AI-Assisted ADHD Diagnosis Through Eye Tracking Technology

Irmak Erkol

Musa Simsek

Sabanci University

CS 549 Human-Computer Interaction

Kursat Cagiltay

1/10/2025

Abstract

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by patterns of inattention, hyperactivity, and impulsivity. Diagnosing ADHD is challenging due to symptom overlap, subjective evaluation methods, and variability in clinical presentations. This study explores the potential of using eye-tracking technology combined with machine learning models to accurately diagnose ADHD. Eye movement metrics, such as saccadic movements, fixation stability, and intrusive saccades, were collected from 8 ADHD patients and 9 healthy participants. Various machine learning classifiers, including Support Vector Machines (SVMs), Logistic Regression, and Random Forest Classifiers, were trained on features captured from the eye-tracking data. SVM with an 'rbf' kernel consistently achieved the highest performance, with a mean F1-score of 0.77 and a standard deviation of 0.04. The findings suggest that eye-tracking data can serve as a reliable basis for ADHD diagnosis. However, the limited sample size indicates the need for larger datasets to improve model performance. This study highlights the promise of integrating AI-driven tools into ADHD diagnostic frameworks to address current limitations and improve diagnostic accuracy.

Keywords: human-computer interaction, machine learning, ADHD, eye tracking, AI in healthcare

AI-Assisted ADHD Diagnosis Through Eye Tracking Technology

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder characterized by persistent patterns of inattention, distractibility, hyperactivity, and impulsivity, which interfere with functioning or development (American Psychiatric Association [APA], 2022). Diagnosis of ADHD is challenging mainly because symptoms overlap with other conditions, variability in clinical presentations, and reliance on subjective evaluation methods such as interviews and behavioral checklists (Kaushal et al., 2024; Grandjean et al., 2025). Current diagnostic frameworks, such as those outlined in the DSM-5, provide criteria but fail to address these limitations thoroughly (Grandjean et al., 2025). As a result, misdiagnosis and overdiagnosis are common, leading to inappropriate treatments (Kaushal et al., 2024). This is particularly concerning given the significant side effects of ADHD medications. To improve diagnostic accuracy and patient outcomes, a data-driven diagnostic approach is needed.

(Munoz et al., 2003) suggested that oculomotor behaviors such as fixation and saccadic movements differ between individuals who have ADHD and individuals who don't have it. Since eye-tracking devices help us to analyze these oculomotor behaviors, we can further investigate them to address the limitations of current diagnostic frameworks. Yoo et al. (2024) collected eye-tracking data from children with ADHD and the control group during the five behavioral tasks measuring selective attention, working memory, and response inhibition and identified 33 eye-tracking features with the potential to distinguish children with ADHD from the control group and the ML model deployed improved the accuracy. Deng et al. (2023) developed a Deep Learning(DL) model pre-trained on a related task and found out that the model outperformed the relevant baselines for the prediction of ADHD. Leet al. (2023) found that Integrating eye-tracking with CPTs resulted in more accurate ADHD diagnosis than CPT alone diagnosis.

In this study, we try to answer the question can a machine learning model accurately diagnose ADHD using only eye-tracking data?

Method

Participants

The study was conducted on 12 ADHD patients and 15 non-ADHD (healthy) control group participants. However, we were not able to process data from some of the participants due to the following reasons: frame rate dropped during recording in the Unity application used for the dot test, resulting in inaccurate outputs. Additionally, some participants' inability to remain still caused the eye-tracking device to lose calibration during the experiment, even in some trials the device completely failed to track the participant's eyes. Furthermore, one participant in the healthy control group had taken medications unrelated to the study, which could have affected the results. As a result of these challenges, the final dataset included data from 8 ADHD patients and 9 healthy participants.

The ADHD group consists of people with ADHD diagnosed by a psychiatrist. They were informed and followed the information and did not take any medication before the experiment. Both the ADHD and healthy control groups were free from any other diagnosed conditions that could potentially influence the study's outcomes. Participants are adults between the age of 18-40, who were either currently enrolled in undergraduate programs, university graduates, or postgraduates. All the participants use a computer at least 2 hours a day. Participation in the study was entirely voluntary, and participants were informed that they could withdraw from the experiment at any time without providing a reason or facing any consequences.

Materials

Stimuli were presented on an MSI Pulse 17 AI C1VGKG laptop with a 17-inch display. Screen resolution was set to 1920 × 1080 pixels, and an eye-to-screen viewing distance was set to approximately 60 cm with head stability maintained throughout the experiment. The Tobii Pro x2-60 eye tracker was used to capture gaze point x, gaze point y, eye movement type, eye movement event duration, fixation point x, and fixation point y, and the screen was also recorded using Tobii Pro Lab. Eye movements were recorded at a frame rate of 30 Hz and 5 ms. Calibration for each participant in Tobii Pro Lab used a nine-point procedure. The experimental environment was designed using Unity.

Procedure

The procedure involves three steps: Participant Setup, Fixation Task, Prosaccade, and Antisaccade Task. During all these tasks, eye-tracking metrics were collected.

Participant Setup: Participants were instructed to sit approximately 60 cm away from the eye screen and this was ensured with the use of Tobii Pro Lab calibration. Then the device was calibrated with the help of the participants' eye movement. Participants were informed to hold their heads and distance still.

Fixation Task: A cross appeared on the center of the screen and participants were instructed to focus on the cross while it was on the screen. The exact amount of time of cross appearance was not informed so that sustained attention and fixation stability could be assessed.

Prosaccade and Antisaccade Tasks: Each trial began with a central cross appearing on the screen for 1500–2500 ms, followed by a brief 200 ms blank interval. A fixation point appeared in green (for prosaccade) or red (for antisaccade), on either the left or right side of the screen. In prosaccade trials, participants were instructed to focus on the green light whereas in the antisaccade trial, participants were instructed to focus on the opposite side of the red light. The light stimulus appeared on random sides, with random colors, and for a random amount of time. A total of 28 trials (14 prosaccade and 14 antisaccade) were presented in random order to reduce predictability and enhance data validity. The full experimental session lasted approximately 6 minutes, including calibration and task completion.

Data Analysis

Based on the eye tracking data of the participants; the mean Simple Reaction Time(SRT), the coefficient of variation of SRT, the percentage of direction errors, and the number of intrusive saccades were computed. SRT was measured as the time from stimulus appearance to the onset of the first saccade in saccade tasks (Munoz et al., 2003). For each saccade, the eye movements in the first 90 milliseconds of the target appearance were eliminated because those movements are called anticipatory movements and may occur independently of the target (Munoz et al., 2003). Using SRT values of saccade trials, we calculated the mean SRT and the coefficient of variation of SRT for participants.

Saccades were scored as correct if the first movement after the appearance of the stimulus was greater than a specified angle, such as 1°, 2°, etc, in amplitude and in the correct direction (towards the stimulus if the light was green, opposite if the light was red) (Munoz et al., 2003). Saccades that did not satisfy these criteria were labeled as incorrect saccades and, consequently, direction errors. Direction errors were also measured only in saccade tasks.

Intrusive saccades are identified as rapid shifts in the eye position that exceed a specified amount of degree such as 2° in amplitude (Munoz et al., 2003). In the analysis part, the number of intrusive saccades that appeared in the fixation task was counted.

Training

For different values of angle threshold for correct saccades and intrusive saccades, the features listed in the Data Analysis part were extracted from the eye tracking data, along with the logs of the application to mark the timestamps of tasks and corresponding responses of the participants. The values used for the angle threshold of correct saccades were 0.5, 1.0, 1.5, 3.0, and 5.0; and the values used for the angle threshold of intrusive saccades were 1.5, 2, 2.5. (Munoz et al., 2003) reported that to distinguish the control group from the ADHD group, the threshold of 5° for saccade correctness and the threshold of 2° for intrusive saccades were sufficient.

Several Machine Learning classifiers were used with the features extracted based on every pair of angle thresholds for saccade correctness and intrusive saccades: Support Vector Machines(SVM), Logistic Regression, and Random Forest Classification. For SVM, different kernels such as ‘linear’, ‘rbf’, and ‘polynomial’ were also tested. Since Logistic Regression uses a built-in ‘sigmoid’ function, SVM was not tested with a sigmoid kernel. Random Forest Classifiers were also implemented with 20 and 100 decision trees. 5-fold Cross Validation was run 100 times with a shuffled version of the data for every model configuration, and the average of F1-macro scores were collected.

Results

Table 1 shows the best model performances for different thresholds of intrusive saccades. The best-performing model among the ones used in this study was SVMs. More specifically, SVM with ‘rbf’ kernel was the overall best performer. Random Forest implementations were not in the top 5 for any configuration, and the models that use a ‘sigmoid’ function such as Logistic Regression were the worst.

Table 1

Best-performing Machine Learning Models for Intrusive Saccades Detection with Varying Saccade Angle Thresholds and Corresponding F1-scores with Standard Deviations.

Intrusive Threshold	ML Model	Saccade Threshold	Mean F1-score	Std(F1-score)
1.5°	SVM with ‘rbf’ kernel	1.0°	0.77	0.04
2.0°	SVM with ‘rbf’ kernel	1.0°	0.77	0.04
2.5°	SVM with ‘rbf’ kernel	1.0°	0.77	0.03

Note. Std(F1-score) stands for the standard deviation of the F1-score

Outcome 1

As seen from Table 1, the machine learning model SVM with ‘rbf’ kernel can accurately diagnose ADHD using only eye-tracking data.

Outcome 2

With varying intrusive threshold values, the F1-score and saccade threshold for best results did not change.

Outcome 3

The average F1-score and the standard deviation of it were constant through different configurations.

Discussion

We argued that SVM with ‘rbf’ kernel was sufficient for the diagnosis of ADHD with a specific configuration. However, Outcome 3 suggests that the dataset might be too small, which could be an indicator of the model’s underfitting. Therefore, the study must be revised with a more accurate and larger dataset before conclusion.

Outcome 2 suggests that the task in which we measured the intrusive threshold was either too easy or too difficult, or we calculated it wrong. A future work could be dedicated to finding an efficient way of distinguishing eye movements of an ADHD patient from a healthy person which relies on intrusive saccades, too.

References

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Publishing.

Deng, S., Prasse, P., Reich, D. R., Dziemian, S., Stegenwallner-Schütz, M., Krakowczyk, D., Makowski, S., Langer, N., Scheffer, T., & Jäger, L. A. (2023). Detection of ADHD Based on Eye Movements During Natural Viewing. *Machine Learning and Knowledge Discovery in Databases*, 403–418.

Lee, D. Y., Shin, Y., Park, R. W., Cho, S.-M., Han, S., Yoon, C., Choo, J., Shim, J. M., Kim, K., Jeon, S.-W., & Kim, S.-J. (2023, September 2). Use of eye tracking to improve the identification of attention-deficit/hyperactivity disorder in children. Nature News. https://www.nature.com/articles/s41598-023-41654-9

Grandjean, M., Hochman, S., & Mukherjee, R. (2023). Malingering in ADHD behavioral rating scales: Recommendations for research contexts. *Frontiers in Psychiatry*. https://www.frontiersin.org/articles/10.3389/fpsyt.2025.1532807/full

Kaushal, A., Kamdar, Y., Nagori, S., Bhandari, A., & Vhatkar, S. (2024). ADHD with depression: Comparing the effectiveness of traditional and computer-based diagnostic methods. *ResearchGate*. https://www.researchgate.net/publication/387067279

Munoz, D. P., Armstrong, I. T., Hampton, K. A., & Moore, K. D. (2003). Altered Control of Visual Fixation and Saccadic Eye Movements in Attention-Deficit Hyperactivity Disorder. *Journal of Neurophysiology*, *90*(1), 503-514. https://journals.physiology.org/doi/full/10.1152/jn.00192.2003

Yoo, J. H., Kang, C., Lim, J. S., Wang, B., Choi, C.-H., Hwang, H., Han, D. H., Kim, H., Cheon, H., & Kim, J.-W. (2024). Development of an innovative approach using portable eye tracking to assist ADHD screening: a machine learning study. *Frontiers in Psychiatry*, *15*, 1337595