REAL-TIME EYE TRACKING USING HEAT MAPS

Chetana Krishnan[1], Vijay Jeyakumar[2*], Alex Noel Jospeh Raj[3]

[1,2]Department of Biomedical Engineering, Sri Sivasubramaniya Nadar College of Engineering, Tamil Nadu, India

[3]Department of Electronic Engineering, College of Engineering, Shantou University, Shantou, China

Email: chetana1808@bme.ssn.edu.in[1], vijayj@ssn.edu.in[2*] (corresponding author), stujalexnoel@gmail.com[3]

DOI: https://doi.org/10.22452/mjcs.vol35no4.3

ABSTRACT

Communication in modern days has developed a lot, including wireless networks, Artificial Intelligence (AI) interaction, and human-computer interfaces. People with paralysis and immobile disorders face daily difficulties communicating with others and gadgets. Eye tracking has proven to promote accessible and accurate interaction compared to other complex automatic interactions. The project aims to develop an electronic eye blinker that integrates with the experimental setup to determine clinical pupil redundancy. The proposed solution comes up with an eye-tracking tool within an inbuilt laptop webcam that tracks the eye's pupil in the given screen dimensions and generates heat maps on the tracked locations. These heat maps can denote a letter (in case of eye writing), an indication to click on that location (in case of gadget communication), or for blinking analysis. The proposed method achieves a perfect F-measure score of 0.998 to 1.000, which is comparatively more accurate and efficient than the existing technologies. The solution also provides an effective method to determine the eye's refractive error, which can replace the complex refractometers. Further, the spatially tracked coordinates obtained during the experiment can be used to analyze the patient's blinking pattern, which, in turn, can detect retinal disorders and their progress during medication. One of the applications of the project is to integrate the derived model with a Brain-computer interface system to allow fast communications for the disabled.

Keywords: Blinking, Heat maps, Lattice coordinates, Refractive error, Scale calibration system, Spatial coordinates

1.0 INTRODUCTION

Eye or pupil tracking is a method of following the direction and the interval distance (between one point to another) to acquire dynamic pupil details for research purposes [1]. Three antagonistic pairs of muscles control eye movements: the lateral and medial rectus muscles, the superior and inferior rectus muscles, and the superior and inferior oblique muscles. Any damage to these muscles can lead to improper internal reflection and dull or blurred vision. Hence, it is important to study these muscles' mechanisms to detect prior chronic disorders. Eye-tracking also helps in modern AI and computer interaction which helps the disabled easily communicate. This application also helps in featuring robotic and prosthetic arm movements [2], gambling-related attentional biases [3] based on the eye location.

The authors of this paper have worked on developing a stand-alone application for eye tracking that not just tracks but also locates, plots, and interprets the eye's various eye movements, such as the saccade, the fixation, and the vergence. The final interpretation will be in the form of equity heat maps that can be further extended to a specific application like blinking rate analysis or autism interpretation. The performance scale of the proposed method was improved compared to existing software in terms of webcam calibration, screen illumination compensation, and end-result algorithm to filter out the required results. The proposed model has another advantage of low cost as it does not require separate single hardware webcams and uses the inbuilt computer webcam [4].

1.1 Eye Mobility

Four basic eye movement types are responsible for producing various stimuli responses [5]. Tracking these movements' peak and transition points will help us understand more about the ocular muscles and help detect the disorders priorly.

- Saccades: The extraocular muscles are responsible for the saccade movement, leading to an abrupt change in the fixation phase in a small interval of time. A delay is introduced between two saccade movements, leading to capturing the images. These movements cause rapid motion and are mostly found during the Rapid eye movement (REM) phase of sleep.

- Smooth Pursuit Movements (SPM): Unlike rapid saccades, SPM is slow and transitions less. This movement mainly involves focusing a moving object at a greater distance. However, in the absence of a moving target, SPM is just a slow abnormal saccade that flattens after a few microseconds. The medial longitudinal fasciculus and the axon communicating muscles are responsible for SPM.

- Vestibulo Ocular Movements (VOM): The VOM is responsible for compensating the "turn error" when the head stays in one position but the eye directs in a different path. Such movements help in detecting the cylindrical error in the eye. Retinal image capture is held intact with VOM peak constants, and blurred vision in the VOM position indicates cylindrical power. The medial rectus and lateral rectus control normal VOM.

- Vergence Movements: Unlike other eye movements conjugate (two-step directive motions), vergence movements are disconjugate in nature (align images from various directions). The horizontal vergence results from the combined activity of six ocular muscles, which originates from three cranial nerves: the abducens nerve, the trochlear nerve, and the oculomotor nerve. The vertical vergence result from the activity of the medial and lateral rectus [6].

Other nominal movements like adduction, abduction, infraduction, cycloduction, supraduction and cyclovergence also contribute to lateral vision, which is not primarily considered during eye-tracking. These movements play a role during refractive and reflective error calculation.

1.2 Available Technologies

Researchers have been developing a portable eye-tracking system to promote telemonitoring; however, conventional systems use a complex mechanism with an externally connected camera [7][8]. Since webcam calibration is an essential parameter in telemonitoring, data-driven technologies post some limitations, and hence calibration fails may result in low calibration features. Gaze direction is the other challenge in eye-tracking. Conventional techniques use tracking sensors to monitor gaze direction, but it fails when the patient is moved or blinks. Besides external webcam tracking, an electrooculogram (a biosignal representing the movements related to the eye due to voluntary and involuntary actions) is widely used to understand the fundamental aspects of eye blinking and movement. However, it fails to calculate the muscle reflex origination, essential to generate refractive and reflective eye errors. To overcome these conventional limitations, the proposed solution develops a 9-to-42-point calibrated system to track the eyes and generate the following parameters.

- Gaze Coordinates: The system determines the exact screen pupil coordinates with the direction displayed in degrees.

- Blinking Pattern: The ocular pattern from the experimental tracking will be recorded to tell us about the rising retinal disorders in patients. It can be calculated by estimating the slope and unilateral value of the Q curve, the curve where the Q curve represents the final pupil movements which can be interpreted to determine the rate of rising and depression of the pupil. Heat Maps: Heat maps are data visualization techniques usually represented in the color code format in the existing technology. The formatted healthy heat map sequence is 184924602 (default), where each digit corresponds to the decoded heat value for a known dimension. Heat maps present the pupil's location within a given dimension along with the time interval. This is clinically applied in artificial eye communication. They are applied in visualizing the distribution of gaze points, and more precision in a single region defines high-pressure areas. Low precision defines low-pressure areas put forth by the pupil.

- Refractive Error: Conventional refractometers work that, on the nulling principle, are bulky and complex to use. The proposed system tends to calculate the approximate refractive error, which will be compared with the traditional ones.

- Area of Interest (AoI): AoI defines the metrics extracted from a particular gaze region. AoI typically ranges from 15 – 60 screen distance, specifying the time interval the patient focuses on at one point.

- Time to First Fixation (TFF): TFF is the time the pupil takes to reach a fixational movement, and TFF is applied in prioritizing moderate pressurized positions.

- Dwell time: It is the time interval between 2 saccades and a fixational movement. A longer dwell time indicates a high-interest metrix and vice versa. It helps calculate the number of blinks between two successive pupil movements.

- Span refers to successive distances between high-interest and low-interest regions calculated from AoI.

- Fixation Sequences (FS): FS refers to the binary decoded value that the computer stores for further digital conversion. They can be used as a command for automatic device control and monitoring muscle actions

Some known eye-tracking protocols are gaze recorders and Tobii [9]. The gaze recorder protocol provides special illumination glasses the user must wear while doing the trial. The glasses supply a part of optic luminance to the eyes, which the user must follow to get it tracked on the screen. Since it directly illuminates the eyeballs, the accuracy is pretty high, but users with astigmatism and other retinal conditions find it difficult to focus on the illuminated screen. Tobii recorders are known for customizing the screen size based on the user's eye (eye width changes due to conditions like cross-eyes). The protocol provides a gaming environment making it easier to assess child users. However, care must be taken to ensure the user focuses appropriately. Heat maps become an easier method to visualize, store and retrieve data. Heatmaps are color code versions of the localized points where the higher density maps correspond to priority focus, and the least dense maps are usually ignored for classification. In the case of autism, equity heat maps are used where every localized point is given equal importance in classification thus producing a single-colored heat map (annotation).

1.3 Technical Challenges and its Mitigation

Table 1 highlights the issues related to conventional eye-tracking systems [10][11] and the mitigations that the proposed project implements to overcome them. The rest of the paper describes the system architecture, methodology, and results, preceded by the conclusion.

Table 1: Challenges and Mitigations

Challenges	Solutions
Both eyes cannot be tracked at the same time	Uses less screen dimension and more area per each eye theorem
Accuracy is less for patients wearing glasses or lens	The reflective phenomena compensate for the error in the tracking during refractive error determination
Gaze direction is not monitored continuously	Uses Viola-Jones algorithm to promote continuous gaze tracking
Factors like lighting and the patient's position will affect the accuracy	The amount of data collected is enormous, and hence use, testing and training data
The gaze span is not uniform throughout the experiment	The interval between saccade and fixation will be made equal at webcam delay points
Current refractometers are non–portable and non-remote	Uses color coding to predict the refractive error, portable, and promotes remote monitoring

2.0 RELATED WORKS

Juhong et al. worked on an eye-tracking system to trace pupil response for different stimuli like band-limited light intensity from 700 to 1000 nanometers. The response will be used in computer-based sketch interaction in medical and robotic applications. However, the system only tracks the difference in response between the varying stimuli, leading to false stimuli distortion [1]. Punde et al. worked on the eye-tracking mechanism for virtual reality applications. The objective of the work was to track the pupil to capture the response in surgical tasks like tonsils removal or kidney stone removal. The images were m acquired at 30 Hertz to obtain a video, but the frame rate was way too slow to record all the maximum gaze directions leading to blurred capture. Similar results were obtained for other frequency ranges [2]. Zhang et al. proposed an eye-tracking system which is a one-time approach. The pupil anatomy, like its length, dimensional direction degrees, and single-point stare coordinates, are captured separately using a stand-alone camera. The captured data will be tested for its gaze points. Since the monitoring is not continuous, limited testing data and location are obtained, which limits the prediction accuracy. Conditions like glaucoma show varied responses in different locations, and failing to

capture them will lead to mispredictions [4]. Vehlen et al. worked on applying eye-tracking software in automatic wheelchairs and device control.

Eyeblink is applied as a command to control the buttons, and the dwell time is taken as a start or stop approach. The system uses stand-alone external devices, and an app lead control mechanism [7]. Peter Kiefer et al. worked on electrooculogram-based eye-tracking systems having invasive protrusions and electrodes. It was a conscious-based approach; hence, false coordinates are recorded [8]. A T Duchowski et al. worked on a wearable brain-machine interface system with flicker classification. The system is a game application where the user has to move the anime characters up and down with eye movements. This can improve the cognitive skills of the disabled and help them communicate [9]. Limitations of the system include delays in the stimuli and response time. Nikhlas et al. worked on deep learning-based video gaze tracking to capture crucial eyelid positions before the eye-blinking event occurs. The authors achieved a correlation between the events before and after blinking [25]. A similar approach was executed by Ildar et al., where neural boundaries captured the relative face and pupil positions while the subjects were asked to follow the mouse pointer [26]. This way, voluntary blinking motions were captured.

Wearables and headbands that operated on infrared transmission were developed to calculate the blinking features, followed by partial eyelid studies in corneal disorder cases. This was unsuccessful, as reviewers reported that the patient was conscious about the band, so the results were always perfect.

3.0 SYSTEM ARCHITECTURE AND ALGORITHM

Figure 1 represents the project's system workflow, which is briefly explained below.

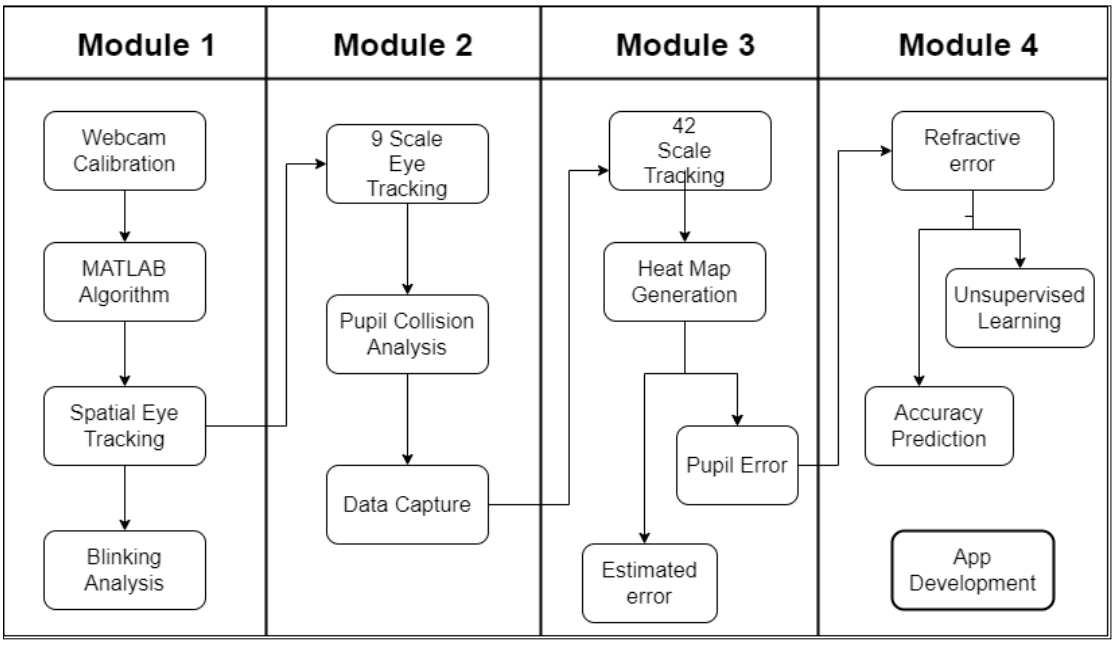

Fig. 1: System Architecture

3.1 Webcam Calibration

The proposed solution employs the user's inbuilt laptop webcam, which needs the following properties to produce accurate results.

- Requires a USB interface option to promote data visualization.
- Requires resolution of 1280 x 720.
- Needs 4500ms of frame rate per 120 image captures. This is important to prevent delays between gaze points.
- Extendible up to the screen dimensions.
- Has an anaglyphic property of 1280 x 450.

The above features can be easily obtained in a Windows 7 operating system. Once the webcam check is made, it is calibrated using computer vision algorithms. Users with less advanced operating systems can use the desktop client version of the project to access the above properties.

3.2 Algorithms

The training data collected is then imported to the MATLAB 2021a tool to acquire the first half of the testing data, which calculates the pupil coordinates in a two-colored (binary) background. The localized binary tracking helps pitch the testing data and segregates it from the lower to higher coordinates. Viola-Jones algorithm [9] effectively obtains the binarized pupil coordinates.

3.3 9 and 42 Scale Tracking

The testing data will be imported to the built website, which is 9 points of equal width and screen distance. The user will then be asked to follow the points to obtain the second half of the testing set. This is followed by a 42-scale tracking system with decreased distance. The locations where the user is tracked are mapped with heat colors to understand the pressure put forth by the user (AoI).

3.4 Blinking and Refractive Analysis

A false eyelid-based fiber blink tracker was developed to analyze eyelids' blink rate and partial closure. A simulation circuit was developed to determine the refractive error in a portable and accessible format.

3.5 App Development

BOLT visualization app was developed using BOLT IoT to enable remote data and results visualization for both the user and the clinician.

3.6 Accuracy

The accuracy of the proposed system was determined using the ball (point) collision method according to Gazer algorithms [12]. The number of balls colliding with each other with the eye will determine the accuracy. The accuracy is backed up by unsupervised learning to perform manual ball collision.

3.7 Work Pattern

The following procedure was adopted to obtain the heat maps and characteristic eye movement rates.

- Instantaneous data capture is a collection of images within a given frame rate where the frame rate and testing time are decided by differential equations (Eqns. 1,2 and 3).
- Webcam calibration based on the user's operating system and extracting the corresponding feed
- Importing training data and localizing pupils using neural lines and plotting the same in the lattice
- 9- and 42-point calibration to track the eye over the given screen dimension
- Generating heat maps of testing data using the results obtained from the testing data

4.0 METHODOLOGY

The proposed project revolves around acquiring three types of training sets - testing data (output acquired from the user), training data (threshold existing conventional data), and garage values (substitute data). The training data consists of parameters (range, threshold, healthy values) of each eye movement are listed in Table 2. These values are based on the quadrant of the screen the user is looking at. There is a slight difference in the original and assumed threshold values [13].

Table 2: Threshold ranges

Movements	Range/ Skew Value (Deg/ No unit)
Saccade – fixation dual	12 – 34 / 130
All types of vergence	-67 – 90 / 230 – 330
Infraversion	-87 / 100
Pursuit Movement	20 – 28 / 85 - 90
Upper-end movement	40-90 (mostly obtuse) / 78
Lower end movement	-40 - -90 (mostly acute) / -90
Rapid – Slow interval	0 – 180 / 210
Torsion	65 – 100 / 40 – 90
Anti torsion	21-56 / 190

The testing data is acquired by an experimental procedure, out of which half are assumed as garage values in case the delay element does not match. The tracking algorithm [14][15] is divided into three equation sets. The first equation, called the offset parameter, describes the effect of the 2D distance between the focal planes and the density receptors.

$$X = \frac{\varphi u}{\omega} \;\; ; \;\; Y = \frac{\varphi u}{\omega} \qquad (1)$$

Where X and Y are screen distances, φ, u, and ω represent the visionary receptor values. The second equation is the skew parameter, an additive of the image plane with the focal plane, and is essential in fixing the webcam frame rate [16].

$$X = \frac{\varphi u}{\omega} + \delta \;\; ; \;\; Y = \frac{\varphi u}{\omega} + \delta \qquad (2)$$

The third equation represents the orientation and position of the camera [17]. They must place the user in the correct position to prevent gaze and lateral inaccuracy.

$$W' = \varepsilon + \frac{\gamma'}{\gamma} + \tau \qquad (3)$$

Where ε, γ, and τ represent the orientation distance from the centre of the screen fixed by the program. The webcam was calibrated by capturing frame rate per millisecond by applying the above equation in the proposed algorithm. From figure 2, it is noted for fixational eye movements, and equation 1 is constant. However, it is varied for other eye movements. If equations 1,2 and 3 are satisfied, the 'True' string is displayed. Along with the webcam feed, the blink count is started to test the time up to which the user can stay without blinking. The green dots are approximate pupil directions.

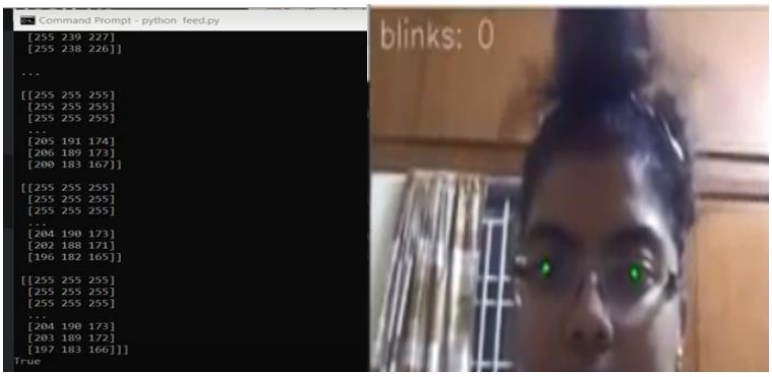

Fig. 2: Live webcam feed and blink count

Once the process is finished, using the blinks and frame rate achieved, the pupil will be located within a small screen dimension at an angle of 45 degrees supraduction position as a part of the training set. Fig.3 shows that the user has blinked 8 times, and the pupils are in supraduction. The focusing screen size is about 0.04 on the X and Y-axis. The focused coordinate at the time t is displayed below the focusing screen. The screen distance can be increased or decreased depending on the blink rate of the user. According to figure 2, if the blink rate is less than 12, between 12 and 20, and above 12 per minute, then the screen distance can be 0.04, 1, and 1.2, respectively. If the green dot disappears, the user is not seated correctly, or the pupil is not in the required position. The delay in the webcam feed was compensated by considering only the peak solutions of the equations (discrete and not continuous).

Fig. 3: Screen Focusing

A single eye tracker succeeds in the screen-focusing process to increase the accuracy of the training set. The Viola-Jones algorithm does this, a cascaded (screened) object recognition framework similar to face detection. Still, it involves segmenting the region of interest using edge processing. The algorithm provided the following features.

- Edge features: Since the algorithm targets the cascaded objects, it fails to detect the secondary pupils (if there are any other people apart from the user before the webcam), i.e., it detects only the pupils which are maximum in size (as secondary pupils are smaller due to webcam angles). The edge mask is given by

1	-j	-j	1
j	0	1	1
j	1	-j	1
1	0	0	0

- The line features: If ABCDEF is the screen dimension where each letter refers to the screen corners, then the algorithm detects the user's position along the slanting lines AC, BD, CE, and BF. The horizontal and vertical lines are detected by adding ‘Haar features.’ The line matrix is given by [18].

-2	-j	-j	1
-j	-2j	1	1
j	1	-j	0
2	1	0	0

- Point Features. Since the number of points in a given space is dimensionless, the algorithm proposes points as circles with small diameters. As the number of points increases, the number of circles can be decreased and vice versa. Unlike lines and edges, point detection helps in segmenting less-density regions. The point matrix is represented by [19][20].

1	-j	-j	1
j	1	1	1
j	1	1	1
1	j	0	1

Apart from the features mentioned above, the algorithm supports custom features based on the screen dimension. Considering figure 3, the following masking equation gives the feature we need to adopt. The hit-and-trial method chooses the values by setting different numbers until we get a proper rectangular feature.

$$(0.5 + 0.4 + 0.5 + 0.6 + 0.4 + 0.7 + 0.5 + 0.4 + 0.4 + 0.5 + 0.6 + 0.8 + 0.5 + 0.7 + 0.6 + 0.6) - (0.1 + 0.1 + 0.2 + 0.2 + 0.2 + 0.1 + 0.2 + 0.2 + 0.2 + 0.3 + 0.2 + 0.1 + 0.2 + 0.3 + 0.2 + 0.2)$$

$$B - W = 8.7 - 3 = 5.7 \quad (4)$$

Where B and W are the base and width of the screen from the centre in mm, the above mask leads to a right rectangular feature as the value is less than 8 and positive. In the case of noisy images, the masking values can be increased, ensuring the result is always positive. For non-uniform masking, mask values greater than one can be considered. Circular and polygonal features can also be extracted; however, they are not accurate for the chosen screen dimension. With a 5.7 rectangular feature, the user sits before the webcam to proceed with the single-eye tracking. As shown in figure 4, the first region of interest – the face is segmented and highlighted, and the second region of interest – is the iris (right eye) and tracking its movements. The segmentation is done using the Hough transformation.

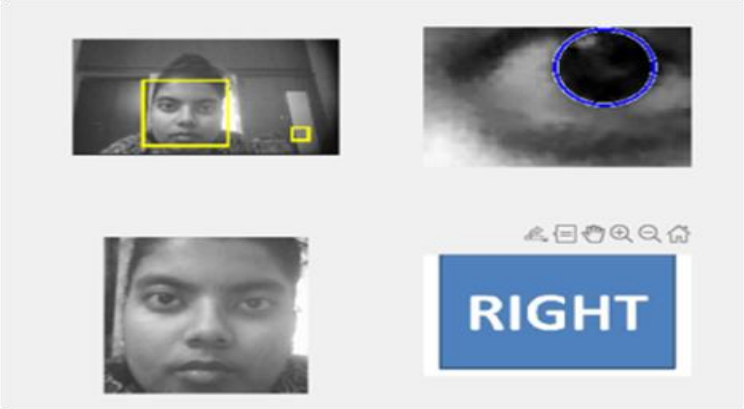

Fig. 4: Single eye-tracking (Each picture dimension represents the screen dimension – ABCDEF)

Since the user's pupils are tracked live, it can be used for continuous and instant monitoring. The images are captured at a 300 x 125 range at a 1280 frame rate for further training data collection. The captured images are then exposed to image dilation and erosion to fill the holes and regions of lower density. The live stream will be binarized to track the pupil in white and the background as black (illustrated in figure 5). the corresponding screen coordinates and the position is also displayed. Till this step, the training data is collected.

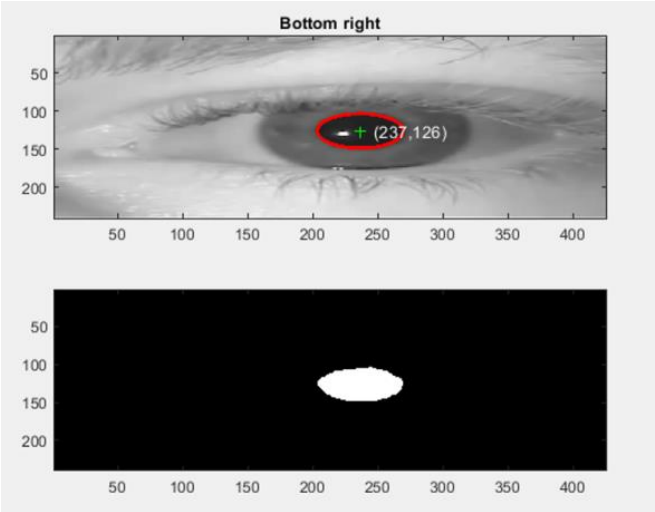

Fig. 5: Binarized Pupil Tracking [X and Y axis are the screen coordinates]

To acquire the testing data for module 1, an interactive environment (website) was created that obeys Hough and Gazer algorithms [21]. The environment has the following features.

- The user calibrates the website when he moves the cursor in random directions. These curser movements are taken as cluster data, followed by the user focusing on the screen in random directions with the eyes. The former and latter data are compared to give calibration accuracy.
- The environment has 9 edge dots of diameter 0.1mm, each placed at a distance of 0.23 of the total screen distance length and breadthwise.
- The center dot is at an equal symmetrical distance from all 9 dots.
- The user's face will be bounded by neural lines to ensure proper seating and lighting conditions.
- The gazer environment will be imported with the training data obtained previously.

4.1 Experimental Procedure for Module 1

The user will be asked to sit at a table distance (ruler distance) from the webcam with the eyes at 180 degrees with the webcam lens with proper lighting conditions. The environment will then do a webcam compatibility and lighting check, and if it is satisfied, a neural boundary is created with green lasers. If the lasers turn red, the user must relocate and seat properly to satisfy the conditions. The length of the neural lines indicates the distance of each face object from the target (pupil), and the dotted regions indicate the efficiency of the user looking at the webcam, as shown in figure 6. The dots disappear if the user is not looking at the webcam.

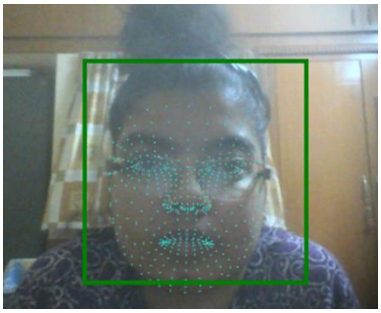

Fig. 6: Green Neural Boundary

The user with the mouse cursor will then focus on the 9 points in the environment. Each point will have an intensity difference from dark red to light yellow, and the user has to keep focusing on each dot until they turn from dark red to light yellow. This color variation corresponds to watershed transformation in determining the complement dot colors and calibrating the environment. Table 3 describes the time taken for the dots to change in different phoretic conditions.

Table 3: Calibration Time

Movement	Condition	Time Taken(s)
Saccade – fixation dual	Crossed eyes	10
All types of vergence	Lazy eye	10.12
Infraversion	Metabolic disorders	12
Pursuit Movement	Nystagmus	10.2
Upper-end movement	Walleye	11.4
Lower end movement	General swollen eye	8.5
Rapid – Slow interval	Eye twitches	8
Torsion	Blinker syndrome	10.5
Anti torsion	Anti-crossed eyes	11

Figure 7 shows the user experiment. Once the user calibrates the 9 points, the centre dot appears, which has to be focused with the eyes for around 5 to 7 seconds, depending on the calibrated time.

Fig. 7: User Performing experiment

While the center dot is focused, the mapped dots (locations wherever the user is focused) will revolve around the center. The closer the peripheral dots are to the centre, the more accurate the calibration. Figure 8 shows the center dot bounded by secondary mapped dots. This principle follows the gazer and complementary algorithm. Figure 9 shows the accuracy of the calibration. The three points on the curve indicate three reference points (screen coordinates Vs. user's eye coordinates) taken as the test data, and it is seen that the test data perfectly coincides with the exponential curve.

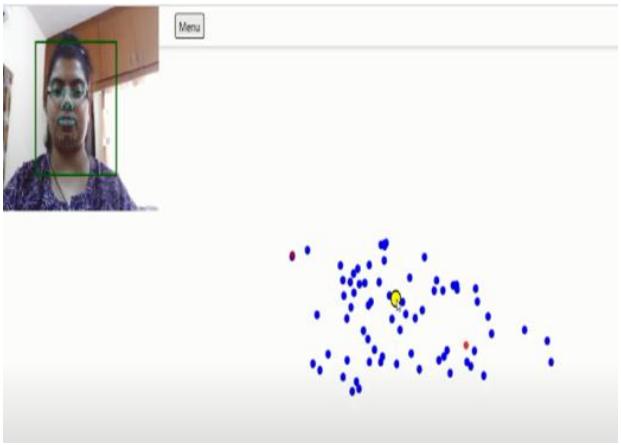

Fig. 8: Centre and Mapped dots

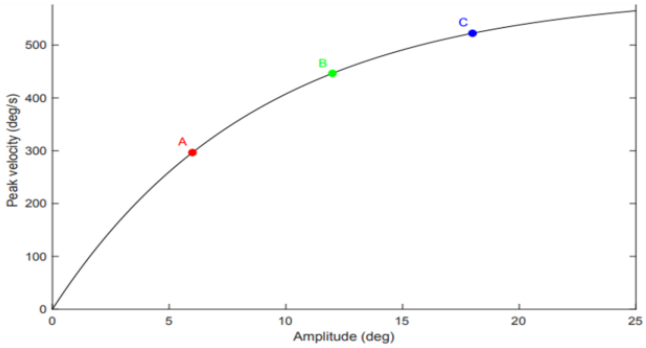

Fig. 9: Accuracy of Calibration

Since the above method uses mapping, it is subjective accuracy (without webcam delay). To perform clinical analysis, we must also consider objective analysis (with webcam delay). For this, the environment will be added with a set of concentric circles with the diameter as the decimal frame rate. There will be two points (primary and sub-main). The leading dot will be fixed in the innermost circle center. The sub-main point is moving and will decide the objective accuracy. When the user focuses on the primary dot, the sub-main point approaches the primary dot, and the more closely it approaches, the more accurate the system is shown in figure 10. The focus time should be at least 4 to 5 seconds for healthy eyes. It is noted from figure 10 that for a 9-point scale experiment, the sub-main dot is pretty far from the primary dot, and the primary dot is not localized in the center of the inner circle. These may lead to mispredictions and count-down errors (errors in dwell time). To overcome these errors, a 42-point scale was developed. This will increase screen time and reduce calibration errors. Improving lighting and neural boundary condition will add to the accuracy.

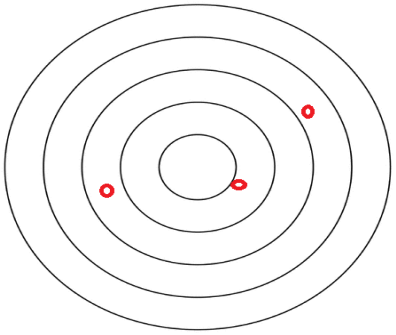

Fig. 10: Objective Accuracy in 9 Scale

4.2 Experimental Procedure for Module 2

Unlike the previous module, the point size increases with a reduced distance between points. The user has to focus on the 42 dots. The following are the features of the environment.

- The environment uses only a single intensity point (red), unlike module 1
- The experiment is divided into two phases – in the first phase, only four dots are focused as calibration, and the second phase has 38 dots.
- The background of the environment changes in color to compensate for lighting errors in available technologies.

The user's face stays inside the green boundary to indicate the correct posture. From figure 11, it is seen that till the 34th point, the environment is light; till the 26th point, the environment is dark; and from the 13th point, it is gray scaled. This will help to compensate for the unsatisfactory lighting conditions.

Fig. 11: Lighting Variation: Depending on the illumination, the background color changes to compensate for the bad lighting delay from white and grey to black.

Once all the 42 points are focused, the end mapped points are focused. The dots bulge and explode upon focusing, indicating that the pupil coordinates are tracked and illustrated in figure 12.

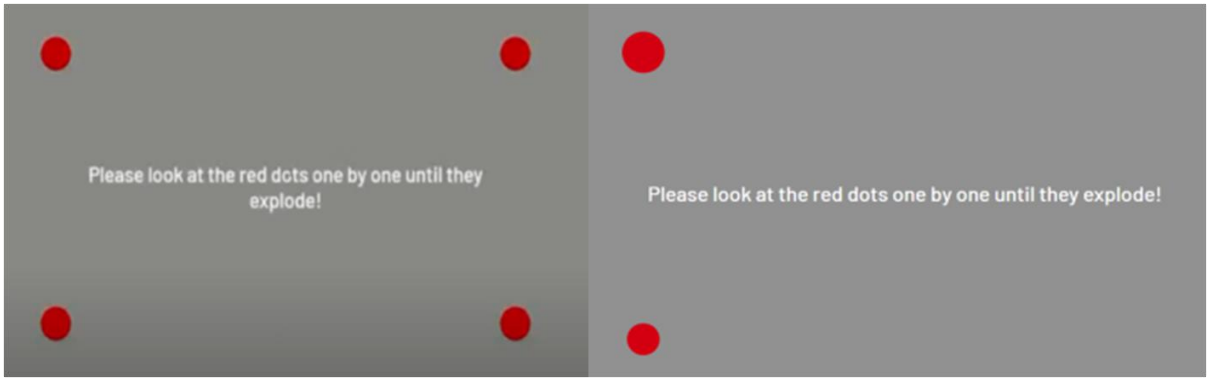

Fig. 12: Exploding Dots: As the user focuses on each dot, they get bigger and explode, thus tracking the pupil

When the objective accuracy of the 42-scaled system is calculated, it is found that the primary dot is nearly at the center of the inner circle. Still, the sub-main point has less precision than the 9-point scale system, as shown in figure 13. This limitation can be mitigated by adding more environmental backgrounds and reducing frame delays.

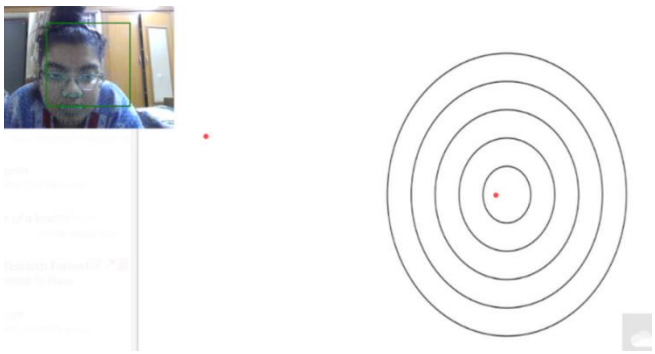

Fig. 13: Objective Accuracy in 42 Scale

4.3 Heat Maps Generation

The acquired testing and training data are imported to realeye.io [22][23] for heat map generation. The user will be asked to surf a website, and the regions the iris focuses on will be heat mapped, as shown in figure 14. The intensity difference in heat maps will be interpreted as a pixel matrix to describe the pressurized and non – pressurized regions. The tracked data can be transferred to MATLAB for further processing. Annotating (equity mapping) helps in understanding the stability of focus at one particular point (fixation) and how linear the user can shift the focus (saccade).

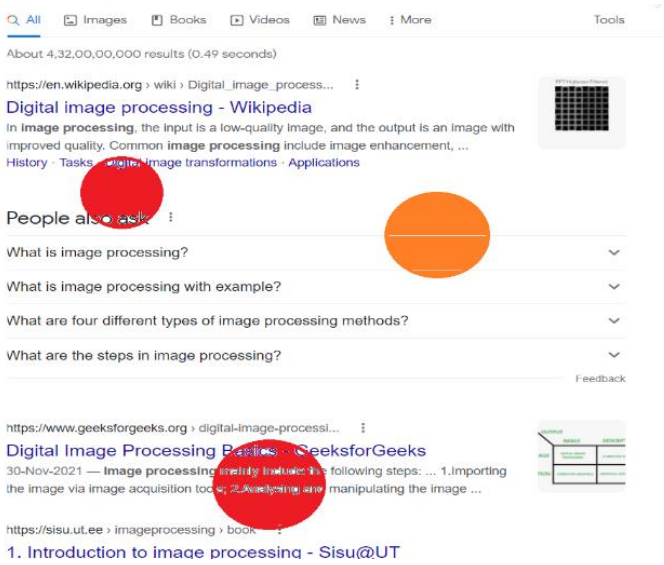

Fig. 14: Website heat maps

The interpretation and code of the above heatmap are as follows.

Table 4: Heatmap Interpretation

Heat Color	Pixel value	Binary Value	No. of Gaze Points
Red	234	101010111	1900
Cyan	1000	0101011101	2389
Blue	345	010110110	1200
Green	98	1111011100	2000
White	256	111111111	50000
Black	0	000000000	0

4.4 Blinking and refractive Error Analysis

The Nose pad of every spectacle encounters the origination of the extraocular nerves, which is essential to analyze the blinking pattern of the eyes. The nose pad connected to the circuit is shown in figure 15. The below schematic can be made into a wearable using conductive fabric attached to a false eyelid. The nose pad will be made of conductive piezo rubber, connected to a two-stage amplifier, followed by filters and acquisition parts. The conductive piezo rubber will change its resistance depending on the force exerted due to the blink. The microprobes are connected to ensure the timing and leakage of the system. (Blue probe – excitation potential due to piezo and green probe – reference potential)

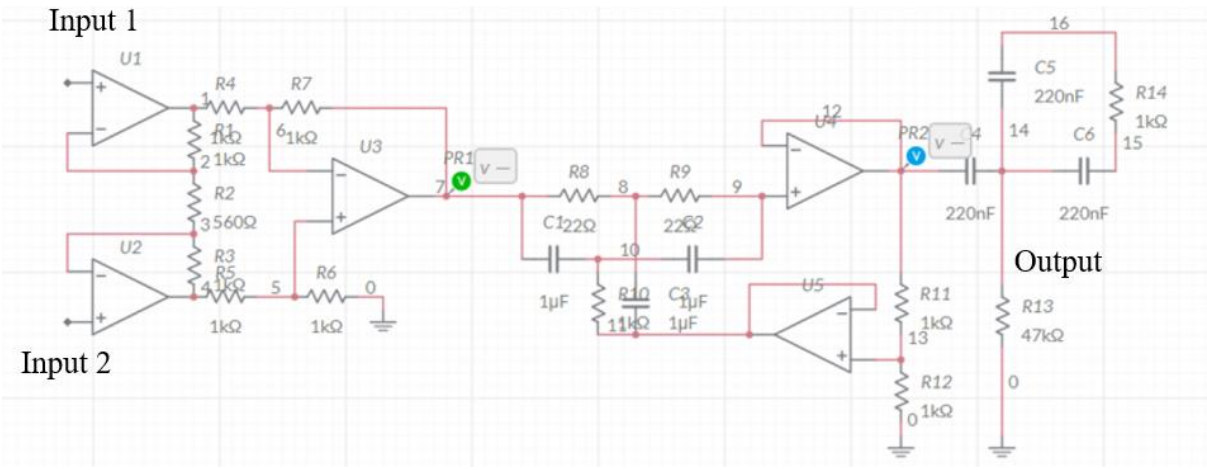

Fig. 15: Blink Circuit

Figure 16 describes the blinking pattern obtained. The graphical data is converted to .txt and imported to MATLAB for further analysis.

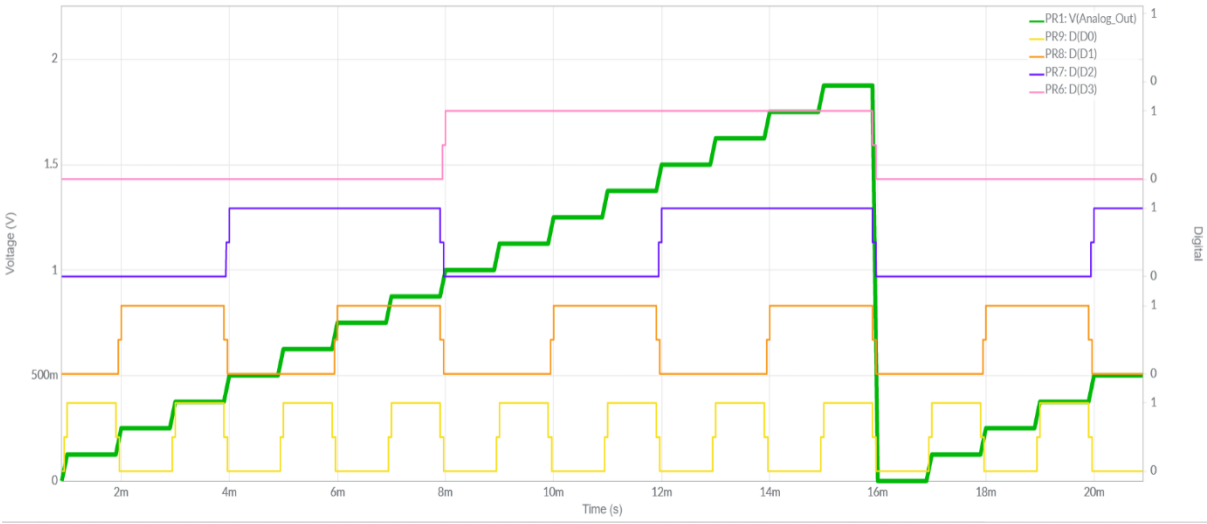

Fig. 16: Blink Pattern

Using the Principle of Tonometry, the refractive error can be determined by the following equation [23][27].

$$\frac{Ehg}{R}.\frac{S}{g}.\left[2-\frac{s}{g}\right] \qquad (5)$$

Where s,g, and h represent peak points indicated by different probes, R is the divisor threshold, and E is the applanation constant. For a healthy individual, the applanation graph is exponential for the given brightness.

4.5 Accuracy of the Proposed System

Simple Gazer algorithms were developed to calculate the precision of the proposed system, as shown in figure 17.

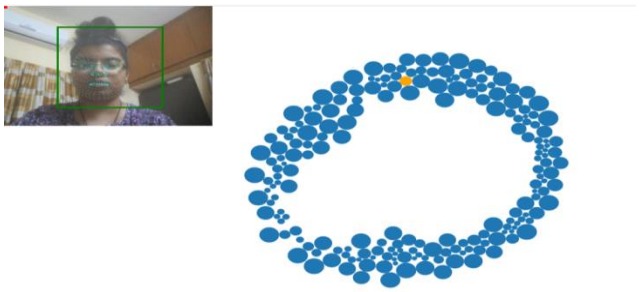

Fig. 17: Precision calculation

It is observed that wherever the eye is focusing, the dots equally distribute around the focused region in an elliptical manner. The non–uniform distribution indicates less precision and vice versa. Poor lighting and webcam delay can affect the precision rate. Ten trials were conducted, and the precision increased from 78% to 89% upon decreasing and increasing the screen size and number of dots, respectively.

5.0 RESULTS AND DISCUSSIONS

The following oculograms were obtained for different waveforms by importing the training and testing data.

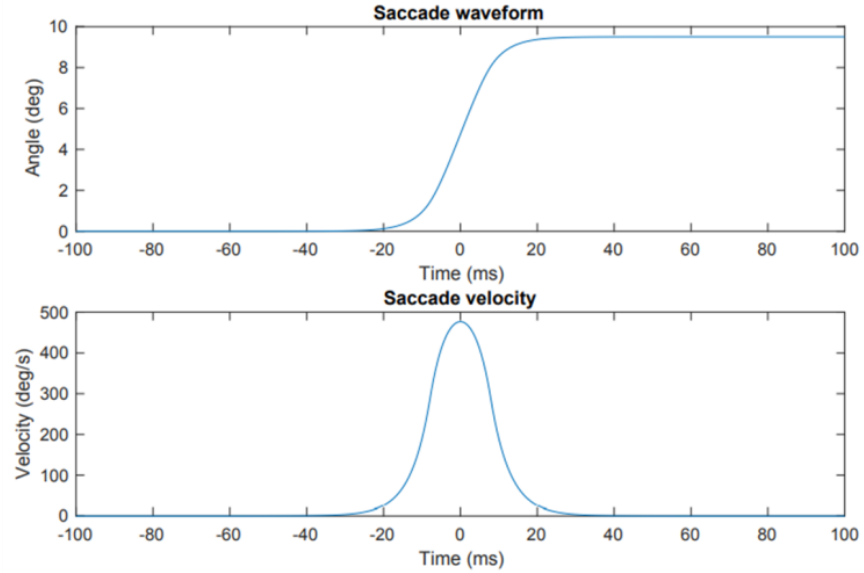

Fig. 18: Saccade Pattern

The saccade waveform undershoots at -20th ms and overshoots after a level interval from +20th ms (illustrated in figure 18). This transition evaluates the user's eyes while reading the activity. The saccade velocity was found to be 500 deg/s for every shoot. When the user is made to read a paragraph continuously, a series of saccade waveforms are obtained and illustrated in figure 19.

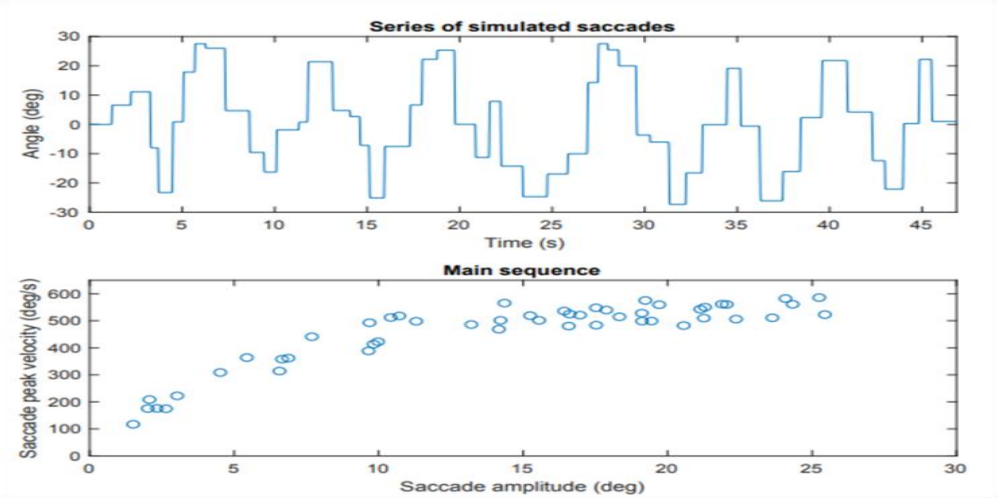

Fig. 19: Series Saccade Pattern

The smooth pursuit pattern for various forced blinking was obtained (in figure 20) where A represents normal blinking and B and C represents forceful blinking at higher and lower levels of the monitor screens. To test the parametric saccade model, the user was asked to blink every 4 seconds and read the paragraphs along with light dimming. Figure 21 shows that an abnormal peak saccade leads to a healthy saccade peak.

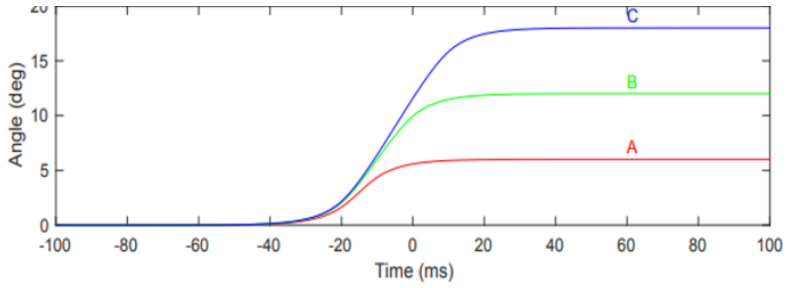

Fig. 20: Smooth Pursuit Pattern

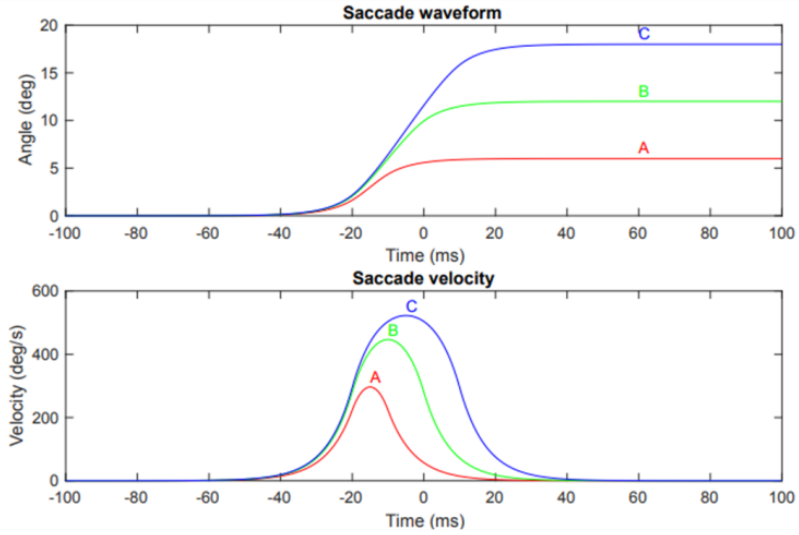

Fig. 21: Abnormal Saccade Pattern

Table 5 highlights the results for each eye location tracked cluster-wise (for the first 18 clusters).

Table 5: Final Results

	Saccade		Fixation		Vergence	
Cluster #	**Peak**	**Blink rate**	**Peak**	**Blink rate**	**Peak**	**Blink rate**
#1	102.3	42.0	171.4	48.7	134.0	118.9
#2	102.0	42.0	135.1	4.9	114.3	7.3
#3	101.4	16.0	119.6	2.1	124.1	4.5
#4	102.0	2.0	112.0	1.5	136.0	1.6
#5	101.1	9.0	121.8	2.9	139.6	3.4
#6	102.6	25.0	261.5	1.9	148.2	5.6
#7	102.1	7.0	161.8	2.1	157.7	3.0
#8	102.0	1,0	110.0	1.0	185.0	1.1
#9	101.5	2.0	132.0	1.5	194.0	1.6
#10	102.0	2.0	122.0	3.0	179.5	1.6
#11	101.7	18.0	223.3	7.8	127.2	4.8
#12	103.0	1.0	160.0	5.0	152.0	1.1
#13	101.0	1.0	120.0	5.0	161.0	1.1
#14	102.0	1.0	150.0	5.0	197.0	1.1
#15	102.1	17.0	114.4	8.3	134.8	4.7
#16	101.5	2.0	210.0	7.0	155.5	1.6
#17	102.3	6.0	511.5	8.5	140.0	2.8
#18	103.0	4.0	113.6	8.5	150.5	2.3

From the oculogrmas extracted, a parametric saccade model was developed at various rates, as shown in figure 22.

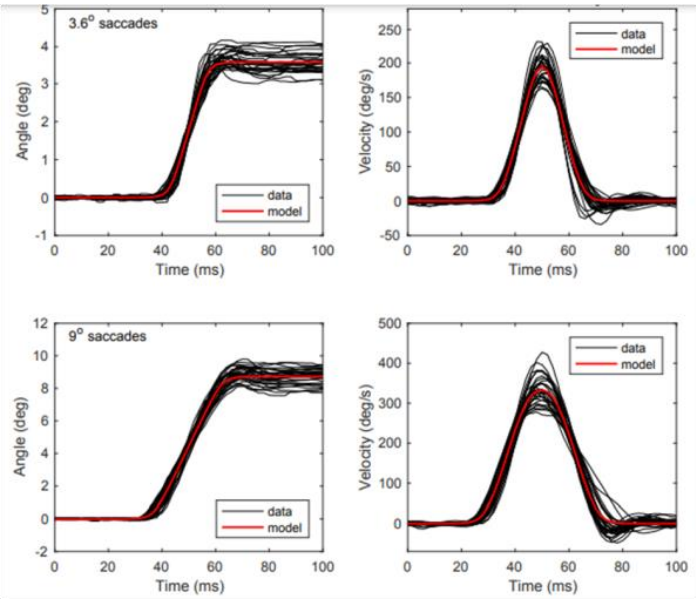

Fig. 22: Varied saccade pattern model

5.1 Precision Measure of the Proposed Tracking System

The term precision refers to the accuracy measurement where the model can provide similar results with approximate efficiency in each trial. The proposed model was tested for 35 trials in a single webcam take, and the precision was calculated. The testing screen was measured for the centre point (yellow in the figure), which acts as the track point. The trial points get mapped based on the results obtained after each trial (blue in the figure). The red dot is the point of action whose distance from the trial points indicates the redundancy error, which can be improved by decreasing the blinking rate. Figure 23 shows the precision results.

Fig. 23: Precision Analysis of the Proposed Model

5.2 F-Score and Recall of Proposed and Standard Algorithms

The points width most frequently localized by the eyes was divided into seven known classes based on the distance from the center of the screen. Each trial's maximum localization was calculated and classified into those classes using the MIFNET algorithm [24]. Figure 24 shows the obtained confusion matrix for the seven classes. The recall and the overall F measure from the confusion matrix can be calculated from equation 6 [25][26].

$$Precision = \frac{TP}{TP+FP} \quad ; \quad Recall = \frac{TP}{TP+FN} \tag{6}$$

$$\text{F-Score} = 2 * \frac{PR}{P+R} \tag{7}$$

Where P and R are precision and recall coefficients, TP and FP are true and false positive values, and TN and FN are true and false negative values. In the confusion matrix, classes with black are true positives, white are true negatives, grey is false positives, and light gray is false negatives.

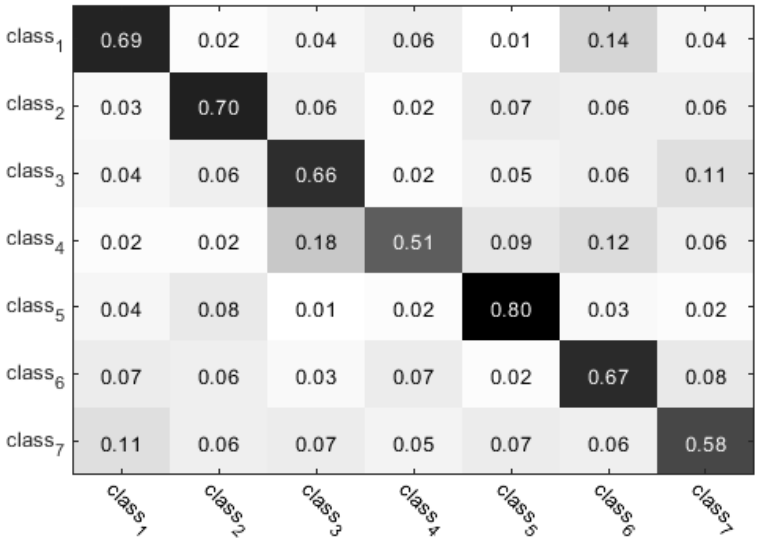

Fig. 24: Confusion Matrix

Some of the standard algorithms currently used are gaze recorder and webeye. Table 6 shows the average precision and recall values for the proposed and benchmark algorithms at different trial points concerning the screen center point and F measure difference. In terms of precision, the total averages of the gaze recorder and webeye were 0.864 and 0.789 on average, respectively. In contrast, the proposed algorithm resulted in a score of 0.978. In terms of recall, the total averages of the algorithms were 0.865 and 0.738, whereas the proposed algorithm resulted in 0.988. Therefore, the average F1 score for the proposed algorithm was up to 0.284 and 0.470 higher than those of the two algorithms. This means that the localization accuracy of the proposed algorithm was higher than that of the standard algorithms.

Table 6: Comparison of proposed method performance measures with the existing models

Distance (mm)	Gaze Recorder [24]			Webeye [25]			Proposed Model		
	Precision	Recall	F-Score	Precision	Recall	F-Score	Precision	Recall	F-Score
1.1	0.712	0.923	0.91	0.725	0.981	0.87	1.00	0.998	0.99
2.1	0.732	0.901	0.90	0.734	0.993	0.89	1.00	0.996	0.98
3.1	0.721	0.913	0.91	0.714	0.911	0.87	1.00	0.987	0.98
4.1	0.714	0.934	0.91	0.789	0.932	0.89	1.00	0.996	0.98
5.1	0.743	0.931	0.90	0.799	0.931	0.87	1.00	0.996	0.97
6.1	0.754	0.914	0.90	0.791	0.914	0.87	1.00	0.997	0.98
7.1	0.775	0.911	0.87	0.734	0.900	0.89	0.997	0.999	0.96

5.3 CONCLUSION AND FUTURE WORKS

Currently, the proposed method shows accurate readings for blinks less than 25 above, and the ‘face not centered’ message pops up. This feature can be extended to nonblinking characteristics to study eye disorders more precisely. The method revolves around only one colored heat map, which will be extended to 5 colored maps to increase the number of features involved. It is observed that for the rate of 3.6 and 9 (up to 10 trials), there is a difference in the peak and the overshoot values, which plays a significant role in iris biometric systems. Illumination less than average room lighting can also affect current readings (but with low error values), which can enhance pipeline measures or laptop flashes. The method targets franchise-level hospitals and home users who can send monitored data via a connected BOLT module.

REFERENCES

[1] Juhong, T. Treebupachatsakul, and C. Pintavirooj, "Smart eye-tracking system", *2018 International Workshop on Advanced Image Technology (IWAIT)*, 2018, pp. 1-4, doi: 10.1109/IWAIT.2018.8369701.

[2] P. A. Punde, M. E. Jadhav and R. R. Manza, "A study of eye-tracking technology and its applications", *2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)*, 2017, pp. 86-90, doi: 10.1109/ICISIM.2017.8122153.

[3] Rachel Emy Straus Takahasi, Hyoun S. Him, Sophie G. Coelho and Hermano Tavares, “A Systematic Review of Eye-Tracking Studies of Gambling-Related Attentional Biases”, *Journal of Gambling Studies*, 2022, https://doi.org/10.1007/s10899-022-10161-3.

[4] Li, Y. Zhang, X. Zheng, X. Huang, S. Zhang and J. He, "A Smart Eye Tracking System for Virtual Reality", *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, 2019, pp. 1-3, doi: 10.1109/IMBIOC.2019.8777841.

[5] X Zhang, H Kulkarni and M.R Morris, "Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities", *CHI Conference on Human Factors in Computing Systems*, 2017.

[6] C Singer and B. Hartmann, "See-Thru: Towards Minimally Obstructive Eye-Controlled Wheelchair Interfaces", *21st International ACM SIGACESS Conference on Computers and Accessibility*, 2019.

[7] Vehlen, A., Spenthof, I., Tonsing, D. et al. “Evaluation of an eye-tracking setup for studying visual attention in face-to-face conversations”, *Sci Rep*, 11, 2661, 2021. https://doi.org/10.1038/s41598-021-81987-x.

[8] Peter Kiefer, Ioannis Giannopoulos, Martin Raubal & Andrew Duchowski, “Eye-tracking for spatial research: Cognition, computation, challenges”, *Spatial Cognition & Computation*, 17:1-2, pp 1-19, 2017. doi: 10.1080/13875868.2016.1254634

[9] Jue Li, Heng Li, Waleed Umer, Hongwei Wang, Xuejiao Xing, Shukai Zhao, Jun Hou, “Identification and classification of construction equipment operators' mental fatigue using wearable eye-tracking technology”, *Automation in Construction*, vol. 109, 2020 ,103000,ISSN 0926-5805.

[10] Clay, V., Konig, P., & König, S, “Eye Tracking in Virtual Reality”, *Journal of eye movement research*, vol. 12, issue 1, 2019, 10.16910/jemr.12.1.3. https://doi.org/10.16910/jemr.12.1.3.

[11] Yirui Wu, Wenxiang Liu, Shaohua Wan, “Multiple attention encoded cascade R-CNN for scene text detection, *Journal of Visual Communication and Image Representation*,vol. 80,2021,103261,, pp. 23-31.

[12] S. Sheela and K. R. Radhika, "Feature-based Methods for Eye Gaze Tracking,"*2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020, pp. 1101-1107, doi: 10.1109/ICECA49313.2020.9297578.

[13] Sauradip Nag, Palaiahnakote Shivakumara, Yirui Wu, Umapada Pal, Tong Lu: “New COLD Feature Based Handwriting Analysis for Enthnicity/Nationality Identification”. *ICFHR* 2018: pp 523-527.

[14] Y Wu, Y He, P Shivakumara, Z Li, H Guo, T Lu, "Channel-wise attention model-based fire and rating level detection in video", *CAAI Transactions on Intelligence Technology*, 2019, pp-117-121.

[15] Charles C.-F. Or, Benjamin K. Goh, Alan L.F. Lee, “The roles of gaze and head orientation in face categorization during the rapid serial visual presentation”, *Vision Research*, vol. 188, 2021, pp 65-73, ISSN 0042-6989.

[16] Liu, T., and Zhu, S., "Eyes Detection and Tracking based on Entropy in Particle Filter", *International Conf. on Control and Automation*, Budapest, 2019, pp. 1002-1007.

[17] S. Akshay, Y. J. Megha and C. B. Shetty, "Machine Learning Algorithm to Identify Eye Movement Metrics using Raw Eye Tracking Data," *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2020, pp. 949-955, doi: 10.1109/ICSSIT48917.2020.9214290.

[18] K. Panetta et al., "ISeeColor: Method for Advanced Visual Analytics of Eye Tracking Data," in *IEEE Access*, vol. 8, 2020, pp. 52278-52287, doi: 10.1109/ACCESS.2020.2980901.

[19] Carr DB, Grover P. “The Role of Eye Tracking Technology in Assessing Older Driver Safety”. *Geriatrics*. 5(2), 2020, 36. https://doi.org/10.3390/geriatrics5020036.

[20] J. Zhu et al., "An Improved Classification Model for Depression Detection Using EEG and Eye Tracking Data," in *IEEE Transactions on NanoBioscience*, vol. 19, no. 3, July 2020, pp. 527-537, doi: 10.1109/TNB.2020.2990690.

[21] Nij Bijvank JA, Petzold A, Balk LJ, Tan HS, Uitdehaag BMJ, Theodorou M, et al., “A standardized protocol for quantification of saccadic eye movements: DEMoNS”, *PLoS ONE*, 13(7), 2018, e0200695.

[22] Bargary G, Bosten JM, Goodbourn PT, Lawrance-Owen AJ, Hogg RE, Mollon JD, “Individual differences in human eye movements: An oculomotor signature?”, *Vision Res*.141, 2017, pp 157–69. pmid:28373058.

[23] Alexander Goettker, Karl R. Gegenfurtner, “A change in perspective: The interaction of saccadic and pursuit eye movements in oculomotor control and perception”, *Vision Research*, vol. 188, 2021, pp 283-296, ISSN 0042-6989.

[24] Kredel R, Vater C, Klostermann A, Hossner EJ, “Eye-Tracking Technology and the Dynamics of Natural Gaze Behavior in Sports: A Systematic Review of 40 Years of Research”, *Front Psychol*.,8, 2017 Oct 17, 1845. doi: 10.3389/fpsyg.2017.01845. PMID: 29089918; PMCID: PMC5651090.

[25] Zdarsky Niklas, Treue Stefan, Esghaei Moein, “A Deep Learning-Based Approach to Video-Based Eye Tracking for Human Psychophysics”, *Frontiers in Human Neuroscience*, vol. 15, 2021. https://doi.org/10.3389/fnhum.2021.685830.

[26] Ildar Rakhmatulin, Andrew T. Duchowski, "Deep Neural Networks for Low-Cost Eye Tracking", *Procedia Computer Science*, vol. 176, 2020, pp 685-694. https://doi.org/10.1016/j.procs.2020.09.041.

[27] Ishtiaque Ahmed, N., Nasrin, F, *"Reducing Error Rate for Eye-Tracking System by Applying SVM"* In: Skala, V., Singh, T.P., Choudhury, T., Tomar, R., Abul Bashar, M. (eds) Machine Intelligence and Data Science Applications. Lecture Notes on Data Engineering and Communications Technologies, vol. 132, 2022. Springer, Singapore. https://doi.org/10.1007/978-981-19-2347-0_4