focused on comparing two algorithms, which were named CHT_TM and CHT_ACM, abbreviated from the core functions: Circular Hough Transform (CHT), Active Contour Models (ACM), and Template Matching (TM). In essence, CHT_TM significantly improved the running speed of the CHT_ACM algorithm, with not much difference in the resource consumption, and improved the accuracy on the x axis. CHT_TM achieved a reduction by 79% of the execution time. CHT_TM performed with an average mean percentage error of 0.34% and 0.95% in the x and y direction across the 19 manually validated videos, compared to 0.81% and 0.85% for CHT_ACM. Different conditions, like manually opening the eyelids with a finger versus without a finger, were also compared across four different tasks. This paper shows that applying TM improves the original eye tracking algorithm with CHT_ACM. The new algorithm has the potential to help the tracking of eye movement, which can facilitate the early screening and diagnosis of neurodegenerative diseases.

Keywords: eye tracking; eye movement; neurodegenerative condition; biosignal processing

1. Introduction

Eye tracking is an increasingly popular methodology across various applications. What captivates numerous researchers is its time-efficiency, cost-effectiveness, reliability, objectiveness, quantifiability, easiness-to-conduct, non-invasiveness, and accessibility. By leveraging techniques such as shape-based, appearance-based, feature-based, and hybrid methods, eye-tracking technologies capture and analyse the movements of key ocular components, including the pupil, cornea, sclera, iris, and retina [1-3]. These movements rely on the physiological mechanism controlled by the central nervous system, involving the brainstem, cerebellum, basal ganglia, and cerebral cortex [4]. Consequently, eye tracking measurements provide a valuable window into the brain function and cognitive processes.

Eye movements are categorized into four basic types: saccades, smooth pursuit movements, vergence movements, and vestibulo-ocular movements [5]. Even during fixation, the eye was not completely still due to fixational eye movements, which are composed of three main kinds of smaller movements: microsaccades, drift, and tremor [6]. Microsaccades and drifts have an average size of approximately 6' arc, with drifts exhibiting an average velocity of about 1' arc per second [7]. These eye movements can influence each other and, sometimes, occur simultaneously. For example, in Miki and Hirata's study [8], microsaccades were observed in parallel with the ongoing vestibulo-ocular reflex. Among the aforementioned eye features, saccades and fixations are the commonly analyzed in eye-tracking studies [9], although pursuit and pupil dilation are also frequently investigated [10-12].

A systematic literature review has summarized the eye tracking methods for enhancing productivity and reading abilities [13], categorizing studies into four areas: general eye metrics and tracking, comprehension measurement, attention measurement, and typography and typesetting. Notably, the eye metrics such as fixation duration, saccade velocity, and blink rates have been linked to the reading behaviours including speed reading and mind wandering. Although findings across studies have limited generalizability, the integration of machine learning with eye-tracking technologies has shown promising potential in understanding reading behavior, attention, and cognitive workload.

1.1. Eye Tracking in Healthcare

Eye tracking has demonstrated significant potential in healthcare applications, particularly in telemedicine [6], assistive technologies (e.g., wheelchair control [14]), the assessment of neurological and vestibular disorders such as acute-onset vertigo [15]. It has been used to evaluate memory in conditions such as Amyotrophic Lateral Sclerosis (ALS) and late stages of Alzheimer's Disease (AD), without relying on intact verbal or motor functions, which may not prevail or be reliable in patients of late stages. By minimizing the need for complex instructions or decision-making, eye tracking provides an accessible means of evaluation [9].

Eye-tracking features have been identified as potential biomarkers for multiple neurodegenerative disorders. For instance, impaired convergence is observed in Parkinson's disease (PD) [16], while atypical saccades characterize preclinical and early-stage Huntington's disease [17, 18]. Neurological abnormalities could also be detected in long COVID patients. García Cena et al. proved that long COVID patients were affected by altered latency for centrally directed saccades, impaired memory-guided saccades, and increased latencies for antisaccades [19].

1.2. Challenges and Limitations

Despite these promising applications, there are still concerns and challenges. The experimental constraints and demands of certain tasks can lead to changes in microsaccade behaviour, potentially producing responses that do not occur in natural conditions. For example, Mergenthaler and Engbert observed more microsaccades in the fixation task than the naturalistic scene viewing task, suggesting those responses might be task-specific [20]. However, in real-world settings, human behavior will involve a range of different tasks (with both fixation and free viewing present), resulting in a combination of different types of eye movements. This highlights the difficulty of replicating experimental conditions outside controlled laboratory settings, with the use of special eye-tracking equipment.

Therefore, barriers to the widespread adoption of these examinations in daily life or at home require patients to visit a clinic or laboratory for eye-tracking assessments. Moreover, the expensive price and high-calibration requirements, which need specific training, present further obstacles [6, 21-23]. While some methods are accurate and effective, they often require skilled professionals to operate, making the process laborious and time-consuming [24].

To address these limitations, researchers have explored the use of consumer-grade electronic devices like smartphones, video recorders, laptops, and tablets. However, these approaches introduce new challenges, particularly regarding head stabilization, which significantly affects tracking accuracy. Solutions such as external head restraints and head-mounted equipment have been proposed to mitigate these issues [6, 21, 25]. Another solution is to use some indicators (e.g., a sticker) to represent the head movement and then subtract it from the eye movement [22].

1.3. State of the Art

Eye tracking is sometimes combined with functional magnetic resonance imaging (fMRI) to investigate neurological deficits [26, 27]. Compared to fMRI, one of the advantages of automated eye feature evaluation is its ability to simultaneously collect multiple measures during the same session

and obtain different information for specific analysis and diagnosis [9]. High-frame-rate cameras further enhance detection capabilities, revealing some covert saccades that are difficult to see with the naked eye.

In order to collect and synthesise the available evidence, in a chapter of 'Eye Movement Research', Hutton [28] introduced the range of eye tracking technologies currently available. There are two common tests to measure eye movement: video recording (video-oculography or VOG) and electromyography (electro-oculography or EOG). The less common approaches are through limbus reflection, Dual Purkinje Image (DPI), and scleral search coils but they will not be discussed in detail here.

EOG uses a pair of electrode patches and measures the activity of two cranial nerves, which respectively connect the brain with the inner ear and the eye muscles. VOG uses a camera to record the eye movements and calibrate the accuracy. According to Corinna Underwood [29], there is a general consensus that the accuracy of directly measuring eye movements is higher than indirectly measuring movements via eye muscle motion. Hutton further highlights that EOG's primary drawback is its susceptibility to spatial inaccuracies due to drift artifacts over time, caused by impedance changes in the electrodes [28]. While most studies favour VOG over EOG, there is insufficient evidence to conclusively recommend one method over the other. VOG, though more precise, is relatively expensive and requires the eyes to remain open. Meanwhile, using EOG requires careful consideration in practice as certain medications (e.g., sedatives) or medical electrical equipment (e.g., cardiac pacemaker) would affect or interfere with the electromyography function [29].

VOG systems require head stability due to vestibulo-ocular movements that cause involuntary eye motion (to maintain the gaze fixation on a location, the head motion would cause the eye movement in the opposite direction [6]). Solutions to this issue fall into two categories: stationary VOG, where the camera remains fixed and external supports such as chin rests stabilize the head, and mobile VOG, where the camera moves with the head (e.g., in wearable eye-tracking glasses). Mobile VOG is generally better suited for real-world tasks, whereas stationary VOG is typically used in research settings that require specific stimuli. However, the need for mobile VOG to be lightweight and wearable imposes limitations on its specifications. As a result, its data quality, sampling rates, and spatial resolution are generally lower than those of high-specification stationary VOG systems [28].

Infrared-based eye tracking, known as infrared oculography (IOG), is a common VOG method that tracks corneal reflection and pupil center [9, 30, 31]. Under infrared illumination, the pupil appears as a comparatively dark circle compared to the rest of the eye (e.g. the iris and sclera) [28]. However, this approach assumes that the pupils dilate and constrict symmetrically around their centre, which is not always accurate under varying luminance conditions according to Wildenmann et al. [32]. Consequently, researchers continue to explore alternative methods that do not rely on infrared illumination.

Table 1 provides a comparative analysis of different eye-tracking technologies, including fMRI, VOG, and EOG, highlighting their respective advantages and limitations [6, 9, 28, 29].

Table 1. A comparison between the advantages and disadvantages of different state of the art technologies.

Technologies			Advantages	Disadvantages
fMRI and other technology			**For specific purposes**	Without multiple eye measures
Eye Tracking	EOG		Cheap, well established, readily available, high temporal resolution, no need for eyes to be open	Low accuracy and safety
	VOG	Stationary	High specifications	Need stimuli (likely in research scenarios)
		Mobile	No need for head stability, real world tasks, lightweight	Low data quality, low sampling rates and poor spatial resolution

Infra-red Cameras (pupil centre)	Not influenced by the environment light levels	Pupil change error (pupils do not dilate or constrict symmetrically around their centre)

1.4. Smartphone Based Eye Tracking

Modern smartphones present an excellent opportunity for accessible and cost-effective eye tracking. Equipped with high-resolution cameras and increasing on-device processing power, they are well-suited for edge computing, i.e., bringing the computation closer to the devices where data is gathered rather than cloud computing. This allows smartphones to function as both data acquisition and processing units, enabling real-time analysis, reducing latency, and preserving user privacy, which are particularly valuable in medical diagnostics. Their internet connectivity further supports remote deployment and monitoring when needed. Previous research has demonstrated the feasibility of using smartphones as an alternative to commercial-grade equipment for pupillometry [23]. Their affordability, ease of use, and resilience to power supply make them an attractive solution for large-scale, low-resource applications.

One challenge with smartphone-based tracking is its small screen size, which can restrict visual stimulus presentation. Many researchers have addressed this by incorporating external displays. For instance, Azami et al. [33] used an Apple iPad display and an Apple iPhone camera to record eye movements in ataxia and PD participants. After comparing different machine learning methods, they decided to use principal components analysis (PCA) and linear support vector machine (SVM), achieving 78% accuracy in distinguishing individuals with and without oculomotor dysmetria. Similarly, before switching to tablets, Lai et al. [21] initially designed their experiment based on a laptop display and an iPhone 6 camera. The iTracker-face algorithm can estimate mean saccade latency with a precision of less than 10 ms error. However, they did not report gaze accuracy. One exception is the study by Valliappan et al. [34], using a smartphone (Pixel 2 XL) as both a display and camera on healthy subjects. The proposed multi-layer feed-forward CNN had an accuracy of 0.6-1 degree, which is comparable to Tobii Pro glasses.

While there is a wide variety of equipment and algorithms, few studies integrate eye tracking with edge computing. Gunawardena et al. [35] found only one study (by Otoom et al. [36]) that combined edge computing with eye-tracking. As for the accuracy, Molina-Cantero et al. [37] reviewed 40 studies on visible-light eye-tracking using low-cost camera, reporting an average visual angle accuracy of 2.69 degrees. Among these, only a subset achieved errors below 2.3 degrees, including the studies by Yang et al. [38], Hammal et al. [39], Liu et al. [40, 41], Valenti et al. [42], Jankó and Hajder [43], Wojke et al. [44], Cheng et al. [45], and Jariwala et al. [46].

1.5. Contribution

Although most studies would unintentionally mix the concept of eye tracking and gaze estimation, it is important to distinguish between these two methods. Eye tracking involves analysing visual information from images or frames in VOG, focusing on identifying facial landmarks, the eye region, and the pupil center. In contrast, gaze estimation further requires stimulus calibration, head orientation detection, and gaze angle calculation. This project aims to develop a stationary VOG application that prioritizes eye tracking over gaze tracking, utilizing smartphone cameras as an alternative to infrared-based systems. The review above of existing literature indicates that the proposed algorithm introduces a novel approach to cross-modal, smartphone-based eye tracking, particularly designed for low-resource settings and edge computing.

This study evaluates the performance of two eye-tracking algorithms—Circular Hough Transform plus Active Contour Models (CHT_ACM) and Circular Hough Transform Template Matching (CHT_TM)—in tracking the iris, particularly in neurodegenerative conditions. The findings will inform the second phase of the project, aimed at developing a novel gaze estimation software that is based on images acquired in the visible-light spectrum and will be validated against a commercial infrared eye tracker.

Modern eye trackers in gaming (e.g., Tobii Eye Tracker 5 [47] or Gazepoint 3 Eye Tracker [48]) achieve an optimistic performance in gaze tracking with unconstrained head movements. However, in line with the United Nations' Sustainable Development Goal 3 (Good Health and Wellbeing) and the principle of frugal innovation, this project aims to develop a cost-effective, smartphone-based eye-tracking solution that eliminates the need for additional accessories, thereby improving accessibility in low-resource settings. This study further contributes to the ongoing development of smartphone-based eye tracking, paving the way for wider adoption in both clinical and non-clinical settings.
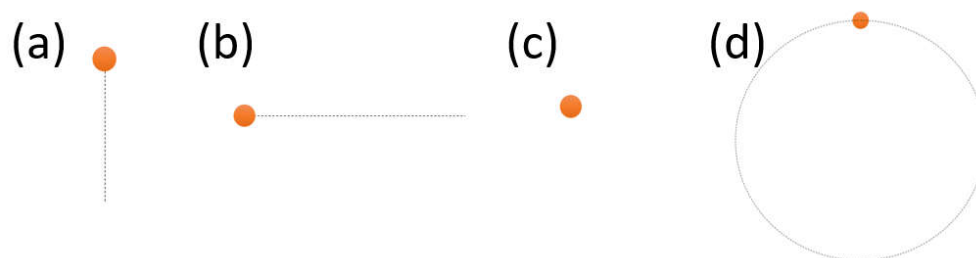
2. Methods

2.1. Experiment Definition

In order to build a dataset to develop and try the proposed algorithm on, an ethical approval application was submitted at the Biomedical and Scientific Research Ethics Committee of the University of Warwick and was granted (Ethical approval number 08/22-23).

This protocol was developed under the supervision of a medical doctor expert in neurorehabilitation and was tested on three subjects with different iris colours, which is to investigate the robustness of the proposed algorithm. The three subjects are with brown, green, and grey eyes, respectively. The experiment took place in the same lab setting and the environmental illumination was measured by a Dr.meter LX1010B digital lux meter.
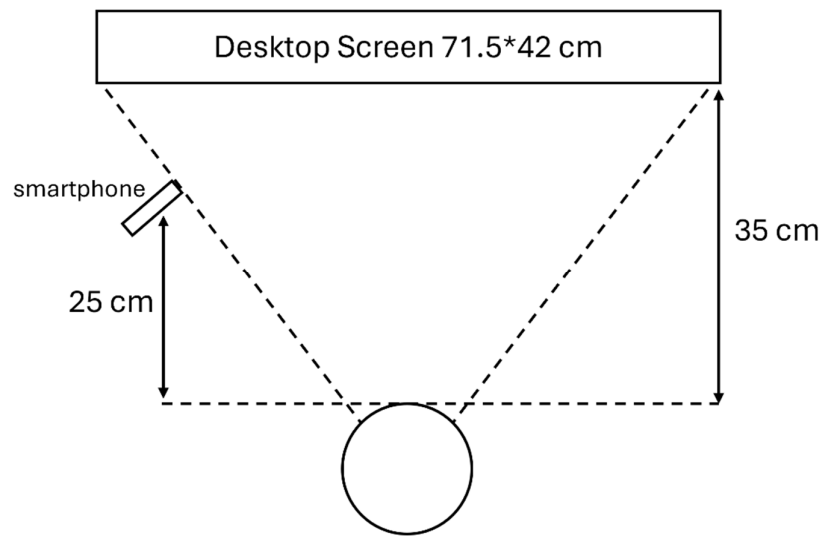
During this, the subject was also recorded in two different conditions, i.e., one with the eyelids naturally maintained open, one where they were asked to keep them open by using their fingers. This was done to understand the influence of blinks and eyelids partial closure on the algorithm performance. The experiment stimuli were designed to have an orange circle as the target and a dotted line to show the trajectory (shown in Figure 1). There were four tasks in total, namely (a) vertical task; (b) horizontal task; (c) fixation task; and (d) circular task. The target would follow the target along the trajectory only once, i.e., one complete round of the circle and back and forth for the linear tasks. The duration of each movement was either 1 or 5 seconds in order to mimic both slow and fast type of movements (e.g., smooth pursuit VS nystagmus). The participant was asked to use fingers to open the eyelids, after one session without doing so.

Figure 1. The initial set of the stimuli in the experiment protocol. (a) vertical task; (b) horizontal task; (c) fixation task; and (d) circular task.

The equipment preparation involved merely a smartphone (iPhone 12, 1920*1080 pixels, 30 frames/second), a desktop display, and a tripod. The participant was asked to sit in a private room in the author's department with a desktop screen right in front of them and a smartphone slightly shifted aside to avoid blocking the screen. The participant was asked to perform these tests without any other tool including the use of chin rest, which is normally used for aiding them to keep their head still. The participants were asked to follow certain instructions on the screen in front, while the different aforementioned tasks were being shown. In the experiment, the smartphone tracker placed

on a tripod was used to track and record the eye movement of the participants. The detailed experimental set-up with the relevant quantifications is displayed in Figure 2.
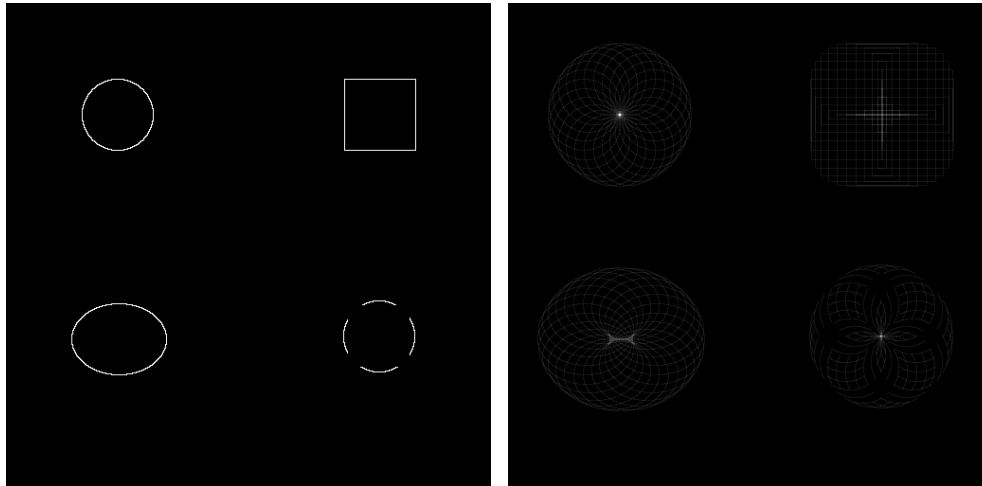

Figure 2. The experimental set-up.

2.2. Algorithm Development

Starting from the acquired videos, which featured the whole face, and after excluding the invalid videos (i.e., those with interfering movements) and manually cropping them to the region of interest (ROI), i.e., the whole eye, the first step in algorithm development entailed the selection of the best pipeline for image pre-processing and iris detection. At this stage, the working platform is Python and the OpenCV library is used for most functions. In terms of image pre-processing, the sampling rate for extracting the frames from the video was every 10 frames. Morphological operations [49] (i.e., erosion and dilation) and binarization were applied. Erosion was used to reduce the influence of the eyelashes on the iris, and dilation was applied to remove the light spots caused by the reflection of light. Consequently, the CHT [50] was applied (minDist: 40; param1: 180; param2: 10; minRadius: 15; maxRadius: 50) in two different cases, one on data pre-processed as above plus ACM, and one on data pre-processed as above plus TM.

CHT is an image analysis technique designed to detect circular shapes within an image, even when the circles are partially occluded or incomplete. It operates by identifying radial symmetry through the accumulation of "votes" in a three-dimensional parameter space defined by the circle's center coordinates and radius. Each edge pixel in the input image contributes votes in a circular pattern for potential circle centers. When multiple edge pixels support a circle of the same radius and center, the votes accumulate at those coordinates. The peaks or local maxima in the voting space indicate the presence and location of the centroid. In real-world scenarios, such as eye tracking, the pupil often appears elliptical due to the angle between the eye and the camera. Consequently, the voting space may produce a cluster of local maxima rather than a single point, as illustrated in Figure 3. This effect can be mitigated by adjusting the voting threshold and defining a minimum distance between detected peaks.

Figure 3. The voting mechanism from the edge pixel in a circular pattern [51].

CHT on its own, despite several adjustments to its parameters to reach optimal performance, was not powerful enough, resulting in too many eligible circle candidates to be the actual iris as output. In order to select the correct circle, binarization [52] was then applied. This allowed for the removal of other spurious round circles. As mentioned above, ACM was also applied as a pre-processing step for one of the algorithms. ACM [53] works by aligning a deformable model with an image through energy minimization. This deformable spine, also called a snake, responds to both constraint and image forces. These forces act collaboratively, pulling the snake toward object contours, while internal forces counteract undesired deformations. Specifically, snakes are tailored to address scenarios where an approximate understanding of the boundary's shape is available.

TM [54], conversely, represents an advanced machine vision technique designed to discern portions of an image that correspond to a predefined template. The process involves systematically placing the template over the image at all conceivable positions, calculating a numerical measure of similarity between the template and the currently overlapped image segment each time. Subsequently, the positions yielding the highest similarity measures are identified as potential occurrences of the template. The computation of the similarity measure between the aligned template image and the overlapped segment of the input image relies on the cross-correlation technique, entailing a straightforward summation of pairwise multiplications of corresponding pixel values from the two images.

Our CHT_TM relies on CHT_ACM only for the first frame. In detail, ACM is used in the first frame to fit an approximate boundary initiated with a circular snake inside of the ROI. With this approximate boundary, the image could be further cropped in order to reduce the noise and enhance the performance. CHT provides the potential circle candidates and binarization selects the best iris circle. After obtaining an intact iris template through CHT_ACM, TM is applied in the following frames to find the position with the highest similarity. The position is translated into the corresponding iris centre and is validated against the manual measurements.

2.3. Algorithm Validation and Error Avoidance

To quantify the speedup, the running speed was calculated by measuring the running time of the same algorithm on 10, 40, 100, and 400 frames' data on the same video. Moreover, in order to perform an initial validation of the algorithm, manual measurements were taken every 10 frames (same as algorithm development) on the 19 acquired videos from nine different stimuli. These stimuli represented various conditions that needed to be analysed, including finger/no finger and quick/slow movement.

Manually annotated iris centers were used as ground truth and comparison for the output of the algorithm. Several measurements of performance including Mean Absolute Error (MAE), Mean

Percentage Error (MPE), Root Mean Square Error (RMSR), and Pearson Correlation Coefficient (PCC) were calculated by the Euclidean distance between the actual location and ground truth in both x and y directions. This validation process compared the difference between several conditions including ones with/without fingers (to keep the eyelid open), 1 s/5 s movement, with/without TM, and all types of eye movements (fixation, vertical, horizontal, circle).

The frames in which the subject was blinking were deemed invalid. Moreover, due to the tendency of CHT_TM of matching the lower half of the non-intact eye (shown in Figure 4 a) resulting into a larger y axis error, a masking technique (shown in Figure 4 b) that only calculates the pixels within the central circle was applied. The larger error on y axis (still quite a low error) also represented the technical issue while developing CHT_TM, that CHT_TM always tended to match the lower half of the non-intact eye (shown in Figure 4 a).
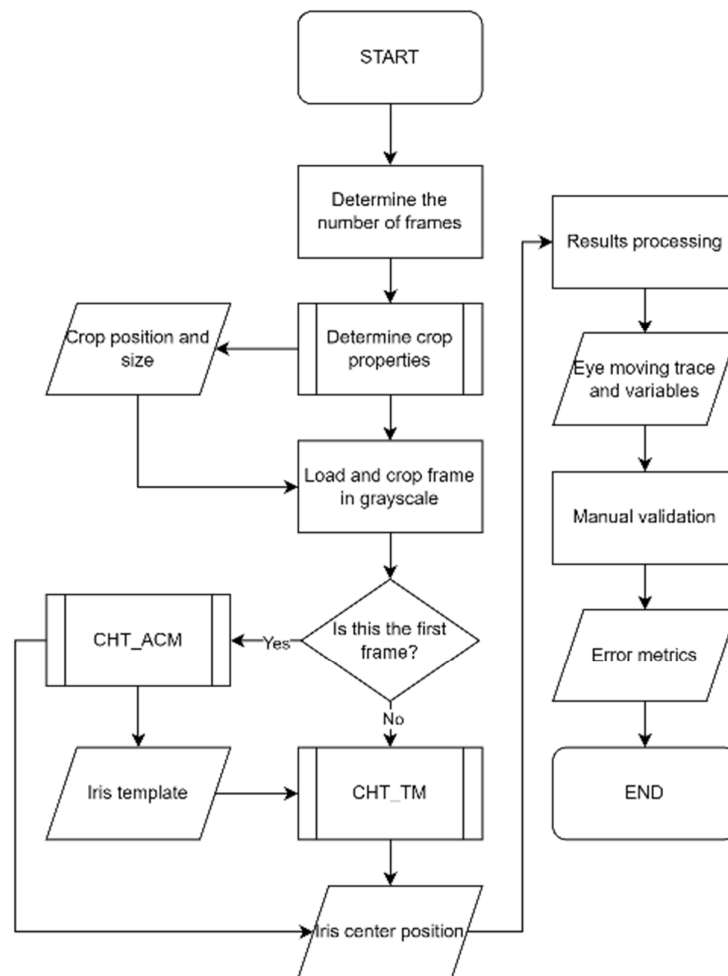

(**a**)   (**b**)

Figure 4. An example of matching the lower half of non-intact eye. (**a**) before using masking; (**b**) after using masking.

3. Results

A total of 19 videos were acquired from three subjects, resulting in a total of 633 analysed frames. The illumination is natural daylight with an average intensity of 1680 ± 282.49 lux. The first subject performed nine videos with 288 images in total and the other two subjects performed five videos with 173 images in total.

Figure 5 describes the steps of the algorithm. The valid frames of the videos were cropped into the eye region and converted into grayscale. After that, the first frame was processed by CHT_ACM. An intact iris template was generated and the iris centre position of the first frame was calculated. With this template, the following frames were processed by CHT_TM, which output the iris centre positions of the frames. The results could be calculated accordingly, and the performance
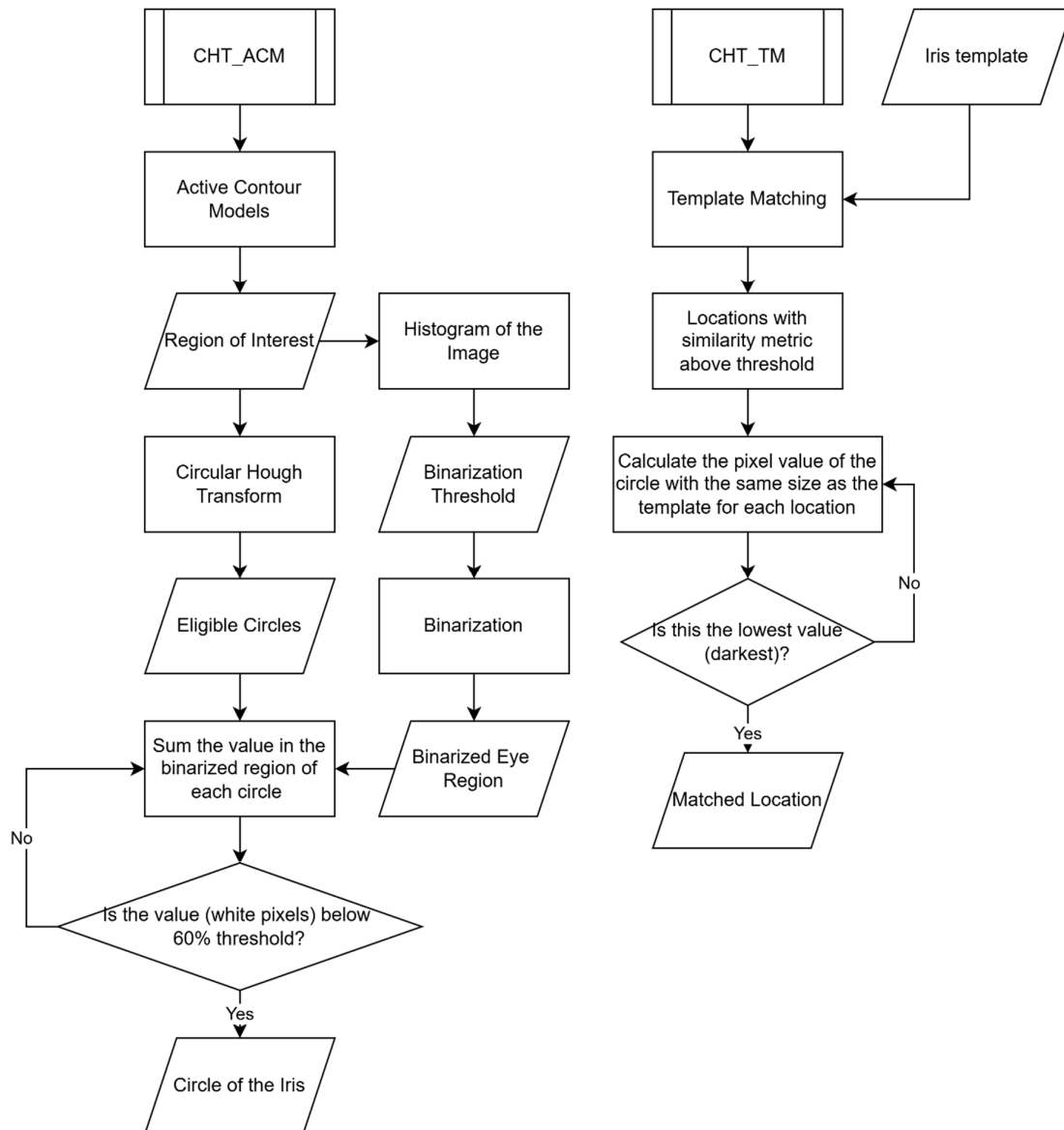
was validated with manual measurements (see Figure 5).

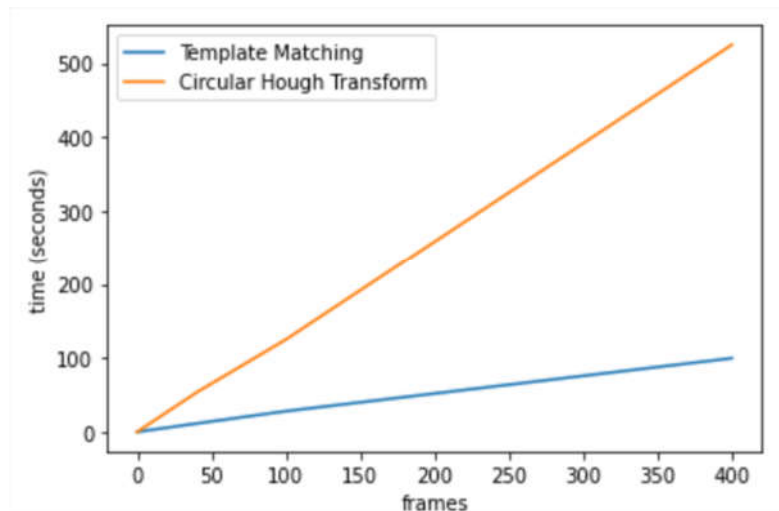**Figure 5.** The flowchart describes the steps of the algorithm.

3.1. CHT_ACM versus CHT_TM

Figure 6 compares the two different algorithms: CHT_ACM and CHT_TM. The CHT_ACM method begins with Active Contour Models to define a region of interest, followed by histogram-based binarization of the image. It then applies the Circular Hough Transform to detect eligible circles and evaluates them by summing white pixel values in their respective binarized regions. A circle is identified as the iris if its white pixel values are below a 60% threshold. In contrast, the CHT_TM method uses a predefined iris template for template matching and identifies locations with a similarity metric above a threshold. For each candidate, it calculates the pixel value within a circular region and selects the darkest (lowest value) location as the final iris center.

Figure 6. The flowchart comparing the two different algorithms in this project.

The TM function would significantly increase the running speed of the CHT_ACM (as shown in Figure 7). The average running time per frame (seconds) is 1.298 (CHT_ACM) and 0.271 (CHT_TM). CHT_TM saved 79% of the execution time, which means around 5 times faster, with a good trade-off between execution time and resource use. Most of the resource consumption is on reading the image so CPU and memory usage is not much different. Therefore, template matching greatly improved the running speed of CHT_ACM.

Figure 7. The running speed of CHT_ACM and CHT_TM.

The MAE of CHT_TM was 1.43 pixel and 1.75 pixel in the x and y axis (shown in Table 2). Meanwhile, the MAE of CHT_ACM was 1.77/2.08 for x/y axis, less accurate than CHT_TM. CHT_TM performed with an average MPE of 1.06% and 1.21% in the x and y direction across the 19 videos, compared to 1.29% and 1.42% for CHT_ACM. Moreover, Pearson's Correlation Coefficients for CHT_TM were extremely high (r_x of 0.993, p value$_x$ less than 0.001; r_y of 0.990, p value$_y$ less than 0.001). The conclusion is that template matching not only improved the running speed but also the accuracy of CHT_ACM.

Table 2. Mean Average Error, Mean Percentage Error, Root Mean Square Error, and Pearson Correlation Coefficient calculated in both x and y directions of all 19 videos. * denotes better performance (lower errors).

	CHT_TM	**CHT_ACM**
MAE for x	1.432*	1.774
MAE for y	1.752*	2.079
MPE for x	1.059*	1.290
MPE for y	1.213*	1.423
RMSE for x	2.362*	2.514
RMSE for y	2.554*	3.069
Pearsons r and pval for x	0.993, <0.0001	0.993, <0.0001
Pearsons r and pval for y	0.990, <0.0001	0.986, <0.0001

3.2. With Fingers versus Without Fingers Condition

In order to investigate the influence of using fingers to open the eyelid, the results of two conditions were analysed, i.e. with finger or without fingers. In Table 3, it is obvious that the with finger condition has lower error and higher Pearson Correlation Coefficient value, except one row, namely "x_CHT_TM". This suggests that using fingers can improve the performance of CHT_ACM and CHT_TM on y axis but it reduces the accuracy of CHT_TM on x axis. PCC_p values of Table 3 are all below 0.00001. Therefore, these values are not reported in Table 3.

Table 3. The comparison of with fingers and without fingers. * denotes better performance (lower errors).

No finger	MAE	MPE	RMSE	PCC_r
x_CHT_TM	0.467*	0.304*	0.748*	0.999
y_CHT_TM	1.697	1.241	2.396	0.984
x_CHT_ACM	1.36	0.904	1.997	0.995
y_CHT_ACM	1.337	0.969	1.876	0.990

Finger	MAE	MPE	RMSE	PCC_r
x_CHT_TM	0.541	0.403	0.849	0.999
y_CHT_TM	0.739*	0.476*	0.995*	0.997
x_CHT_ACM	0.901*	0.653*	1.252*	0.998
y_CHT_ACM	1.009*	0.653*	1.760*	0.993

3.3. Comparison Among Tasks

Table 4 presents the performance of the two algorithms CHT_ACM and CHT_TM on the four tasks, i.e. Vertical task, Horizontal task, Circular task, and Fixation task. This is used to analyse the characteristics of each algorithm on different tasks, which helps identify the specific improvements of CHT_TM compared to CHT_ACM. Moreover, it provides evidence of what is sacrificed by applying TM. It is worth noticing that CHT_ACM performs better on x axis in the Fixation task but is worse in the Horizontal task and Vertical task regardless of axis. On the other hand, CHT_TM is more accurate on the y axis but has larger error on the x axis in the Circular task and Fixation task. PCC_p values of Table 4 are all below 0.00001. Therefore, these values are not reported in Table 4.

Table 4. The comparison of different tasks. * denotes the better performance (lower error).

Vertical	MAE	MPE	RMSE	PCC_r
x_CHT_TM	1.291	0.911	2.175	0.995
y_CHT_TM	1.659	1.153	2.190	0.992
x_CHT_ACM	1.665	1.147*	2.237*	0.996
y_CHT_ACM	2.015	1.367	2.786	0.984
Horizontal	**MAE**	**MPE**	**RMSE**	**PCC_r**
x_CHT_TM	1.020*	0.726*	1.553*	0.997
y_CHT_TM	1.672	1.184	2.520	0.964
x_CHT_ACM	2.128	1.516	3.014	0.988
y_CHT_ACM	2.298	1.588	3.489	0.934
Circular	**MAE**	**MPE**	**RMSE**	**PCC_r**
x_CHT_TM	1.755	1.260	2.925	0.985
y_CHT_TM	2.054	1.386	3.039	0.993
x_CHT_ACM	1.747	1.273	2.416	0.990
y_CHT_ACM	2.009	1.347*	3.086	0.993
Fixation	**MAE**	**MPE**	**RMSE**	**PCC_r**
x_CHT_TM	1.537	1.353	2.182	0.999
y_CHT_TM	1.276*	0.922*	1.665*	0.947
x_CHT_ACM	1.428*	1.186	2.246	0.996
y_CHT_ACM	1.997*	1.427	2.685*	0.850

The best and the worst performing applications for both algorithms were summarized in Table 5. The Fixation task is the most accurate on the y axis regardless of CHT_ACM or CHT_TM and the PCC is the lowest. The Horizontal task can be well tracked on the x axis by CHT_TM but is the worst for CHT_ACM no matter the axis. The Vertical task is best measured on the x axis by CHT_ACM. Circular task has the largest error in general, which means the hardest to track, especially by CHT_TM. However, the PCC of the Circular task is the highest so the direction is similar between the real output and the manual validation result. PCC_p values of Table 5 are all below 0.00001. Therefore, these values are not reported in Table 5.

Table 5. The best performer/the worst performer among the four tasks for all the metrices. The best performer means the lowest error or the highest Pearson Correlation Coefficient value and the worst performer is on the contrary.

	MAE	MPE	RMSE	PCC_r
x_CHT_TM	Horizontal/Circle	Horizontal/Fixation	Horizontal/Circle	Fixation/Circle
y_CHT_TM	Fixation/Circle	Fixation/Circle	Fixation/Circle	Circle/Fixation
x_CHT_ACM	Fixation/Horizontal	Vertical/Horizontal	Vertical/Horizontal	Vertical/Horizontal
y_CHT_ACM	Fixation/Horizontal	Circle/Horizontal	Fixation/Horizontal	Circle/Fixation

3.4. Comparison Among Iris Colours

To understand the robustness of the proposed algorithm, the error metrics from subjects with different iris colours were presented in Table 6. The performance table indicates that there is a slight decrease in the accuracy of the algorithm for subjects with light iris colour but the final error is still pixel-level. In specific, subject 1 with brown iris colour had the lowest error, especially on x axis, and subject 2 with green iris colour had the highest error. Subject 3 with grey iris colour showed a higher error on x axis than y axis, which was different from the other two subjects. The accuracy of CHT_TM is better than CHT_ACM across all three subjects.

Table 6. The performance of both algorithms on subjects with different iris colours. * denotes the better performance (lower error).

Subject 1	MAE	MPE	RMSE	PCC_r
x_CHT_TM	0.497*	0.343*	0.789*	0.999
y_CHT_TM	1.325*	0.945*	1.974	0.992
x_CHT_ACM	1.182*	0.807*	1.746*	0.996
y_CHT_ACM	1.210*	0.847*	1.832*	0.993
Subject 2	MAE	MPE	RMSE	PCC_r
x_CHT_TM	2.791	1.985	3.732	0.972
y_CHT_TM	2.814	1.759	3.686	0.973
x_CHT_ACM	2.528	1.802	3.291	0.981
y_CHT_ACM	3.689	2.329	4.695	0.924
Subject 3	MAE	MPE	RMSE	PCC_r
x_CHT_TM	1.822	1.493	2.512	0.988
y_CHT_TM	1.415	1.143	1.957*	0.991
x_CHT_ACM	2.148	1.712	2.823	0.990
y_CHT_ACM	2.037	1.583	2.734	0.980

4. Discussion

This project aimed at developing and validating an effective eye tracking algorithm to be used on visible-light images captured by a smartphone camera, in order to unlock more affordable and user-friendly technologies for eye tracking. Two algorithms, named CHT_TM and CHT_ACM, were compared in terms of performance and computational efficiency. The selection of these algorithms was based on their potential to enhance accuracy and speed without increasing resource consumption.

Comprehensively, CHT_TM demonstrated improved runtime and superior performance in vertical eye movement tracking (y-axis), although CHT_ACM outperformed it in horizontal tracking (x-axis) in two out of four tasks. No matter CHT_ACM or CHT_TM, larger errors were consistently observed along the y-axis from Table 2. This can be attributed to the anatomical reality that the upper and lower regions of the iris are more likely to be covered by the eyelids, especially during upward or downward gaze, or when participants are fatigued and the eyes are half-closed. This introduces inaccuracies in iris center detection.

From the results comparing with/without fingers conditions, using the fingers to open the eyelid seems to improve accuracy in most cases. This supports the hypothesis that using fingers can help the algorithm diminish the error of non-intact iris as the eyelid is no longer covering the iris. While CHT_TM performed better along the x-axis even without finger assistance, the benefit of finger usage was more significant on the y-axis, where eyelid interference is typically greater on the upper edge or the lower edge. Despite this, the use of fingers was reported as uncomfortable for participants and is therefore not advisable in future studies. Alternative non-invasive strategies or postprocessing solutions are recommended.

Task-wise, the Circular task produced the largest errors, suggesting that the primary challenge lay in the instability of the task rather than the algorithm itself. CHT_ACM remains less accurate than CHT_TM. As the mean absolute errors of both CHT_ACM and CHT_TM are on pixel level, these errors are overall very small also when taking into account the limits of the system for manual

measurements with a precision of 0.5 pixels. Fixation tasks, being the most stable, resulted in the lowest tracking errors for both algorithms, which means there was no drastic head movement during the experiment. CHT_TM is better at tracking Fixation task on the y axis. Interestingly, CHT_TM improved x-axis tracking in the Horizontal task, likely due to its robustness in recognizing elliptical iris shapes during lateral gaze. In contrast, CHT_ACM retained an advantage on the x-axis for the Fixation task. CHT_TM performed more reliably across tasks, especially in preventing tracking loss.

Comparing subjects with different iris colour, the algorithm showed the best performance in a subject with dark iris colour. Nonetheless, both algorithms produced pixel-level errors across all subjects, with CHT_TM consistently outperforming CHT_ACM. This indicates that the iris colour would influence the accuracy of both algorithms simultaneously, especially on x axis.

No significant differences were found between the fast and the slow experiment.

Despite minor differences in error rates, it is noteworthy that both algorithms achieved high accuracy, with an average error of just 1.7 pixels (1.2%) across 19 videos. These results underscore the feasibility of both methods for reliable iris center detection. However, the most significant difference lay in computational performance, with CHT_TM offering faster processing times.

Comparing the proposed method with existing smartphone-based approaches proved challenging due to a lack of validated benchmarks and methodological transparency in the literature. Many studies fail to disclose algorithmic details and rely instead on vague references to platforms like OpenCV or ARKit, thereby hindering reproducibility. In contrast, this study prioritizes transparency and reproducibility by making both the data and algorithms publicly available.

Some may question the absence of machine learning or deep learning in this work, especially given their strong performance in image analysis tasks. However, the lack of a suitable public dataset, particularly one containing data from neurodegenerative patients, prevented the use of AI-based models. This is mainly due to the constraint of Ethical Approval and data privacy regulations that no identifiable data from patients should be disclosed. Therefore, this barrier makes it difficult to adopt or adapt open-access datasets, considering the final goal of distinguishing patients and healthy subjects. In fact, data collection from patients would still be inevitable but replicating the same experiment settings as the used public dataset would be hindered as explained above.

Additionally, using AI trained on different hardware and image conditions (e.g., infrared cameras) would compromise compatibility with the smartphone-based setup employed here. For instance, the structure of this algorithm is inspired by and is similar to the one proposed by Zhang et al. [55]. However, they used a CNN to condense the video and their data was collected from a portable infrared video goggle instead of the smartphone intended in this paper. Not only is the video grayscale but also the distance from the camera to the eye is different, making it impossible to use their data in this experiment or to develop the same algorithm based on varied data.

Beyond data limitations, deploying AI models on smartphones presents practical challenges. Deep learning methods typically require powerful processors or graphics processing units (GPUs) designed for computer systems, which is not the best fit for smartphones. In this case, it is necessary to upload the video data to the cloud server and use cloud computing. This reliance introduces new issues such as internet connectivity (not always available in rural areas or within low-resource settings), delayed response times, and potential data privacy risks (identifiable data like face videos). In contrast, a self-contained, built-in algorithm avoids these complications and better serves low-resource environments.

Nonetheless, AI remains a promising avenue for future work. Studies have shown a similar or even better performance with a CNN using the front-facing camera of Pixel 2 XL phone [34] or the RealSense digital camera [56] than a commercial eye tracker. Once a sufficiently large and diverse dataset is collected, including both healthy individuals and patients, AI models could be trained to refine or replace parts of the current algorithm. Such models could automate preprocessing or minimize tracking errors through learned feature extraction.

The future plan for this study is to develop a refined experimental protocol in collaboration with medical professionals, followed by validation against a commercial infrared eye tracker. Then

experiments will be carried out at hospital level on patients affected by neurodegenerative conditions. The Ultimate goal is to design a smartphone-compatible eye-tracking toolkit and AI-based system for the early screening of neurodegenerative diseases.

4.1. Limitations

One limitation of the current protocol is the visible trace of the target during the task (see Figure 1), which allows participants to predict the target's trajectory. As a result, their eye movements may precede rather than follow the target. Additionally, the absence of a headrest introduces variability due to head movement, which can compromise signal quality. To address this, future studies will explore the feasibility of using a sticker placed in a fixed location as a reference point to track and compensate for head motion, allowing reconstruction of more accurate eye movement data.

Although the study aligns with the principle of frugal innovation and avoids using additional apparatus, a tripod is currently used as a temporary substitute for a user's hand or arm. One thing worth noticing is the differences in participant height, which can affect the camera angle towards the eye and may contribute to varying errors among subjects.

Another limitation is the small sample size, as this pilot study was primarily intended to demonstrate feasibility. Manual validation was used to assess the algorithm's performance. This method, while effective for small datasets, lacks the efficiency and scalability of automated validation methods. This limits the generalizability of the results and the potential for large-scale application.

At this stage, comparisons were made between algorithm outputs and manual annotations of actual eye movement centers, rather than estimated gaze points. Each video was relatively short and did not include significant head movement, so the ROI of the eye was manually cropped and fixed at a constant image coordinate. Consequently, all movement was referenced to the same top-left corner of the cropped ROI (coordinate [0, 0]).

5. Conclusion

This paper shows that applying TM improves the original eye tracking algorithm with CHT_ACM. The improved algorithm shows potential for supporting eye movement tracking in early screening and diagnosis of neurodegenerative diseases. Its integration into a non-invasive, cost-effective telemedicine toolbox could improve the diagnostic reliability and accessibility as well as the electronic transferability, in a cost-effective way. Such innovations sit perfectly within the current healthcare landscape, which has limited funding and is challenged by health emergencies (e.g., COVID-19 pandemic). Our research contributes to the development of such telemedicine tools by offering a reliable and accessible method for eye tracking, which can be integrated into telemedicine platforms for remote screening and diagnosis. As highlighted above, the study has several limitations including reliance on a traditional image processing approach rather than AI, visibility of the target trace, and the lack of head movement correction. Future work will aim to refine the experimental protocol with guidance from medical experts, validate the algorithm against commercial benchmarks, and expand testing to patients with neurodegenerative symptoms. The integration of AI techniques will also be explored as a potential enhancement to the system.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Biomedical and Scientific Research Ethics Committee of the University of Warwick (Ethical approval number 08/22-23 and was granted on 18 January 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets used and/or analysed during the current study are available in the supplementary materials or from the corresponding author on reasonable request.

Acknowledgments: None.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AD   Alzheimer's Disease
AI   Artificial Intelligence
ALS   Amyotrophic Lateral Sclerosis
BSREC   Biomedical and Scientific Research Ethics Committee
CHT   Circular Hough Transform
CHT_ACM CHT plus Active Contour Model
CHT_TM   CHT plus Template Matching
CNN   Convolutional Neural Network
DPI   Dual Purkinje Image
EOG   Electro-Oculography
fMRI   Functional Magnetic Resonance Imaging
IOG   Infrared Oculography
MAE   Mean Absolute Error
MPE   Mean Percentage Error
PCC   Pearson Correlation Coefficient
RMSR   Root Mean Square Error
ROI   Region of Interest
VOG   Video-Oculography

References

1. Mengtao, L.; Fan, L.; Gangyan, X.; Su, H. Leveraging eye-tracking technologies to promote aviation safety- A review of key aspects, challenges, and future perspectives. Safety Science 2023, 168, 106295, doi:https://doi.org/10.1016/j.ssci.2023.106295.
2. Hansen, D.W.; Ji, Q. In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. IEEE Transactions on Pattern Analysis and Machine Intelligence 2010, 32, 478-500, doi:10.1109/TPAMI.2009.30.
3. Majaranta, P.; Bulling, A. Eye Tracking and Eye-Based Human–Computer Interaction. In Advances in Physiological Computing, Fairclough, S.H., Gilleade, K., Eds.; Springer London: London, 2014; pp. 39-65.
4. Pouget, P. The cortex is in overall control of 'voluntary' eye movement. Eye (Lond) 2015, 29, 241-245, doi:10.1038/eye.2014.284.
5. Purves D, A.G., Fitzpatrick D, et al., editors. Types of Eye Movements and Their Functions. In Neuroscience. 2nd edition; Sinauer Associates: Sunderland (MA), 2001.
6. Alexander, R.G.; Macknik, S.L.; Martinez-Conde, S. Microsaccades in Applied Environments: Real-World Applications of Fixational Eye Movement Measurements. J Eye Mov Res 2020, 12, doi:10.16910/jemr.12.6.15.
7. Mahanama, B.; Jayawardana, Y.; Rengarajan, S.; Jayawardena, G.; Chukoskie, L.; Snider, J.; Jayarathna, S. Eye movement and pupil measures: A review. frontiers in Computer Science 2022, 3, 733531.

8.   Miki, S.; Hirata, Y. Microsaccades generated during car driving. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013-07-01, 2013.

9.   Bueno, A.P.A.; Sato, J.R.; Hornberger, M. Eye tracking–The overlooked method to measure cognition in neurodegeneration? Neuropsychologia 2019, 133, 107191.

10. Gooding, D.C.; Miller, M.D.; Kwapil, T.R. Smooth pursuit eye tracking and visual fixation in psychosis-prone individuals. Psychiatry Research 2000, 93, 41-54.

11. Garbutt, S.; Matlin, A.; Hellmuth, J.; Schenk, A.K.; Johnson, J.K.; Rosen, H.; Dean, D.; Kramer, J.; Neuhaus, J.; Miller, B.L. Oculomotor function in frontotemporal lobar degeneration, related disorders and Alzheimer's disease. Brain 2008, 131, 1268-1281.

12. Van Gerven, P.W.M.; Paas, F.; Van Merriënboer, J.J.G.; Schmidt, H.G. Memory load and the cognitive pupillary response in aging. Psychophysiology 2004, 41, 167-174.

13. Arnold, L.; Aryal, S.; Hong, B.; Nitharsan, M.; Shah, A.; Ahmed, W.; Lilani, Z.; Su, W.; Piaggio, D. A Systematic Literature Review of Eye-Tracking and Machine Learning Methods for Improving Productivity and Reading Abilities. Applied Sciences 2025, 15, 3308.

14. Iturrate, I.; Antelis, J.M.; Kubler, A.; Minguez, J. A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation. IEEE transactions on robotics 2009, 25, 614-627.

15. Chang, T.P.; Zee, D.S.; Kheradmand, A. Technological advances in testing the dizzy patient: the bedside examination is still the key to successful diagnosis. Dizziness and Vertigo across the Lifespan 2019, 9-30.

16. Kang, S.L.; Shaikh, A.G.; Ghasia, F.F. Vergence and strabismus in neurodegenerative disorders. Frontiers in neurology 2018, 9, 299.

17. Golding, C.; Danchaivijitr, C.; Hodgson, T.; Tabrizi, S.; Kennard, C. Identification of an oculomotor biomarker of preclinical Huntington disease. Neurology 2006, 67, 485-487.

18. Hicks, S.L.; Robert, M.P.; Golding, C.V.; Tabrizi, S.J.; Kennard, C. Oculomotor deficits indicate the progression of Huntington's disease. Progress in brain research 2008, 171, 555-558.

19. García Cena, C.; Costa, M.C.; Saltarén Pazmiño, R.; Santos, C.P.; Gómez-Andrés, D.; Benito-León, J. Eye Movement Alterations in Post-COVID-19 Condition: A Proof-of-Concept Study. Sensors (Basel) 2022, 22, doi:10.3390/s22041481.

20. Mergenthaler, K.; Engbert, R. Microsaccades are different from saccades in scene perception. Experimental Brain Research 2010, 203, 753-757, doi:10.1007/s00221-010-2272-9.

21. Lai, H.Y.; Saavedra-Peña, G.; Sodini, C.G.; Sze, V.; Heldt, T. Measuring Saccade Latency Using Smartphone Cameras. IEEE Journal of Biomedical and Health Informatics 2020, 24, 885-897, doi:10.1109/JBHI.2019.2913846.

22. Sangi, M.; Thompson, B.; Turuwhenua, J. An Optokinetic Nystagmus Detection Method for Use With Young Children. IEEE Journal of Translational Engineering in Health and Medicine 2015, 3, 1-10, doi:10.1109/jtehm.2015.2410286.

23. Piaggio, D.; Namm, G.; Melillo, P.; Simonelli, F.; Iadanza, E.; Pecchia, L. Pupillometry via smartphone for low-resource settings. Biocybernetics and Biomedical Engineering 2021, 41, 891-902, doi:https://doi.org/10.1016/j.bbe.2021.05.012.

24. Ilesanmi, A.E.; Ilesanmi, T.; Gbotoso, G.A. A systematic review of retinal fundus image segmentation and classification methods using convolutional neural networks. Healthcare Analytics 2023, 4, 100261, doi:https://doi.org/10.1016/j.health.2023.100261.

25. Newman, J.L.; Phillips, J.S.; Cox, S.J.; Fitzgerald, J.; Bath, A. Automatic nystagmus detection and quantification in long-term continuous eye-movement data. Computers in Biology and Medicine 2019, 114, 103448, doi:10.1016/j.compbiomed.2019.103448.

26. Alichniewicz, K.K.; Brunner, F.; Klünemann, H.H.; Greenlee, M.W. Neural correlates of saccadic inhibition in healthy elderly and patients with amnestic mild cognitive impairment. Frontiers in Psychology 2013, 4, 467.

27. Witiuk, K.; Fernandez-Ruiz, J.; McKee, R.; Alahyane, N.; Coe, B.C.; Melanson, M.; Munoz, D.P. Cognitive deterioration and functional compensation in ALS measured with fMRI using an inhibitory task. Journal of Neuroscience 2014, 34, 14260-14271.

28. Hutton, S.B. Eye tracking methodology. Eye movement research: An introduction to its scientific foundations and applications 2019, 277-308.

29. Underwood, C. Electronystagmography (ENG). Available online: https://www.healthline.com/health/electronystagmography (accessed on 31 Jan).

30. Cornsweet, T.N.; Crane, H.D. Accurate two-dimensional eye tracker using first and fourth Purkinje images. JOSA 1973, 63, 921-928.

31. Guestrin, E.D.; Eizenman, M. General theory of remote gaze estimation using the pupil center and corneal reflections. IEEE Transactions on biomedical engineering 2006, 53, 1124-1133.

32. Wildenmann, U.; Schaeffel, F. Variations of pupil centration and their effects on video eye tracking. Ophthalmic and Physiological Optics 2013, 33, 634-641.

33. Azami, H.; Chang, Z.; Arnold, S.E.; Sapiro, G.; Gupta, A.S. Detection of Oculomotor Dysmetria From Mobile Phone Video of the Horizontal Saccades Task Using Signal Processing and Machine Learning Approaches. IEEE Access 2022, 10, 34022-34031, doi:10.1109/access.2022.3156964.

34. Valliappan, N.; Dai, N.; Steinberg, E.; He, J.; Rogers, K.; Ramachandran, V.; Xu, P.; Shojaeizadeh, M.; Guo, L.; Kohlhoff, K.; et al. Accelerating eye movement research via accurate and affordable smartphone eye tracking. Nature Communications 2020, 11, 4553, doi:10.1038/s41467-020-18360-5.

35. Gunawardena, N.; Ginige, J.A.; Javadi, B. Eye-tracking Technologies in Mobile Devices Using Edge Computing: A Systematic Review. ACM Comput. Surv. 2022, 55, Article 158, doi:10.1145/3546938.

36. Otoom, M.; Alzubaidi, M.A. Ambient intelligence framework for real-time speech-to-sign translation. Assistive Technology 2018, 30, 119-132.

37. Molina-Cantero, A.J.; Lebrato-Vázquez, C.; Castro-García, J.A.; Merino-Monge, M.; Biscarri-Triviño, F.; Escudero-Fombuena, J.I. A review on visible-light eye-tracking methods based on a low-cost camera. Journal of Ambient Intelligence and Humanized Computing 2024, 15, 2381-2397, doi:10.1007/s12652-024-04760-8.

38. Yang, B.; Zhang, X.; Li, Z.; Du, S.; Wang, F. An accurate and robust gaze estimation method based on maximum correntropy criterion. IEEE Access 2019, 7, 23291-23302.

39. Hammal, Z.; Massot, C.; Bedoya, G.; Caplier, A. Eyes Segmentation Applied to Gaze Direction and Vigilance Estimation. Berlin, Heidelberg, 2005; pp. 236-246.

40. Liu, Y.; Lee, B.-S.; Rajan, D.; Sluzek, A.; McKeown, M.J. CamType: assistive text entry using gaze with an off-the-shelf webcam. Machine Vision and Applications 2019, 30, 407-421.

41. Liu, Y.; Lee, B.S.; Sluzek, A.; Rajan, D.; Mckeown, M. Feasibility Analysis of Eye Typing with a Standard Webcam. Cham, 2016; pp. 254-268.

42. Valenti, R.; Staiano, J.; Sebe, N.; Gevers, T. Webcam-Based Visual Gaze Estimation. Berlin, Heidelberg, 2009; pp. 662-671.

43. Jankó, Z.; Hajder, L. Improving human-computer interaction by gaze tracking. In Proceedings of the 2012 IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom), 2012; pp. 155-160.

44. Wojke, N.; Hedrich, J.; Droege, D.; Paulus, D. Gaze-estimation for consumer-grade cameras using a Gaussian process latent variable model. Pattern Recognition and Image Analysis 2016, 26, 248-255.

45. Cheng, S.; Ping, Q.; Wang, J.; Chen, Y. EasyGaze: Hybrid eye tracking approach for handheld mobile devices. Virtual Reality & Intelligent Hardware 2022, 4, 173-188.

46. Jariwala, K.; Dalal, U.; Vincent, A. A robust eye gaze estimation using geometric eye features. In Proceedings of the 2016 Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC), 6-8 July 2016, 2016; pp. 142-147.

47. Tobii. Tobii Eye tracker 5. Available online: https://gaming.tobii.com/product/eye-tracker-5/ (accessed on 18 June).

48. Gazepoint. GP3 Eye Tracker | Hardware Only. Available online: https://www.gazept.com/product/gazepoint-gp3-eye-tracker/?v=79cba1185463 (accessed on 18 June).

49. Batchelor, B.G.; Waltz, F.M. Morphological Image Processing. In Machine Vision Handbook, Batchelor, B.G., Ed.; Springer London: London, 2012; pp. 801-870.

50. Cherabit, N.; Chelali, F.Z.; Djeradi, A. Circular hough transform for iris localization. Science and Technology 2012, 2, 114-121.

51. ctrueden. Hough Circle Transform. Available online: https://imagej.net/plugins/hough-circle-transform (accessed on 9 April).

52. Yousefi, J. Image binarization using Otsu thresholding algorithm. Ontario, Canada: University of Guelph 2011, 10.

53. Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. International Journal of Computer Vision 1988, 1, 321-331, doi:10.1007/BF00133570.

54. Brunelli, R. Template Matching Techniques in Computer Vision: Theory and Practice; Wiley Publishing: 2009.

55. Zhang, W.; Wu, H.; Liu, Y.; Zheng, S.; Liu, Z.; Li, Y.; Zhao, Y.; Zhu, Z. Deep learning based torsional nystagmus detection for dizziness and vertigo diagnosis. Biomedical Signal Processing and Control 2021, 68, 102616, doi:https://doi.org/10.1016/j.bspc.2021.102616.

56. Hassan, M.A.; Yin, X.; Zhuang, Y.; Aldridge, C.M.; McMurry, T.; Southerland, A.M.; Rohde, G.K. A Digital Camera-Based Eye Movement Assessment Method for NeuroEye Examination. IEEE Journal of Biomedical and Health Informatics 2024, 28, 655-665, doi:10.1109/JBHI.2023.3285940.