

Augmenting Early Stroke Diagnosis With an Eye-Tracker

Mohamed Abul Hassan

hassanutp@gmail.com

University California Davis <https://orcid.org/0000-0002-3076-8075>

Yan Zhuang

University of Virginia

Mohammed E-Rabbi

University of Virginia

Chad Aldridge

University of Virginia Health System

Andrew Southerland,

University of Virginia Health System

Gustavo Rohde

University of Virginia

Article

Keywords:

Posted Date: August 8th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4656842/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations:

(Not answered)

Competing interests: The authors declare no competing interests.

Augmenting Early Stroke Diagnosis With an Eye-Tracker

^{a,b} Mohamed Abul Hassan, ^c Yan Zhuang, ^c Mohammad Shifat-E-Rabbi, ^d Chad M. Aldridge, ^d Andrew M. Southerland, ^{*,b,c} Gustavo K. Rohde

^aDepartment of Biomedical Engineering, University of California Davis, USA

^bImaging and Data Science Lab, Department of Biomedical Engineering, University of Virginia, USA

^cImaging and Data Science Lab, Electrical and Computer Engineering, University of Virginia, USA

^dDepartment of Neurology, University of Virginia, USA

Abstract—Posterior circulation stroke (PCS) presents significant diagnostic challenges due to poorly localizing and non-specific symptoms, such as dizziness, nausea, and headache, which are often misattributed to benign conditions. This study introduces an innovative diagnostic tool that utilizes a machine learning algorithm-driven eye tracker to enhance early diagnosis of PCS. Our approach involves analyzing eye movements during three standard neurological eye examinations: the Dot Test, H Test, and Optokinetic Nystagmus (OKN) Test. The Discrete Radon Cumulative Distribution Transform (DRCDT) and nearest subspace (NS) classification methods were employed to distinguish between PCS patients and healthy controls by identifying specific eye movement patterns. Results demonstrate that the ensemble model combining the three tests achieved the highest sensitivity and accuracy, with a sensitivity of 96% and an accuracy of 88%, in diagnosing PCS. This study's findings underscore the potential of an eye-tracker-based diagnostic tool to support a more accurate and efficient diagnosis, particularly for non-neurology trained providers, which would improve patient outcomes with more timely and appropriate treatment. The proposed tool offers a practical solution to the limitations of current diagnostic methods, such as the need for calibration and reliance on highly trained specialists, and can be seamlessly integrated into clinical settings to support emergency medical services (EMS) and emergency department (ED) triage.

I. INTRODUCTION

Stroke is a significant contributor to mortality and long-term disability worldwide, with anterior circulation stroke (ACS) being the most prevalent type and posterior circulation stroke (PCS) being less common and often misdiagnosed [1]–[3]. PCS can have non-specific and variable symptoms, such as dizziness, nausea and vomiting, and headache, similar to other more commonly presenting conditions such as migraine, metabolic disturbances, infections, and peripheral vestibulopathies [4], [5]. In contrast, the symptoms of ACS, including facial asymmetry, arm weakness, and speech disturbances, are more specific and easier to recognize as signs of stroke, particularly for non-neurology trained providers [6].

The HINTS (Head Impulse, Nystagmus, Test of Skew) examination is a bedside diagnostic approach to help differentiate PCS from an inner ear disorder in patients presenting

with acute vestibular syndrome [7]. However, the accuracy of the HINTS exam depends on a certain level of training and experience, which is often not the case for lay providers in the emergency setting [8], [9], [10], [11]. Thus, there is a need to develop augmented diagnostic tools not reliant on provider experience and training to help detect PCS in the emergency setting.

PCS may be differentiated from a peripheral disorder by abnormal patterns of eye movements, which are often subtle and difficult to distinguish from benign conditions. Commonly observed abnormalities in eye movements include nystagmus, fixed gaze deviation, and dysconjugacy, which neurologists and ophthalmologists are trained to examine at the bedside. For an automated diagnostic tool to accurately diagnose PCS, it would need to replicate the ability of a specialist to correctly discern these oculomotor abnormalities in real time.

Eye tracking has emerged as a promising tool for diagnosing neuro-ocular disorders. Eye tracking may be used to characterize various neurological conditions, including traumatic brain injury, Parkinson's disease, and multiple sclerosis [12]–[17]. Previously, we demonstrated that a non-calibrated eye tracker can be used to assess eye movement symmetry and variability in PCS patients [18]. The calibration procedure posed an obstacle to the translation of the eye tracker for clinical settings, as patients affected by stroke and other neurological conditions may have difficulty fixating on a target or following a moving object; i.e., the prerequisite for calibrating an eye tracker.

This study proposes a novel approach to augment the diagnosis of PCS using a commercial, off-the-shelf eye tracker. By analyzing eye movements during computer-adapted versions of three standard bedside oculomotor tests (i.e., the Dot-Test, H-Test, and OKN-Test), we aim to identify patterns to aid in the early and accurate diagnosis of PCS. Additionally, acquiring data from PCS patients is challenging due to the nature of their illness, which can affect patient compliance during the examination. To address these challenges, we will use a custom-built mobile rig equipped with compliance protocols for data acquisition. The study involves recruiting patients with PCS and healthy controls and recording their eye movements using an eye tracker during three oculomotor tests. These tests

investigate gaze-evoked nystagmus, smooth pursuit, saccades, and the vestibular-ocular reflex. Given that PCS presents less commonly than anterior circulation stroke, our machine-learning algorithms must be data-efficient. The resulting eye-tracking data will be analyzed using advanced pattern recognition and machine learning algorithms to identify specific eye movement patterns characteristic of PCS.

Our approach employs the Discrete Radon Cumulative Distribution Transform (DRCDT) and nearest subspace (NS) classification model. The DRCDT is applied to gaze points from both left and right eye movements to maximize information on eye conjugacy while being invariant to deformation caused by non-calibrated eye tracking data. This method generates a convex set of gaze point distributions in the DRCDT space, making them linearly separable and thus suitable for classification using the NS model. These factors will contribute to the robustness and applicability of our work.

The overall diagnostic tool comprised an ensemble model based on three separate DRCDT-NS classifiers for the individual tests: the Dot-Test, H-Test, and OKN-Test. Each of these neurological eye examinations interrogates different ocular motor abnormalities, and a separate diagnosis of PCS per test may result in the potential for false positives or negatives (see Fig. 1). Therefore, the ensemble model allows for a more accurate and efficient diagnosis by combining the diagnostic power of the three neurological eye examinations. The results are compared to the final diagnosis of a vascular neurologist to assess the accuracy and efficiency of the proposed approach.

To summarize, this work makes significant contributions:

- Proposing a novel diagnostic tool that utilizes a non-calibrated eye tracker combined with machine learning algorithms to accurately distinguish between posterior circulation stroke (PCS) patients and healthy controls. This tool addresses the limitations of current diagnostic methods, such as the need for calibration and the reliance on highly trained specialists.
- The systematic validation of the tool on a cohort of PCS patients and healthy controls demonstrated high sensitivity (96%) and accuracy (88%) to detect PCS, underscoring its potential for early and precise diagnosis.
- This diagnostic tool offers a practical solution that can be seamlessly integrated into clinical settings, supporting emergency medical services (EMS) and emergency department (ED) triage. By enhancing the accuracy and efficiency of PCS diagnosis, this tool holds promise to improve patient outcomes through timely and appropriate stroke treatment.

II. RESULTS

A. Fully Automated PCS Diagnosis Tool Using Eye-tracker

The validation of the method for early-stage stroke diagnosis encompassed two sets of results: (1) Cohort 1: Stroke patients with abnormal eye movement ($n = 24$) vs. healthy controls ($n = 18$), (2) cohort 2: Stroke patients with abnormal eye movement ($n = 24$) vs. healthy controls ($n = 18$) and stroke patients with no abnormal eye movement ($n = 3$) (see Table. I). The

model for cohort 1 is trained with healthy controls and patients with central neurological symptoms (i.e., exhibiting abnormal eye movements). The analysis from the 10,000 bootstrapped participant-level classification performance of cohort 1 shows that the ensemble model using data from the eye tracker of all three neuro-ophthalmological tests reported the highest sensitivity and accuracy in diagnosing posterior circulation stroke, with a sensitivity of 0.96 and accuracy of 0.88 (see Table. I). The analysis from cohort 2, trained and validated with patients with abnormal eye movement as the positive label and healthy controls along with stroke patients with peripheral neurological symptoms (no abnormal eye movement) reported no change in sensitivity for the ensemble model. However, reported a minor decrease in accuracy and specificity 0.87 and 0.76 respectively.

The path plot and heatmap for the three neurological exams Fig. 2 (a) and 2 (b) illustrate the contrast in gaze data from patients with abnormal eye movement and normal eye. The true positive classification of predicting abnormal eye movement as positive in a patient Fig. 2 (b), and the true negative classification of predicting a patient with normal eye movement as negative. The path plot of the true positive instance with a high degree of variation in the horizontal axis with two distinctly different shapes and orientations of clusters (much pronounced in the H-test) illustrates the cause of pathology during the eye examinations. The Fig. 2 (c) illustrates a false negative classification instance of a patient with abnormal eye movement as negative. The path plot, and heat map for the three neurological exams showed normal conjugate eye movement as referenced to 2 (a) and overall normal eye movement distribution of this study. This led to the point that the pathology of this patient was not detected by the eye examinations performed during this study. The false positive prediction originating primarily from the Dot-test Fig. 2 (d), 2 (e) and H-test Fig. 2 (e). The gaze data from these exams shows eye conjugacy where both eyes move together. However, the gaze data from the control and patient did not follow the overall normal distribution as shown from the instance Fig. 2 (a). Therefore the classification model classified the gaze pattern out of the normal distribution as abnormal eye movement. This probes the question of compliance during the eye examination.

The overall results irrespective of the participant cohort indicate that the ensemble model comprising models specific to each eye examination reported the highest classification performance. Considering the individual tests, the Dot Test reported the highest sensitivity and accuracy after the ensemble model, with a sensitivity of 0.83 and an accuracy of 0.83. The individual tests reported a higher specificity compared to the ensemble model. Clinicians often prioritize higher sensitivity over specificity, as avoiding a false negative diagnosis is more important than a false positive in the clinical scenario. A false negative diagnosis may lead to delayed treatment, increased morbidity, and mortality, while a false positive diagnosis can be corrected with further testing or follow-up. These results further underscore that the combination of our proposed pattern recognition approach with the eye tracker has the potential to provide a more accurate and efficient diagnosis.

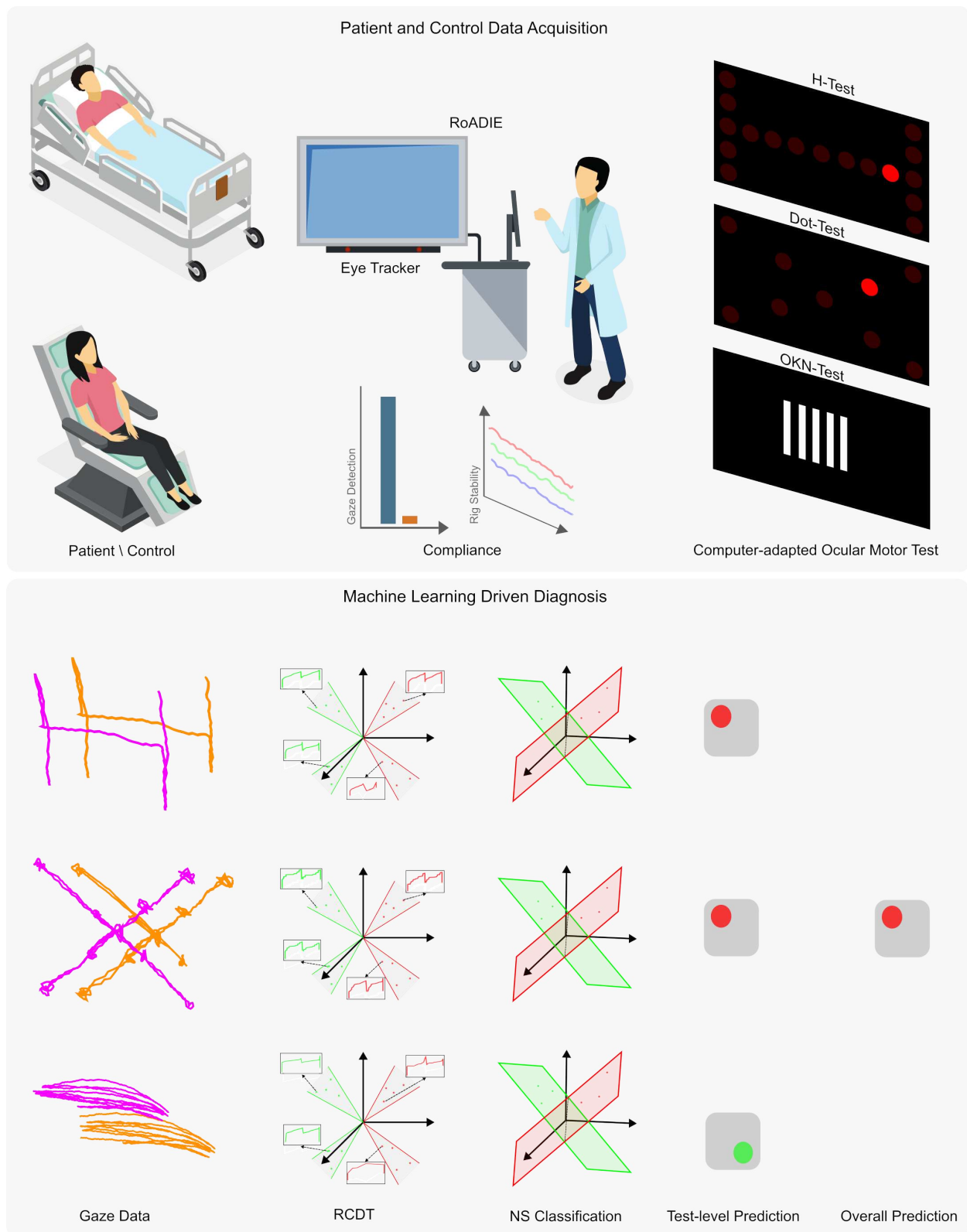


Fig. 1. Overview of the Proposed Diagnostic Tool Using an Eye Tracker. The upper panel presents the data acquisition process, while the lower panel outlines the machine learning adaptation for automatic diagnosis from non-calibrated gaze data. Data were collected from patients and controls either reclined or seated. The Rolling Apparatus to Detect Impairment of the Eyes (RoADIE) is equipped with a screen-based eye tracker to monitor gaze during computer-adapted neurological eye examinations on an extended screen. Compliance with gaze detection and rig stability requirements is ensured (as shown in the upper panel, second column). The upper panel's third column illustrates the digitally adapted bedside ocular motor tests: H-test, Dot-test, and OKN-test. In the lower panel, the first column displays the gaze path captured by the eye tracker during the ocular motor tests. The second column shows the transformation of the gaze point data's 'x' and 'y' coordinates to the RCDT space. The third and fourth columns illustrate the three near-subspace classification models and their predictions for each ocular motor test. Finally, the fifth column depicts the ensemble model that combines the three individual models to provide a binary diagnosis.

TABLE I
MEAN WITH 95% CONFIDENCE INTERVAL (CI) OF CLASSIFICATION PERFORMANCE FOR EACH TEST AND ENSEMBLE MODEL FROM 10,000 BOOTSTRAPPED PARTICIPANT-LEVEL SENSITIVITY, SPECIFICITY, AND ACCURACY

Model	Accuracy	Sensitivity	Specificity
Patients with abnormal eye movement vs Healthy controls			
Dot	0.83 [0.71 - 0.93]	0.83 [0.67 - 0.96]	0.83 [0.65 - 1.0]
H	0.67 [0.52 - 0.81]	0.5 [0.29 - 0.7]	0.89 [0.72 - 1.0]
OKN	0.71 [0.57 - 0.86]	0.54 [0.33 - 0.74]	0.94 [0.81 - 1.0]
Ensemble Model	0.88 [0.79 - 0.98]	0.96 [0.86 - 1.0]	0.78 [0.57 - 0.95]
Patients with abnormal eye movement vs Healthy controls and Patients with no abnormal eye movement			
Dot	0.82 [0.71 - 0.93]	0.83 [0.67 - 0.96]	0.81 [0.63 - 0.96]
H	0.69 [0.56 - 0.82]	0.5 [0.3 - 0.7]	0.9 [0.76 - 1.0]
OKN	0.73 [0.6 - 0.84]	0.54 [0.33 - 0.74]	0.95 [0.85 - 1.0]
Ensemble Model	0.87 [0.76 - 0.96]	0.96 [0.86 - 1.0]	0.76 [0.57 - 0.94]

B. Ablation study and comparison of model performance

In our ablation study, we examined cohort 1, consisting of stroke patients with abnormal eye movements, and compared them to healthy individuals. Our focus was on evaluating two versions of the proposed DRCDT-NS model: one that incorporates deformation modeling and one that does not. Our findings indicate that omitting the deformation vector from the model's learning parameters led to a decrease in the model's accuracy. Specifically, the version of the DRCDT-NS model that omitted deformation modeling showed reduced sensitivity and specificity across tests, with a notably lower specificity of 0.55 in the combined analysis of the three tests (i.e., ensemble).

This drop in performance was attributed to a higher occurrence of false positives, suggesting that the model struggles to distinguish between abnormal and normal eye movements when deformation is not considered. This outcome aligns with the results from the comparative analysis involving another deep learning model, which similarly showed low specificity in its ensemble predictions. However, it's worth mentioning that models like the 1D-VGG and 1D-ResNet demonstrated promising results for individual assessments, such as the H-test and the Dot-test.

The aggregate prediction is crucial for the accurate detection of disease, as it leverages the unique ocular motor functions assessed by each test, which may be affected by the condition in question. A model that minimizes false positives is key to developing a diagnostic tool with significant practical utility. In this context, the DRCDT-NS model, when it includes deformation modeling, proves to be particularly effective, offering high classification accuracy.

III. DISCUSSION

Clinically, PCS is three times more likely to be misdiagnosed than anterior circulation stroke [2]. In our study, we demonstrated that a non-calibrated eye tracker, augmented with ML can effectively diagnose PCS by analyzing eye movements during neurological examinations. This method was applied to a population of stroke patients exhibiting abnormal eye movements compared with healthy controls and stroke patients without abnormal eye movements. These results are particularly significant as the requirement for calibration

TABLE II
ABLATION STUDY AND COMPARISON OF MODEL PERFORMANCE FOR PATIENTS WITH ABNORMAL EYE MOVEMENT VS HEALTHY CONTROLS

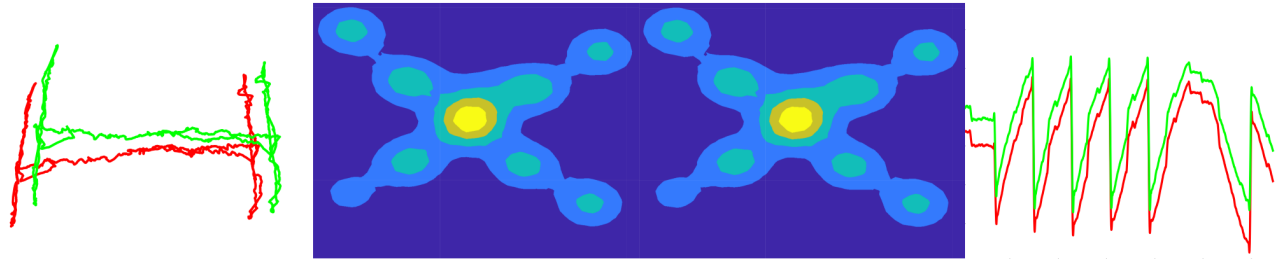
Model	Test	Accuracy	Sensitivity	Specificity
1D VGG	Dot	0.711	0.852	0.500
	H	0.738	0.833	0.611
	OKN	0.690	0.792	0.556
	Ensemble	0.643	0.958	0.222
1D ResNet	Dot	0.578	0.741	0.333
	H	0.714	0.792	0.611
	OKN	0.595	0.958	0.111
	Ensemble	0.643	0.958	0.222
LSTM	Dot	0.689	0.926	0.333
	H	0.643	0.833	0.389
	OKN	0.667	0.833	0.444
	Ensemble	0.595	1.000	0.056
DRCDT-NS	Dot	0.711	0.704	0.722
	H	0.744	0.721	0.778
	OKN	0.786	0.792	0.778
	Ensemble	0.733	0.851	0.555
DRCDT-NS + Deformation Modeling	Dot	0.822	0.814	0.833
	H	0.622	0.444	0.889
	OKN	0.666	0.481	0.944
	Ensemble	0.866	0.925	0.778

has traditionally been a major barrier to the adoption of eye-tracking technology in clinical diagnosis. Previous studies exploring the application of eye tracking in various neurological disorders have often had to exclude patients unable to complete the calibration process, either fully or partially [13], [17], [19]–[24]. Our approach addresses this limitation, broadening the potential for eye tracking in clinical settings.

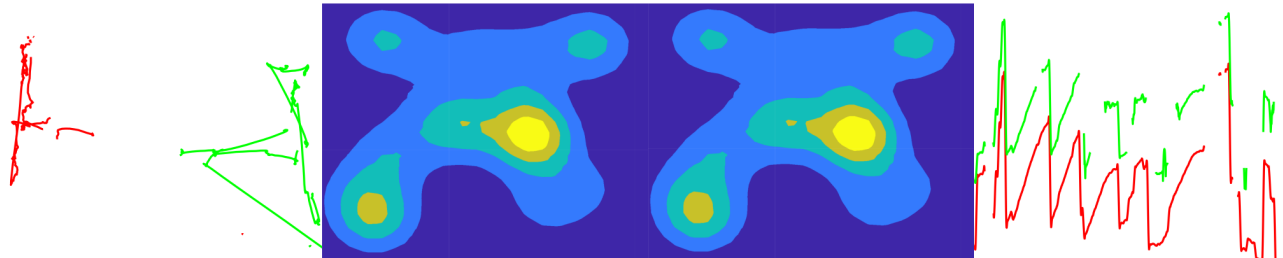
The proposed approach overcomes the calibration by measuring the conjugacy of the eye movement instead of measuring the point of gaze coordinates concerning the screen coordinates. The lack of calibration introduces deformation to the measurement translation, scaling, rotation, and shearing. These deformations pose challenges during discriminative analysis. The DRCDT accounts for these deformations by modeling them as a deformation vector derived from the probability distribution of gaze point gradients. Excluding the deformation vector as a learning parameter degrades the robustness of the classification model. Augmenting early detection of PCS using an non-calibrated eye tracker enables paramedics and other emergency medical providers to more accurately perform early stroke screening and improve rapid stroke triage and treatment.

Vascular neurologists in our study diagnosed patients with

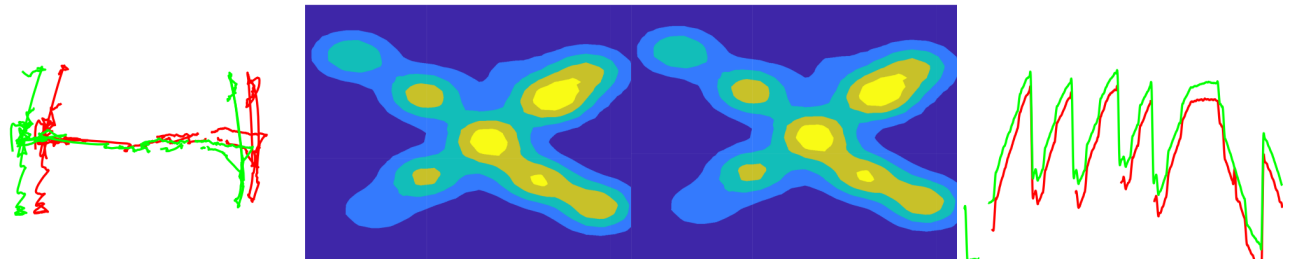
(a) True Negative: Patients with no abnormal eye movement



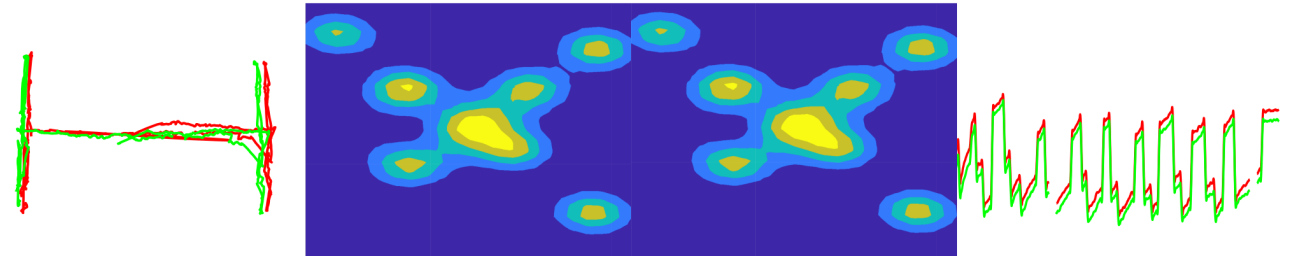
(b) True Positive: Patients with abnormal eye movement



(c) False Negative : Patients with abnormal eye movement



(d) False Positive: Control



(e) False Positive: Patients with no abnormal eye movement

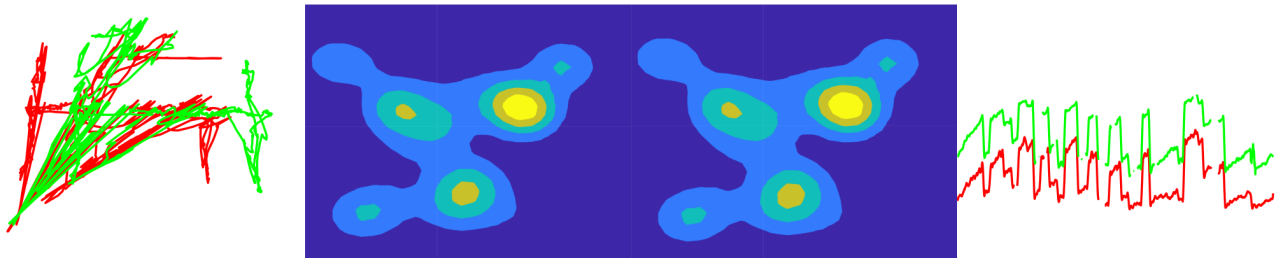







Fig. 2. Multifaceted Eye-Tracking Data Visualization for Classification Prediction Analysis. The columns of the figure represent each neurological eye examination and the rows depict instances of gaze patterns, integrating True Positives, True Negatives, False Positives, and False Negatives in the context of patient and control characteristics. The first column depicts the path plot from the 'H-test'. Here gaze data is transformed into screen coordinates measured in pixels (px). The central columns feature heatmaps for the 'Dot-test', with the left and right eyes' fixation densities portrayed through a color spectrum, where warmer colors denote higher fixation densities and cooler colors indicate less frequent fixations. The final column provides relative angle plots from the 'OKN test', sequentially plotted over time, to reveal the rhythmic eye movements characteristic of nystagmus.

a range of signs including impaired convergence, oculomotor palsies, skew deviation, abnormal patterns of nystagmus, impaired vestibulo-ocular reflex, saccadic pursuits, and acute hearing loss. Given the complexity of these variable presentations, with 48% of patients exhibited multiple symptoms with abnormal nystagmus being the most prevalent, an ensemble model that combines multiple tests is essential for diagnostic robustness. The neurological eye examination adopted in this study assesses both singular deficits and syndromes with multiple deficits. Relying solely on the results of a single test could be inadequate due to the need to fully characterize and localize a PCS syndrome. This is underscored by the low sensitivity of the H-test and OKN-tests alone, as detailed in Table I. Furthermore, the classification prediction analysis illustrated in Figure 2 shows that models based on a single test may yield false negatives when the manifested symptom does not align with the specific test used.

The computer-adapted tests provide a controlled, repeatable, and precise method to assess complex eye movements and coordination, which are critical in diagnosing and monitoring the nervous system. The integration of these tests into the study's methodology allows for detailed analysis of oculomotor function in a way that is not feasible with traditional bedside tests alone. The examinations included tests for visual fields and eye movements, such as the ability to maintain a steady gaze and detecting abnormal patterns of nystagmus. However, our methodology did not encompass specific tests for diplopia, or double vision, that often arises from strokes affecting the cranial nerves or brainstem. We incorporated the Head Orientation (HO) test, which required participants to alternately tilt their heads to the left and right. This test is known to exacerbate vertical misalignment due to an ocular tilt reaction. Nonetheless, the HO test was ultimately excluded from our methodology because it showed poor gaze correlation coefficients among both patients and healthy participants and exhibited higher variance compared to the other three tests used [18]. Moreover, our observations indicated that screen-based eye-tracking technology was inadequate for this test, as it failed to detect gaze accurately when the head's orientation was rotated along the y-axis.

Rigorous quality control measures during data acquisition required multiple test attempts, especially within the patient cohort, highlighting the challenges of maintaining consistency and reliability in a clinical setting. Despite these difficulties, strict adherence to the protocol was crucial for securing high-quality data. This underscores the efficacy of the integrated technological approach employed in our study.

Data efficiency is critical in solving the clinical problem as clinical data for PCS is challenging to acquire due to its limited incidence. Furthermore, acquiring a clinical dataset is more expensive than other domains' datasets. We evaluated the data efficiency of the DRCDDT-based classification method by performing a k-fold cross-validation by varying the training sample. The k-folds are randomly drawn from the original dataset. The experiment for each k-fold was repeated five times, and the results were averaged. Fig. 3 highlights the computational efficiency of the classification model as it maintains classification accuracy over the varying number of training

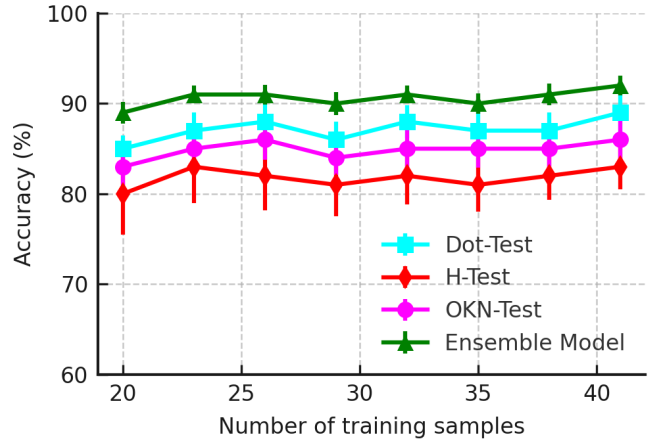


Fig. 3. Accuracy as a function of a number of training samples used to evaluate the classifier performance.

samples. This is feasible since the classifier utilizes a transport-based generative model to define the classification problem and uses the mathematical properties of the DRCDDT to render the problem more accessible in the transform domain. This approach allows one to solve nonlinear classification problems using linear classifiers.

We observed that 81% of participants in our study presented with dizziness, a non-specific symptom widely manifested in other conditions. Non-specific symptoms like dizziness, nausea, and blurred vision increase the likelihood of diagnostic errors in PCS [25]. Diagnostic error is exacerbated by the fast-paced, high-pressure environment of the emergency department, leading clinicians to rely on heuristics and rapid decision-making prone to cognitive biases [25]–[27]. Diagnostic error rates in PCS can reach up to 52%, with non-specific symptoms like dizziness and visual disturbances being common reasons for misdiagnosis [28], [29]. Furthermore, despite the introduction of more advanced imaging techniques, diagnostic errors still contribute to significant clinical and socioeconomic consequences, including higher rates of disability, increased mortality, and longer hospital stays [25], [29]. This study is proof of concept that a non-calibrated eye tracker using machine learning techniques could augment detection of PCS by non-neurology providers and emergency personnel, improve patient care and outcomes through early recognition, triage, and more effective stroke treatment.

IV. MATERIALS AND METHODS

A. Human Data Acquisition

The experimental protocol was approved by the University of Virginia's Institutional Review Board. As such, the protocol complies with all national ethical research standards and in accordance with the Declaration of Helsinki. Written informed consent was obtained prior to subject enrollment and testing of hospitalized patients and healthy participants.

Nineteen healthy participants were enrolled initially, with a mean age (range) of 40 (25–62) years, and 79% were female. The racial distribution of the healthy participants was 81%

TABLE III

PARTICIPANT DEMOGRAPHICS, INCLUDING THE NUMBER OF INDIVIDUALS SCREENED AND ENROLLED, MEAN AGE, GENDER, AND RACIAL DISTRIBUTION FOR HEALTHY PARTICIPANTS AND PATIENTS

Group	Healthy Participants	Patients
Screened	19	29
Enrolled	18	27
Mean Age (Range)	40 (25-62)	61.8 (19-96)
Female (%)	79	53
White (%)	81	82
Asian (%)	14	0
American Indian or Alaska Native (%)	5	0
Black (%)	0	12
Other (%)	0	6

White, 14% Asian, and 5% American Indian or Alaska Native. One healthy control was excluded since the participant failed to follow instructions during the experiment and moved out of the required testing position, which prevented eye gaze capture, resulting in 18 healthy participants included in the analysis.

Comparatively, twenty-nine patients were screened, and of those, 27 enrolled. The patients consisted of 53% females with a mean age (range) of 61.8 (19-96) years. The patients had a racial distribution of 82% White, 12% Black, and 6% Other. Two patients were excluded from the analysis (acute ischemic stroke $n=1$ and acute vestibular syndrome peripheral origin $n=1$) due to malfunction of the hardware/software of RoADIE, resulting in 27 patients included in the analysis (see Table. III).

Patient enrollment and data collection were completed within a median of 3 days, ranging from 0 to 37 days of symptom onset. Acute ischemic stroke accounted for 22 patients, with the rest having a diagnosis of Acute Vestibular Syndrome of suspected peripheral origin ($n=3$), Multiple Sclerosis ($n=1$), and Vestibular Neuritis ($n=1$). Diplopia, or double vision, was the primary presenting symptom in six patients, with nystagmus or vertigo being the presenting characteristic in seven patients. Head impulse testing showed that seven patients had normal responses, while four had an impaired vestibulo-ocular reflex. The head impulse test was not performed for six patients (see Table. IV).

None of the patients had acute hearing loss. Only four patients had impaired smooth pursuit.

TABLE IV

DISTRIBUTION OF PATIENT DIAGNOSES AND PRESENCE OF ABNORMAL EYE MOVEMENT FOR ENROLLED PARTICIPANTS USED IN THE ANALYSIS

Diagnosis	Neurological Symptoms	Population	Abnormal Eye Movement
Acute ischemic stroke	Central	22	Yes
Hemorrhagic stroke	Central	1	Yes
Multiple sclerosis	Central	1	Yes
Acute vestibular syndrome	Peripheral	3	No
Healthy	None	18	No

B. Rolling Apparatus to Detect Impairment of the Eyes

The study employed a custom-built mobile rig designed to facilitate the integrated acquisition of diverse sensory data within a clinical environment [18], [30]. The central component of this rig was the Tobii Pro Fusion Eye Tracker [31], which provided gaze estimates at a sampling frequency of 120 Hz. The setup also included a secondary screen for visual stimulation and a RealSense camera that captured RGB, infrared, depth, and motion data via an integrated accelerometer (see Fig. 1).

To accommodate different patient conditions, the mobile rig was designed for easy transportation to a patient's bedside, allowing for data collection while the patient was either reclined or seated. For consistency in baseline measurements, participants from the control cohort were seated during data acquisition.

The data acquisition protocol was defined to ensure the reliability and precision of the collected data. It required the eye tracker to detect at least one eye and capture at least 90% of gaze data during neurological eye examination tests. Furthermore, the protocol leveraged the accelerometer to monitor the rig's stability, with a directive to halt data collection upon detecting any significant motion. In post-processing, the series of gaze points for each exam, $\mathbf{x}_k \in \mathbb{R}^2, k = 1, \dots, N$ is smoothed using a Savitzky-Golay filter with a window size of 30 points to remove artifacts including those caused by blinks and minor involuntary movement.

C. Computer-adapted Neurological Eye Examinations

To quantitatively assess ocular motor function, three standard bedside ocular motor tests were adapted into computerized formats: the Dot Test, the H Test, and the Optokinetic Nystagmus (OKN) Test. These tests were designed to mimic the visual tracking and coordination tasks typically conducted by clinicians, with digital enhancements to measure and record eye movement accurately [18].

Dot Test: This test evaluates the quality of eye coordination in performing volitional saccades. Participants were instructed to shift their gaze from one dot to another as these appeared on the screen. This test replicates the clinical assessment of saccade accuracy, where a clinician observes if the eyes move sharply and stop precisely at the target. Key metrics for analysis include instances where the eyes undershoot or overshoot the target, or fail to initiate a saccade.

H Test: In this test, eye movements are evaluated as participants track a visual target moving in a pattern that resembles the letter "H." This pattern requires the eyes to follow the target through the four cardinal ocular directions and quadrants, simulating the movement of following a clinician's finger during a physical examination. The test particularly assesses the smoothness of pursuit and the ability of the eyes to initiate and maintain motion in all directions. Abnormal findings are characterized by a lack of motion, delayed initiation, or the use of compensatory small saccades to maintain target tracking.

OKN Test: Measures the participant's ability to switch from a smooth pursuit to a saccade in order to fixate on the next visual target after the first target disappears. The visual

stimulus is typically comprised of vertical bars with a high contrast to the background. The bars move at a quick and constant pace from right to left. This is done for each opposite direction. OKN typically remains preserved in individuals with occipital lobe infarcts, although impairment of a visual field may limit the amplitude of saccadic fixation. Asymmetry in the performance of the eyes or poor ability to generate saccades during the OKN test may suggest damage to the ocular, brainstem, or cerebellar nuclei or tracts.

D. Discrete Radon Cumulative Distribution Transform (DRCDT) for Eye Movement Classification

The Discrete Radon Cumulative Distribution Transform (DRCDT) [32] for pattern recognition and classification of eye movements using the nearest subspace (NS) classification approach. By analyzing the gaze points from both the left and right eyes, DRCDT maximizes the information on eye conjugacy while being invariant to the deformation caused by non-calibrated eye tracking data. This invariance allows for the generation of a convex set of gaze point distributions within the DRCDT space, which are then linearly separable and can be effectively classified using the NS model.

1) *Mathematical formulation of DRCDT*: Let Ω_2 refer to a set of coordinates in \mathbb{R}^2 and $s : \Omega_2 \rightarrow \mathbb{R}$ be a mapping from Ω_2 to a set of real numbers. The discrete point-set distribution P_s then can be defined as

$$P_s := \frac{1}{|\Omega_2|} \sum_{\mathbf{x} \in \Omega_2} \delta_{s(\mathbf{x})} \quad (1)$$

where $|\cdot|$ denotes the cardinality of a set. The DCDT transform of a distribution P_s is defined as

$$\mathcal{F}(P_s) = \mathcal{P}[s(\mathbf{x}_1), s(\mathbf{x}_2), \dots]^T = [\tilde{s}(\mathbf{x}_1), \tilde{s}(\mathbf{x}_2), \dots]^T \quad (2)$$

where \mathcal{P} is a permutation matrix such that $\tilde{s}(\mathbf{x}_1) \leq \tilde{s}(\mathbf{x}_2) \leq \dots$. The Radon transform of a distribution P_s is defined as

$$P_{s\theta} := \frac{1}{|\Omega_2|} \sum_{\mathbf{x} \in \Omega_2} \delta_{s(\mathbf{x} \cdot \mathbf{w}_\theta)} \quad (3)$$

where $\mathbf{w}_\theta = (\cos \theta, \sin \theta)^T$ is a unit vector in the direction of θ . The DRCDT transform of a distribution P_s is defined as

$$\begin{aligned} \hat{P}_s(\theta) &= \mathcal{F}(P_{s\theta}) = \mathcal{P}_\theta[s(\mathbf{x}_1 \cdot \mathbf{w}_\theta), s(\mathbf{x}_2 \cdot \mathbf{w}_\theta), \dots]^T \\ &= [\tilde{s}(\mathbf{x}_1 \cdot \mathbf{w}_\theta), \tilde{s}(\mathbf{x}_2 \cdot \mathbf{w}_\theta), \dots]^T \end{aligned} \quad (4)$$

where \mathcal{P}_θ is a permutation matrix such that $\tilde{s}(\mathbf{x}_1 \cdot \mathbf{w}_\theta) \leq \tilde{s}(\mathbf{x}_2 \cdot \mathbf{w}_\theta) \leq \dots$.

2) *Classification Process*: The classification of unknown test samples of gaze distributions uses the DRCDT space. During the training phase, the DRCDT transform of each training sample is calculated and used to approximate the subspace for each class: $\{P_{s_1}^{(1)}, P_{s_2}^{(1)}, \dots\}$ (class 1), $\{P_{s_1}^{(2)}, P_{s_2}^{(2)}, \dots\}$ (class 2) and the subspace can be approximated as,

$$\mathbb{V}^{(k)} = \text{span} \left(\left\{ \hat{P}_{s_1}^{(k)}, \hat{P}_{s_2}^{(k)}, \dots \right\} \cup \mathbb{U}_T \right) \quad (5)$$

where $\mathbb{U}_T = \{\mu_1(n, \theta), \mu_2(n, \theta)\}$ with $\mu_1(n, \theta) = \cos \theta$, $\mu_2(n, \theta) = \sin \theta$ is the spanning set corresponding to the deformation modeling.

In the testing phase, the class of an unknown test distribution P_s is obtained as

$$k^* = \arg \min_k d(P_s, \mathbb{V}^{(k)}) \quad (6)$$

where $d(\cdot, \cdot)$ denotes the distance of a test sample from a trained subspace in the DRCDT transform space, and k refers to the class index.

E. Evaluation

We evaluated our model's ability to detect abnormal eye movement versus health controls and patient sub-cohort that present with normal eye movement using the following metrics: Accuracy, Sensitivity, and Specificity. We did this for each Neurological eye examination test (Dot, H, and OKN) as well as the overall classification of abnormal eye movement versus healthy control using all the information from these tests. This approach allowed us to determine how much value each test has in the aim of discriminating between abnormal eye movement versus healthy controls. Additionally, we studied the classification performance with several time series neural network methods: 1D Visual Geometry Group (1D-VGG) [33], 1D Residual Network (1D-ResNet) [33], [34], Long Short Term Memory (1D-LSTM) [33].

V. ACKNOWLEDGEMENT

This work was supported by the AHA/ASA under Grant 19IPLOI34760692

REFERENCES

- [1] Steven Go, "Posterior circulation ischemic stroke," *Missouri medicine*, vol. 112, no. 3, pp. 192, 2015.
- [2] "Types of stroke," <https://www.stroke.org/en/about-stroke/types-of-stroke>, Jan 2021.
- [3] Silvia Olivato, Silvia Nizzoli, Milena Cavazzuti, Federica Casoni, Paolo Frigio Nichelli, and Andrea Zini, "e-nihss: an expanded national institutes of health stroke scale weighted for anterior and posterior circulation strokes," *Journal of Stroke and Cerebrovascular Diseases*, vol. 25, no. 12, pp. 2953–2957, 2016.
- [4] S Kmetonyova, J Paulasova Schwabova, T Sramkova, M Dankova, A Olserova, M Petrzalka, A Tomek, and J Jerabek, "Posterior circulation stroke diagnosis in unselected group of acutely dizzy patients," *Clinical Neurology and Neurosurgery*, vol. 224, pp. 107541, 2023.
- [5] Yan Zhuang, Mohamed Hassan, Chad Aldridge, Xuwang Yin, Timothy McMurry, Andrew Southerland, and Gustavo Rohde, "A pilot study on camera-based neurological deficit detection," in *ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE '20), December 16–18, 2020, Arlington, VA, USA*, 2020, pp. 16–17, ACM / IEEE.
- [6] Alexander Salerno, Davide Strambo, Stefania Nannoni, Vincent Dunet, and Patrik Michel, "Patterns of ischemic posterior circulation strokes: A clinical, anatomical, and radiological review," *International Journal of Stroke*, vol. 17, no. 7, pp. 714–722, 2022.
- [7] Jorge C Kattah, Arun V Talkad, David Z Wang, Yu-Hsiang Hsieh, and David E Newman-Toker, "Hints to diagnose stroke in the acute vestibular syndrome: three-step bedside oculomotor examination more sensitive than early mri diffusion-weighted imaging," *Stroke*, vol. 40, no. 11, pp. 3504–3510, 2009.
- [8] John H Pula and Carlen A Yuen, "Eyes and stroke: the visual aspects of cerebrovascular disease," *Stroke and vascular neurology*, vol. 2, no. 4, pp. 210–220, 2017.
- [9] Kiersten L Gurley and Jonathan A Edlow, "Avoiding misdiagnosis in patients with posterior circulation ischemia: A narrative review," *Academic Emergency Medicine*, vol. 26, no. 11, pp. 1273–1284, 2019.

- [10] Millie Nakatsuka and Emma E Molloy, "The hints examination and standing algorithm in acute vestibular syndrome: A systematic review and meta-analysis involving frontline point-of-care emergency physicians," *Plos one*, vol. 17, no. 5, pp. e0266252, 2022.
- [11] Kailash Krishnan, Kerolos Bassilious, Erik Eriksen, Philip M Bath, Nikola Sprigg, Sigrun Kierulf Brækken, Hege Ihle-Hansen, Morten Andreas Horn, and Else Charlotte Sandset, "Posterior circulation stroke diagnosis using hints in patients presenting with acute vestibular syndrome: a systematic review," *European Stroke Journal*, vol. 4, no. 3, pp. 233–239, 2019.
- [12] Yasuo Terao, Hideki Fukuda, and Okihide Hikosaka, "What do eye movements tell us about patients with neurological disorders?—an introduction to saccade recording in the clinical setting—," *Proceedings of the Japan Academy, Series B*, vol. 93, no. 10, pp. 772–801, 2017.
- [13] Deepesh Kumar, Anirban Dutta, Abhijit Das, and Uttama Lahiri, "Smart-eye: developing a novel eye tracking system for quantitative assessment of oculomotor abnormalities," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 24, no. 10, pp. 1051–1059, 2016.
- [14] Akane Oyama, Shuko Takeda, Yuki Ito, Tsuneo Nakajima, Yoichi Takami, Yasushi Takeya, Koichi Yamamoto, Ken Sugimoto, Hideo Shimizu, Munehisa Shimamura, et al., "Novel method for rapid assessment of cognitive impairment using high-performance eye-tracking technology," *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.
- [15] Daniel S Asfaw, Pete R Jones, Laura A Edwards, Nicholas D Smith, and David P Crabb, "Using eye movements to detect visual field loss: a pragmatic assessment using simulated scotoma," *Scientific reports*, vol. 10, no. 1, pp. 1–13, 2020.
- [16] William Rosengren, Marcus Nyström, Björn Hammar, and Martin Stridh, "A robust method for calibration of eye tracking data recorded during nystagmus," *Behavior research methods*, pp. 1–15, 2019.
- [17] Alessandro Grillini, Daniel Ombelet, Rijul S Soans, and Frans W Cornelissen, "Towards using the spatio-temporal properties of eye movements to classify visual field defects," in *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, 2018, pp. 1–5.
- [18] Mohamed Abul Hassan, Chad M Aldridge, Yan Zhuang, Xuwang Yin, Timothy McMurphy, Gustavo K Rohde, and Andrew M Southerland, "Approach to quantify eye movements to augment stroke diagnosis with a non-calibrated eye-tracker," *IEEE Transactions on Biomedical Engineering*, 2022.
- [19] William Rosengren, Marcus Nystöm, Björn Hammar, and Martin Stridh, "Suitability of calibration polynomials for eye-tracking data with simulated fixation inaccuracies," in *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, 2018, pp. 1–5.
- [20] Zhe Zeng, Felix Wilhelm Siebert, Antje Christine Venjakob, and Matthias Roetting, "Calibration-free gaze interfaces based on linear smooth pursuit," *Journal of Eye Movement Research*, vol. 13, no. 1, pp. 3, 2020.
- [21] Angela Jinsook Oh, Tiffany Chen, Mohammad Ali Shariati, Naz Jehangir, Thomas N Hwang, and Yaping Joyce Liao, "A simple saccadic reading test to assess ocular motor function in cerebellar ataxia," *Plos one*, vol. 13, no. 11, pp. e0203924, 2018.
- [22] JA Nij Bijvank, LJ Van Rijn, LJ Balk, HS Tan, BMJ Uitdehaag, and A Petzold, "Diagnosing and quantifying a common deficit in multiple sclerosis: Internuclear ophthalmoplegia," *Neurology*, vol. 92, no. 20, pp. e2299–e2308, 2019.
- [23] Jenny A Nij Bijvank, Axel Petzold, Danko Coric, H Stevie Tan, Bernard MJ Uitdehaag, Lisanne J Balk, and Laurentius J van Rijn, "Quantification of visual fixation in multiple sclerosis," *Investigative ophthalmology & visual science*, vol. 60, no. 5, pp. 1372–1383, 2019.
- [24] JA Nij Bijvank, A Petzold, LJ Balk, HS Tan, BMJ Uitdehaag, M Theodorou, and LJ Van Rijn, "A standardized protocol for quantification of saccadic eye movements: Demons," *Plos one*, vol. 13, no. 7, pp. e0200695, 2018.
- [25] Carolin Hoyer and Kristina Szabo, "Pitfalls in the diagnosis of posterior circulation stroke in the emergency setting," *Frontiers in Neurology*, vol. 12, pp. 682827, 2021.
- [26] David E Newman-Toker, Alexander C Schaffer, Christina W Yu-Moe, Negin Nassery, Ali S Saber Tehrani, Gregory D Clemens, Zhen Wang, Yunfeng Zhu, and Daniel Siegal, "Serious misdiagnosis-related harms in malpractice claims: The 'big three' - vascular events, infections, and cancers," *Diagnosis*, vol. 6, no. 3, pp. 227–240, 2019.
- [27] Mark L Graber, Nancy Franklin, and Ruth Gordon, "Diagnostic error in internal medicine," *Archives of internal medicine*, vol. 165, no. 13, pp. 1493–1499, 2005.
- [28] Ali S Saber Tehrani, Huan Lee, Stephen C Mathews, Adam Shore, Martin A Makary, Peter J Pronovost, and David E Newman-Toker, "25-year summary of us malpractice claims for diagnostic errors 1986–2010: an analysis from the national practitioner data bank," *BMJ Quality & Safety*, vol. 22, no. 8, pp. 672–680, 2013.
- [29] Tracy E Madsen, Jane Khoury, Roberto Cadena, Opeolu M Adeoye, Kathleen A Alwell, Charles J Moomaw, and Dawn O Kleindorfer, "Potentially missed diagnosis of ischemic stroke in the emergency department in the greater cincinnati/northern kentucky stroke study," *Academic Emergency Medicine*, vol. 23, no. 10, pp. 1128–1135, 2016.
- [30] Mohamed Abul Hassan, Xuwang Yin, Yan Zhuang, Chad M Aldridge, Timothy McMurphy, Andrew M Southerland, and Gustavo K Rohde, "A digital camera-based eye movement assessment method for neuroeye examination," *IEEE Journal of Biomedical and Health Informatics*, 2023.
- [31] Tobii AB, "Tobii pro fusion," Tech. Rep., Tobii Pro Inc., 2019.
- [32] Yan Zhuang, Shiyang Li, Xuwang Yin, Abu Hasnat Mohammad Rubaiyat, Gustavo K Rohde, et al., "Local sliced-wasserstein feature sets for illumination-invariant face recognition," *arXiv preprint arXiv:2202.10642*, 2022.
- [33] Brian Kenji Iwana and Seiichi Uchida, "An empirical survey of data augmentation for time series classification with neural networks," *Plos one*, vol. 16, no. 7, pp. e0254841, 2021.
- [34] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller, "Data augmentation using synthetic data for time series classification with deep residual networks," *arXiv preprint arXiv:1808.02455*, 2018.