Predicting ADHD in Children Using Eye-Tracking Data and Mathematical Modeling

Christos Papanikolaou

Thesis submitted for the degree of
Master in Applied Computer and Information Technology - ACIT
(Mathematical Modelling and Scientific Computing)
30 credits

Department of Computer Science
Faculty of Technology, Art and Design

Oslo Metropolitan University — OsloMet

Spring 2023

Predicting ADHD in Children Using Eye-Tracking Data and Mathematical Modeling

Christos Papanikolaou

Predicting ADHD in Children Using Eye-Tracking Data and Mathematical Modeling

http://www.oslomet.no/

Printed: Oslo Metropolitan University — OsloMet

Abstract

Attention-Deficit/Hyperactivity Disorder (ADHD) is a prevalent neurobiological condition with far-reaching implications for a child's academic achievement, well-being, and social interactions. Despite its impact, current diagnostic methods, largely reliant on behavioral observations and subjective symptom reporting, may result in over-diagnosis or underdiagnosis. This thesis explores the potential of eye-tracking metrics and mathematical modeling as more objective and precise tools for ADHD diagnosis.

We utilized a dataset from a visuospatial memory test involving children both with and without ADHD. The methodology encompasses a two-fold approach. Initially, we establish a benchmark classification model using non-eye-tracking metrics. Subsequently, we incorporate eye-tracking metrics derived from two mathematical models; the "Feed and Fly" and Lévy Flight models, contributing to our understanding of ADHD's complex nature.

Our findings reveal that while the "Feed and Fly" model did not yield significant differentiation between ADHD and non-ADHD groups, the Lévy Flight model, especially when applied in an event-based analysis, showed potential in enhancing the benchmark model's performance. The final model refined to only three key metrics, each representing different analytical domains - WISC tests, eye-tracking data, and memory tests - emphasizing the necessity of a multi-faceted approach in diagnosing ADHD.

Despite the limitations of the study, including questionable quality of the eye-tracking data and a small dataset size, our research highlights the promising potential of eye-tracking data and mathematical modelling in ADHD diagnosis, laying the foundation for future research in this field. We anticipate that this exploration will inspire further studies into more objective diagnostic tools, not only for ADHD but also for a wider range of cognitive disorders.

Acknowledgements

I would like to express my sincere thanks to my two supervisors, Professor P. Lind and Doctoral Researcher P. Lencastre, for their invaluable support and guidance throughout my research. Their feedback and advice have been instrumental in shaping my thesis.

I would also like to extend my gratitude to my family for their continuous love and encouragement during my academic pursuits. Without their unwavering support, I would not have been able to achieve this accomplishment.

Finally, I would like to acknowledge the contributions of my friends and colleagues who have supported me throughout my research journey.

Thank you to everyone who has contributed to my success.

Contents

Chapter 1

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is the most common neurobiological disorder in children and adolescents. Being one of the main causes of consultation in both Mental Health Units and schools [1], and can profoundly affect children's academic achievement, well-being, and social interactions [2].

It was first introduced in the American Psychiatric Association's (APA) "Diagnostic and Statistical Manual of Mental Disorders" (DSM) in its third edition, released in 1980, as attention deficit disorder (ADD) with two variations: the ADD with hyperactivity and the ADD without hyperactivity. The APA renamed it as attention deficit hyperactivity disorder (ADHD) in 1987, encompassing inattentiveness, impulsivity, and hyperactivity. In 1994, the APA's DSM 4^{th} edition listed three types of ADHD: mostly inattentive; mostly hyperactive and impulsive; and a combined type. In the most recent edition, published in 2013, the APA designated the three types as the three "presentations" of ADHD, which means the ways the disorder may affect people [3].

According to the Centers for Disease Control and Prevention (CDC), the estimated number of children aged 3–17 years ever diagnosed with ADHD in the US is 6 million using data from 2016-2019 [4], while in a meta-analysis of 175 research studies worldwide on ADHD prevalence in children aged 18 and under, was found an overall pooled estimate of 7.2% of the global children population [5].

Data from the National Survey of Children's Health (NSCH) and the National Health Interview Survey (NHIS) reveal a higher ADHD prevalence among older age groups, with boys having approximately double the ADHD diagnoses compared with girls [6]. However, girls with ADHD displayed greater intellectual impairment, lower levels of hyperactivity, and lower rates of other externalizing behaviors [7], which resulted in the fact that many girls are not diagnosed until middle school or later when learning tasks become more complex [8].

As already mentioned, it is characterized by persistent inattention or hyperactivity–impulsivity, or a combination of the two, that interfere with functioning [9]. No gene or structural abnormality of the brain accounts for the diversity of the ADHD spectrum; instead, it is believed to result

from the complex interaction of genetic, biological, and environmental risk factors [8]. While Swanson J. et al. [10], in their study evaluated one of the prevailing theories, the dopamine deficit theory, and identify critical issues about etiologic subtypes of ADHD that may involve dopamine.

According to Frazier T. et al. [11] study, their meta-analytic results indicated a moderate to the large discrepancy in academic achievement between individuals with ADHD and typical controls. This outcome substantiates the significant impact of ADHD symptoms on academic performance and reveals a pattern of impairments beyond the achievement test decrements identified in previous studies.

That fact makes it even more crucial to have an accurate diagnosis to ensure appropriate help for students in need and to remove the risk of misdiagnoses for non-disabled students. However, as is true with most developmental disorders, there is no definite test for ADHD, and both diagnosis and classification rely on the description. This lack of definite diagnostic tools has led to significant variations in ADHD prevalence, which has been reported from 2.2% to 17.8% [12].

Currently, diagnosis is being performed solely based on observed behavior and reported symptoms, as well as other subjective measures, such as clinical interview and rating scales which carries the potential risk of over-diagnosis or under-diagnosis, given the heterogeneity of ADHD symptoms [13].

Consequentially, there are ongoing efforts to identify stratification markers that may one day aid in early diagnosis or help predict prognosis and treatment response in ADHD. Oculomotor studies have been utilized to investigate aspects of both motor and cognitive control in ADHD, as well as the potential value of oculomotor markers in predicting treatment response.

Understanding the mechanisms between oculomotor control, neurocognitive traits, and broader ADHD symptoms requires a comprehensive understanding of the spectrum of oculomotor characteristics in ADHD. Wainstein et. al. [14], conducted a visuospatial experiment, based on a memory test, on children (subjects) with and without ADHD to determine if pupil size can be used as a biological marker that could help objectively define the disorder and provide information about its pathophysiology.

In this thesis, we will utilize their dataset and base our analysis on it, extending their work by examining the subject's eye movement during memory tests from a different perspective to gain additional insights from these data. Our motivation for this thesis is to explore the potential of eye movement characteristics as a more objective and accurate diagnostic tool for ADHD, and to highlight the importance of a comprehensive, multi-domain approach to ADHD diagnosis. By investigating the relationship between oculomotor control and ADHD symptoms, we aim to contribute to the ongoing efforts in identifying stratification markers that could aid in early diagnosis, prognosis, and treatment response prediction.

We will employ various eye-tracking techniques and algorithms, including the "Feed and Fly" and Lévy Flight models, to analyze the eye movement data, focusing on specific eye movement metrics like fixation

durations, saccade frequency, and others to differentiate between ADHD and non-ADHD subjects. The mathematical modeling of visual scanpaths plays a crucial role in understanding these differences and is an integral part of our analysis.

We will formulate two assumptions throughout this study. The first would be that the eye movements can be described as an intermittent process from which the characterization of eye-movements as two alternating processes (saccades and fixations) is a natural consequence, while the second would be that the visual scanpath follows the pattern of a Lévy flight. Both assumptions will be examined and their derived metrics will be evaluated upon their capability to differentiate between the ADHD and non-ADHD group.

Several challenges in analyzing eye-tracking data arise due to the inherent complexity of human eye movements and the need for accurate detection and classification of events. Our methodology aims to address these challenges and enhance the analysis of eye movement data for ADHD research.

Our initial hypothesis is that there could be a significant difference in eye movement characteristics, such as fixation duration and saccade frequency, between children with and without ADHD. This assumption is based on prior research indicating that youths with ADHD perform abnormally in these tasks. However, our exploration is open-ended, acknowledging the possibility that these differences might not be universally significant, and the importance lies in understanding the potential of eye-tracking data.

In the upcoming chapters, we will delve deeper into the methodology and results of our analysis.

Chapter 2, "Background and State of the Art," will provide a comprehensive review of the background and state-of-the-art techniques in vision systems, eye movements, eye tracking, and eye movement analysis. We will also explore the mathematical modeling of visual scanpaths, different types of eye movements, and various event detection methods. The mathematical background of different event-classification algorithms, such as velocity-based, dispersion-based, and area-based algorithms, will be discussed. Additionally, we will delve into eye tracking techniques, the latest technologies and systems, and the recent advancements in detecting ADHD through eye-tracking data analysis.

Chapter 3, "Methodology," will focus on our methodology, including data collection and dataset description, methodology procedure, and data analysis methods. We will also discuss the benchmark model and parameter selection process, as well as the assumptions and models used in our eye-tracking data analysis.

Chapter 4, "Results," will present our results, including the performance of our benchmark model and the eye-tracking metrics generated using different models. We will discuss the incorporation of these metrics into the benchmark model and provide an event-based analysis.

Finally, in Chapter 5, "Conclusion," we will summarize our study, discuss the results and their implications, and suggest future research

directions.

While our study hypothesizes certain outcomes, it's important to note that our findings might yield a mix of results. These outcomes, whether they confirm or challenge our initial hypothesis, will provide invaluable insights into the potential use of eye-tracking data in diagnosing ADHD.

Our research aims to advance our knowledge about ADHD and foster more objective, precise diagnostic tools. Enhanced diagnostic accuracy can lead to better outcomes for children with ADHD, improving their academic achievement, social interactions, and overall quality of life.

Chapter 2

Background and State of the Art

2.1 Vision System and Fundamentals of Eye Movement

Vision is the prominent sense among the five primary senses that give people a rich perceptual experience of the outside world. The primary sensory system involved in vision is the eye. Through the pupil, light waves enter the eye after traveling through the cornea, Figure 2.1. The eye's clear exterior is called the cornea. It helps focus light waves that enter the eye and acts as a boundary between the inner eye and the outside world. The pupil is the tiny opening in the eye through which light travels, and both light intensity and emotional arousal can affect the pupil's size.

The pupil will dilate, or widen, to allow more light to enter the eye when light levels are low. When there is much light, the pupil will contract, or shrink, to allow less light to enter the eye, [15]. The muscles associated with the iris, the colorful part of the eye, regulate the size of the pupil.

Light travels via the lens, a curved, transparent device that acts as an extra focus after passing through the pupil. In order to help focus the light

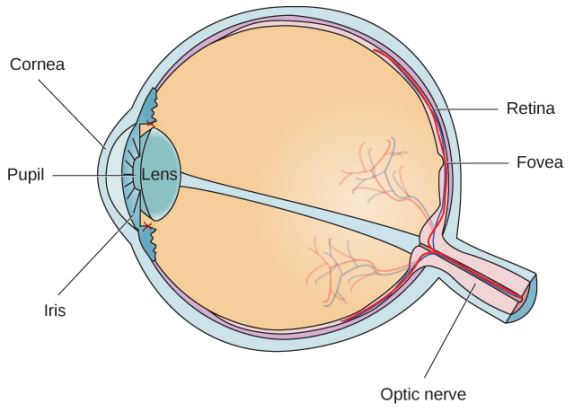

Figure 2.1: The anatomy of the eye, taken from [15].

reflected from close or far objects, the lens is linked to muscles that can alter its shape. A normal-sighted person's lens will focus images precisely on the fovea, a small depression in the retina, the light-sensitive lining of the eye located at the back of the eye. There are specialized photoreceptor cells packed closely together in the fovea. These cone-shaped photoreceptor cells are light-detecting cells and are known as cones.

Cones are particular kinds of photoreceptors that function best in bright light. Cones have outstanding spatial resolution and are highly sensitive to minute details. They also play a crucial role in how we perceive color. The cones are concentrated in the fovea, where images typically focus. The other type of photoreceptors, rods, are distributed across the rest of the retina, [16]. Although they lack the cones' spatial precision and color function, rods are specialized photoreceptors performing well in low light. They are essential in our vision in dimly lit surroundings and our perception of movement on the periphery of our visual field.

We, therefore, adapt our eyes so that the light reflected from the object falls onto the fovea in order to view an object with 100 percent sharpness. Although our foveal vision is clear and vibrant, it only encompasses a small portion of our visual field—about 2 degrees, Figure 2.2. Visual clarity rapidly decreases away from the gaze's center. For instance, at 5 degrees, our visual acuity is approximately 50%. Our retina is mainly employed to detect movements outside of our "useful" visual field of roughly 30 degrees, [17].

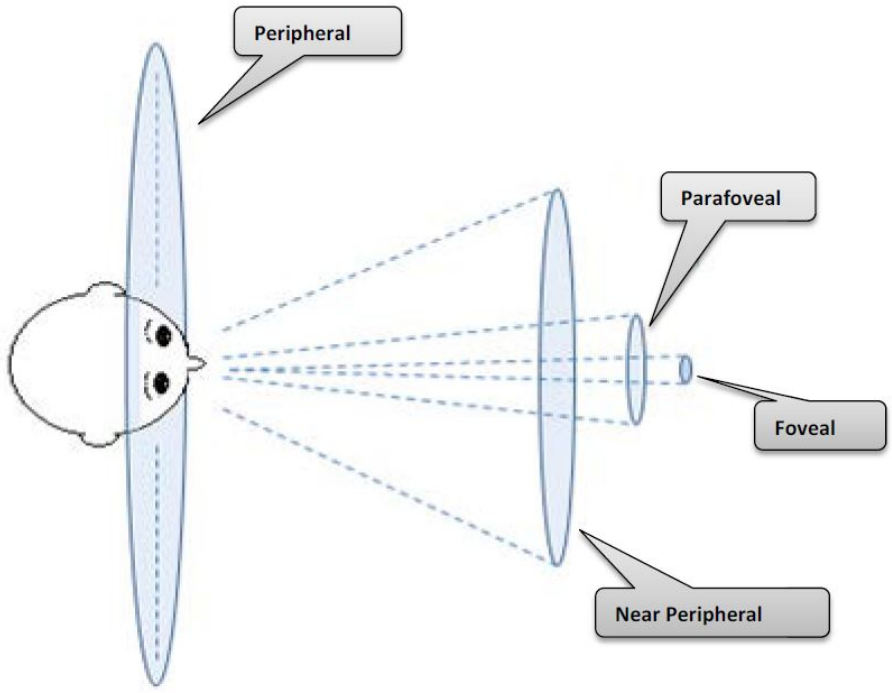

Figure 2.2: Cone of vision, taken from [17].

We continuously scan our visual field with quick ballistic eye movements to make up for our limited foveal vision. The oculomotor system facilitates this dynamic interaction with the surroundings. The oculomotor system regulates gaze-orienting movements as well as maintaining visual stability. It comprises the vestibular system and the visual system's efferent limb. The efferent limb carries out eye movements and maintains eye posture, [18].

On the other hand, the vestibular system provides our brain information about motion, head position, and spatial orientation, which in turn supports motor tasks, including balance, stability when moving, and posture. There are five fundamental types of eye movements based on their function: three gaze-orienting movements, saccadic, smooth pursuit, vergence, and two gaze-stabilizing motions, vestibulo-ocular (VOR) and optokinetic nystagmus (OKN), [19].

More specifically, saccadic eye movements might be small (i.e., cover short distances) or large (i.e., cover larger distances), as in reading (e.g., looking around a room). We do not process visual information when our eyes shift quickly from one region to another during saccades, [17]. Instead, we process visual information between saccades when we maintain a relatively stable gaze for brief periods to reposition a new image onto the fovea or capture an object's "foveal snapshot". Most of our viewing time is spent during the brief fixations that occur in between saccades (about 90 percent),[19]. On the other hand, smooth pursuit refers to the movement generated when our eyes are following a moving object, while vergence is the movement of eyes in the opposite directions.

In particular, the human visual system utilizes saccades to actively shift fixations towards areas of interest, allowing for the extraction of detailed information from the visual environment. This sequence of saccades and fixations is commonly known as a scanpath [20].

Overall, it is safe to claim that our ability to interact with a complex and dynamic environment through the seamless coordination of our visual and oculomotor systems gives us our subjective sense of a stable world with uniform clarity.

Having established the fundamentals of the vision system and the various types of eye movements, the next step is to analyze these eye movements and their underlying patterns. This analysis will involve the mathematical modeling of scanpaths, which represent sequences of saccades and fixations that our visual system uses to actively explore areas of interest in the visual environment. By examining these scanpaths, we can gain insights into an individual's cognitive processes and attentional mechanisms.

In the following section, we will discuss the mathematical modeling of scanpaths and delve deeper into the specific eye movement types, as well as the event detection techniques used to identify and analyze these eye movements. This information will provide a solid foundation for understanding how eye movement analysis can be utilized in detecting ADHD using eye-tracking data.

2.2 Eye Movement Analysis

2.2.1 Mathematical Modelling of Visual Scanpaths

The investigation of visual scanpaths is shaped by three primary factors. First, top-down processes encompass the observer's individual characteristics, which influence their eye movements. These can be intentional, such as completing a task, or unintentional, stemming from factors like the observer's background, age, or gender. Second, bottom-up processes are associated with the visual stimulus and involve basic image features like color and motion. Lastly, the third factor pertains to the eye movement system itself, including the natural inclination to concentrate on the center and the patterns of eye movements, such as saccades, [21].

In every instance, mathematical modeling of visual scanpaths relies on the Markov process. This stochastic model describes a sequence of potential events, where the probability of each event is dependent solely on the state reached in the preceding event. It is a stochastic process featuring a series of alternating states, [22].

Coutrot A. [21] outlines three Markov process-based approaches: the Hidden Markov Model (HMM), the Brownian Motion, and the Lévy Flight. All three are employed to represent random movement in various fields, such as physics, finance, and biology. Brownian motion is observed in the motion of pollen grains on still water or movement of dust in a room and pollutants in the air, [23], while HMMs have been employed to analyze the movements of large carnivores, such as African wild dogs, in relation to human-altered landscapes and their associated effects on connectivity between isolated populations [24]. Lévy flights are observed in various natural phenomena, such as the foraging patterns of animals [25], the search behavior of microorganisms, the strategies of hunter-gatherers [26], the dynamics of financial markets [27] and the visual gaze.

Brownian Motion

The early attempts to model visual scanpaths were based on the assumption that eye movements during visual search follow a random walk, which is often modeled using Brownian motion. Brownian motion is a simple and widely studied random process that exhibits normal diffusion, and it provided a reasonable starting point for modeling eye movements during visual search. It is characterized by linear-in-time mean square displacement [29]. This continuous-time stochastic process can model the random motion of a viewer's gaze or visual scanpath when observing a scene. The motion adheres to a Gaussian distribution, implying that the gaze moves in small, random steps with zero mean ($\mu = 0$) and a fixed standard deviation (σ), reflecting the Gaussian nature of the displacement [20]. The displacement from the starting point typically grows as the square root of time, signifying a diffusive exploration of the visual scene. This gaze behavior model represents a viewer's attention dispersing over time without any specific pattern or direction.

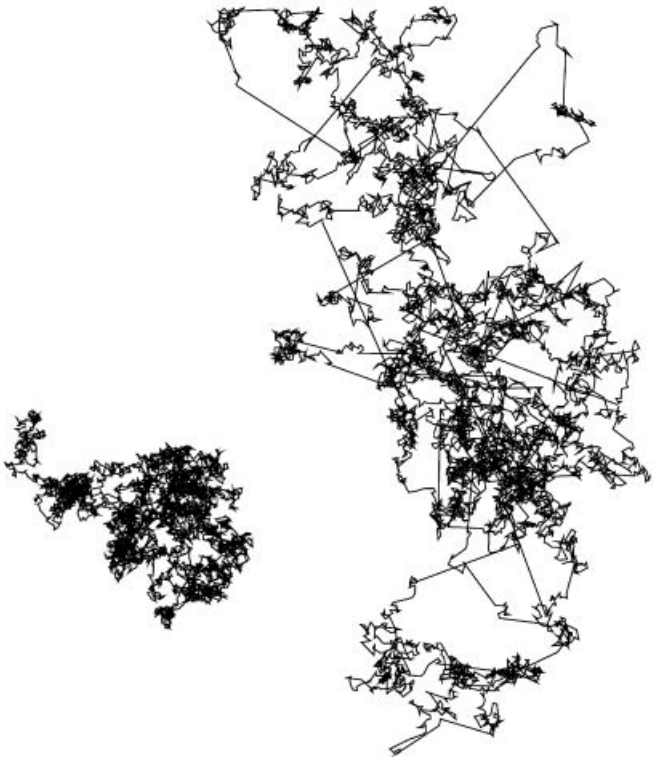

Figure 2.3: Comparison of the trajectories of a Brownian or subdifusive random walk (left) and a Lévy flight with exponent 1.5 (right). Both walks are drawn for the same number of steps (approx. 7000), taken from [28].

Lévy flights

While Brownian motion exhibits some similarities with Lévy flights, their primary distinctions lie in their underlying probability distributions and the nature of the steps they take. As presented in Figure 2.3 even though both trajectories are statistically self-similar, the Lévy flight trajectory possesses a fractal dimension, characterising the island structure of clusters of smaller steps, connected by a long step, [28].

Recent research in cognitive science suggests that standard diffusion processes may not be suitable models for human looking behavior. Specifically, experimental findings affirm that superdiffusive Lévy-type dynamics emerge in this context, [29]. This is something that Dirk Brockmann and Theo Geisel highlighted in their study "The ecology of gaze shifts", [30] more than 20 years ago. They suggested that visual scanpaths generated under natural circumstances are similar in their nature to Lévy fights, meaning that they possess a power law dependency in their magnitude distribution.

More specifically, a Lévy flight is a stochastic process used to model the random motion of a viewer's gaze or visual scanpath when observing a scene. This motion follows a heavy-tailed, power-law distribution [31], as opposed to the Gaussian distribution observed in Brownian

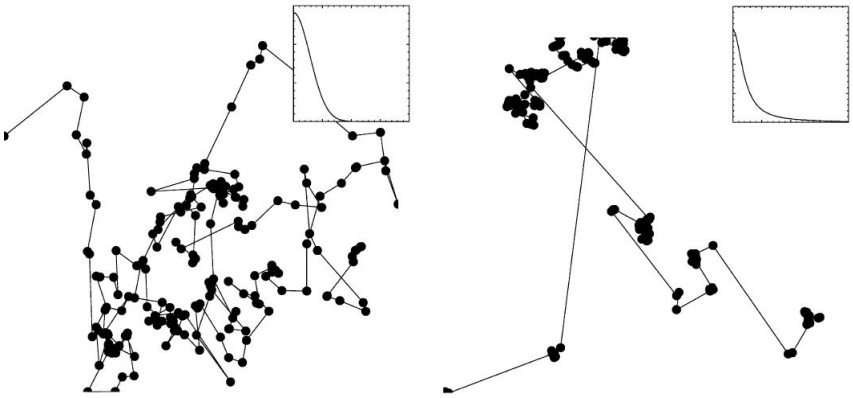

Figure 2.4: Left: Regular difusion process generated from a single-step distribution with finite second moment. Right: A scale free Lévy flight generated from a single-step distribution with an algebraic tail, taken from [30].

motion, Figure 2.4. Lévy flights are characterized by random jumps with lengths distributed according to a heavy-tailed, power-law density, and discontinuous trajectories [32]. These jumps allow for long-range movements, as they are more likely to include larger steps than the average. The impact of the long-tail behavior of the single-step distribution p(x) is depicted in Figure 2.4. Furthermore, Lévy flights have infinite propagation velocities [32], as each jump, regardless of its size, takes one unit of time according to Viswanathan [33] and Shlesinger [34]. Consequently, the ratio of step size to time (i.e., velocity) would also follow a power-law distribution, resulting in a heavy-tailed distribution of velocities in Lévy flights.

In the context of visual scanpaths, this indicates that a viewer's attention may occasionally make extensive shifts across the scene, intermixed with smaller, local movements. Lévy flights, as already mentioned, are typified by power-law scaling, with the tails of the distribution adhering to a power-law, $P(x) \propto |x|^{-\alpha}$, with $1 < \alpha \leq 3$. The parameter α determines the heaviness of the tails, [35]. When $\alpha > 3$, the process resembles a Gaussian random walk [36], while values $\alpha \leq 1$ do not correspond to normalizable probability distributions. For $1 < \alpha \leq 2$, the distribution has no mean and no variance and for $2 < \alpha \leq 3$, the distribution has a mean but no variance [37]. The impact of the exponent α is presented in the Figure 2.5.

It has been shown that Lévy flights reproduce well human sacaddic movements over a saliency map [38]. The saliency is associated with the deterministic part of the dynamics and it drives the gaze to the most relevant points in the image, while the stochastic component, associated with large shifts, is responsible for scene exploration.

Building upon this understanding, it is important to note that foraging models typically exhibit a scaling exponent around 2. It is hypothesized that visual and oculomotor systems would adapt to a similar Lévy flight model [39]. While there is limited literature on this aspect, eye-tracking

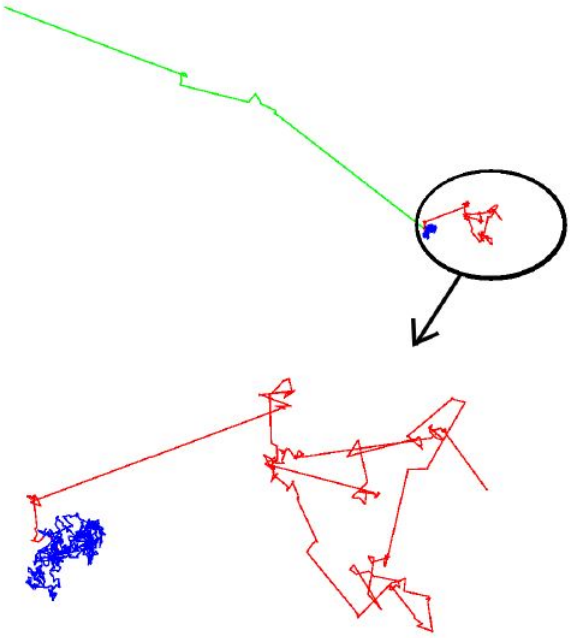

Figure 2.5: Impact of the exponent value, with $\alpha = 1.5$ (green, not present on the zoom), $\alpha = 2$ (red), $\alpha = 3$ (blue), taken from [37].

experiments conducted by Credidio H. et. al. [40] found the exponent to be approximately 2.9, regardless of the level of difficulty of the visual task. This finding supports the notion that the Lévy flight model can be applied to the study of eye-tracking patterns and visual scanpaths in various contexts, and it emphasizes the potential relevance of the scaling exponent in characterizing eye movements.

Hidden Markov Model

The Hidden Markov Model (HMM) is another mathematical approach used to describe visual scanpaths. HMM is a statistical model that assumes an underlying Markov process with hidden or unobserved states, such as "saccade" and "fixation" in the context of eye movements. These hidden states are not directly observed but inferred from the observable data, namely gaze position and velocity, through a series of algorithms [41]. HMMs have been widely used in various applications, including speech recognition, computational biology, and natural language processing [41].

"Feed and Fly" Model

Boccignone and Ferraro proposed a novel model known as the "feed and fly" or IHCS model to simulate the visual scanpaths [20]. The model addresses the variability in visual exploration patterns among different observers and even within the same individual, an aspect often overlooked in other foveation models. The IHCS model generates a series of fixations and gaze shifts under the influence of an information

foraging mechanism that alternates between two distinct states: the "feed" state, which corresponds to fixations when relevant visual information is encountered, and the "fly" state, akin to saccades, which involves seeking significant visual stimulus areas. Additionally, the model features a unique internal gaze-shift simulation step that helps estimate the optimal motor parameters for the actual shift. The simulations indicate that the approach is not dependent on specific features for deriving saliency and can effectively handle both bottom-up and top-down semantic cues. Furthermore, the "feed and fly" model can be readily adapted to incorporate object-based paradigms, offering insights into individual differences in scanpath patterns, an essential aspect of cognitive science.

Considering the "feed and fly" model, which effectively captures the distinct phases of fixations and saccades in visual exploration patterns, it is worth revisiting the earlier discussion on Lévy flights. In cases where the anticipated log-normal distribution [40] does not occur and a power-law distribution is observed instead, the distinction between fixations and saccades becomes more challenging. This power-law distribution suggests the presence of Lévy flights, which complicates the separation between fixations and saccades, unlike the clear differentiation in the "feed and fly" model. This crucial difference emphasizes the distinct characteristics of each modeling approach and warrants further examination of their respective implications.

In light of these differences between the "feed and fly" model and Lévy flights, understanding the implications of such discrepancies will be a vital aspect of this thesis. As we progress through the following chapters, we will delve deeper into the significance of these distinctions and their effects on the analysis of visual scanpaths and eye movements.

2.2.2 Eye Movement Types

In Section 2.1, we provided a brief overview of the fundamentals of eye movement. These fundamentals of eye movements are often characterized by various parameters that can be measured using eye-tracking systems, as demonstrated in Table 2.1.

In this section, we will examine the concepts of fixation and saccade more thoroughly, and introduce both the concept of event detection and the taxonomy of different classification algorithms used in eye movement analysis.

A fixation is a movement that occurs while the eye is essentially stationary and focuses on an object, Figure 2.6. For clear vision, the fixation movement stabilizes the object on the fovea. Tremor, gradual drift, and microsaccades are three different tiny movement types that may be present during fixation events. The eye motion known as a tremor has a frequency below 150 Hz and an amplitude of about 0.01. Tremors' specific purpose still needs to be discovered. When drift and tremor occur together, the eye slowly moves away from the fixation point. With a period of roughly 25 ms, a microsaccade is the fixational eye movement that moves the fastest. A microsaccade movement's purpose is swiftly repositioning the eye to its

Table 2.1: Most common eye movements measured with eye-tracking, adjusted from [42].

Eye movements types	Definition
Fixation	The time the visual gaze maintains on a single location
Dwell	One visit to a relevant part of a stimulus from entry to exit can consist of several fixations
Saccade	Quick movement between fixations, which relocates the focus of attention to a new location
Anti-saccade	A voluntary eye movement made in the direction opposite to the side where a stimulus is presented
Smooth Pursuit	Slowly moving fixation occurring only on moving stimuli
Pupil diameter	The diameter of the pupil which reflects the activity of the autonomic nervous system
Blink	Briefly closures of the eyes

starting point. [43]

On the other hand, a saccade is a fast eye movement between two fixation points. A typical saccade lasts between 30 and 80 ms and moves at a speed between 30 and 500 m/s. A saccade's duration, amplitude, and velocity are all related. Due to this link, it is possible that longer and faster saccades have bigger velocities. Around 200 ms pass between the stimulus's start and the eye movement's beginning. This interval is known as the saccadic latency. It includes the time needed for the brain to decide whether to start a saccade or not, determine how far the eye should move, and send neural pulses to the muscles that move the eyes. Since the human brain is thought to not "see" the image during a saccade, accurate saccade detection is crucial. The saccadic suppression is the name of this occurrence. [44]

At this point, it worth to be mentioned that saccade velocity is task but also subject dependable. The saccade velocity has been found to increase with increasing task difficulty, increasing intrinsic value of visual information, and increasing task experience, [45].

Finally, the saccade velocity depends also on the condition of the subject both physical, mental and psychological. Becker and Fuchs, [46], presented in their study that fatigue and alertness can have a significant effect on eye movement trajectories, while Russo et.al. [47], demonstrated a sensitivity of saccade velocity to sleepiness. In their study Boxer A. et.al., [48], showed that neurodegenerative conditions like frontotemporal dementia and Alzheimer disease were decreasing the visually guided saccade velocity, while Baroni N., et. al., [49], presented that in HIV positive subjects, the accuracy of saccades was also significantly reduced.

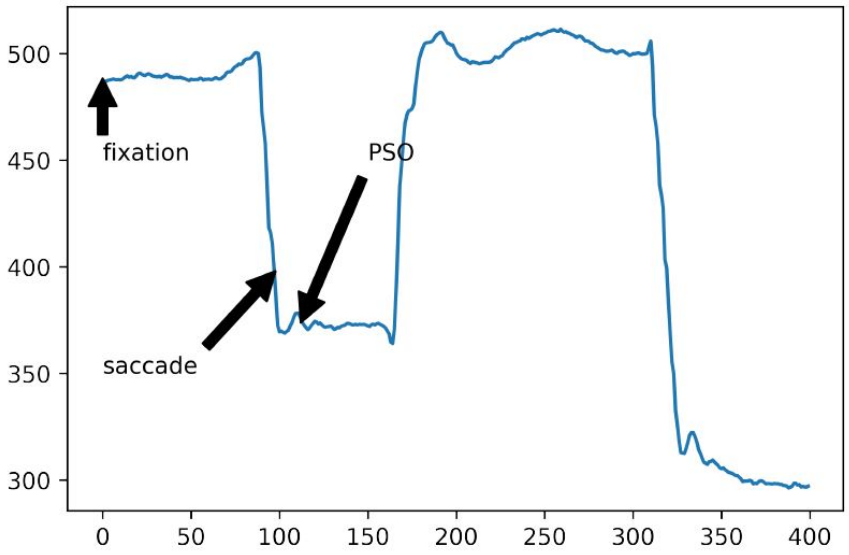

Figure 2.6: Graphical presentation of eye movement events for the horizontal axis, taken from [44].

More in depth correlation of the saccade velocity with mental disorders, and more specifically ADHD, will be presented in the following chapter 2.5.

In the graphical representation, Figure 2.6, the post-saccadic oscillations (PSO) are also presented. The PSO are rapid oscillatory movements or instabilities that may occur immediately after the saccade. However, for the purposes of this thesis we will focus only on the fixation and saccades.

The fixation, saccades and PSO, along with smooth pursuit and blinks, are called events, and the process of distinguish between them is called event detection, which will be presented more analytically in the following chapter.

2.2.3 Event Detection

In eye movement research, event detection aims to precisely and consistently extract events—such as fixations and saccades—from a stream of raw eye movement data. Based on various assumptions about fixation durations, saccadic amplitudes, and saccadic velocities, the acquired raw data are split into events. When eye movement events are separated from the raw eye-tracker data, eye movement analysis is made easier. In the past, event detection was done manually, which was time-consuming. In 1948, Hartridge and Thomson [50] created a method to analyze eye movements at a pace of 10,000 s (almost three hours) of analysis time for 1s of recorded data, while in 1975, Monty [51] noted that it is customary to spend days processing data gathered merely in minutes. Computers, however, have drastically changed how eye movement data are analyzed in the current day. Currently, the only technique employed for event identification is to apply a detection algorithm to the raw gaze data.

Two main categories of algorithms were employed to detect eye move-

ment events. The first category consists of dispersion-based algorithms, which identify fixations and consider the remaining movements as saccades. The second category comprises velocity-based algorithms, which identify saccades and consider the remaining movements as fixations, [52].

In a more detailed taxonomy carried out by Salvucci, [53], fixation identification algorithms were classified with respect to spatial and temporal characteristics, Table 2.2.

Salvucci defined three criteria—velocity-based, dispersion-based, and area-based—that set the three main types of algorithms apart in terms of their spatial properties.

Because fixation locations have low velocities and saccade points have high velocities, velocity-based algorithms highlight the velocity information in the eye-tracking procedures. Dispersion-based algorithms emphasize fixation point dispersion (or spread distance) based on the presumption that fixation points are typically located close to one another. Area-based algorithms locate the relevant visual targets inside defined areas of interest (AOIs). These methods offer both lower-level identification and higher-level assignment of fixations to AOIs, in contrast to the other techniques.

Two criteria are included for temporal characteristics: whether the algorithm employs duration information and whether it is locally adaptive. The fact that fixations are very rarely less than 100ms and frequently in the range of 200–400ms provides guidance for the usage of duration information. Incorporating local adaptivity enables the interpretation of one data point to be affected by the interpretation of temporally neighboring points; this is helpful, for example, to account for variations between "steady-eyed" people and those who exhibit big, frequent eye movements. [53]

Salvucci, focused only on 5 different algorithms, the Velocity-threshold identification (IVT), the Hidden Markov model identification (I-HMM), the Dispersion-Threshold Identification (I-DT), the Minimum Spanning Trees identification (I-MST) and the Area-of-Interest Identification (I-A OI). Table 2.2 shows main temporal and spatial characteristics, which will be considered in the next section.

We will see in the following section that there are several other algorithms. Richard Andersson in his paper, [54] , evaluated ten different

Table 2.2: Taxonomy of fixation identification algorithms, adjusted from [53].

Criteria		Algorithms				
		IVT	IHMM	IDT	IMST	IAOI
Spatial	Velocity-based	x	x			
	Dispersion-based			x	x	
	Area-based					x
Temporal	Duration sensitive			x		x
	Locally adaptive		x	x	x	

Table 2.3: Classification Algorithms, taken from [54].

Algorithm	Events				
	Fixation	Saccade	PSO	Smooth Pursuit	Blink
(Humans)	x	x	x	x	x
CDT	x				
EM		x			
IDT	x	x			
IKF	x	x			
IMST	x	x			
IHMM	x	x			
IVT	x	x			
NH	x	x	x		
BIT	x				
LNS		x	x		

algorithms, as presented in the Table 2.3 . These algorithms are the following: Fixation Dispersion Algorithm based on Covariance (CDT), Engbert and Mergenthalser (EM), Identifiacation by Dispersion-Threshold (IDT), Identification by Kalman Filter (IKF), Identification by Minimal Spanning Tree (IMST), Identification by Hidden Markov Model (IHMM), Identification by Velocity Threshold (IVT), Nyström and Holmqvist (NH), Binocular-Individual Threshold (BIT), and Larsson, Nyström and Stridh (LNS).

In the following section, we will review more thoroughly the most important classification algorithms as well as their mathematical background.

2.3 Mathematical Background of Event-Classification Algorithms

2.3.1 Velocity-based Algorithms

Identification by Velocity Threshold (IVT)

The easiest identification approach to comprehend and use is the velocity-threshold fixation identification (IVT). By comparing the point-to-point velocities of fixation and saccade locations, the velocity-based IVT approach may distinguish between them. For each sample of an eye position, the velocity value is calculated and then put up against a threshold, [53].

Fixations are defined as segments of samples with point-to-point velocities below the predefined velocity threshold, and saccades are defined as segments of samples with velocities above the threshold. Fixations and saccades are distinguished by this fixed velocity threshold. Often, other algorithms are built around this fundamental velocity criteria. The traditional IVT method divides all input data from eye-tracking into fixations and saccades only. The other event kinds, such as smooth pursuits, post-saccadic oscillations, and noises, are not considered, [54].

The algorithm's primary flaw is that it only considers the gaze velocity while ignoring other factors like signal acceleration, the direction in which the gaze moves, the distance between the eye and the camera, etc. As a result, the velocity ranges of the fastest "slow" eye movements and the slowest portions of saccades may overlap, leading to incorrect classification of events. It would seem that utilizing additional eye movement data like acceleration, amplitude, and position could enhance the outcome.

The performance of the event detection algorithms is impacted by changing the threshold values because there is no single optimal threshold velocity value. For these reasons, several researchers have designed and assessed the effectiveness of IVT algorithms using various threshold levels. Due to this heterogeneity, it is challenging to compare different studies of threshold-based event detection algorithms, [44].

However, as mentioned by Salvucci, [53], if angular velocities can be computed (i.e., the distance from eye to visual stimuli is known), the point-to-point velocity threshold can be approximated from a reasonable angular velocity threshold. On that basis Sen and Megaw, [55], in order to register a saccade, the peak velocity must have exceeded 20 deg/s for a minimum of 10 ms, while if it was reduced below that, then PSO would be identified as saccades.

Engbert and Mergenthaler (EM) Algorithm

The algorithm employed by Engbert and Mergenthaler, [56] is an extension of the algorithm employed by Engbert and Kliegl, [57]. Similar to its predecessor, this algorithm uses a velocity threshold to detect saccades, but the threshold is calculated for each recording based on an estimate of the data's noise level. In addition, this algorithm enforces a minimum saccade duration to mitigate the effects of noise.

Its original function was to detect microsaccades, which are erratic, miniature eye movements that occur during visual fixation on a stationary target. In contrast to the prevalent belief that microsaccades were randomly distributed in time, they suggested that microsaccades were triggered dynamically. As a result of this dynamic triggering mechanism, the fractal dimension of trajectories could predict the individual microsaccade rate. Their proposed algorithm for detecting microsaccades could also detect voluntary (larger) saccades, [54].

Identification by Hidden Markov Model (IHMM)

A more advanced iteration of the IVT model that incorporates a probabilistic depiction of human visual systems is called the hidden Markov algorithm (I-HMM). It employs probabilistic analysis to determine the most likely identifications for a specific protocol. For example, the domains of speech and handwriting recognition have extensively used probabilistic finite state machines called hidden Markov models (HMMs).

As Salvucci and Goldberg, [53], describe the two-state HMM presented in Figure 2.7 is the core of I-HMM. Both observation and transition

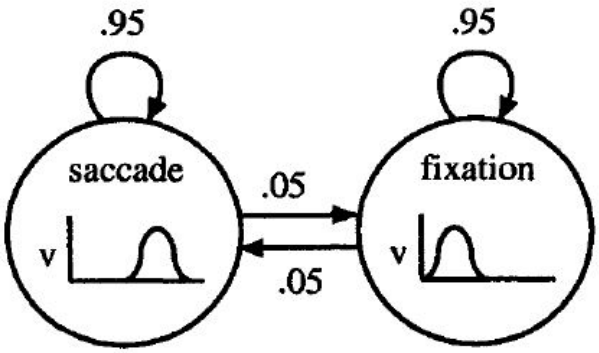

Figure 2.7: Sample two-state HMM, taken from [53].

probabilities are included in the HMM. The distribution of anticipated velocities in each state is represented by the observation probabilities for each state. The distribution in the first state, which corresponds to saccade points, is centered around greater velocities; in the second state, which corresponds to fixation points, it is centered around lower velocities. The transition probabilities show the possibility of staying in one state or transitioning to another for each state (the arrows leaving the states). According to Figure 2.7 transition probabilities, there is a high possibility of staying in each state (.95) and a low likelihood of transitioning (.05). The observations (i.e., velocities) produced during saccadic eye movements are thus represented probabilistically by the HMM.

Following up on the above analysis, Komogortsev [58], describes the process more explicitely while introducing the viterbi algorithm as described by Forney [59]. He notes that the I-HMM uses three crucial stages of the procedure. Each eye position sample in the first stage is categorized as either a fixation or a saccade depending on the velocity threshold, just like in the IVT and similar to what we mentioned above. According to the probabilistic parameters (initial state, state transition, and observation probability distributions) of the model, the Viterbi sampler defines the second stage, where each eye position is reclassified as a fixation or saccade. Given the probabilistic parameters of the model, the Viterbi sampler aims to optimize the probability of the state assignment. Usually, the I-HMM's probabilistic parameters are not at their optimal level and need to be improved. As a result, the Baum-Welch re-estimation method determines the third and final stage of the I-HMM. In an effort to reduce state assignment mistakes, this approach re-estimates the initial probabilistic parameters. The Baum-Welch can re-estimate a parameter numerous times if necessary.

Identification by Binocular-Individual Threshold (BIT)

Another velocity-based algorithm is the Binocular-Individual Threshold (BIT) algorithm. It is superior to existing velocity algorithms in three distinct ways. First, it accommodates binocular viewing and identifies

fixations and saccades using information about covariations between the movements of both eyes. Even though it is difficult to distinguish minor saccades from background noise, it makes sense to exploit the fact that the eyes are frequently directed at the same object. Therefore, if the left eye travels toward an item, the right eye should also do so. In this way, it is easier to identify whether a velocity peak is due to an actual movement or noise, as both eyes should display this velocity peak simultaneously. Second, it estimates rather than pre-sets the velocity threshold to identify fixations and saccades, allowing the threshold to vary among eye-movement directions, tasks, and users. Thirdly, it accounts for the intrinsic randomness of eye movements so that not every record that exceeds the threshold is identified as a saccade. [60]

This algorithm was proposed by Ralf van der Lans, Michel Wedel and Rik Pieters, and in their paper [60], the note that in order to determine individual, task, and eye-specific thresholds, they use techniques from Robust Statistics which allows them to estimate velocity thresholds based on individual-level variability of the eye movement recordings within a fixation. Subsequently these velocity thresholds are input to Shewhart quality control chart procedures, that classify the point of regard (POR) as a saccade or a fixation. The intuition of the algorithm is that it determines the variability of the POR of both eyes within a fixation for a specific individual, and then determines when the velocity of the POR exceeds the within-fixation variability, and label the corresponding epochs as saccades.

Nyström and Holmqvist (NH) Algorithm

Nystrom and Holmqvist's, [61] approach was the first algorithm to specifically recognize post-saccadic oscillations in addition to fixations and saccades. It is an adaptive method since it modifies the velocity threshold dependent on the data's noise level. The algorithm is based on Smeets and Hooge's saccade detection velocity algorithm and does not utilize positional data directly. Bahill et al. recommended using velocity data instead of position data to identify saccades because it offers more accurate and intuitive information regarding the precise onset and offset of a saccade. The proposed technique is comprised of five main steps: filtering and denoising, peak saccade detection, saccade onset/offset detection, post-saccade oscillations detection, and fixation detection. In addition to explicit post-saccade oscillations detection, the algorithm's contributions include

- an adaptive, data-driven peak saccade detection threshold,
- a new way for defining saccade onsets and offsets,
- and the usage of thresholds driven by physiological restrictions of eye movements.

Figure 2.8 presents how the adaptive velocity threshold is being defined. An initial threshold PT1 is selected. For all samples with

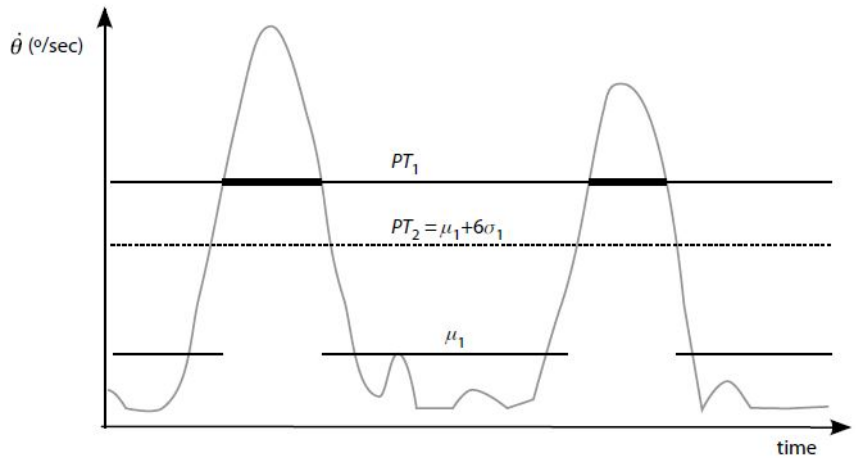

Figure 2.8: Principle for iterative estimation of saccadic velocity threshold, taken from [61].

velocities below PT1, the average velocity $\mu 1$ and standard deviation $\sigma 1$ are calculated. The next threshold, PT2, is then set as $\mu 1 + 6\sigma 1$. The iterative process is continued until |PTn - PTn-1| is smaller than 1°/sec, and PTn is selected as the final velocity threshold. This iteration can be done separately for each participant or even for each trial.

Larsson, Nyström and Stridh (LNS) Algorithm

LNS is another velocity-based algorithm capable of detecting post-saccadic oscillations, [62]. The algorithm is adaptive, but what is novel about it is that it was designed to detect saccades and post-saccadic oscillations in signals recorded when viewing both static and dynamic scenes. This indicates that it is able to detect saccades and post-saccadic oscillations even when eye pursuit movements are present.

Smooth pursuit movements generate velocities that algorithms that maintain standard velocity thresholds handle inconsistently. Because the velocity of a fast smooth pursuit movement overlaps with the velocity range of a slow saccade, it is challenging to establish a velocity threshold for discriminating between these two types of eye movements.

Consequently, there is typically no clear classification of these movements as fixations or saccades; their classification depends on the particular smooth pursuit movement and the algorithm thresholds in use. Since the acceleration of saccades is greater than that of smooth pursuit movements, the acceleration signal has been used to detect saccades reliably in signals containing smooth pursuit movements. This algorithm achieves this and provides more accurate estimates of the saccades' on- and offsets, outperforms a previously proposed estimation method for PSO, and enables modeling of the PSO.

2.3.2 Dispersion-based Algorithms

Identification by Dispersion-Threshold (IDT)

The Identification by Dispersion-Threshold (IDT) algorithm is one of the most popular algorithms for detecting fixations. In contrast to velocity-based identification algorithms, the dispersion-threshold identification algorithm takes advantage of the fact that fixation points tend to cluster close together due to their low velocity. According to Salvucci and Goldberg [53], it is based on Widdel's algorithm, [63] for data reduction.

The IDT algorithm combines x and y data with two fixed thresholds: the maximum fixation dispersion threshold and the minimum fixation duration threshold. IDT identifies fixations as clusters of consecutive points within a given dispersion or maximum separation. Because fixations typically last at least 100 ms, dispersion-based identification techniques typically incorporate a minimum duration threshold between 100 and 200 ms.

For a data sample to be considered a fixation, it must be contained within a spatial region that does not exceed the dispersion threshold and span at least enough time to satisfy the duration threshold. The samples that meet these criteria are identified as belonging to a particular fixation, [54].

In particular, the IDT algorithm employs a moving window that spans successive data points to check for potential fixations. The moving window commences at the beginning of the procedure and spans an initial minimum number of points determined by the given duration threshold and sampling frequency. The dispersion of the points within the window is determined by calculating the sum of the differences between the maximum and minimum x-values and the maximum and minimum y-values. In simpler terms, the dispersion D is calculated as the sum of [max(x) - min(x)] and [max(y) - min(y)], [53]. If the dispersion is greater than the dispersion threshold, the window does not represent a fixation and shifts to the right by one point. If the dispersion is less than the threshold, the window represents a fixation. In this instance, the window is expanded (to the right) until its dispersion exceeds the threshold value. The final window is registered as a fixation at the window's centroid with the specified onset and duration. This procedure continues with the window moving to the right until its conclusion.

The centroid and diameter are employed to characterize fixations. Typically, a circular area is assumed, and the average distance between each sample and the fixation centroid is used to estimate the radius. The dispersion threshold can be set to include between 1/2 and 1 degree of visual angle if the eye-to-screen distance is known, or it can be estimated through exploratory data analysis. The duration threshold is typically set between 100 and 200 milliseconds based on the processing demands of the task.

Fixation Dispersion Algorithm based on Covariance (CDT)

The Fixation Dispersion Algorithm based on Covariance (CDT) by Veneri et al. [64] is an improvement over the fixation dispersion algorithm based on F-tests (FDT) that the same authors, [65], had previously developed. The improvement consists of incorporating co-variance calculations on the x and y coordinates of the gaze into their previous algorithm, which was based on the F-test.

The FDT algorithm identified fixations using the F-test of equal variance. During fixations, this technique was based on the assumption that the variance along the x was not significantly different from that along the y. In general, the FDT algorithm was able to identify more fixations than the I-DT algorithm, [53], which according to Shic et. al. [66] was accounted to be the most robust method in his evaluation.

However, in clinical cases where eye movements were affected by brain diseases, assumptions such as the normality of distribution and the number of data points reduced the algorithm's effectiveness. The F-test was extremely sensitive to deviations from the normality assumption. Veneri et al. used a mixed method based on covariance and the F-test for equal variance to develop the CDT algorithm, which employs variance and covariance thresholds in addition to a duration threshold.

Identification by Minimal Spanning Tree (IMST)

Identification by Minimal Spanning Tree is another technique for event detection (IMST). MST identification is based on minimum spanning trees (MSTs), which are trees that connect a set of points in a way that minimizes the total length of the tree's line segments. MSTs can provide a highly adaptable and controllable representation for dispersion-based identification of fixations, [54].

The algorithm aims to capture the data comprehensively while minimizing branching. It prioritizes creating branches that effectively separate samples from two distinct clusters into separate nodes higher up in the tree, rather than forcing extensive branching towards a single node at a lower level, Figure 2.9. A two-step approach is required, first the construction of the MST and then a search of the MST. The building process uses Prim's algorithm [67].

One and only one MST exists for a set of points. The advantage of an MST data representation is the degree of control, flexibility, and local adaptation for dispersion analysis and the improvement of subsequent characterizations of defined fixations.

Following is the procedure for the IMST algorithm: Utilizing Prim's algorithm, construct MST from protocol data points initially. The maximum depth for each MST point is then determined using a depth-first search. Consequently, the branching depths below a defined setpoint define locations near the edge of the MST that are unsuitable for separating fixations. Saccades can be identified if the edges connecting each endpoint exceed the minimum branching depth. The mean m and standard deviation σ of edge

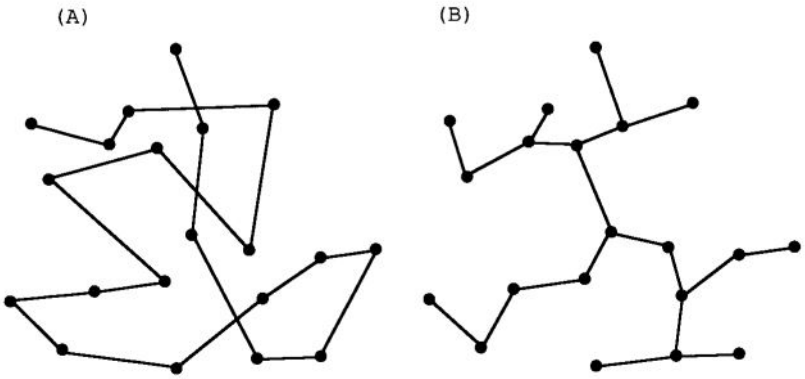

Figure 2.9: Two interpretations of an identical sample of eye-gaze locations: (A) temporal scanpath record, and (B) minimum spanning tree from graph theory, taken from [67].

lengths provide a locally adaptive comparison for the separation of fixations, which are identified as clusters of points not separated by saccades. Comparing the edge under consideration to both m and σ of neighbor edge lengths can result in separation, [53].

The MST format allows for additional fixation characterization parameters. For instance, the MST length is determined by the graph's longest path. Areas can define critical paths with a minimal degree of branching. These parameters and others allow for an estimation of the shape and possibly the direction of fixations, going beyond simple centroid and size descriptions.

2.3.3 Area-based Algorithms

The previous dispersion-based identification methods can identify fixations at any location within the visual field. In contrast, area-of-interest fixation identification (IAOI) only identifies fixations that occur within a target area. The target areas are rectangular regions of interest in the visual field representing information units. These target regions, typically used in subsequent analyses such as tracing, keep identified fixations close to relevant targets. On this basis, Buurman [68], conducted an analysis to determine whether the changing size of the target window affected the students' reading behavior.

IAOI utilizes a duration threshold to differentiate between fixations and saccades in target regions. IAOI begins by associating data points with target areas, labeling points within a target area as fixation points for that target, and points outside all target areas as saccades. IAOI then combines successive fixation points for the same target into fixation groups and discards saccade points. It then excludes fixation groups that fall below a specified duration threshold and converts each fixation group into a fixation tuple, [53].

Overall IAOI can provide a more comprehensive picture than fixations

alone. Similar to a fixation, dwell time within an area has a beginning and an end. However, because fixations can serve as input data for AOI determination, the time between these dwells is insufficient to describe saccadic behavior. Instead, multiple saccades and fixations may be interspersed between AOIs. Although AOI-based dwell-time algorithms are not fixation algorithms in and of themselves, they are useful for describing higher-level collections of fixations organized around visual targets and areas.

Astar Lev utilized the above concept of AOI in his work, [69] by integrating an eye tracker with a neuropsychological performance test that assessed cognitive impairment related to ADHD.

His study aimed to evaluate the utility of eye movement measures for distinguishing ADHD patients from healthy controls. Using face-valid gaze direction measures, it was discovered that ADHD patients spent more time gazing at irrelevant regions, both on and off-screen, than healthy controls, Figure 2.10.

The group differences appear to be driven by the inability to suppress spontaneous eye movements toward different distractors, which corresponds to the distractibility that is characteristic of ADHD.

In the following chapter we will focus on applications of eye-tracking

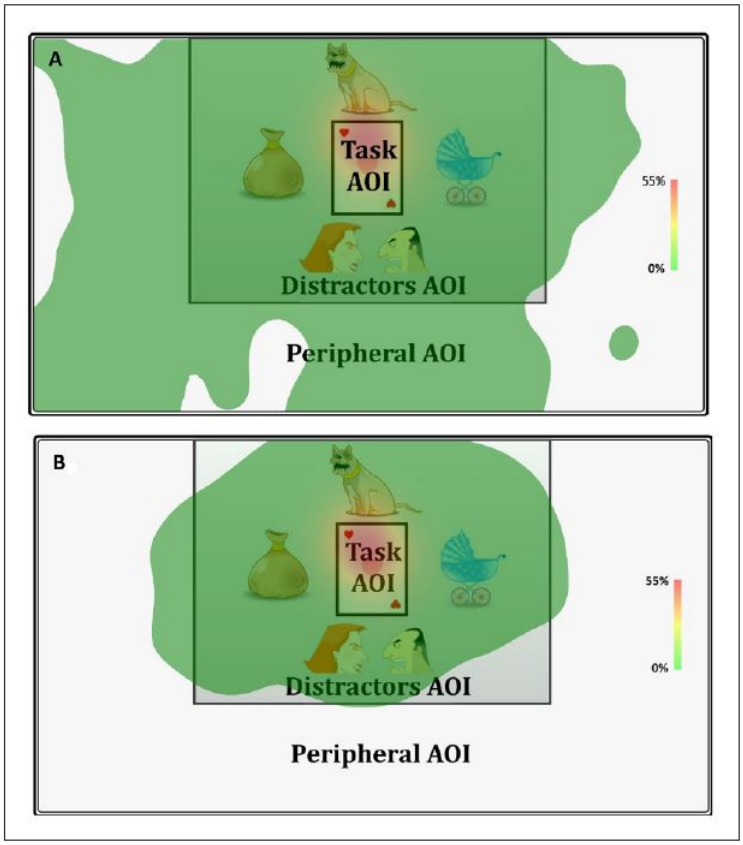

Figure 2.10: Relative gaze duration heat maps, ADHD patients (A) and healthy controls (B), taken from [69].

technology in the field of neurodevelopmental disorders, especially in studying the clinical characteristics of ADHD. We will review the latest studies and focus on what is considered to be the State-of-the-Art techniques for the detection of ADHD.

2.4 Eye Tracking

Eye tracking is the process of recording a viewer's gaze at a specific moment of a stimulus. Usually, this is accomplished by watching a viewer's eye movements. Eye-tracking technologies use measurements of the user's pupil size, movement, and position to identify areas of interest. The early observations of eye movements were made by the investigator looking within or utilizing a mirror, telescope, or peephole to observe the subject's eye, [70].

Eye tracking was frequently a laborious process that required the individual to bite on a stationary object or have their head held perfectly still using a chin rest, which limited its application to research labs, [71]. These techniques were often unreliable because any feature of the eye being researched may be obscured by the eye performing the study. They were also frequently intrusive because a portion of the measurement apparatus was in direct contact with the subject. Therefore, the first significant development was the development of mechanical tools that could convert the eye's movements into long-lasting, impartial records of its motion.

Modern eye trackers are far more adaptable, and non-intrusive techniques have been created to enable the use of the device in settings that are closer to nature.

Different fields, including neuroscience, marketing, advertising, and computer science, have utilized eye-tracking technology. In tandem with new technological advancements, new applications are emerging in a vast array of research fields.

To capture all these new opportunities, several different eye-tracking techniques have been developed throughout the years. These techniques we will analyze more thoroughly in the following chapter.

2.4.1 Eye Tracking Techniques

Eye tracking in literature often involves one of four techniques: the scleral coil, electro-oculography (EOG), photo-oculography (POG)/infrared-oculography (IOG), and video-oculography (VOG), [72].

In the scleral coil approach, the eye is directly covered by a contact lens that has a mechanical or optical reference object, Figure 2.11. Robinson initially presented it in 1963. When placed in a magnetic field, a coil linked to the contact lens induces an electric potential that can be used to determine the position of the eyes. The scleral coil method is the most precise approach for determining the eyes' location. In addition, eye trackers may accurately move an image in line with the eye movements

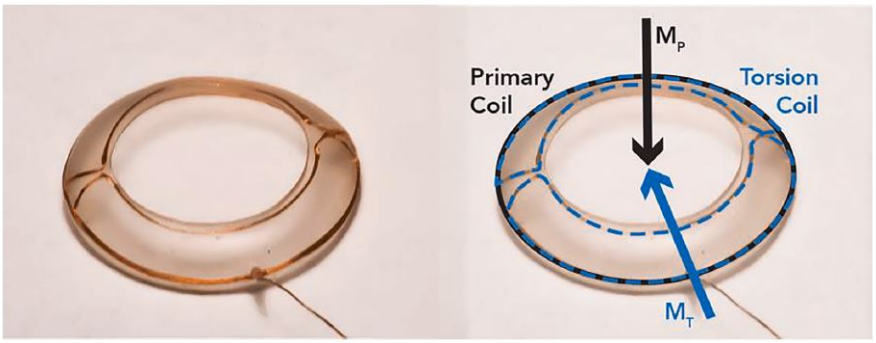

Figure 2.11: Scleral Search Coil contact lenses, taken from [73].

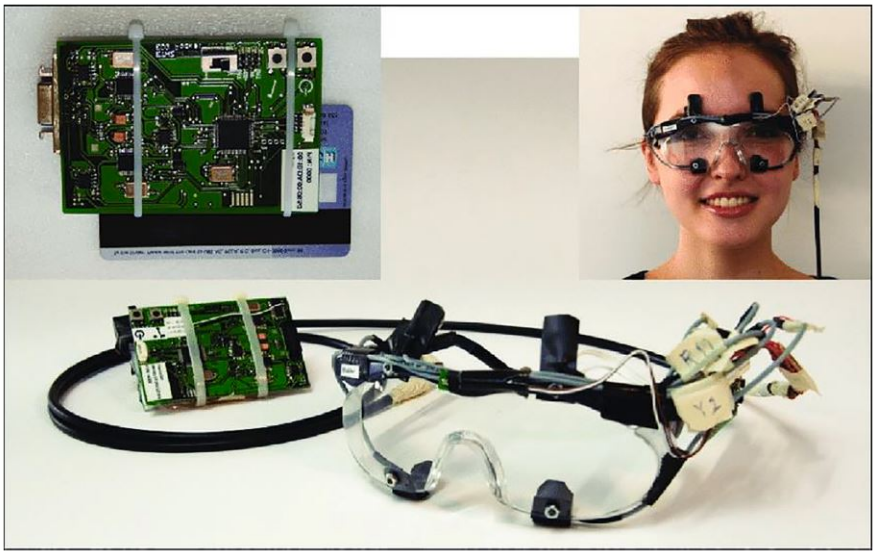

Figure 2.12: Electro-oculography device, taken from [73].

by monitoring the infrared wavelength range reflected by the mirror and recording the eye movements, [73].

This technique also establishes where the eye is in relation to the head. Both the head posture and the eye position are necessary for gaze estimation. The lack of head posture data constrains its application for point gaze estimation. However, the significant discomfort brought on by its intrusive method precludes its usage in actual applications.

EOG is a practical and affordable method for human-computer interaction, Figure 2.12. The EOG method measures the electric potential variations in the skin's surface using electrodes positioned near the eye. The head posture cannot be inferred from the eye movements captured by EOG. Although EOG is not a technique for everyday usage, medical settings and labs can benefit from its use. This method can track head movements and is linearly related to eye movements. However, eye drifts and eye problems may restrict the use of EOG.

POG is based on the measurement of the corneal light's deviation from

Figure 2.13: Infrared oculography approach setup, taken from [73].

the pupil's center. It measures eye characteristics like corneal refractions from the light source, limbus sclera boundary, and pupil shape, [74]. As the recording apparatus can be made of two near-infrared cameras watching each eye by reflection over a heated mirror transparent to visible light, this method does not require any attribute on the subject's head or eye. It measures eye alignment independent of head translations and distinguishes between lateral and rotary motion. On the other hand, infrared light can also be created using a pair of spectacles, Figure 2.13. Algorithms for detecting light and the pupil are primarily employed in this method.

Video or image-based trackers use camera and image processing techniques to locate the gaze point in real-time. Infrared or visible lighting options are available for the VOG method. The VOG is a non-invasive technique that conducts remote eye tracking. Based on the number of cameras utilized, there are two approaches to implementing video eye tracking: the first approach employs a single camera, while the other uses numerous cameras.

Other video-based eye trackers include head mount, table mount, and tower mount trackers. Due to changes in head position, there is a significant disadvantage when employed in Human-Computer interaction systems. Using two stereo cameras, Figure 2.14, or one wide-angle camera to look for the person in front and another to point at the person's face and zoom in can overcome this problem for remote trackers.

According to published research, image-based eye-tracking techniques are appropriate for real-world use because they are non-intrusive and non-contact. In addition, the effectiveness of image-based eye trackers is improved by the constant advancements in computing power and camera quality, [72].

In the following chapter we will present the State-of-the-Art eye tracking systems and methods, as well as more details regarding the specific eye tracker that was utilized to collect the data used in this thesis.

Figure 2.14: Video oculography using two cameras, taken from [73].

2.4.2 State of the Art Technologies and Systems

In this chapter, we will focus on three of the most prominent systems that are currently available: the EyeLink eye-tracker provided by SR Researchm, which is also the one utilized in the present thesis, the Tobii eye-tracking system, and the open source eye tracking toolbox Pygaze. For these three technologies we will review in the following paragraphs, their embedded classification algorithms, so to get a better understanding what differentiates them and makes the state-of-the-art technologies.

SR Research - EyeLink Eye-Tracking System

The eye tracker that was utilized to collect the data on which this thesis is based on, was the EyeLink 1000 by SR Research Ltd., Mississauga, Ontario, Canada, [14].

Eyelink 1000 is considered to be one of the most precise and accurate video-based eye tracker. With several mount options, replaceable lenses, and head-fixed and head-free tracking modes, it is very adaptable. The Desktop Mount was the chosen mount type. The Desktop Mount, which is positioned below the tracked area, is ideal for conventional screen-based eye tracking. It makes it ideal for research using EEG or TMS because the camera may function at a range of 40-70 cm, and no electronics are required to be close to the participant's head. It was paired with the SR Research Head Support, allowing for more comfort and improved head placement by allowing for separate height adjustments for the chin, forehead, and chin

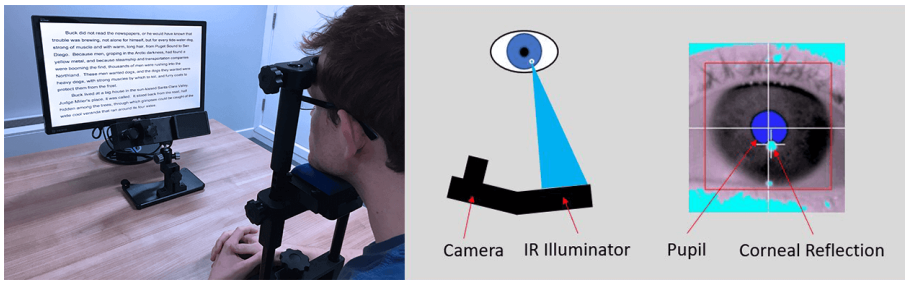

Figure 2.15: EyeLink 1000 configuration, taken from [76].

supports, [75].

The EyeLink 1000 incorporates an exceptional high-speed camera as its core component, enabling the recording of eye movements at a remarkable rate of up to 2000 frames per second [76]. Within a mere three milliseconds of capturing the participant's eye image, EyeLink systems swiftly identify the precise location on the screen where the participant is looking and promptly transmit this information back to the computer responsible for managing stimulus presentation. The eye-tracking software employs image processing algorithms to locate two critical areas on each image supplied by the eye-tracking camera: the pupil's center and the corneal reflection's center. The infrared illuminator, which is positioned adjacent to the camera, is the fixed light source that is reflected by the cornea, Figure 2.15.

The eye-tracking mechanism operates based on the pupil-corneal reflection (P-CR) phenomenon. As the eye rotates, the center of the pupil moves across the camera sensor. Conversely, the position of the corneal reflection (CR) on the camera sensor remains relatively constant when the head is stabilized since the source of the reflection remains stationary relative to the camera. In Figure 2.16, depicted below, you can observe the camera's perspective as an eye turns to one side and subsequently rotates in the opposite direction. Notably, the center of the pupil undergoes a noticeable shift, while the center of the CR remains relatively fixed in terms of camera pixel coordinates.

SR Research develops and sells high-speed, video-based eye-tracking equipment. The company has expertise in all areas of eye-tracking technology, including hardware development, manufacturing, and software de-

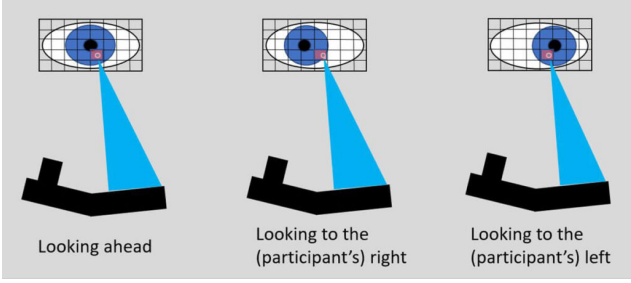

Figure 2.16: Pupil-Corneal Reflection (P-CR) Eye Tracking, taken from [76].

velopment, [75].

The EyeLink 1000 Host PC performs real-time eye tracking at speeds of 250, 500, 1000, or 2000 samples per second with no loss of spatial resolution while also computing the subject's true gaze position on the display. Eye-motion events such as saccades and fixations are detected and analyzed in real-time. These events can be saved in a data file on the Host PC, sent with minimal delay over the Ethernet link to the Display PC, or output as analog signals (if the analog/digital I/O card is installed). The operator can perform subject setup, monitor performance, and communicate with applications running on a Display PC from the Host PC.

The stored samples are time-stamped in milliseconds and contain monocular or binocular eye-position data in eye-rotation angle or display-gaze coordination. Pupil sizes such as area or diameter can also be recorded. Samples may also include eye-movement resolution (used to calculate true velocity or saccadic amplitudes), button presses, or the status of digital inputs. Fixations, blinks, and saccades are examples of eye-movement events detected by the EyeLink tracker's on-line parser. The beginning and end of these events are marked, allowing samples to be assigned to eye-movement periods without using complex algorithms. Important data for analysis, such as average position for fixations and peak velocity for saccades, is also recorded in the events. Other events record subject responses (such as button presses), synchronization, and data messages from applications. These can be used to record the time of a change in the display or an experimental condition.

For saccade detection, three thresholds are used: velocity ($^\circ/sec$), acceleration ($^\circ/sec^2$), and motion ($^\circ$). The velocity threshold is the speed at which the eye must move for a saccade to be detected. With a velocity threshold of 22 degrees per second, saccades as small as 0.3° can be detected, making it ideal for smooth pursuit and psychophysical research. A conservative threshold of 30°/sec is preferable for reading and cognitive research, as it reduces saccades while increasing fixation durations. The higher the threshold, the fewer microsaccades detected, and thus the fewer short fixations (less than 100 msec in duration) in the data. Short fixations (2% to 3% of total fixations) are to be expected, and most researchers simply ignore them.

Eye-movement acceleration is critical for detecting small saccades, particularly in smooth pursuit. Because acceleration data contains far more noise than velocity data, thresholds of $4000^\circ/sec^2$ are recommended for small saccade detection and $800^\circ/sec^2$ for reading and cognitive research. False saccade reports will result from lower acceleration thresholds. The EyeLink tracking system's acceleration data and thresholds sometimes may be higher than those reported for analog eye trackers. For noise reduction, these systems employ multi-pole filters, which add delay and smooth the data, significantly lowering the measured acceleration.

The saccadic motion threshold is used to keep a saccade from starting until the eye has moved significantly. Saccades can be reduced by a threshold of 0.1° to 0.2°. Larger values should be used with caution to eliminate short saccades: a threshold of 0.4°, for example, will always

merge fixations separated by 0.5° or less but may also eliminate some 1° saccades. However, it should be noted that for non-cognitive research or when statistics such as saccadic duration, amplitude, and average velocity are required, the threshold should be set to zero.

However, in the case of other type of eye-movements, like smooth pursuit and nystagmus, saccades must be detected against a background of smooth eye motion as fast as $70^{\circ}/sec$. While acceleration can be used to detect these saccades, velocity data must also be used to ensure that all saccades are detected reliably. During pursuit, the EyeLink 1000 parser raises the saccadic velocity threshold by the average velocity over the last 40 milliseconds. This is dependable and has no negative impact on parser performance during non-pursuit eye movements.

Tobii Eye-Tracking System

Tobii is the world's leading provider of eye-tracking solutions, with the most patents and a diverse product portfolio. That makes them a one-stop shop for eye-tracking tools, providing eye-tracking hardware, analysis software, and research consulting, [77]. Tobii's eye trackers combine high accuracy and precision with a wide tolerance for large head movements and a wide range of environments. As a result, they are suitable for a wide range of applications. The company, however, primarily focuses on three areas: scientific research, marketing and user research, and human performance.

In her 2012 publication [78], Anneli Olsen outlines the fundamental principles underlying an I-VT fixation filter and details its implementation in the Tobii I-VT Fixation Filter. Olsen also highlights the primary challenges encountered during the eye-tracking process, summarizing them into four categories: noise, missing data, eye selection, and sequences of fixations/saccades disrupted by very short saccades/fixations.

Addressing the first challenge, noise represents a pervasive issue in eye-tracking data collection. It can originate from design flaws within the system itself and can also stem from environmental factors that introduce influences and disturbances during the measurement process. For example, minor eye movements such as tremors and microsaccades can be seen as noise in eye tracking research, where the fixation is the eye movement of interest. I-VT filters calculate velocity, in their most basic form, by dividing the angle measured from the eye between two consecutive sample points by the sampling frequency. This works very well if the sampled gaze data is noise-free.

The higher the sampling frequency the smaller the eye movement between two samples at the same speed. This means that if the eye tracker makes even minor errors in the direction of the gaze or detects minor eye movements such as tremors or microsaccades, this will appear as significant noise in the velocity calculations when using data collected with a high-speed eye tracker, Figure 2.17.

However, there is another type of issue with lower-frequency eye trackers. Noise introduced by measurement issues will typically have the

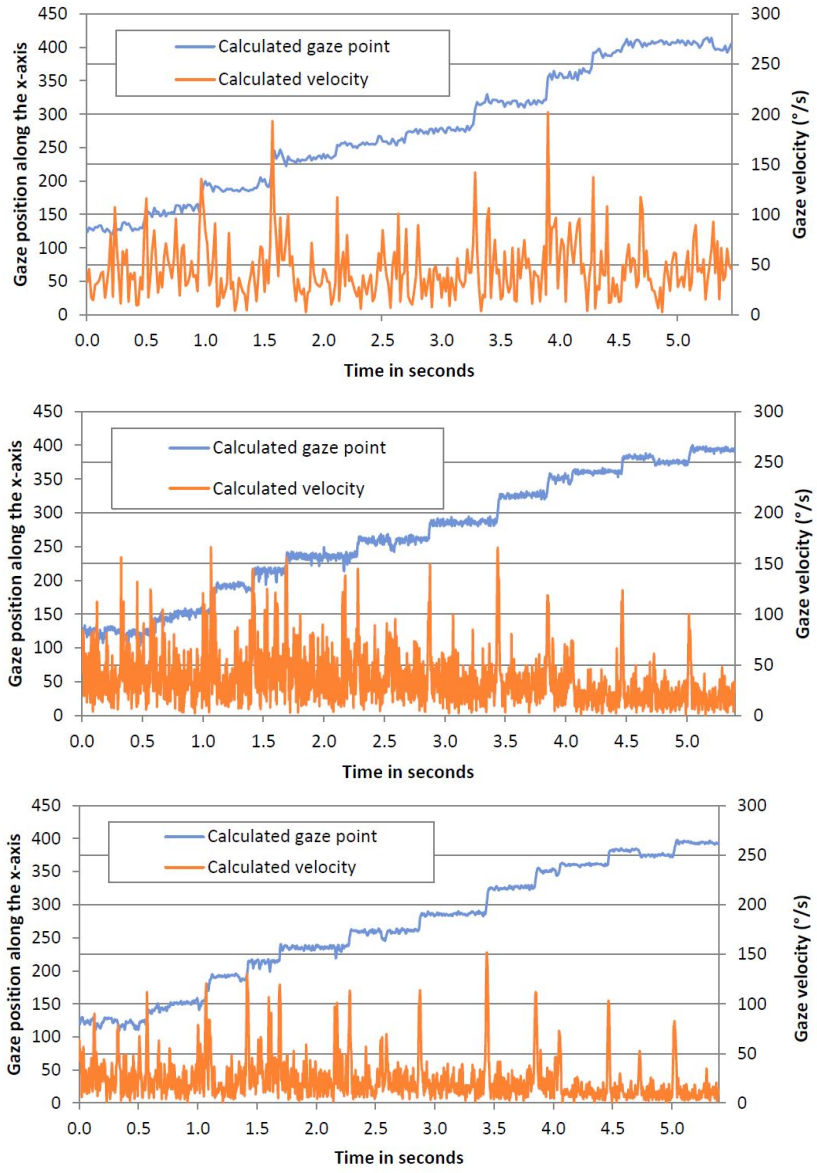

Figure 2.17: Charts showing noisy data sampled at 60Hz (top), 300Hz (middle) and after a moving average noise reduction algorithm has been applied (bottom). The calculated gaze points are along the x-axis. Taken from [78].

same amplitude in the gaze data as for high-frequency eye trackers, but filtering out that noise introduces the risk of severely modifying the actual real gaze data as the time between each sample increases.

Although noise in a high-frequency eye tracking data set can introduce amplified noise in the calculated velocity, this can often be removed using either a low pass filter which when applied smoothes out and reduces the noise spikes by removing high frequency signals, a noise reduction algorithm; however, with noisy low-frequency data, this must be done with extreme caution, Figure 2.17. In the Tobii software there are two different

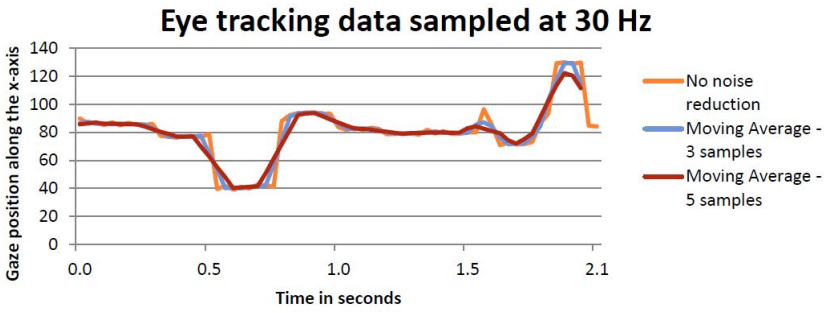

Figure 2.18: Chart showing raw eye tracking data sampled at 30Hz as well as data subjected to the Moving Average noise reduction function with a window size of 3 and 5 samples, taken from [78].

noise reduction functions: Moving average and Median.

In signal processing terms, the noise reduction algorithm implemented in the moving average noise reduction function is known as a non-weighted moving-average filter. This means that the signal's output for each sample is an average of a specified number of samples preceding and following the current sample. Because the previous and subsequent samples should have the same weight on the average, an equal number of previous and subsequent samples should be included in the average. Therefore, the average includes the current sample as well. This implies that the total number of samples included in such an averaging window must be an odd number. The 'Window size' parameter allows you to specify the size of the averaging window.

The smaller the window, the faster shifts between the sample being calssified as a saccade and a fixation will be retained, allowing the onset of a saccade or fixation to be recognized closer to the recording time. A smaller window permits more noise in the data; thus, false saccades and fixations may be spotted. Larger windows smooth data. After noise reduction, the steep fixation transitions in the raw data will become smoother. Saccade velocity will be much lower, so the velocity threshold should be modified. Figure 2.18 shows that as the "Window size" increases, saccade duration increases too which subsequently decreases the fixation duration. This means that saccades will appear longer and slower, making short fixations followed by prolonged saccades even shorter and perhaps rejected if their duration is less than the duration threshold parameter. The default value for the 'Minimum fixation duration' parameter is 60ms as this commonly used in the existing literature. Overall, all the above dictates the importance of a correct and careful selection of the "window size".

A median noise reduction algorithm, like the moving average algorithm, iterates a sliding window through the data stream, but the middle data point is replaced with coordinates that are the median values for the set of points in the window. When compared to a moving average algorithm, a median noise reduction algorithm produces less smoothed data while removing the most prominent noise, Figure 2.19.

The second challenge of eye-tracking data, is the data loss which is

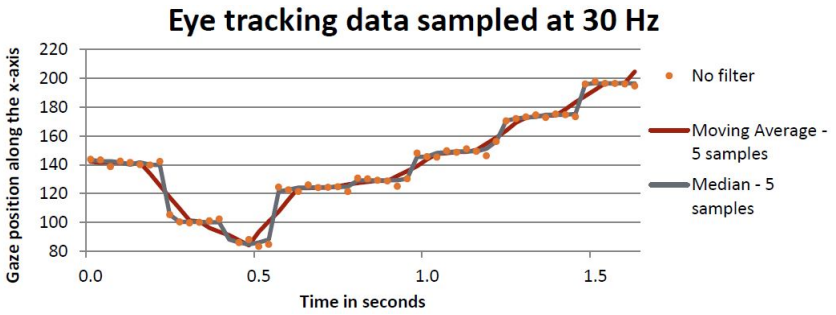

Figure 2.19: Chart showing raw eye tracking data sampled at 30Hz subjected to the Moving Average noise reduction function and the Median noise reduction filter. Both filters were set to use a window of 5 samples. Taken from [78].

frequently caused by the participant blinking, looking away, or putting something between the eye tracker and the eye, obscuring the eye tracker's view of the eyes, resulting in gaps of a hundred milliseconds or more. However, data loss can also occur due to other factors, such as delays in data transfers within hardware systems, temporary hardware malfunctions, time-out issues, temporary reflections in prescription glasses that make it impossible for the eye tracker to identify the eyes, and so on. In these cases, the data loss is typically much shorter than caused by the previous reasons. Suppose the data loss occurs in the middle of a fixation. In that case, the fixation classification algorithm may interpret the fixation as two separate fixations if valid data do not replace the lost data. As a result, a gap-filling algorithm is required.

The following three steps should be taken by this algorithm. To begin, a scaling factor is created. The scaling factor is calculated by dividing the timestamp of the sample being replaced by the total duration of the gap by the timestamp of the last valid sample before the gap. That would be as follows:

$$SF = \frac{\text{timestamp(sample to be replaced) - timestamp(1st sample after gap)}}{\text{timestamp(last sample prior gap) - timestamp(1st sample after gap)}}$$

The scaling factor is then multiplied by the position data from the first valid sample after the gap. Finally, the result is combined with the position data of the last valid sample before the gap, Figure 2.20.

It is critical to remember that it should not fill in blinks or other gaps caused by the participant or researcher, such as the participant turning away from the eye tracker or the eye tracker's view of the participant's eyes being obstructed. As a result, the value chosen should be shorter than a standard blink. 75ms is a value used by, for example, Komogortsev et al. [58], and it is also the value chosen by Tobii as the default value for this parameter. This value is also consistent with results from other eye-tracking studies.

Another aspect of gap fill-in interpolation is that it can either increase

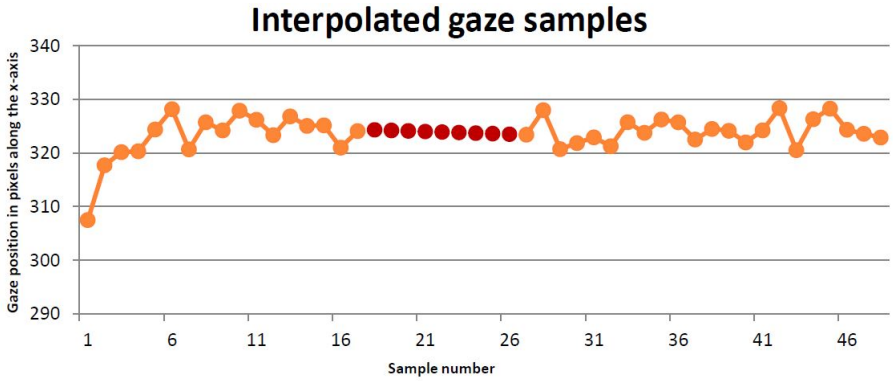

Figure 2.20: Missing or invalid samples have been replaced by interpolated samples. The interpolated samples are highlighted in red. Taken from [78].

or decrease the duration of fixations and saccades. Because of the linear interpolation, the measured speed between two samples within the gap will be the same. If that velocity is less than the 'Velocity threshold' parameter in the I-VT classification filter, all samples within the gap will be classified as fixation samples. The position data of the last valid sample before the gap and the first valid sample after the gap determine whether the samples in the gap are saccade or fixation. If the samples are close, they will be classified as fixation samples. Otherwise, the samples will be labeled as saccades.

The third challenge addresses the issue of the intrinsic differences in our eyes. Even though it is assumed that our eyes are identical, that is not the case, and it can happen that there is difference between one eye and the other when it comes to start time and end time of fixations as well as for blinks. For that reason, if the eye tracker is recording both eyes, there should be an algorithm that merges the data into a single stream of data which can be further classified.

Finally, the last challenge which is very crucial, is the assumption by the algorithms that the data are perfect. That results in situations that the algorithm classifies, what should have been a long fixation, into 2 short fixations inflitrated by a very short saccade in between. For that reason the algorithm should be "smart" enough in the post-processing of the data to identify and merge those fixations that are very close in time and very close in space. Same challenge is also when a long saccade is interupted and split into two sequences, by a very short fixation moment. In that case the algorithm should have a filter to remove data points or groups of data points that are labeled as fixations but last too short a time for the visual input to be registered by the brain.

Pygaze

The PyGaze toolbox is a free and open-source software package for the Python programming language. It is intended for creating eye-tracking experiments in Python syntax with as little effort as possible,

and it provides programming ease and script readability without limiting functionality and flexibility, [79].

PyGaze can be used to present visual and auditory stimuli, collect responses via keyboard, mouse, joystick, and other external hardware, and detect eye movements online using a custom algorithm. Eye trackers from various brands (EyeLink, SMI, and Tobii systems) are supported. PyGaze's uniqueness lies in its provision of an easy-to-use layer on top of the numerous software libraries required for implementing eye-tracking experiments. PyGaze is a software bridge for eye-tracking research.

The creators of PyGaze in their publication, [80], mentioned that the algorithm for online saccade detection is similar to the Kliegl [57] algorithm for (micro)saccade detection in that it identifies saccades by calculating eye movement velocity across multiple samples. However, because the current algorithm was designed for online saccade detection, events should be detected as soon as possible. As a result, eye movement velocity is calculated using the fewest number of samples possible, which is two: the most recent and the previous sample.

The Kliegl algorithm, on the other hand, "looks ahead" to two samples and uses a total of five samples. Furthermore, the current algorithm uses eye movement acceleration as a saccade indicator. Because the acceleration in the previous sample is based on the speed in that sample and the sample before it, the sample window for calculating acceleration is actually 3. Because this algorithm was designed for speed and responsiveness, it is less reliable than more advanced algorithms for offline event detection. As a result, researchers are advised to analyze the raw eye-movement data using an offline event detection algorithm, such as those described by Engbert and Kliegl [57] or Nyström and Holmqvist [61], as one would typically do when analyzing eye-movement data.

2.5 Detecting ADHD through Eye-Tracking Data Analysis: Background and Latest Improvements

Attention-Deficit/Hyperactivity Disorder (ADHD) stands as one of the most common neuro-developmental disorders affecting children. It is distinguished by symptoms such as inattention, impulsivity, and hyperactivity, which manifest across different environments and significantly impact daily functioning [42].

High rates of comorbidity with other disorders and high intra-disorder heterogeneity make it difficult for clinicians to diagnose ADHD and predict treatment responses, [81]. ADHD is typically diagnosed based on behavioral observations and symptoms reported by parents, teachers, and the child himself or herself. Unfortunately, all of these evaluations are characterized by a high degree of subjectivity, which carries the risk of over- or under-diagnosis and the possibility of inappropriate interventions.

Efforts have been made to develop more objective and dependable diagnostic assessment tools, but results remain inconsistent. Efforts are ongoing to identify stratification markers that may one day aid in the early

diagnosis of ADHD or in predicting prognosis and treatment response, [82].

The use of eye movements, as measured by eye-tracking systems, has garnered increasing focus in recent years. In addition, oculomotor studies have been utilized to investigate aspects of both motor and cognitive control in ADHD, as well as the potential value of oculomotor markers in predicting treatment response.

The significant advances in non-invasive, high-precision video-oculography methods during the last twenty years have greatly facilitated the assessment of oculomotor control in ADHD, especially in child populations. In addition, they have provided novel insights into the neurophysiology and neurocognitive profiles of ADHD that traditional neuropsychological assessments or tasks may not capture, [83].

Using a variety of behavioral tasks that assess different aspects of cognitive control and studying the range of ocular motor responses elicited by these tasks can help in comprehending the heterogeneity of hyperactive, impulsive, and inattentive behaviors among individuals with ADHD.

Maron D. et. al. [83] in their systematic review identified 275 unique studies which they are combining eye-movement terminology with ADHD. However, after filtering through several eligibility criteria of theirs, such as the study to include a sample of individuals with a formal diagnosis of ADHD and the oculomotor metrics to be measured using simple reflexive, volitional or pursuit eye movement tasks, along with an explicit study of the saccades, fixation or pursuit the number of studies was reduced down to only 27.

More than 70% of the above studies were conducted prior to 2010. Their participants spread from children and young adults, up to people in their late 30s. On the other hand, Levantini V. et. al. [42] are focusing on their paper, to summarize the findings of eye-tracking studies which carried out on youths only.

Combining these two very interesting reviews we provide in the following paragraphs a chronological summary of what has been done till today in the field of eye-tracking for youths with ADHD, and what constitutes the state-of-the-art methodologies.

2.5.1 Historical Review

Ross in 1994, [84] evaluated a group of 23 children (13 with ADHD and 10 the control group) on their reaction in an oculomotor delayed response task. In a task like that the subject is cued as to where he/she should look but must delay a short period and then shift gaze to the location where the cue previously existed but no longer exists. Children with ADHD showed, relative to normal controls, deficits on inhibiting response during the delay period. As presented in the Figure 2.21, the child with ADHD exhibited a saccade during the delay period (a premature saccade). However, no differences in latency (preparation of motor response) or accuracy of visuospatial memory were detected.

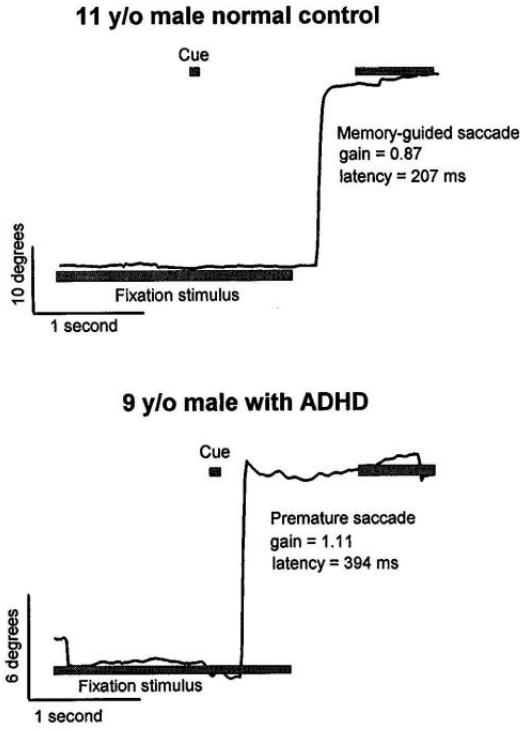

Figure 2.21: Delayed oculomotor response task, taken from [84].

To arrive to these conclusions they recorded raw data consisting of eye position and target position for each millisecond of tracking and then divided the eye movements into discrete segments. The saccades were identified on the basis of peak velocity greater than 30°/second, an initial acceleration greater than $2000^\circ/sec^2$, and a minimum duration (9 msec). Finally the saccades were further classified based on time relative to cue activiry, Figure 2.21. Those saccades that occurred from 80 msec after cue onset to 80 msec after fixation point extinction, i.e., during the time when the subject was supposed to maintain gaze on the fixation stimulus were classified as premature saccades.

In 2000, Castellanos F. X. et. al. [85], tested fifty two girls (32 with ADHD and 20 age-matched, normal control) on a variety of oculomotor tasks requiring attention, working memory, and response inhibition, which included smooth pursuit, delayed response, and go/no-go tasks.

The results showed that the group of the girls with ADHD performed the delayed response task with significantly lower scores than the control group, confirming the fact that ADHD impairs the executive function. The setup was similar with the experiment performed by Ross [84], where a cue appeared and disappeared, and the evaluation was on the basis if the premature saccade occured, middle pannel at the Figure 2.22. In such cases the trial was considered as failed. Also in the case there was no memory guided saccade at all, third panel in the Figure 2.22 it was also scored as an error. On the other hand, at the smooth pursuit experiment, the performance was equivalent across the two groups.

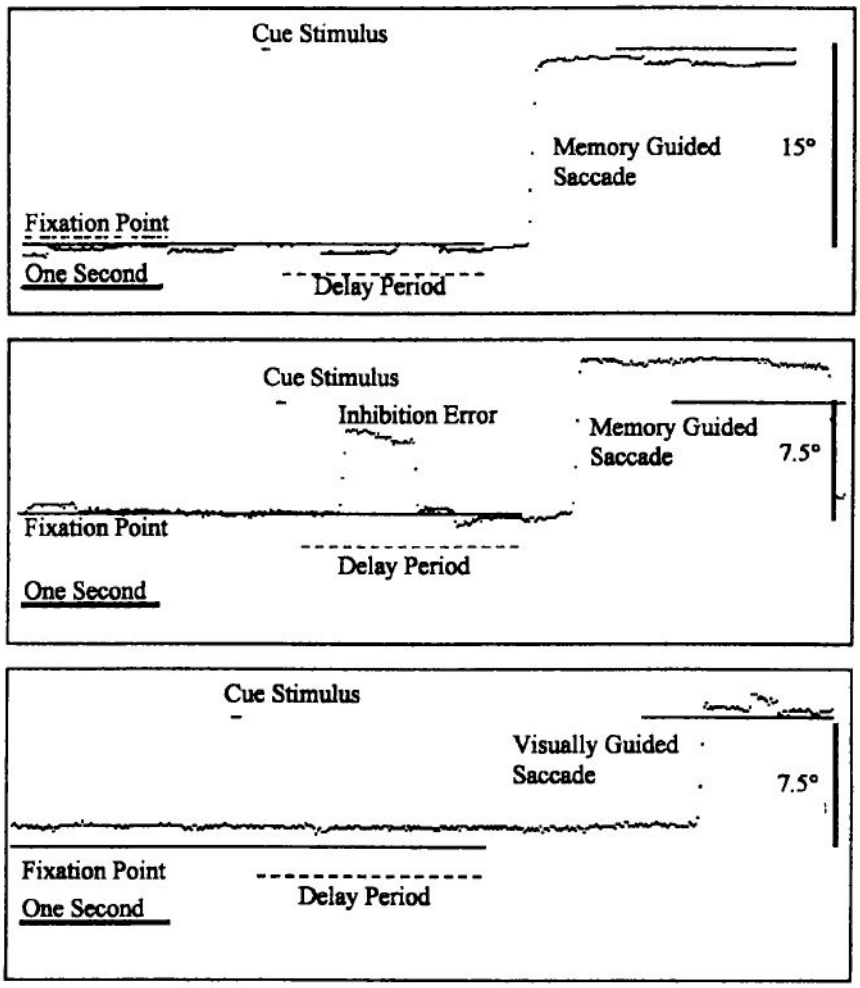

Figure 2.22: Representative tracing from delayed response task, taken from [85].

Gould et. al. [86] in their study, in 2001, compared 53 children with ADHD to 44 healthy control children on a 21-second fixation task. They found out that children with ADHD made significantly more intrusive saccades during fixation than did the control subjects. These data dictate that children with ADHD not only have difficulty on inhibiting response but also on maintaining fixation in the absence of any external or internal distractor, possibly reflecting an intrinsic neurological dysfunction.

In this experiment the task was that the subject would fixate at a dot showing in the center for about 30 sec, and then as the dot will bounce back and forth across the screen the subject would follow it with his eyes. Eye position data automatically clasiffied to saccades if they were meeting the follwoing criteria: peak velocities greater than 25°/sec, an initial acceleration of greater than $1500°/second^2$, and a minimum duration of greater than 8 msec.

However, Gillian O'Driscoll et. al. [87] focused not only on the differences between the ADHD and the control group, but also

within ADHD group itself. They had two different ADHD groups, covering two types, the ADHD-inattentive, ADHD-I, which has only inattention symptoms, and the ADHD-combined, ADHD-C, which has both inattention and hyperactivity/ impulsivity symptoms. Both ADHD groups and the control group were evaluated on four different tasks, Saccade Control Task, Motor Planning, Response Inhibition (Antisaccade Task), and Task Switching (Saccade–Antisaccade Mixed Task).

Throughout their results they found that ADHD-I and ADHD-C subtypes differed in measures of executive function, with ADHD-C but not ADHD-I showing deficits in motor planning and response inhibition, while neither of these two subtypes was impaired in task switching. With these results it is obvious how much difficult it can be to classify ADHD diagnosed subjects, when even on a subtype level there are such significant differentations.

The classification of the eye movement data was carried out with a semi-automated custom analysis software package by SR Research. However, the data were also visually inspected. In terms of the criteria the saccades were identified with velocity greater than 22°/sec and acceleration greater than $4000°/second^2$ criteria.

In 2006, Hanisch et. al. [88] evaluated 44 children, 22 with ADHD and 22 as a control group, on a series of different tasks covering fixation task, prosaccade task, countermanding task, and antisaccade task. Throughout their experiment, the ADHD children showed a lower number of successfully inhibited saccades than controls, while they also took longer to stop a saccade into the right than the left hemi-field. They concluded that there is a specific oculomotor inhibition deficit for initiated and ongoing responses in children with ADHD, which confirms the observation of an ADHD-related impairment in cognitive inhibition that has been associated with prefrontal lobe functions.

In 2008, Rommelse et. al. [89] assesed in their paper the visuo-spatial working memory and inhibition in children with ADHD by recording saccades during a memory-guided saccade task with two types of delays. The novel thing in their study was that apart from the ADHD group and the control group, they introduced a third group consisted of the non-affected brothers of the ADHD group.

The experiment setup was as follows: the participants were instructed to fixate at a central fixation point (a light gray dot on a black background) until it disappeared and then to move their eyes to the memorized location. After 1200 msec, a light-gray circle was presented for 50 msec, positioned on one of four corner locations of an imaginary square. Each target location was equally probable. This circle was indicating the location to which a saccade had to be made after a variable delay (the memory location).The delay was either 3000 or 7000 msec and it was signaled by the removal of the central fixation point. The sequence of 40 trials was counterbalanced and randomized for each participant.

The task variables to assess the visuo-spatial working memory were the accuracy of saccades toward the memorized location of the target, the latency of the correctly performed memory saccades, the percentage of

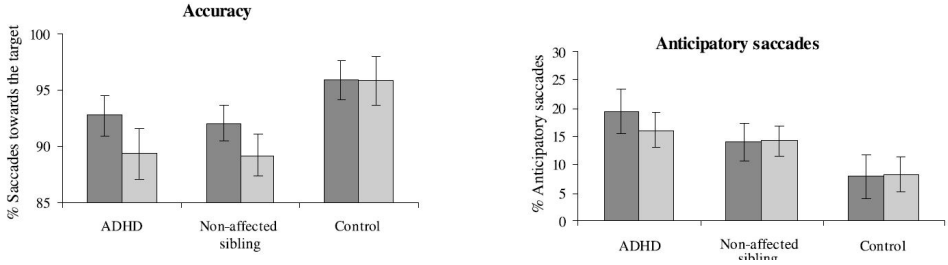

Figure 2.23: Visuo-spatial working memory: The means are adjusted for the covariate age and error bars represent 1 standard error. left bar; 3 second delay, right bar; 7 second delay, taken from [89].

anticipatory saccades, i.e. saccades towards the target during the delay period, the percentage of intrusive saccades, i.e. saccades directed towards another location than the target during the delay period, the peak velocity, the tendency to under- versus overshoot saccades and the duration.

The criteria for an eye movement to be classified as a saccade were either when the movement velocity exceeded 35°/sec or when the movement acceleration exceeded $9500°/second^2$.

Their findings showed that the memory-guided saccade deficits may relate to a familial predisposition for ADHD, especially regarding the variables of accuracy of visuospatial working memory and anticipatory saccades, Figure 2.23.

A year later, two more studies were published in the Journal of American Academy of Child and Adolescent Psychiatry. The first one was by Loe et. al. [90], which had as a goal to evaluate the differences in cognitive control in children with ADHD compared with controls using oculomotor tests of executive function. Their test covered the following three main tasks; The fixation, where the subject had to look straight ahead despite the appearance of peripheral distractors, and was assessing its ability to inhibit a response toward the distractors; the antisaccade which was requiring the subject to inhibit a reflexive response to a suddenly appearing stimulus and instead, to look to the mirror location and it was assessing the voluntary response suppression or response inhibition which consitutes a core component of executive function; and finally the memory-guided saccade task which was requiring the subject to look to the location of a previously presented visual target and it was assessing the maintenance of spatial working memory. The raw data that were recorded from thei eye tracker were classified into saccades based on a velocity algorithm using a 30°/s criterion.

Their findings showed that, the inability of children with ADHD to complete tasks or assignments was also confirmed during the antisaccade task trials which was lasting only for a few seconds, as it is showen in the Figure 2.24. Children with ADHD made the same proportion of errors on the antisaccade task regardless of the duration of fixation, while the performance of the control group was improving as the fixation time was increasing.

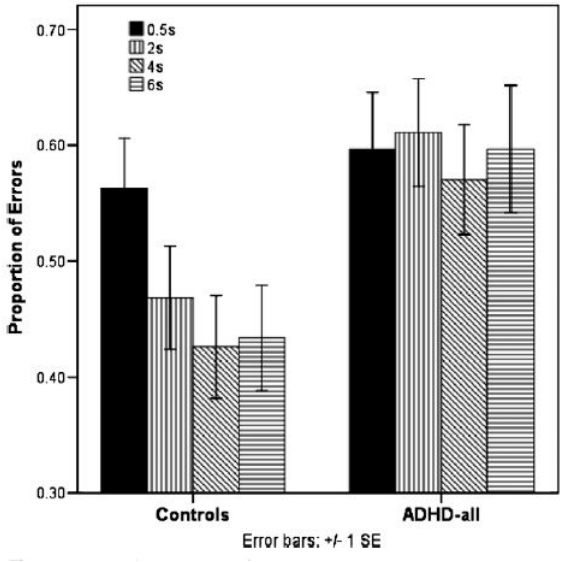

Figure 2.24: Antisaccade proportion of errors, taken from [90].

The second study was carried out by Mahone M. et. al. [91], and its objective was to examine patterns of executive and oculomotor control in a group of children with ADHD. In an experiment including 120 children, 60 with ADHD and 60 as a control group, going through different types of tasks, visually guided saccades (VGS), antisaccades, memory-guided saccades, and a go/no-go test, they concluded that the children with ADHD demonstrated significant deficits in oculomotor response preparation (VGS latency and variability) and response inhibition but not on working memory. In this case the eye movements were recorded by using bitemporal electrooculography, while the saccade onset was defined when the eye velocity was reaching 25°/s.

In 2010, Carr, Henderson and Nigg [92] addressed in their paper the question whether ADHD is related to early- or late-stage attentional control mechanisms and whether this differentiates a nonhyperactive subtype (ADD). Similarly to O'Driscoll's, [87], it was again confirmed once again that ADHD does not only differentiate from the control group, but also within different subtypes itself. In this case it was presented that the ADHD-combined showed greater weakness in response inhibition compared to the nonhyperactive subtype (ADD).

At the same year, Karatekin C. [93] with his experiment to ADHD children, challenged the theories and the models that posit inhibition as a core cognitive impairment. While evaluating not only children with ADHD but also with psychosis, he concluded that the ADHD group did not have elevated antisaccade error rates, suggesting that improving topdown control of attention over time should be a target of cognitive remediation efforts in ADHD examinations. However, this was in contradiction of what Goto Y. et. al. [94] suggested, as according to their analysis the ADHD group showed significantly higher percentage of anticipatory errors in memory-guided saccade task and percentage of direction errors rates in

antisaccade task than the control group.

In 2012 Pishyareh E. et. al. [95], conducted a study to evaluate if children with ADHD react explosively and inappropriately to emotional stimuli, due to some impairment in attending to emotional cues. They compared two groups of 30 children each, one group with ADHD and one control group. The participants were presented with pairs of emotional and neutral scenes selected from the International Affective Picture System - IAPS. Their findings showed that children with ADHD spent less time on pleasant pictures than the normal group and more time on unpleasant pictures compared to the normal group, which leads to their emotional reactivity. It could be interpreted that children with ADHD since they do not spent much time to look at pleasant pictures, they may not recognize pleasant pictures properly compared to normal children.

In this case the methodology was not based on velocity algorithms, and evaluation of saccades, like we have seen in the majority of the previous studies, but on the area of interest approach. The stimuli was presented on a 19 inch monitor with a resolution of 1400 by 900 pixels and if the eye position was remaining within a 50 pixel area for more than 100ms was considered as fixation. The variables were the mean number and the duration of first fixation on each emotion as well as the duration of first gaze which was the time taken to fixate on one picture before moving to another picture.

2.5.2 Recent Developments

In more recent developments, we have the study in 2015 by Matsuo et. al. [96] which had as main goal to examine whether children with ADHD exhibit abnormalities during a visually guided pro-saccadic eye-movement and to clarify the neurophysiological mechanisms associated with their behavioral impairments. They evaluated 125 children out of which thirty seven had ADHD and eighty eight were in the control group, and they concluded that the ADHD group had a significantly longer reaction time than the control group.

The experimental setup was requiring the children to maintain their eyes on the fixation point and to move their eyes toward the peripheral stimulus in response to its appearance as quickly and accurately as possible. Two different conditions were probable to occur. In the first, the step condition, the peripheral stimulus appeared at the same time as the fixation point disappeared, while in second, the gap condition, the peripheral stimulus appeared 200 ms after the fixation point disappeared.

The study team concluded that also the gap effect, which is the difference in the reaction time between the two conditions was markedly attenuated for the ADHD children, especially in some ages, Figure 2.25.

The eye movement data were collected by the eye tracker and they were classified into saccades based on a radial eye velocity criteria threshold of 30°/s. The detected saccades were further analyzed if their amplitude is greater than 2°, their peak velocity greater than 50°/s, and their duration greater than 20 ms.

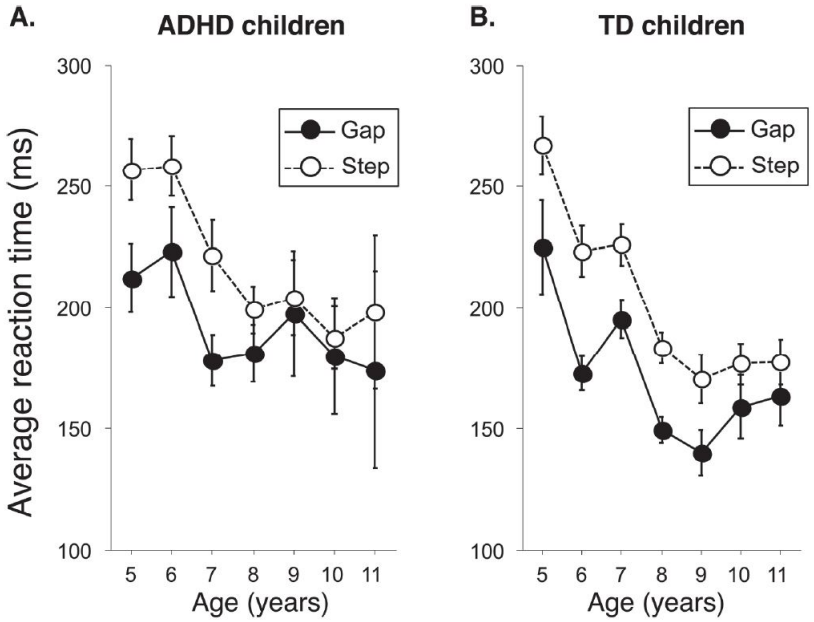

Figure 2.25: Saccade reaction time during gap and step trials, taken from [96].

A year later, Turkan et. al. [97], performed a study on 48 children, half of which had ADHD and the other half was the control group, in order to evaluate change detection performance and visual search patterns between the two groups. In order to evaluate the change detection performance they used the Flicker paradigm, [98], where an original image is alternated with its modified image and a blank field is placed between successive images. This change cycle repeats until either the subject responds or an elapse of 60s.

They concluded that the change detection performance, associated with visual attention and memory, was worse in children with ADHD, while regarding the eye movement analysis, the control group fixated mostly on the changed area whereas ADHD children fixated on the whole scene. However, the fixation duration of the ADHD group was lower than of the control group, Figure 2.26. However, no notable differences were observed in terms of reaction time.

In 2017, two studies were published regarding the identification of ADHD children with eye-tracking methodologies. The first was from Bucci et. al. [99] which evaluated the difference in the elicited saccades from different paradigms on children with and without ADHD.

They evaluated 62 children half of which had ADHD, in three paradigms which were used to stimulate horizontal visually guided saccades (gap, step, and overlap paradigms) and one for antisaccades, Figure 2.27. The classification of the data was done automatically by the software of the eye tracking system. Also a fixation paradigm was performed, where the children had to fixate on a target for 30s.

In this study, the researchers concluded that children with ADHD

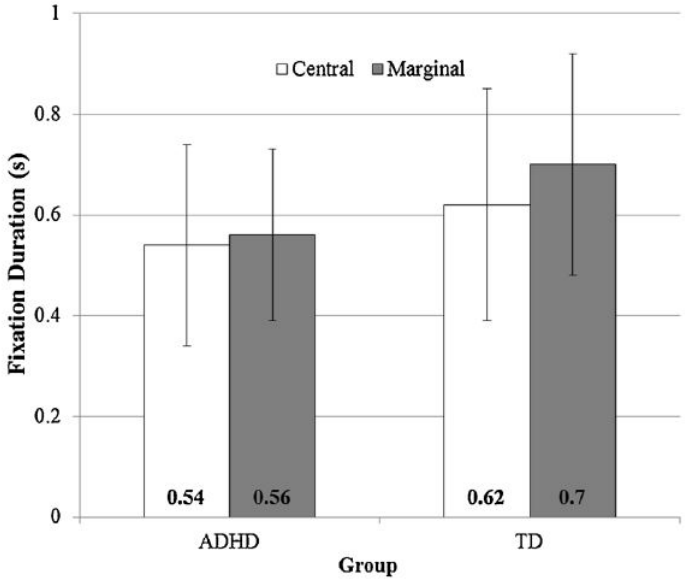

Figure 2.26: Mean fixation duration on the changed area, taken from [97].

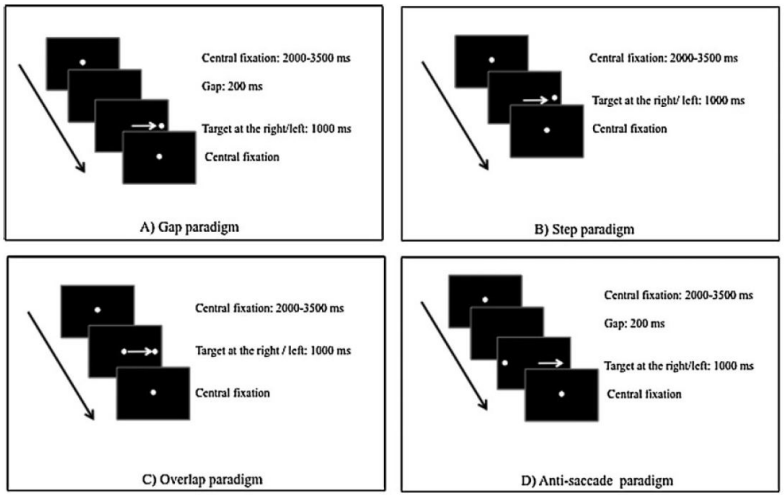

Figure 2.27: Schematic diagram of the temporal and spatial arrangement used in different paradigms, taken from [99].

exhibited a general decline in oculomotor ability, as observed through the examination of saccades, antisaccades, and fixation paradigms. The ADHD group demonstrated a higher number of saccades during the fixation paradigm and a higher error rate during the antisaccade paradigm. These findings indicated a deficiency in inhibitory control and executive resources, particularly attention, dedicated to the task. Refer to Figure 2.28 for a graphical representation of the performance.

The second study that was published the same year, was by Wainstein et.al. [100], and evaluated how the pupil size diameter can be utilized as a biological marker for ADHD detection among children. This study was

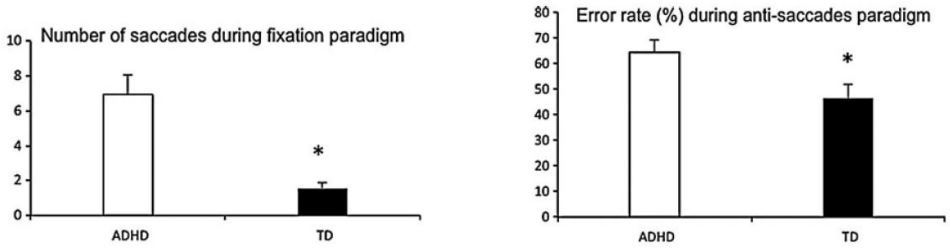

Figure 2.28: Means of error rate during antisaccade paradigm (left graph), of number of saccades during fixation paradigm (right graph) in ADHD and TD (control) children. Vertical bars indicate the standard error. *p < 0.05. ADHD. Taken from [99].

based on the same dataset that we will utilize in this thesis. However, the analysis that was carried out by Wainstein et. al. was focused mainly on the impact of the ADHD on the pupil diameter of the subject. They monitored pupil size from an ADHD group and a control group, during a visuospatial working memory task, and concluded that pupil size is correlated with the subjects' performance and reaction time variability, which both are indicators of attention.

The subjects had to follow a sequence of presentation of different dot arrays, followed by a distractor, while in the end of the trial they had to answer if a dot array that was presented to them it was the same with one of those presented in the beginning of the trial. The pupil diameter data were gathered by an eye-tracker and then normalized by means of a z-score, separately for each trial. Apart from the control group and the ADHD group, a subset of the ADHD group was re-evaluated once again but under medication.

Figure 2.29 illustrates the connection between changes in pupil diameter and behavioral performance markers in ADHD, highlighting the contrasting associations observed between ADHD and non-ADHD children. These findings indicate that alterations in pupil diameter during a visual-spatial working memory task hold potential as a valuable biological marker for ADHD.

Similar to the work already presented by Pishyareh E. et. al. [95], who evaluated the correlation between the emotional stimuli and the ADHD, in 2018 Serrano et. al. [101] compared in their study the viewing patterns for emotion stimuli between children with ADHD and a control group and examined the relationship of the ADHD symptoms, with viewing patterns, emotion knowledge accuracy and response time. The study involved 45 children, out of which 26 with ADHD.

The eye tracking system was collecting the data while the subjects ware going through the task trial. Face and situation images were displayed on a computer screen for 4 and 7 seconds respectively, and then were replaced by a screen showing the response options. In the whole task trial process a total of 38 facial expression and 20 situation images were shown. A fixation was defined as a relatively stable eye position (within 6 degrees

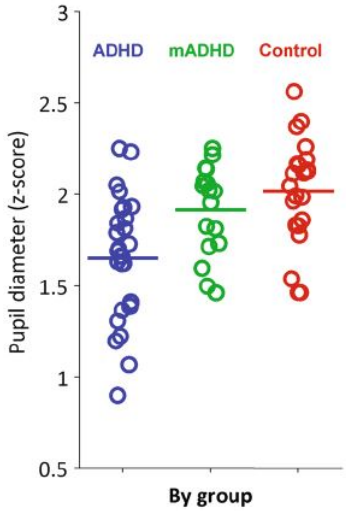

Figure 2.29: Maximum pupil diameter. Each dot represents the mean session value for a subject. Horizontal bars correspond to session averages across subjects. Taken from [100].

horizontally and 4 degrees vertically) for a minimum of 100 ms.

In this study the protocol was based on the Areas of Interest. The eye-tracking dependent measure was the proportion of total fixation duration on all areas of interests, i.e. the total AOI fixation duration divided by the total duration of all fixations on each image.

Even though the viewing patterns of children with and without ADHD were fairly similar, the children with ADHD spent less time viewing relevant areas of images and took longer to respond, i.e., detect an emotion in comparison to the children without ADHD, Figure 2.30.

One year later, Cladani et. al. [102] published their study in which they explored the corellation between the oculomotor behavior and executive motor control with ADHD through an oculomotor and postural dual task. They evaluated 42 children out of which half of them had ADHD.

They evaluated the children in two visual paradigms; a simple fixation case, where the children had to to fixate on the target appearing in the center of the black screen for 30 s, while the second paradigm was a fixation with distractors where the child had to maintain fixation on the central target and to inhibit saccades toward the distractors, a white smile target appearing for a random duration from 500 to 2000 ms, Figure 2.31.

The novelty of this experiment was that it was combined with three postural conditions for the children; a simple sitting condition, a complex standing on stable platform condition and a complex standing on unstable platform condition.

The analysis of the raw data was done automatically by the built-in saccade algorithm of the MeyeAnalysis software, and all the saccades equal or greater than 2 degrees were counted.

The authors concluded that the children with ADHD had poor fixation capabilities compared to the control group, while they also showed

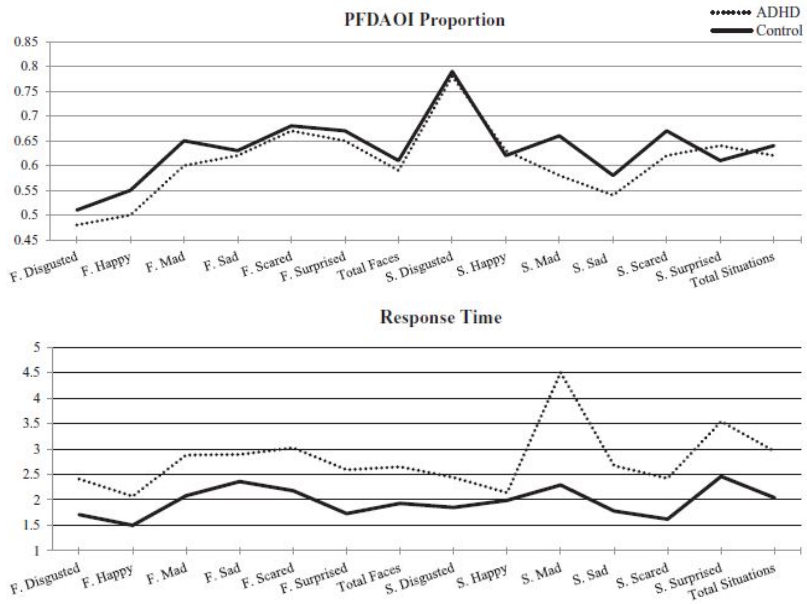

Figure 2.30: Profile plots for proportion of fixation duration on Area of Interst, and response time for children with and without ADHD. Taken from [101].

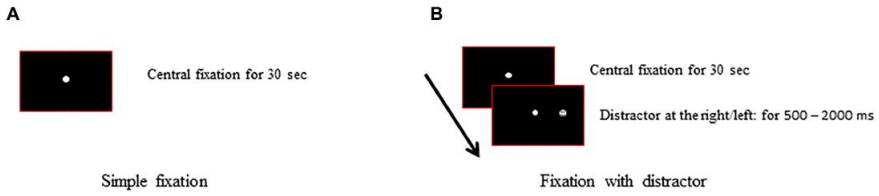

Figure 2.31: Temporal arrangement of simple fixation (A) and of fixation with distractor (B) task, taken from [102].

more difficulties in the simple fixation visual task. Even though all the children had more difficulty standing on the unstable platform, those with ADHD showed poor postural stability also in the other cases too. Finally the number of sccades performed by the children was increased for both groups while standing on the unstable platform than in the sitting condition.

Finally, the last two most recent publications were carried out by Fernandez-Ruiz et. al. [103], and Huang et. al. [104], both published in 2019.

In the first study, Fernandez-Ruiz and his team, evaluated 42 children out of which the twenty two were diagnosed with ADHD. The task they had to undergo was a prosaccade/antisaccade task. The trial was comprised with a preparatory stage that cued a prosaccade or an antisaccade without the presentation of a peripheral target. This allowed testing inhibitory control without the confounding activation from an actual response. The study concluded that the adhd group showed longer

reaction time compared to the control group. Also they had more direction errors during the antisaccade task. These findings align with previous studies in which the oculomotor inhibitory control was evaluated.

On the other hand, Huang et. al. [104], examined in their study 24 children out of which had ADHD. These children were evaluated in a visual guided saccade and in anti-saccade task. The visual guided saccade requires the capability of orientation of attention and sequential procedure, while the anti-saccade task could evaluates the response inhibition. In the first, latency and accuracy values were collected, while in the latter the percentage of direction errors was also recorded. The recording and classification of the data was performed by the built-in software of the eye-tracker.

More specifically, the visual guided saccade task was starting with a fixation point at the center of the screen, which was disappearing after 1000 or 1500 ms, randomly set. Then a gray screen was presented followed by a visual target which was appearing in the horizontal direction left or right position from the center of fixation point, either 7° or 15° from the center of the fixation point, as presented in the Figure 2.32.

In the anti-saccade task, the participants were instructed not to look at the target but to look at a position in the opposite direction at a similar distance from the target. Once the visual target disappeared, a black dot appeared in the correct position where the participants were expected to look, providing in this way a response feedback to the participants.

Their main findings in this study were the following. Firstly, during the visual guided saccade task, that the latency time of the children with ADHD and the control group in the 7.5° target showed statistically significant differences. However, in the case of the 15° target, there was no significant difference, indicating that small saccade amplitude tasks could induce some difficulty in regulating processes of saccade initiation of children with ADHD. Moreover, the ADHD children had a significantly larger ratio of saccade amplitude over target eccentricity compared to the

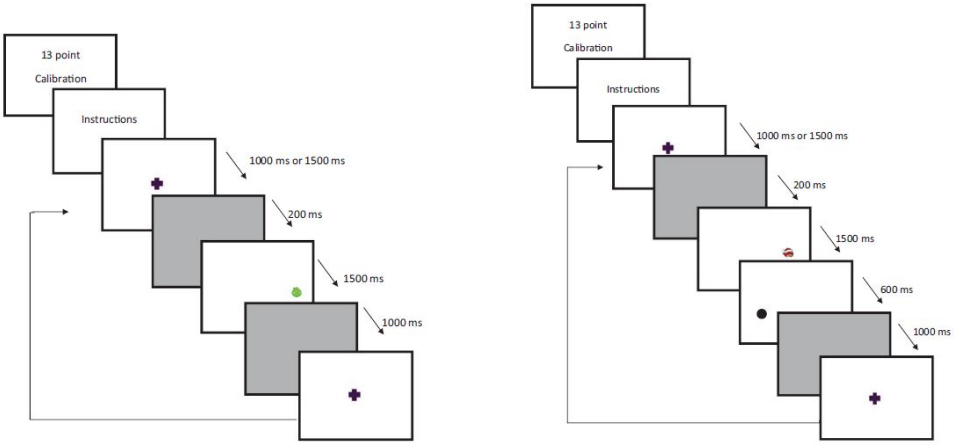

Figure 2.32: The procedure of the visual guided saccades task (left image) and anti-saccade task (right image), taken from [104].

control group, presenting a significantly less saccade precision.

During the anti-saccade task, different saccade amplitude targets may influence saccade latency and accuracy values. The study concluded that short latency in children with ADHD on anti-saccade tasks indicates difficulty inhibiting prepotent response, resulting in faster but less accurate responses. This finding also implies that ADHD children were unable to inhibit their initial prepotent responses to an external event, resulting in stimulus-driven behavior. This dysfunction may contribute to ADHD children's impulsive and hyperactive behavior, which is aligned with the higher percentage of direction errors of the ADHD children in the anti-saccade tasks that was showed in this study.

In conclusion the ADHD children presented weaknesses in the recruitment of visual attention resources and limited saccade precision, as well as deficit of inhibition and impulsive control, compared to the control group.

2.6 Eye-Tracking Criteria to Differentiate ADHD and Non-ADHD Subjects

In this chapter, we synthesize the findings from the literature review presented earlier and emphasize the key criteria for differentiating ADHD and non-ADHD children using eye-tracking data. Based on the reviewed literature, a comprehensive list of potential criteria is compiled and presented in Table 2.4.

Although each criterion in the table has demonstrated potential utility in distinguishing between ADHD and non-ADHD children, our analysis primarily focuses on measures that can be derived from the x, y coordinates and timestamps, which are the primary data utilized in this thesis.

Based on the focus of this thesis, the following criteria have been identified as most relevant and will be further evaluated for their applicability:

- Saccade Frequency. Saccade frequency, defined as the number of saccades per unit of time, has been found in multiple studies to

Table 2.4: List of eye-tracking criteria that can be used to differentiate children with ADHD.

No.	Eye-Tracking Criteria
1	Fixation duration
2	Saccade frequency
3	Saccade amplitude
4	Interference saccade frequency
5	Smooth pursuit accuracy
6	Blink rate
7	Pupil dilation during cognitive tasks
8	Visual search efficiency
9	Anticipatory saccade frequency
10	Microsaccade frequency

be higher in children with ADHD compared to their non-ADHD counterparts.

- Fixation Duration. Fixation duration, representing the duration of time during which the eyes remain stationary in a fixation, has been observed to be shorter in children with ADHD compared to non-ADHD children in various studies.

These quantitative measures have demonstrated the potential to differentiate between ADHD and non-ADHD children through the analysis of eye-tracking data. It is important to note, however, that the indicative values derived from previous studies may vary depending on the specific eye-tracking tasks employed and the sample size and characteristics of the participants. Consequently, these criteria may be more effective for pattern recognition rather than absolute basis comparison.

In the subsequent chapter, we will introduce an additional metric, derived from the novel IVT Optimization algorithm, 3.4.1, which will complement the aforementioned eye-tracking criteria in distinguishing between ADHD and non-ADHD subjects.

Chapter 3

Methodology

In the previous chapter, a comprehensive literature review was conducted, examining various eye-tracking algorithms, state-of-the-art technologies, and their corresponding software. It was found that the majority of studies primarily relied on the embedded algorithms within the eye-tracking commercial software for data analysis, with only a few cases utilizing more in-depth mathematical analysis.

This thesis aims to address this issue by establishing a stronger connection between the test data and the mathematical models governing the eye-tracking system. The ultimate goal is to gain a deeper understanding of the correlations between ADHD status, performance test results, and eye-tracking data.

In the upcoming sections, we will explore the dataset utilized in this thesis, the experimental setup on which it was based, and the proposed methodology we will employ for our analysis. This will include a discussion of the algorithms and mathematical models we will use to better integrate and analyze the data, with a focus on uncovering novel insights into the relationship between ADHD and eye movement patterns.

3.1 Dataset Description

The dataset was generated from an experiment reported in [14] on visuospatial working memory conducted on participants with and without ADHD. Before examining the dataset in detail, it is essential to understand the experimental design and its framework.

The participants completed a Sternberg-type delayed visuospatial working memory (WM) task, which was adapted from Dolcos & McCarthy's study, [14]. The stimuli used in the task were 1- or 2-dot arrays, with the dots positioned in any of the sixteen locations within a 4×4 grid, as illustrated in Figure 3.1. In each trial, the participants were instructed to fixate on a black cross at the center of the screen and, after 500ms, a dot array presentation (load) began. Three different dot arrays were presented in succession, with delay periods in between each presentation.

Following the final delay period, a distractor image was displayed for 500ms, after which a 'probe' dot was presented for 1.5 seconds. The probe

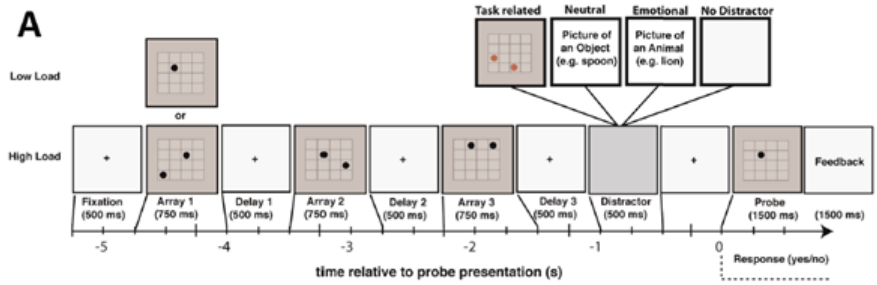

Figure 3.1: Schematic representation of a single trial, taken from [100].

dot was located within the grid, and the participants had to respond 'yes' if the probe dot had appeared in any of the previous arrays in the trial, or 'no' if it had not appeared, [14]. A total of 160 trials were presented on each session, separated in 8 blocks of 20 trials. Sessions usually lasted 30 minutes.

Based on the aforementioned experiment, Rojas-Líbano D. et. al. [14] compiled a dataset that contained various pieces of information regarding the participants and their performance. The original format of the dataset was structured in MATLAB arrays, as shown in the Figure 3.2, which we needed to analyze and filter in order to generate specific (.csv) files that could be used for our project, which was conducted using Python.

More specifically, the 'Task data' field in the dataset contains a table array that includes all the raw data from each participant and session. The data are organized into four columns, with one row for each timestamp. The columns are as follows: Time, Diameter, Position, and Events. The Time column offers precise timestamps of each session in milliseconds, sourced from the Eyelink recording. The Diameter column contains

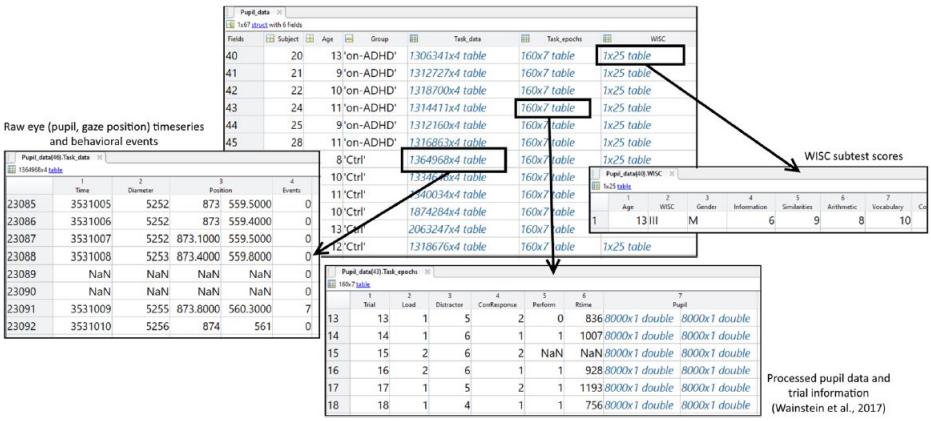

Figure 3.2: Screenshot samples of data structure Pupil data within the Matlab® environment. The file contains a structure array with six fields ('Subject', 'Age', 'Group', 'Task data', 'Task epochs', and 'WISC'), taken from [14].

Table 3.1: List of Events in the Dataset.

Event	Description
1	Presentation of one-dot memoranda
2	Presentation of two-dot memoranda
3	Presentation of neutral distractor
4	Presentation of empty-grid distractor
5	Presentation of task-related distractor
6	Presentation of emotional-image distractor
7	Presentation of probe array
8	Presentation of feedback screen
9	Occurrence of button press with participant answering 'yes'
10	Occurrence of button press with participant answering 'no'

the participant's pupil diameter in arbitrary units, also recorded by the Eyelink equipment. The Position column is divided into two sub-columns representing the x- and y-coordinates of gaze positions in pixels, relative to the screen size in pixels (1920 × 1080). Lastly, the Events column records relevant task and behavioral events using integers from 1 to 10. These integers correspond to specific events such as the presentation of memoranda, distractors, probe arrays, and the participant's response, Table 3.1.

Although the 'Task data' field provides a comprehensive event-by-event breakdown for each participant, the primary focus of our analysis will be on the complete dataset for each participant, rather than a granular event analysis. By examining the overall patterns and trends in the data, we aim to gain a broader understanding of the participants' performance and the potential impact of ADHD on visuospatial working memory.

In this context, it is important to note that the dataset encompasses a sample of 50 participants, each characterized by their age and gender. This sample is divided into two distinct groups: the non-ADHD group, consisting of 22 participants, and the ADHD group, with 28 participants. Furthermore, the ADHD group contains two subgroups representing off- and on-medication states, the latter of which includes a subset of 17 participants.

Initially, we aimed to explore the potential of both on- and off-medication subgroups to provide additional insights. We performed an ANOVA test, a statistical technique used to determine whether the means of two or more groups differ significantly from one another, with our null hypothesis being that the means between the two groups were the same or at least did not have any significant difference, [105].

By performing the ANOVA test, Figure 3.3, we found that the p-value was greater than 0.05, indicating that we could not reject the null hypothesis, and that the means between the two groups were not significantly different from each other, [106]. Therefore, we decided to disregard the on-medication subgroup in our analysis.

Moreover, the dataset includes behavioral data on the participant

Anova: Single Factor						
SUMMARY						
Groups	*Count (number of subjects)*	*Sum of the average performances*	*Average Performance*	*Variance*		
ADHD off-medication	28	17.770	0.635	0.020		
ADHD on-medication	17	11.803	0.694	0.014		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.0376	1	0.0376	2.1146	0.1532	4.0670
Within Groups	0.7654	43	0.0178			
Total	0.8031	44				

Figure 3.3: ANOVA analysis for on- and off-meds groups.

based on the Weschler Intelligence Scale for Children (WISC), [107], which provides information on their intelligence test results, including Full-Scale IQ and other metrics that we will evaluate for their applicability in this analysis.

Finally, as already mentioned, the dataset includes performance data on the different subtasks, which we will analyze on an average basis to generate a benchmark for our analysis with the eye-tracking data.

3.2 Methodology Overview

Our proposed methodology entails a two-fold approach, which seeks to comprehensively evaluate the impact of incorporating eye-tracking metrics into the classification model for ADHD diagnosis.

In the first stage, we will establish a classification model based on non eye-tracking related metrics, including the age of the children, their performance on the memory task test, and relevant metrics from their WISC test results. The accuracy of this initial model will be used as a benchmark to evaluate the extent to which the subsequent inclusion of eye-tracking metrics can enhance its diagnostic power.

In the second stage, we will investigate two assumptions regarding the eye-tracking data: 1) the visual scanpath follows a "feed and fly" model, [20], and 2) the visual scanpath follows a Lévy flight model, [31]. For the "feed and fly" model, we will apply corresponding eye-tracking algorithms to generate additional metrics. For the Lévy flight model, we will derive the metric from the velocity distributions of the participants. In both cases, the resulting metrics will be incorporated into the classification model. We will then re-evaluate the accuracy of the model, with the aim of assessing whether the inclusion of eye-tracking data can result in a more robust and accurate classification of children with ADHD. Our goal is to demonstrate that by leveraging eye-tracking metrics, we can enhance the classification model's sensitivity and specificity, and ultimately improve our ability to diagnose ADHD in children.

Overall, this dual approach methodology will provide a more compre-

hensive and nuanced understanding of the diagnostic value of eye-tracking data, and help to identify the potential for novel diagnostic and treatment approaches for ADHD.

3.3 Benchmark Model and Parameter Selection

In this section, we will discuss in more detail the three types of parameters used for the benchmark model: age, WISC metrics, and participant test performance.

3.3.1 Non-Eye Tracking Parameters

Age

The first parameter will be the age of the children. All the children vary from ages from 8 years old up to 13 years old. However the age distribution within the two groups is not equal, Figure 3.4. This discrepancy makes the age parameter crucial for our analysis.

WISC Metrics

The second parameter is the WISC metrics. The Wechsler Intelligence Scale for Children (WISC) is a widely used cognitive test designed to measure the intellectual ability of children between the ages of 6 and 16, [107]. In our dataset, all children were tested individually on the WISC-III or the WISC-R (3 children) by trained neurologists or neuropsychologists from the clinical team at Universidad Católica de Chile, [14].

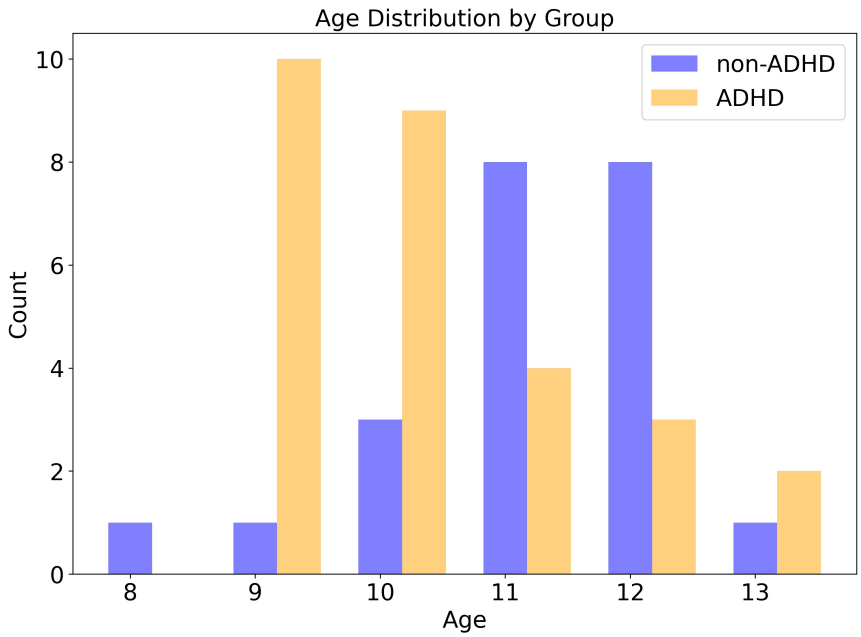

Figure 3.4: Age Distribution.

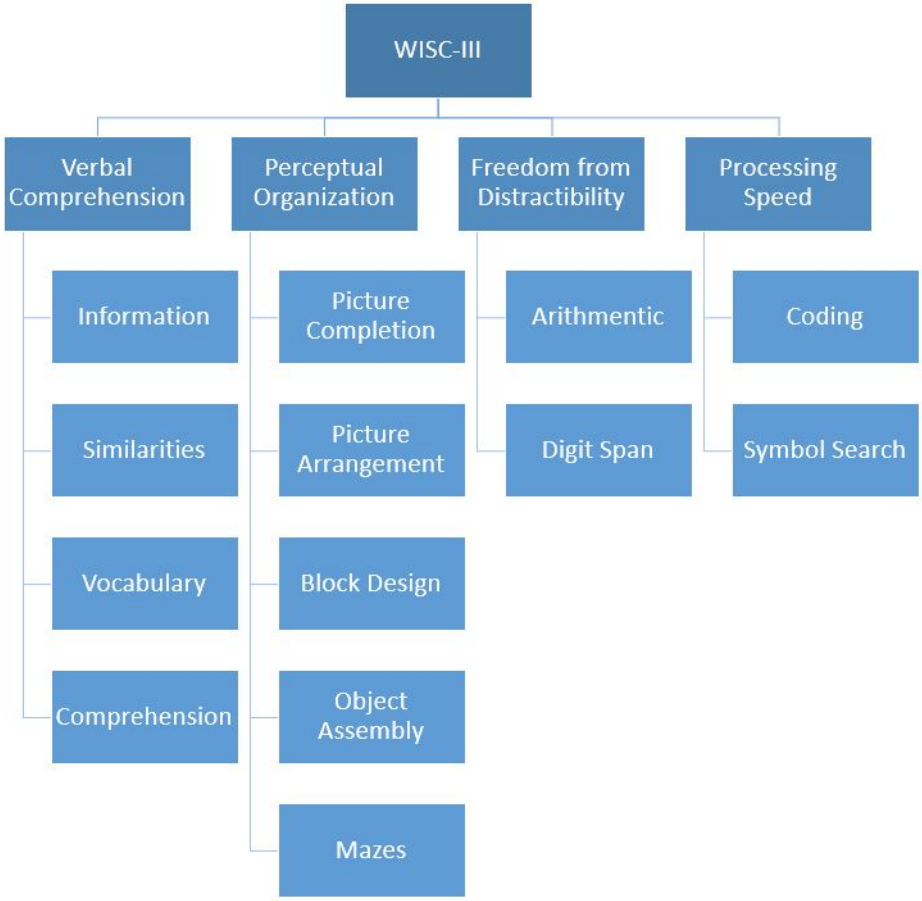

Figure 3.5: Theoretical structure of the WISC-III, taken from [108].

The WISC-III is a comprehensive measure of cognitive abilities and includes 13 subtests that assess different aspects of intelligence, including Verbal Comprehension, Perceptual Organization, Processing Speed, and Freedom from Distractibility, Figure 3.5. These subtests are combined to provide a Full-Scale IQ score, as well as scores for Verbal IQ and Performance IQ.

Numerous studies in the literature indicate that children with ADHD exhibit specific cognitive deficits, in addition to symptoms of inattention and hyperactivity-impulsivity. In their study, Moura et al. [109] found that WISC-III Full-Scale IQ was significantly lower in children with ADHD than in the non-ADHD group. On the other hand, A. Wood et al. [110] concluded that lower IQ does not account for the key cognitive impairments observed in ADHD.

For this reason, it is essential to evaluate other WISC metrics as well, in order to have a combination of them in our analysis.

Susan Dickerson Mayes and Susan L. Calhoun [111] found that children with ADHD have lower mean scores on WISC-III Freedom from Distractibility index and Processing Speed index than on Verbal Comprehension index and Perceptual Organization index. This suggests

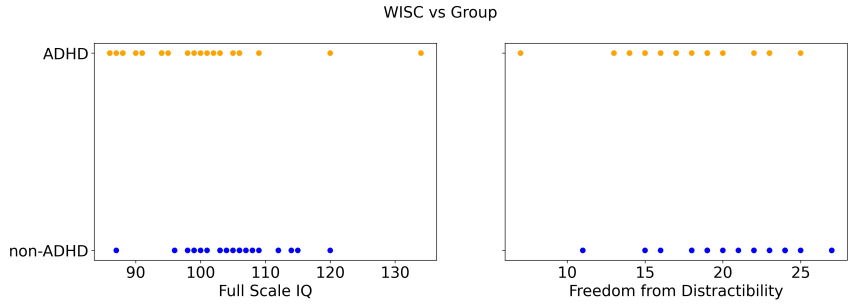

Figure 3.6: WISC metrics distribution per group.

that low scores on the subtests of the first two indices may be indicators of neurological dysfunction due to their consistent and reliable association with neurological disorders. This aligns with what Moura et al. [109] suggested regarding the importance of the Freedom from Distractibility Index as being the index score most impaired in children with ADHD.

Mealer et al. [112] also found that the ADHD group scored significantly lower than the non-ADHD group on Freedom from Distractibility. Similar findings were presented by Lopes et al. [113], who identified that WISC-III Digit Span and Arithmetic subtests could aid in the diagnosis of ADHD.

Based on these findings, we decided to include two parameters from the WISC test in our benchmark model: the Full-Scale IQ and the Freedom from Distractibility, as shown in Figure 3.6. Since data for the Freedom from Distractibility index is not available for all children, we decided to use the sum of the specific subtests of that index, Arithmetic and Digit Span, which we will refer to as "FD" as we want to keep the correlation with their Index.

As shown in the Figures 3.6, there is an apparent trend between the ADHD group and the non-ADHD group when it comes to the Full-Scale IQ. However, as mentioned earlier, the age distribution between the two groups is not equal. Thus, this trend becomes more apparent when the IQ distribution is combined with the age distribution, as seen in Figure 3.7.

Performance on the Memory Task Test

The final parameter for the benchmark model is the children's performance on the memory task test. As mentioned in paragraph 3.1, each child completed 160 trials. Each trial was either a low-load trial, where only one dot was presented on each image, or a high-load trial, where two dots were presented on each image, combined with a distractor image.

For the purposes of this study, we decided not to distinguish between the different combinations of loads and distractors. Instead, we will calculate the average performance across the complete set of 160 trials for each child.

Since the performance on each trial was binary (1 if the child was correct and 0 if the child was incorrect), taking the average of the complete set of 160 trials will yield a linear score ranging between 0 and 1.

Figure 3.8 ppresents the distribution of children based on their group

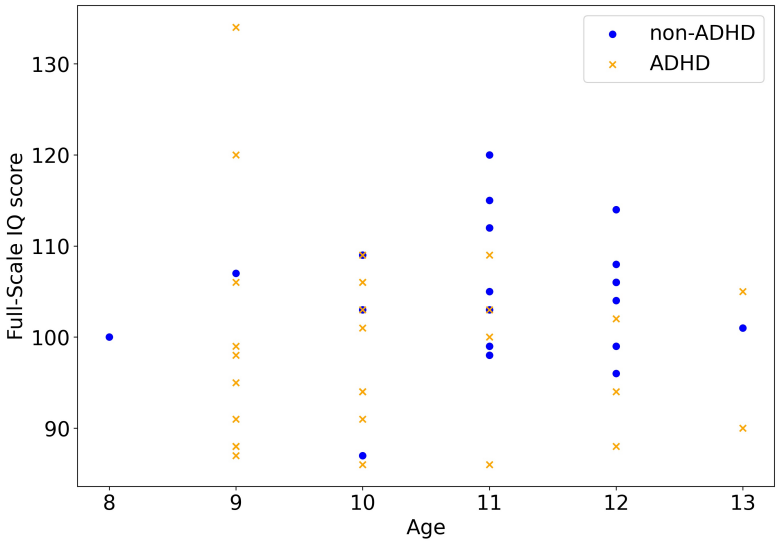

Figure 3.7: Full-Scale IQ vs. Age for both Groups.

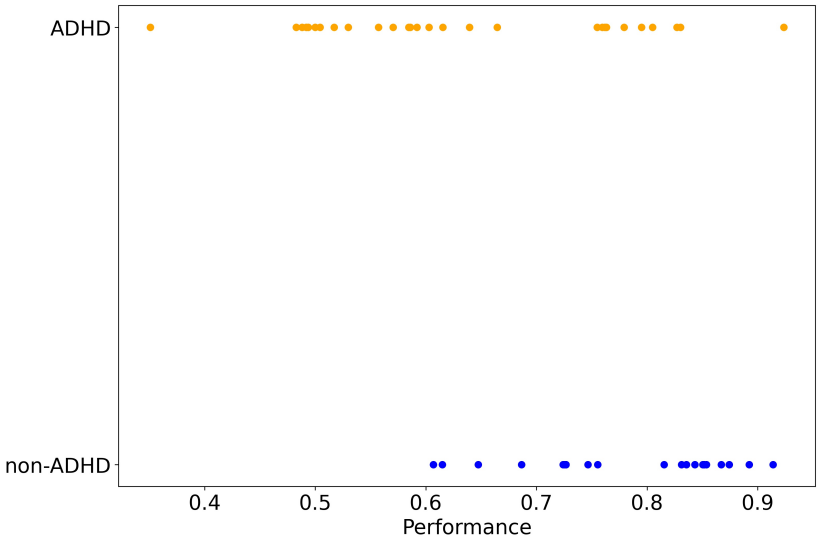

Figure 3.8: Performance distribution for both Groups.

and average performance. The non-ADHD group appears more concentrated at higher performance levels, while the ADHD group seems to be more scattered across a broader performance range. We expect that accounting for age and IQ will make the distinction between the two groups more explicit.

3.3.2 Classification Algorithms for Benchmark Model

In this chapter we will establish a classification model based on non-eye-tracking related metrics, as described in Section 3.3.1.

To accomplish this, we will evaluate several different classification algorithms to determine which one provides the best performance in our

specific context. Some of the classification algorithms that we plan to evaluate include logistic regression, support vector machines, decision trees, and random forests. These algorithms were selected because they have been shown to perform well in other classification tasks and are commonly used in the field of machine learning.

In the following paragraphs, we provide a more detailed description of the algorithms that we used, along with a brief explanation of how they work, their main options, and the importance of cross-validation and hyperparameter tuning in optimizing their performance. These techniques ensure that the selected model is robust and accurately generalizes to unseen data.

Logistic Regression

The logistic regression method estimates the relationship between a dependent variable and one or more independent variables but predicts a categorical variable instead of a continuous variable. True or false, one or zero, yes or no, and so on are all examples of categorical variables. The unit of measure differs from linear regression in that it generates a probability, whereas the logit function transforms the S-curve into a straight line, Figure 3.9, [114].

Logistic regression models are classified into three types based on categorical response.

- Binary logistic regression: The response or dependent variable in this approach is dichotomous in nature, with only two possible outcomes (e.g., 0 or 1). This is the most common approach within logistic regression, and it is also one of the most common classifiers for binary classification in general. And this is the one we'll be using for this project.

- Multinomial logistic regression: The dependent variable in this type of logistic regression model has three or more possible outcomes, but the order of these values is not specified.

- Ordinal logistic regression: When the response variable has three or more possible outcomes, but these values have a defined order, this type of logistic regression model is used. Ordinal responses include grading scales ranging from A to F and rating scales ranging from 1 to 5.

The logistic function is defined by the expression:

$$\rho(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

where β_0 is known as the intercept, and β_1 as inverse scale parameter or rate parameter.

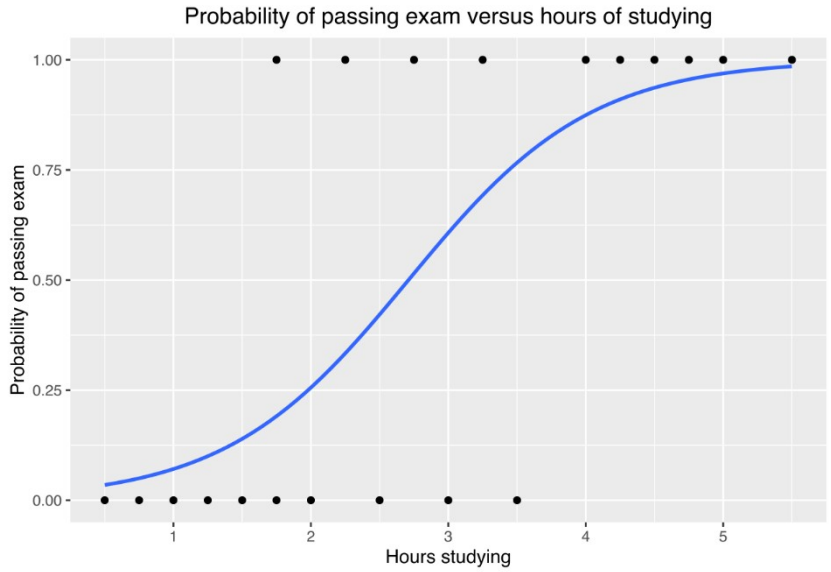

Figure 3.9: Graph of a logistic regression curve fitted to the (xm,ym) data, taken from [115].

Support Vector Machine / C-Support Vector Classification

The Support Vector Machine (SVM) is a machine learning technique that can be used for classification and regression tasks. It is a powerful method because it is able to predict outcomes with high accuracy while avoiding the problem of overfitting, which occurs when a model becomes too complex and fits the training data too closely. The SVM algorithm is particularly well-suited to analyzing datasets with a large number of predictor fields [116].

The goal of the Support Vector Machine (SVM) algorithm is to locate a hyperplane within an N-dimensional space that distinctly classifies data points into separate groups, Figure 3.10. The dimensionality of this hyperplane is contingent upon the number of features, translating to a line for two features, a 2-D plane for three features, and higher dimensional spaces for an increased number of features [117].

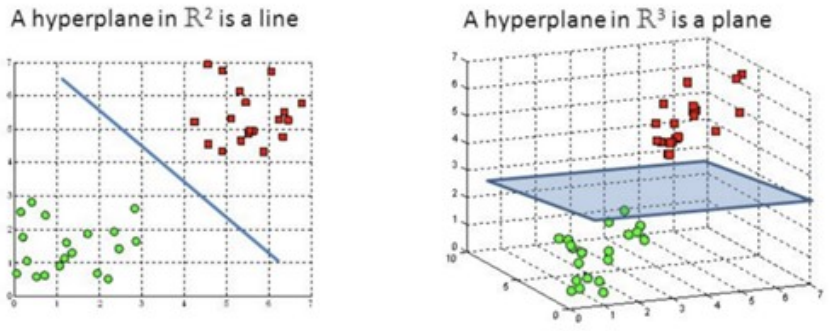

Figure 3.10: Hyperplanes in 2D and 3D feature space, taken from [118].

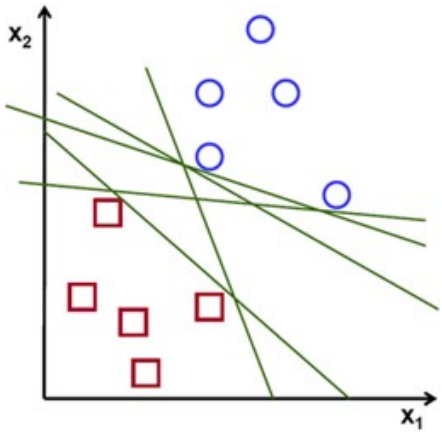

Figure 3.11: Possible hyperplanes, taken from [118].

SVM endeavors to identify the optimal hyperplane by amplifying the margin, defined as the distance between the nearest data points belonging to the two distinct classes, Figure 3.11. This strategy instills greater confidence in the classification of future data points. By maximizing the space between the two classes, the probability of misclassifying newly introduced data points is reduced [118].

Decision Tree Classifier

A decision tree is a non-parametric method utilized in supervised learning, suitable for both classification and regression tasks. It employs a tree-like structure, depicted in the Figure 3.12, comprising a root node, internal nodes, branches, and leaf nodes. Each decision tree commences with a root node devoid of any incoming branches. The root node subsequently branches out to internal nodes, otherwise referred to as decision nodes. These nodes, both root and internal, scrutinize the provided features to yield homogeneous subsets which are then demarcated as leaf nodes or terminal nodes. These terminal nodes represent all potential outcomes derived from the dataset [119].

There are two main types of decision trees used in data mining, the classification tree and the regression tree.

The classification tree analysis is used when the expected outcome is the class (discrete) to which the data belongs, and the regression tree analysis is used when the predicted outcome may be regarded a real number (e.g., the price of a house or the duration of stay in a hospital).

Random Forest Classifier

The random forest method is a popular machine learning technique that mixes the output of numerous decision trees to produce a single conclusion. Its ease of use and versatility, as well as its ability to tackle classification and regression challenges, have boosted its popularity [120].

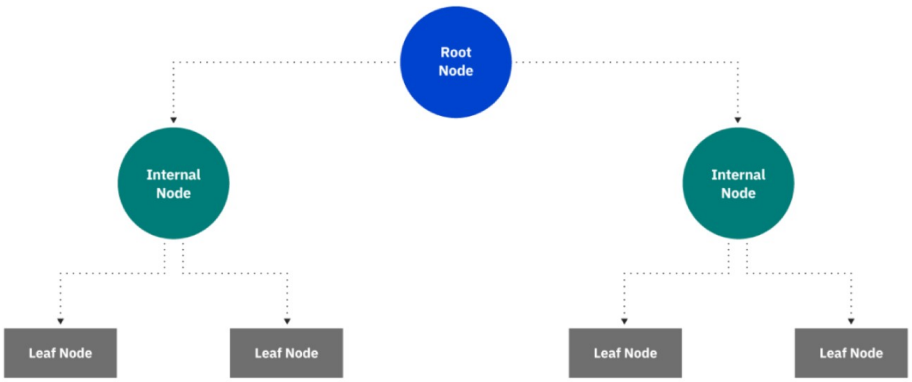

Figure 3.12: Schematic of a Decision Tree, taken from [119].

It is made up of a huge number of independent decision trees that work together as an ensemble. Figure 3.13 shows how each individual tree in the random forest generates a class prediction, and the class with the highest votes becomes our model's forecast [121].

The requirements for the random forest to perform well are that there is some actual signal in our features so that models created using those features outperform random guessing and that the predictions (and hence mistakes) generated by the individual trees have minimal correlations with one another.

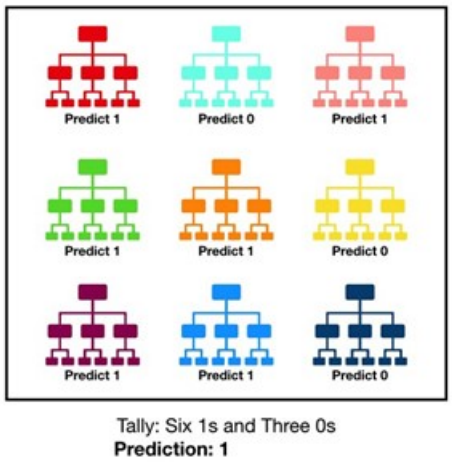

Figure 3.13: Visualization of a Random Forest Model Making a Prediction, taken from [119].

K-Nearest Neighbors Algorithm

The k-nearest neighbors algorithm, often known as KNN or k-NN, is a non-parametric, supervised learning classifier that employs proximity to classify or predict the grouping of a single data point. While it may be used for either regression or classification issues, it is most commonly utilized as a classification technique based on the idea that similar points can be discovered nearby [122].

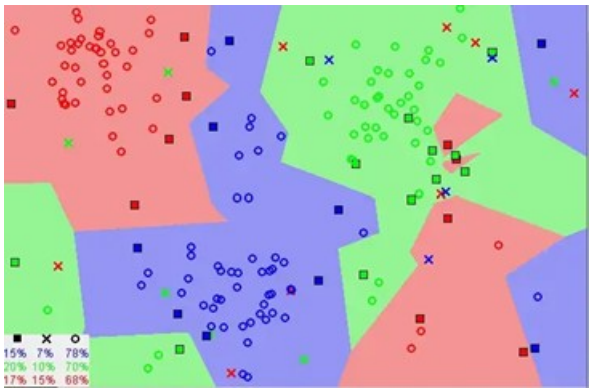

Figure 3.14: Image showing how similar data points typically exist close to each other, taken from [123].

Figure 3.14 shows that similar data points are usually adjacent to one other. The KNN algorithm is based on the premise that this assumption is true enough for the algorithm to be beneficial. KNN captures the concept of similarity (also known as distance, proximity, or closeness) with certain mathematics we may have learned as children, such as calculating the distance between the points on a graph.

Before concluding this section, it is crucial to discuss three concepts that are highly relevant to the classification algorithms described above: cross-validation, automated feature selection methods and hyperparameter tuning.

Cross-Validation

Cross-validation is a prevalent technique in machine learning for assessing the performance of classification algorithms. It is a statistical method that assists in estimating a model's accuracy by evaluating its generalization capabilities to unseen data. Cross-validation is employed to counteract overfitting, a phenomenon where the model demonstrates excellent performance on the training dataset but underperforms on new, unseen data.

This technique facilitates hyperparameter tuning, algorithm selection for a specific problem, and overall performance assessment of a classification algorithm. Cross-validation is a critical step in constructing robust and accurate machine learning models.

Cross-validation entails partitioning the dataset into multiple parts or

"folds." The most common method is k-fold cross-validation, wherein the data is divided into k equally-sized parts. In each iteration, one fold serves as the validation set, while the remaining k-1 folds form the training set. The classification algorithm is trained on the training set and tested on the validation set. This procedure is repeated k times, with each fold used as the validation set exactly once. The algorithm's performance is averaged across all k iterations, providing an overall accuracy estimate.

In this study, two distinct cross-validation methods were utilized to evaluate the performance of classification algorithms: Stratified K-fold Cross-Validation and Leave-One-Out Cross-Validation (LOOCV). The former, an extension of the standard k-fold cross-validation method [124], maintains the ratio of target classes in each fold consistent with the full dataset (i.e., 56% ADHD individuals and 44% non-ADHD individuals), optimizing the bias-variance tradeoff in the testing process. The latter, LOOCV, is a special case of cross-validation where the number of folds matches the number of instances in the dataset [125], offering minimal bias in the testing process but introducing substantial variance due to using a single data point for testing in each iteration. This method is employed due to the small size of the dataset.

By comparing the outcomes of these methods, a comprehensive evaluation of the classification algorithm's performance can be achieved, facilitating the development of a robust and accurate model.

Automated Feature Selection Methods

Automated feature selection techniques aim to identify the most pertinent features for a model, reducing the likelihood of overfitting and enhancing computational efficiency. In this study, we utilized three automated feature selection approaches: Recursive Feature Elimination (RFE), Recursive Feature Elimination with Cross-Validation (RFECV), and SelectKBest.

RFE is a method that iteratively fits a model to the data, assesses the significance of each feature, and removes the least important features. This process continues until the desired number of features is achieved.

RFECV is an enhancement of RFE that integrates cross-validation to further optimize the feature selection process. It expands upon RFE's iterative procedure by evaluating the model's performance and ranking the features based on their importance in each cross-validation fold. By removing the least significant features and reconstructing the model, RFECV offers a more reliable assessment of the model's performance and improves generalization by retaining influential, independent features while discarding superfluous and weak ones [126].

SelectKBest, conversely, is a univariate feature selection technique that chooses the top k features according to their statistical relationship with the target variable[127].

Although automated feature selection methods offer potential advantages, they may not consistently outperform manual feature selection. Manual feature selection, which relies on domain expertise and informed

judgment, can occasionally generate better models by incorporating contextual information that automated methods might miss.

Hyperparameter Tuning and Model Training

Hyperparameter tuning is an essential aspect of machine learning, as it significantly influences a model's performance. The process involves adjusting specific parameters that guide the model's learning process, ultimately aiming to optimize its performance on unseen data. In this study, we concentrate on several key hyperparameters for logistic regression. As we will discuss later in the results section, logistic regression will be chosen as the preferred method for this study. The following parameters are considered ([128]):

- ***Penalty (or regularization)*:** This is designed to minimize the error in model generalization and serves as a control mechanism to prevent and manage overfitting.
- ***C (or regularization strength)*:** This parameter works in tandem with the penalty to manage overfitting. Lower values indicate stronger regularization, while higher values instruct the model to assign significant weight to the training data. It is essential that this value is a positive floating-point number.
- ***Solver*:** This is the algorithm employed to solve the optimization problem. The available options include:
  - ***Lbfgs*:** This solver tends to perform well compared to others and is memory-efficient. However, it can sometimes struggle with convergence issues.
  - ***sag*:** This is more efficient than other solvers when dealing with large datasets, particularly when both the number of samples and features are substantial.
  - ***Saga*:** This solver is the recommended choice for sparse multinomial logistic regression, and it is also suitable for very large datasets.
  - ***newton-cg*:** This solver can be computationally costly due to the calculation of the Hessian Matrix.
  - ***liblinear*:** This is suggested for high-dimensional datasets, as it efficiently handles large-scale classification problems.

Finally, to find the optimal values for these hyperparameters and optimize the accuracy of our logistic regression models, we will use GridSearchCV. GridSearchCV is an exhaustive search technique that evaluates a model with various combinations of hyperparameter values. It systematically explores the search space of possible hyperparameter values, selecting the combination that yields the best performance based on a specified evaluation metric.

3.4 Eye-Tracking Data Analysis

Eye-tracking data analysis is an essential step in understanding the relationship between a participant's ADHD status and their eye movements. This section aims to detail the methods and techniques employed to process and analyze the eye-tracking data collected during the experiments. We will discuss various algorithms, their characteristics, and their application in our study. Furthermore, we will outline the specific steps taken to process the data and extract meaningful insights from it.

By leveraging these techniques and the insights obtained from the data, we aim to enhance the diagnostic power of our Benchmark model and better understand the behavioral patterns exhibited by participants with ADHD.

In this study, as already mentioned in the methodology overview (Section 3.2), we will make two main assumptions regarding the eye-tracking data, which are based on different models describing the visual scanpath of the participants. These assumptions will guide our choice of metrics and algorithms used to analyze the eye-tracking data.

3.4.1 Assumption I: Feed and Fly Model

The main aspect of this model is that the two phases are discrete. This means that, with the appropriate fixation-saccade detection algorithm, we will be in a position to define specific metrics that distinguish these two events. In the following sections, we present the chosen algorithm, which is the Identification by Velocity Threshold (IVT), as well as an optimization for it, while defining the derived metrics.

Identification by Velocity Threshold (IVT)

The Identification by Velocity Threshold (IVT) as already presented in Section 2.3.1 is a commonly used approach for the identification of fixations and saccades in eye-tracking studies. The core of the IVT algorithm is rooted in geometry and calculus, allowing us to determine the instantaneous velocities of eye movements.

Firstly, the Cartesian coordinates of two consecutive gaze points are used to calculate the Euclidean distance between them. If the coordinates are (x_1, y_1) and (x_2, y_2), then the straight-line or Euclidean distance between the points d, i.e., the length of the vector, is computed using the Pythagorean theorem:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\,. \tag{3.1}$$

After calculating the Euclidean distance, the time difference between consecutive points is determined by subtracting the timestamp of the previous point from that of the current point. With this information, we can compute the instantaneous velocities between consecutive gaze points using the basic distance divided by the time equation. This calculation is

derived from the concepts of differentiation in calculus and approximates the derivative of the position function concerning time.

Once we have the point-to-point velocities for all gaze points, a threshold is applied to classify the data as fixations (low velocities) or saccades (high velocities). A fixed threshold may not provide optimal results, as it can impact the detection capability of the algorithm. To address this issue, researchers have explored dynamic velocity thresholds, as described in Section 2.3.1.

In the present study, our approach will incorporate a novel algorithm called the Fixation-Saccade Transition Ratio Algorithm (FSTR) algorithm, which will be presented and thoroughly described in the following chapter. This algorithm aims to improve the mathematical modeling of eye movements by refining the velocity-based classification of fixations and saccades, enhancing the overall detection capability. By highlighting the mathematical principles behind the IVT algorithm, we emphasize the role of mathematics in understanding and analyzing eye movement data to detect ADHD.

Fixation-Saccade Transition Ratio Algorithm for IVT Optimization

The FSTR algorithm is an advanced, adaptive, and versatile tool for meticulously analyzing eye-tracking data, specifically tailored for identifying the optimal velocity threshold necessary to differentiate between fixations and saccades. By testing a range of thresholds, the FSTR algorithm adapts its sensitivity to individual differences in eye movement patterns, ensuring a comprehensive understanding of the unique ways in which participants process visual information.

In the initial stage of the analysis, the FSTR algorithm detects fixations and saccades by employing a specific velocity threshold. It is based on the assumption that saccades and fixations can be distinguished by their respective velocities, with saccades having a higher velocity than fixations. Furthermore, it assumes that both saccades and fixations last longer than the sampling frequency of the data, implying that a saccadic or fixational data point is likely to follow from one in the same category.

This assumption directly addresses the challenge identified by Olsen [78], where traditional algorithms often misclassify long fixations or saccades that are interrupted by very short saccades or fixations. By making this assumption, the FSTR algorithm aims to provide a more robust and accurate identification and differentiation of these eye movements. The assumption sets the stage for the calculation of the FSTR value, as it establishes the foundation for the calculation of empirical and theoretical (independent) probabilities of transitions between fixations and saccades.

The FSTR value is a specialized metric that quantifies the proportion of observed transitions, empirical, between saccades and fixations compared to the expected number of transitions if the occurrence of one does not affect the probability of the other, independent. The FSTR value is calculated using the following formula:

$$FSTR = \frac{P_{emp}(S \rightarrow F)}{P_{ind}(S \rightarrow F)}\,. \tag{3.2}$$

Here, $P_{emp}(S \rightarrow F)$ represents the empirical probability of observing a transition from a saccade (S) to a fixation (F) from the measured/collected data. Conversely, $P_{ind}(S \rightarrow F)$ represents the theoretical probability of observing the same set of points if the occurrence of one does not affect the probability of the other, independent probability. Mathematically, $P_{ind}(S \rightarrow F)$ can be calculated as:

$$P_{ind}(S \rightarrow F) = P_{ind}(F \rightarrow S) = P_{emp}(S)P_{emp}(F)\,. \tag{3.3}$$

In this equation, $P_{emp}(S)$ and $P_{emp}(F)$ represent the probabilities of empirically observing a single point as a saccade and fixation, respectively.

To calculate the FSTR value, Equation 3.2, the algorithm following the detection of fixations and saccades, generates a binary vector wherein fixations are coded as 0 and saccades as 1. This binary vector is crucial for creating a 2x2 transition frequency matrix contingent on the chosen velocity threshold. This matrix effectively illustrates the frequency of transitions between fixations and saccades throughout the eye-tracking task, serving as a valuable resource for understanding the temporal dynamics of visual attention.

Each element of the 2x2 transition matrix represents the frequency of a specific transition:

- Element (1,1) represents the frequency of transitions from a fixation to another fixation (F $\rightarrow$ F).
- Element (1,2) represents the frequency of transitions from a fixation to a saccade (F $\rightarrow$ S).
- Element (2,1) represents the frequency of transitions from a saccade to a fixation (S $\rightarrow$ F).
- Element (2,2) represents the frequency of transitions from a saccade to another saccade (S $\rightarrow$ S).

The elements of the matrix are used to calculate the empirical and independent probabilities required for the FSTR formula, Equation 3.2. To calculate the probabilities, the following steps are performed:

1. Divide each element of the matrix by the sum of all elements to obtain the empirical transition probabilities:
   $P_{emp}(F \rightarrow F)$, $P_{emp}(F \rightarrow S)$, $P_{emp}(S \rightarrow F)$, and $P_{emp}(S \rightarrow S)$.
2. Calculate the marginal probabilities of observing a fixation ($P_{emp}(F)$) and a saccade ($P_{emp}(S)$):
   $P_{emp}(F) = P_{emp}(F \rightarrow F) + P_{emp}(F \rightarrow S)$
   $P_{emp}(S) = P_{emp}(S \rightarrow F) + P_{emp}(S \rightarrow S)$

The FSTR value, therefore, provides insight into how well the chosen threshold can separate saccades and fixations based on their velocities. A lower FSTR value (closer to 0) indicates better separation between fixations and saccades, as fewer transitions between the two states mean that the algorithm is more successful in identifying and maintaining the two different states without frequently switching between them. A higher FSTR value, on the other hand, suggests a poor separation between fixations and saccades, implying that the chosen threshold is not effective in differentiating the two types of eye movements. This metric provides valuable information on the effectiveness of the chosen threshold in delineating eye movement patterns.

In addition to the FSTR value, the algorithm also calculates the log p-value using Fisher's exact test to assess the statistical significance of the association between fixations and saccades for each velocity threshold. Fisher's exact test is applied to the 2x2 transition matrix and computes the probability of observing the given matrix (or a more extreme one) under the null hypothesis, which assumes that the row and column variables (fixations and saccades) are independent. A lower log p-value indicates stronger evidence against the null hypothesis, suggesting that the observed association between fixations and saccades is not due to random chance and that the given velocity threshold effectively separates these eye movements.

This information can be used in conjunction with the FSTR value to help determine the optimal threshold that best differentiates between fixations and saccades based on their velocities.The identification of this optimal threshold enables for gaining deeper insights into the complex processes underlying visual information processing and environmental navigation through eye movements.

In summary, the FSTR algorithm is a cutting-edge analytical tool for the in-depth examination of eye-tracking data, designed to uncover the optimal threshold for separating fixations and saccades. Its adaptive nature accommodates individual differences in eye movement patterns, ensuring a comprehensive understanding of visual attention and information processing. By calculating the FSTR value and log p-value, the algorithm provides a robust evaluation of the quality of separation between these eye movements.

Ultimately, by determining the participant-specific threshold that offer the most accurate grouping of fixations and saccades, the FSTR algorithm paves the way for a more precise analysis of the participants' visual perception and attention. This adaptability and precision are particularly beneficial for studies focusing on attentional disorders, such as ADHD.

Fixation-Saccade Transition Ratio Algorithm and Derived Metrics

The FSTR Algorithm allows us to derive several valuable metrics that can be used to examine the differences in eye movement patterns between ADHD and non-ADHD participants. In the following sections, we will discuss three specific metrics derived from the FSTR algorithm, which

include:

1. Minimum FSTR Value: A metric that highlights the differences in eye movement patterns between ADHD and non-ADHD participants by examining the minimum FSTR values.
2. Cumulative Sum FSTR Value: A metric that aggregates the FSTR values across the tested threshold range, magnifying the distinction between ADHD and non-ADHD participants.
3. Saccade Frequency and Fixation Duration Metrics: Metrics that use the velocity threshold corresponding to the minimum FSTR value to calculate the saccade frequency and fixation duration for each participant.

Each of these metrics will be explored in detail in the subsequent sections, focusing on their potential to differentiate between ADHD and non-ADHD participants based on their eye movement patterns.

Minimum FSTR Value

Capitalizing on the FSTR algorithm's strengths, we can explore an interesting hypothesis within the context of attentional disorders, such as ADHD. We propose that participants with ADHD may exhibit distinct eye movement patterns, characterized by more frequent and shorter interfering saccades. These unique patterns could make it challenging to differentiate between fixations and saccades, potentially leading to higher minimum FSTR values for participants with ADHD compared to non-ADHD participants. The rationale behind this hypothesis is that the algorithm may struggle to find an optimal threshold that effectively separates the eye movements in ADHD participants due to their atypical patterns.

To evaluate the validity of this hypothesis, we can apply the FSTR algorithm to both ADHD and non-ADHD groups to determine if there is a significant difference in the minimum FSTR values between the two populations. If our hypothesis is supported, this finding would suggest that the FSTR algorithm captures a distinctive aspect of eye movement patterns in participants with ADHD. This information could contribute to more accurate assessments of visual attention and information processing and potentially enhance the Benchmark model's performance, as discussed in Section 3.3.

Cumulative Sum FSTR Value

Building upon the hypothesis that participants with ADHD may exhibit higher minimum FSTR values, examining the cumulative sum of FSTR values across the tested threshold range could further emphasize this difference, making it even more evident.

The cumulative sum represents the aggregated FSTR values throughout the tested threshold range, which could help accentuate the distinction

between ADHD and non-ADHD participants. Incorporating the cumulative sum analysis within the FSTR algorithm could offer an additional layer of precision when differentiating between ADHD and non-ADHD participants.

By uncovering unique aspects of attentional dynamics in the ADHD population, this combined approach could lead to a more accurate and comprehensive understanding of visual attention and information processing in participants with ADHD. Ultimately, this could enhance the performance of the Benchmark model, as discussed in Section 3.3.

Utilizing Velocity Threshold for Saccade Frequency and Fixation Duration Metrics

After obtaining the velocity threshold corresponding to the minimum FSTR value from the FSTR Algorithm, we can effectively differentiate between fixations and saccades for each participant. By utilizing this threshold, we will calculate the saccade frequency and fixation duration metrics, as introduced in Section 2.6. These metrics will then be assessed for their potential to differentiate between ADHD and non-ADHD participants based on the eye-tracking data, potentially providing valuable insights into the attentional differences between these groups.

3.4.2 Assumption II: Lévy Flight Model

In this section, we will investigate the Lévy flight model's ability to characterize the eye-tracking data by focusing on the power-law distribution, a key aspect of the model, as already described in the Section 2.2.1. To do so, we will evaluate the distribution of the velocities for each participant, following a similar process to the one used in the Identification by Velocity Threshold (IVT) section.

Initially, we prepare the data by loading the eye-tracking data for all participants. The x and y positions, along with the time values, are extracted from the data, and the velocities for each participant are calculated using the same method as in the IVT Section 3.4.1. These results are then stored in appropriate data structures.

Next, we analyze the distribution of velocities by creating histograms for each participant. To better visualize and fit a power-law model to the velocity histograms, we transform the data to a log-log scale and calculate the log-bin centers. This step allows for a clearer representation of the data distribution when dealing with heavy tails and a wide range of values.

Following the plotting of histograms, we fit a linear regression model to the log-transformed histograms, estimate the scaling exponent of the power-law distribution, and calculate the coefficients of determination (R-squared values) for each participant. These results can help us assess the goodness of fit for the Lévy flight model and provide insights into the eye-tracking patterns of the participants.

Lévy flights are characterized by power-law scaling, with the tails of the distribution adhering to a power-law, $P(x) \propto |x|^{-\alpha}$, with $1 < \alpha \leq 3$,

as described in Section 2.2.1. The exact value of the exponent α varies depending on the specific system being modeled. In our analysis, we estimate the scaling exponent α for each participant to assess the goodness of fit for the Lévy flight model and provide insights into their eye-tracking patterns.

As mentioned in Section 2.2.1, literature on expected values of the scaling exponent for visual scanpaths is limited, apart from the study by Credidio H. et. al. [40] which reported a value for the scaling exponent close to 2.9, while the general exponent value on foraging models is approximately 2.

In this study, we would expect that the exponent values of all the participants in general should be within that range. We hypothesize that participants with ADHD may have smaller scaling exponents compared to those without ADHD, as the Lévy flights with smaller exponents exhibit more diffusive behavior and longer steps. This is based on the assumption that individuals with ADHD may exhibit more frequent and longer gaze shifts due to their difficulty in maintaining attention. Additionally, since Lévy flights consider each step, regardless of its size, to occur in the same unit of time, these longer gaze shifts would correspond to larger velocities.

To test this hypothesis, we will compare the scaling exponents between the two groups and explore any potential differences in their eye-tracking patterns. The scaling exponent will be used as a metric in the benchmark model with the aim of improving the classification accuracy between participants with ADHD and those without ADHD. By incorporating this metric, we hope to develop a more accurate and robust classification model that can better distinguish individuals with ADHD from those without the condition.

Chapter 4

Results

In this Chapter, as already described in the Methodology Procedure Section 3.2, we will evaluate the impact of incorporating eye-tracking metrics into the classification model for ADHD diagnosis on children. This will be achieved in two steps.

In the first step, we will create a classification model solely dependent on non eye-tracking related metrics, 3.3, i.e. the age of the children, the specific metrics from their WISC test, and their test performance. The accuracy of this model will serve as the benchmark for our analysis.

In the second step, we will revise our classification model with eye-tracking metrics this time, and evaluate if its accuracy has been increased. The eye-tracking metrics will be based on both assumptions regarding the visual scanpath, i.e. that it follows either a "feed and fly" model or a Lévy flight pattern.

Throughout the following sections, we will first present the results of the classification model that relies solely on non-eye-tracking related metrics, providing a benchmark for our analysis. Next, we will revise our classification model by incorporating eye-tracking metrics and evaluate whether its accuracy has been increased. Finally, we will thoroughly analyze and discuss the results.

4.1 Benchmark Model

As presented in the Section 3.3.1, the selected parameters for the Benchmark model are the following:

- Age
- Full Scale IQ (IQ)
- Arithmetic and Digit Span (FD)
- Test Performance (Perf.)

The cross validation results are presented in the Table 4.1, and in both methods the Logistic Regression has the highest accuracy.

Table 4.1: SKF and LOOC cross validation results.

Method	Logistic Regression	SVC	Random Forest	K-Neighbors	Decision Tree
SKF (k = 5)	0.720	0.540	0.680	0.540	0.611
LOOC	0.720	0.540	0.720	0.540	0.596

4.1.1 Random State Analysis and Model Training

We would split the data into a training set and a testing set. We will use a ratio of 4:1, aligning with the 5-fold cross validation from the previous paragraph. An important parameter in the train_test_split method, is the random state. In Scikit-learn, the random state hyperparameter takes one of the following values:

- *None*: This is the default value. This allows the function to use the global random state instance from np.random. Calling the same function multiple times with random_state=None, will produce different results across different executions.
- *int*: We can use an integer for random_state. Any positive integer including 0. When this option is used, the function will produce the same results across different executions. The results are only changed if the integer value is changed. [129]

It is expected that for different random states, the accuracy of the model will vary slightly. However, as presented in Figure 4.1, our model's accuracy variation is large, with mean value at 0.7 and a standard deviation of 0.14. The main reason is that the dataset is very small, only 50 points, and with a 4:1 split, the 40 datapoints is not enough to train the model

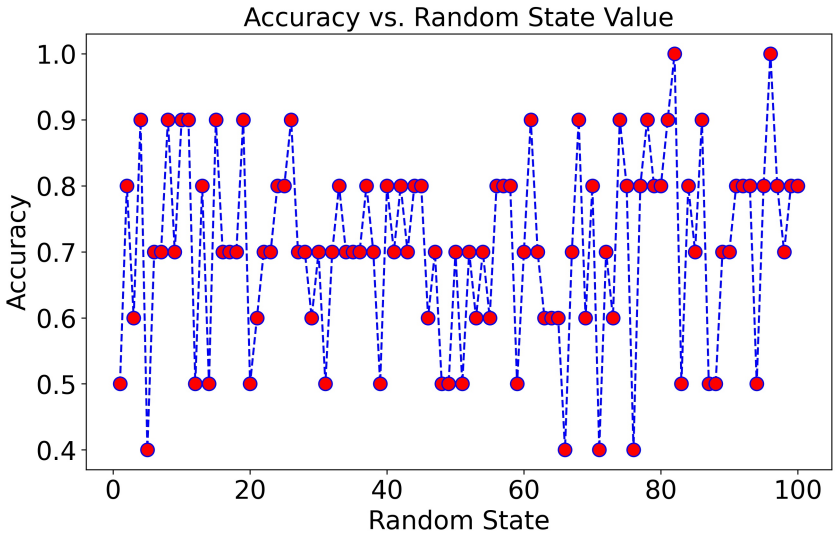

Figure 4.1: Model Accuracy for various random state values.

accurately. Therefore, the model is not stable enough to proceed any further.

4.1.2 An Alternative Approach – A Set of Models

Our new approach will be the following. We will use the same concept with the Leave-One-Out Cross-Validation, to create a model, using 49 datapoints as train dataset, and 1 datapoint for testing. We will repeat this process 50 times, to cover all possible combinations. Therefore, we will have a set of 50 trained models, Figure 4.2, which they will differentiate from each other only on a single datapoint in their training dataset.

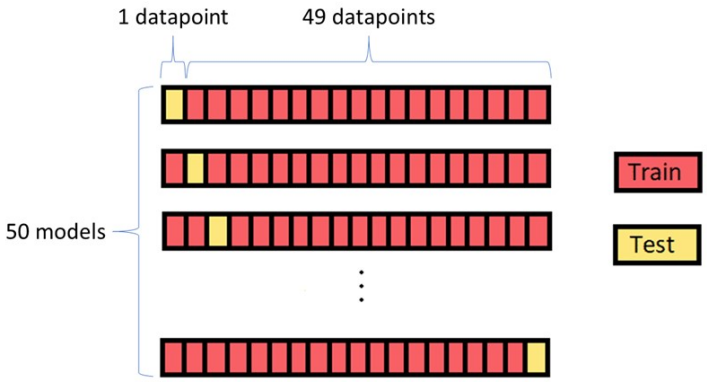

Figure 4.2: Alternative Approach – A Set of Models.

In our approach, we make the following critical assumption:

Critical Assumption: We assume that the average accuracy of the set of the 50 models is representative of the individual accuracy of any of the 50 models.

This assumption is key to understanding how the average accuracy of the set of models can be used to predict the status (ADHD or not) of a new child in our dataset.

Of course, similarly to the LOOCV, with this training approach we cannot achieve the bias-variance tradeoff in our testing process. Each individual model will have low bias, and high variance, towards the testing data. However, as a collective set of all 50 models, we will overcome that problem, and get a good estimation of the prediction accuracy.

Our steps forward will be first to perform a hyperparameter tuning to optimize our model settings and then to evaluate the set of the 50 models with a confusion matrix.

The result from the GridSearchCV on the train data, for the LeaveOneOut() method, is presented in Table 4.2. We will apply the above parameter settings, to all 50 models.

4.1.3 Model Accuracy

Now we can run all 50 models, in sequence, and generate a collective confusion matrix, Figure 4.3. Out of the set of the 50 models, 16 of them

Table 4.2: Tuned Hyperparameters and Accuracy.

Hyperparameter	**C**	**Penalty**	**Solver**
Value	10.0	l2	newton-cg
Accuracy	0.755		

predicted that an individual belongs to the non-ADHD group ("0") and they were correct, and 22 predicted that an individual belongs to the ADHD group ("1") and they were also correct. However, six models predicted that an individual belongs to the ADHD group while it belonged to the non-ADHD group and another six models predicted the individual belongs to the non-ADHD group while it belonged to the ADHD group.

Based on the results presented in the collective confusion matrix, Figure 4.3, our accuracy will be:

$$Accuracy = \frac{38}{50} = 0.76 \tag{4.1}$$

This accuracy represents the accuracy of the set of the 50 regression models, which based on our assumption from Section 4.1.2, will also be the individual accuracy of any of these 50 models. Therefore, if we need to predict the status of any new entry, using any of these 50 models we will have a 76% probability to predict it accurately.

Based on the collective confusion matrix, we can also calculate precision, recall, and specificity [130] to further evaluate the performance of the set of 50 models. Precision measures the proportion of true positive predictions among all positive predictions made by the model. In our case, precision is equal to 0.79, which means that out of all the children predicted to have ADHD by the model, 79% actually had ADHD, while the remaining 21% were falsely identified as having ADHD.

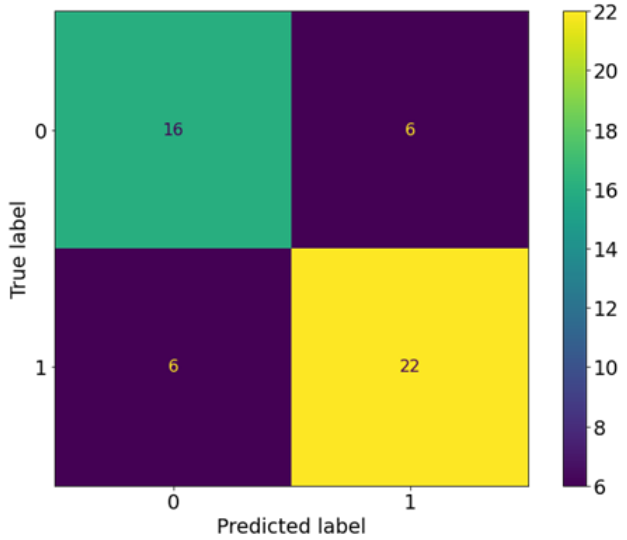

Figure 4.3: Collective Confusion Matrix for the set of 50 Models.

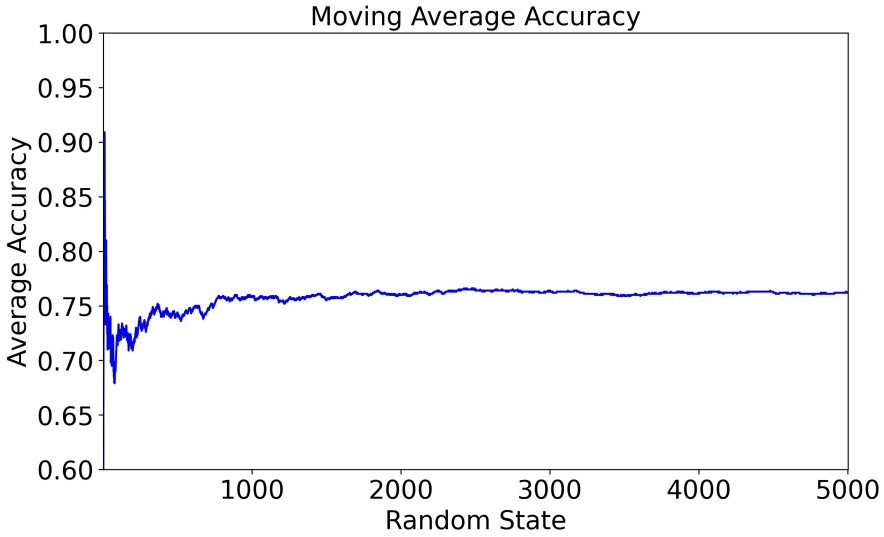

Figure 4.4: **Model$_{1/50}$** accuracy moving average.

Recall, on the other hand, measures the proportion of true positive predictions among all actual positive instances in the dataset. Our recall is also equal to 0.79, indicating that out of all the children in the dataset who actually had ADHD, the model correctly identified 79% of them, while the remaining 21% were falsely identified as not having ADHD.

Finally, specificity is the proportion of true negative predictions among all negative predictions made by the model. Our specificity is equal to 0.73, meaning that out of all the children predicted by the model to not have ADHD, 73% of them actually did not have ADHD, while the remaining 27% were falsely identified as having ADHD.

These metrics suggest that the logistic regression models have a good ability to identify children with ADHD, but may also falsely identify some children as having ADHD when they do not. Therefore, additional analysis may be needed to determine the root causes of these false predictions and how to improve the models further.

We can now re-check the accuracy of one out of the 50 models, **Model$_{1/50}$**, by running it through a large range of random states and calculate its moving average. As shown in Figure 4.4, the average accuracy of the model stabilizes at the same level as calculated from the collective confusion matrix.

The Logistic regression equation of the **Model$_{1/50}$** is the following:

$$\begin{aligned} \text{logit}(p) = 17.3079 - 0.6343 \times Age - 0.0349 \times IQ \\ -0.2417 \times FD - 3.0135 \times Perf. \end{aligned} \tag{4.2}$$

Of course, the above coefficients will slightly vary between the 50 models, but this variation is marginal because, as it has already been mentioned, the difference between the 50 models is only one datapoint in their training set.

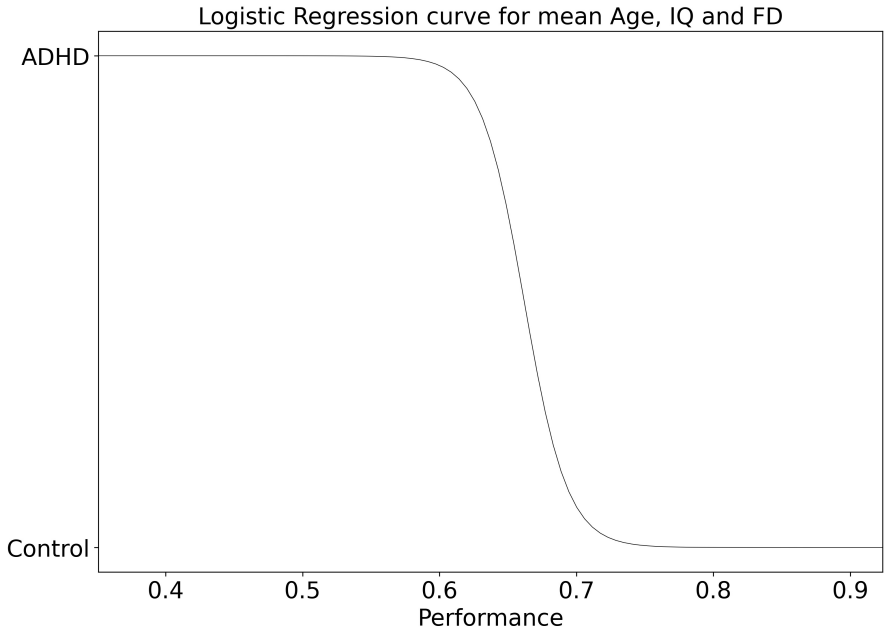

Figure 4.5: Logistic Regression Curve.

Based on the coefficients of each parameter, it is obvious that the main impact on the prediction of the status of a individual (ADHD or not) derives from its average performance.

On the other hand, between the Full-Scale IQ and the Arithmetic and Digit Span (FD), it is obvious that the latter has the higher impact which aligns with what we described in the Section 3.3.1.

Finally, the age appears also to have some impact on the individual's status.

We can visualize the correlation of ADHD and average task performance, where the other three parameters, Age, Full-Scale IQ (IQ) and Arithmetic and Digit Span (FD), are kept constant, Figure 4.5.

In this visualization we can see the performance level/ threshold, at which an individual of average age and IQ, will shift from the non-ADHD to the ADHD group, and vice versa. The curve is slightly steeper than a typical sigmoid curve due to the value of the hyperparameter C which is equal to 10, Table 4.2, while in a typical sigmoid curve it is equal to 1. It is clear, that the task performance can be a good indicator of ADHD for children.

Now that we have set up our Benchmark Model, we can proceed to the evaluation of the eye-tracking data, in order to generate those metrics that could potentially increase the accuracy of our model. The subsequent analysis and discussion of these metrics will inform us whether they are useful in differentiating between ADHD and non-ADHD populations, and if so, they can be incorporated into our model for improved performance.

4.2 Eye Tracking Data

In this section, we will conduct a Data Pre-Analysis to evaluate the quality and reliability of the eye-tracking dataset before diving into the investigation of the two assumptions and their derived metrics. Our analysis will include visualizations such as mean velocity for each participant, standard deviation, kurtosis, and skewness. These visualizations will help us identify any potential issues with the dataset and provide insights into the overall distribution of the eye-tracking data.

One critical issue we need to highlight is the missing 'Y' coordinate data for two participants, one with ADHD and one without. The absence of this data significantly limits our ability to conduct a complete and comprehensive analysis for these individuals, as both the pre-analysis and the main analysis are based on velocity, which inherently requires both 'X' and 'Y' coordinates for calculation. This missing data might introduce some bias in our results, affecting the overall generalizability of our findings.

In addition to the visualizations, we also need to address a concern

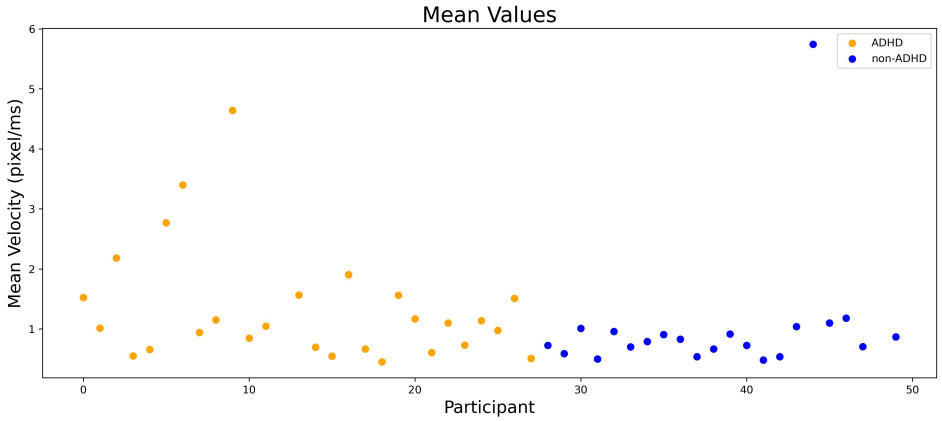

Figure 4.6: Mean Velocity for all participants.

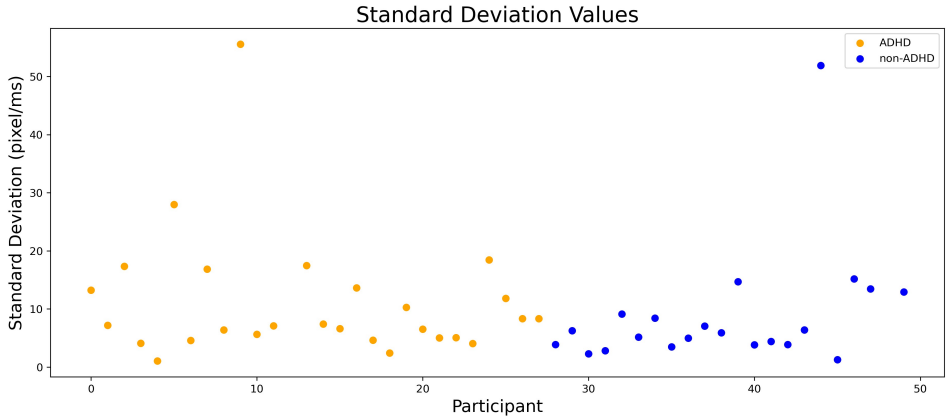

Figure 4.7: Standard Deviation of Velocity for all participants.

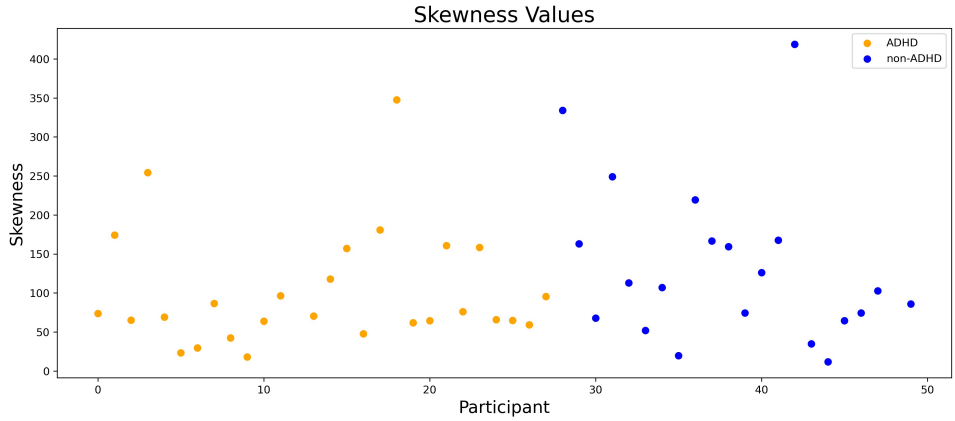

Figure 4.8: Skewness of Velocity for all participants.

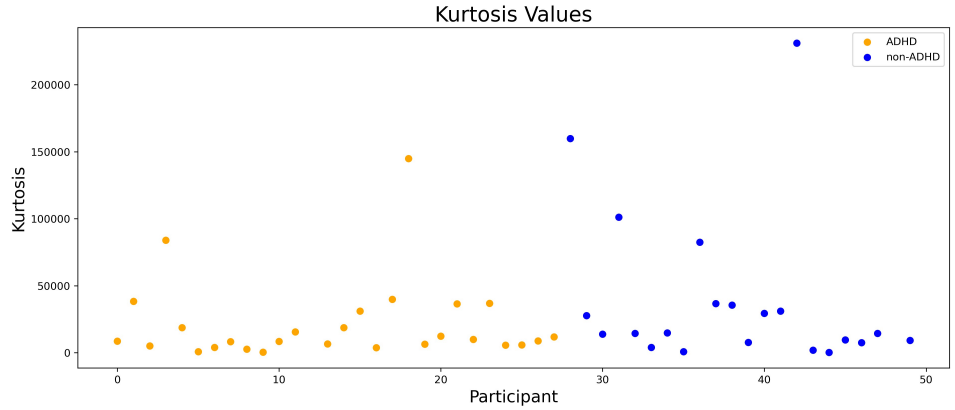

Figure 4.9: Kurtosis of Velocity for all participants.

regarding the consistency of the experimental setup reported in the dataset paper [14] and subsequent publications [100]. More specifically, there are discrepancies in the information provided about the distance to the monitor and the size of the monitor used in the experiments. It is known that eye-tracker data is very sensitive and should be collected with care and such inconsistencies may introduce potential biases impacting the validity of the derived metrics and, consequently, our results.

To account for these concerns, we will carefully examine the available information and, if necessary, consider the potential impact of these inconsistencies on our analysis. By addressing these concerns in the Data Pre-Analysis section, we can ensure that our subsequent investigation of the two assumptions and their derived metrics is based on a solid foundation, ultimately contributing to a more accurate and reliable understanding of the relationship between eye movement characteristics and ADHD diagnosis.

4.3 Feed and Fly Model: Eye-tracking Metrics

In this section, we present and discuss the results obtained from the analysis of eye-tracking data using the FSTR Algorithm, as described in Section 3.4.1. Our analysis focuses on four key metrics: the minimum FSTR value, the Cumulative Sum FSTR value, the saccade frequency, and the fixation duration, as outlined in Section 3.4.1.

Through our analysis, we found no statistically significant differences between the ADHD and non-ADHD groups for these metrics, suggesting that they may not be sufficient to differentiate between individuals with and without ADHD based on their eye-tracking data.

4.3.1 Minimum FSTR Value and Cumulative Sum FSTR Value

The FSTR value, derived directly from the optimization algorithm, is visually represented in the graph below, Figure 4.10. As previously mentioned, the lower FSTR value indicates better separation between fixations and saccades, suggesting that the algorithm has successfully identified distinct groups of these eye movements with minimal overlap or ambiguity. In the context of our hypothesis discussed in Section 3.4.1, we expect the FSTR value to differentiate between ADHD and non-ADHD groups.

Next, we explore the validity of our hypothesis by comparing the minimum FSTR values between the two populations. If supported, this finding would suggest that the FSTR algorithm captures a distinctive aspect of eye movement patterns in individuals with ADHD, which could be used as an additional parameter to improve our Benchmark model outlined in Section 3.3.

Similarly, we do also for the Cumulative Sum FSTR value, which according to our hypothesis would magnify the difference betweeen the groups even further.

Before performing statistical tests, visual inspection of the data is often

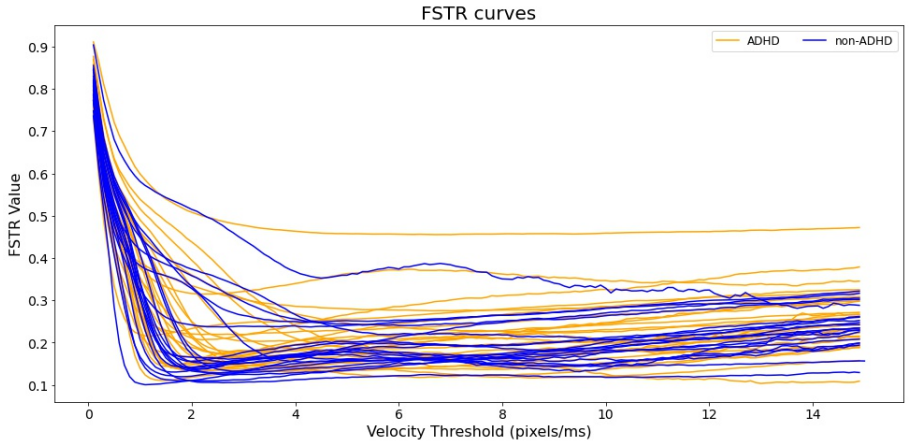

Figure 4.10: FSTR vs. Velocity Threshold for all participants.

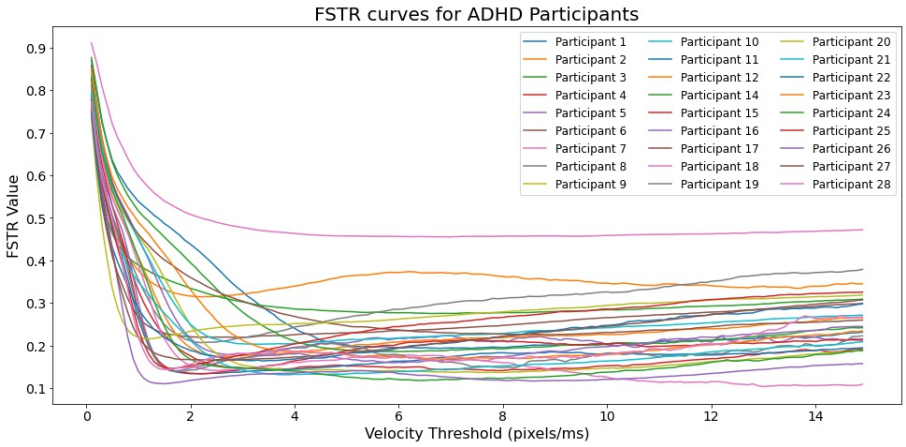

Figure 4.11: FSTR vs. Velocity Threshold for ADHD participants.

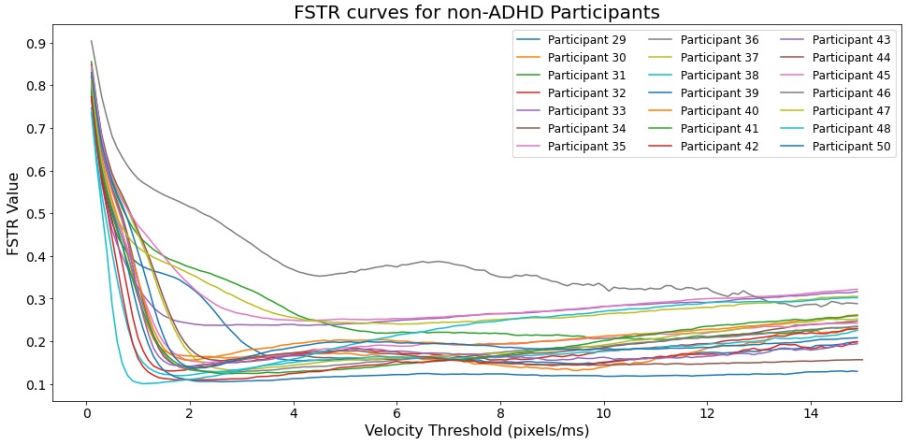

Figure 4.12: FSTR vs. Velocity Threshold for non-ADHD participants.

a useful first step in analyzing eye-tracking data. Upon examining the graphs for the two groups in Figures 4.11 and 4.12, it becomes evident that there is no visible difference between the shapes of the curves or the lower minimum FSTR values for both ADHD and non-ADHD participants. In both graphs, there are FSTR curves with more erratic behavior, and at first glance, there is no actual differentiation between the curves corresponding to the ADHD and non-ADHD groups.

To further clarify this visually, a summary of the corresponding minimum FSTR values is presented in Figure 4.13.

The same result we have also for the Cumulative Sum FSTR, as presented in the Figure 4.14, and more explicitely in the Figure 4.15

Next we proceed to investigate whether there are significant differences between the two groups, we performed independent samples t-tests on the 'Minimum FSTR Value' and 'Cumulative Sum FSTR Value' scores.

The t-test is a commonly used statistical test that allows us to determine if there is a significant difference between the means of two groups. The t-statistic measures the size of the difference between the two groups

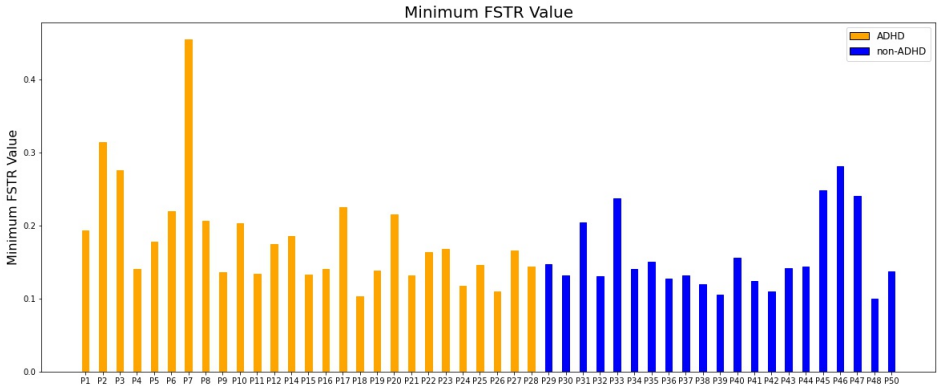

Figure 4.13: Minimum FSTR value for all participants.

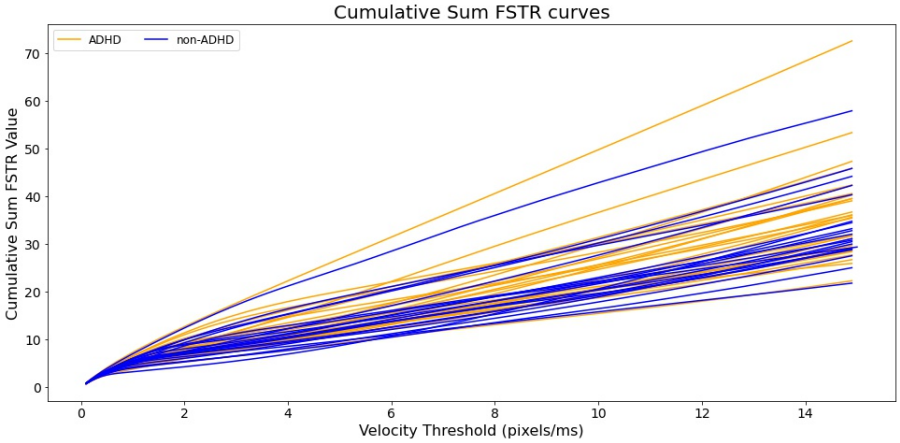

Figure 4.14: Cumulative Sum FSTR vs. Velocity Threshold for all participants.

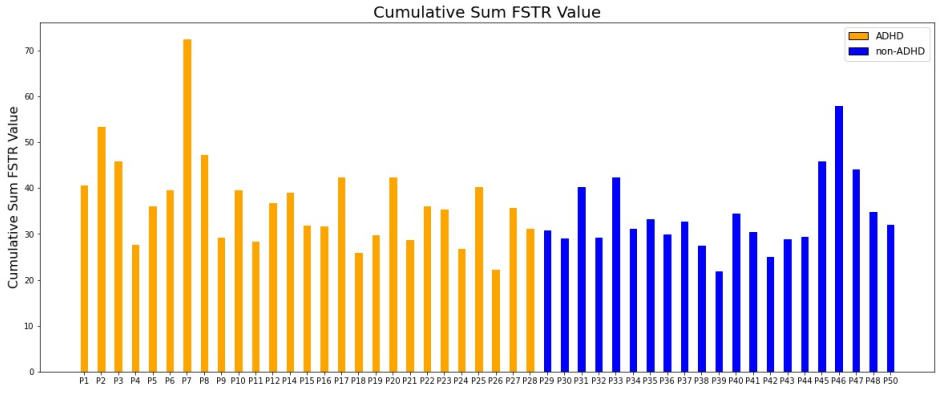

Figure 4.15: Cumulative Sum FSTR value for all participants.

relative to the variability in the sample data, and the p-value represents the probability of obtaining a result as extreme as the one observed, assuming that there is no real difference between the two groups.

Based on the results of the t-tests, there was no significant difference

between the means of the 'Minimum FSTR Value' and 'Cumulative Sum FSTR Value' scores for participants with ADHD and those without ADHD. The t-tests yielded t-statistics of 1.31 (p-value = 0.1967) for the 'Minimum FSTR Value' and 1.11 (p-value = 0.2720) for the 'Cumulative Sum FSTR Value'.

These results suggest that the FSTR algorithm was equally effective in identifying distinct groups of fixations and saccades with minimal overlap or ambiguity, regardless of the presence of ADHD. However, the absence of significant differences between the two groups does not necessarily mean that there are no differences between their eye movement patterns. Additional analyses using different metrics or measures may be necessary to more accurately differentiate between participants with and without ADHD based on their eye-tracking data.

Overall, the results of the t-tests suggest that the 'Min FSTR Value' and 'Cumulative FSTR Value' may not be strong enough metrics to differentiate between participants with and without ADHD based on their eye-tracking data. However, it is important to consider that these metrics are only one aspect of the eye-tracking data and may not be representative of the full range of differences in eye movement patterns between the two groups.

For that reason we will incorporate these two metrics on our Benchmark model to see if the benchmark accuracy will improve.

4.3.2 Saccade Frequency and Fixation Duration

In this section, we discuss the calculation of saccade frequency and mean fixation duration for each participant, based on the participant-specific threshold obtained using the FSTR algorithm. The participant-specific threshold, 4.16 corresponds to the minimum FSTR value and yields the most accurate separation of fixations and saccades for each individual's data. These metrics provide additional insights into the eye movement patterns of individuals with and without ADHD.

It would have been beneficial to compare the thresholds calculated using our FSTR algorithm with the commonly accepted fixed thresholds suggested by literature, typically at 30°/s, as discussed in Section 2.5. How-

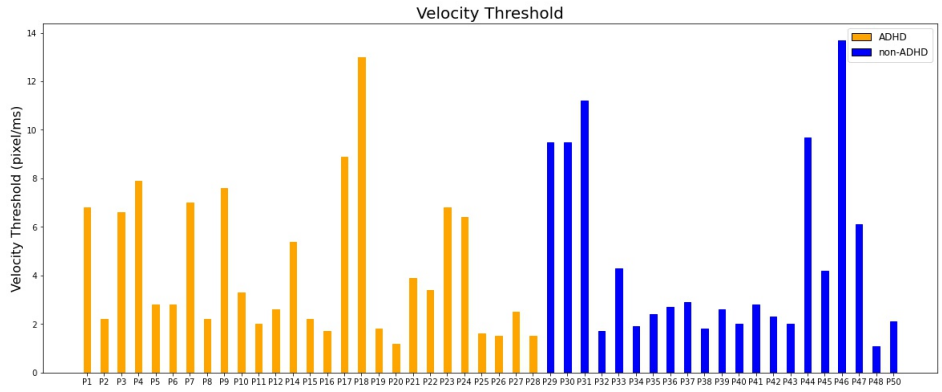

Figure 4.16: Velocity Threshold for all participants

ever, to convert the velocity from degrees/second to pixels/second, one needs to know the participant's distance from the monitor and the monitor size. Unfortunately, given the inconsistencies in the provided information about the experimental setup, particularly the distance from the monitor and the size of the monitor, we are unable to perform this comparison. This missed comparison opportunity may have offered additional insights and contributed to a more comprehensive understanding of our findings. Despite this, the results obtained using the FSTR algorithm still provide valuable insights into the differentiating eye movement patterns between the ADHD and non-ADHD groups.

Before we dive into the metrics, the scatter plots in Figure 4.17 show the eye movement data for selected ADHD and non-ADHD participants during the first 1-second duration of the eye-tracking task. Fixations and saccades are differentiated by color, providing a visual representation of the eye movement patterns for each participant.

Saccade frequency, as already presented in Section 2.6 is the number

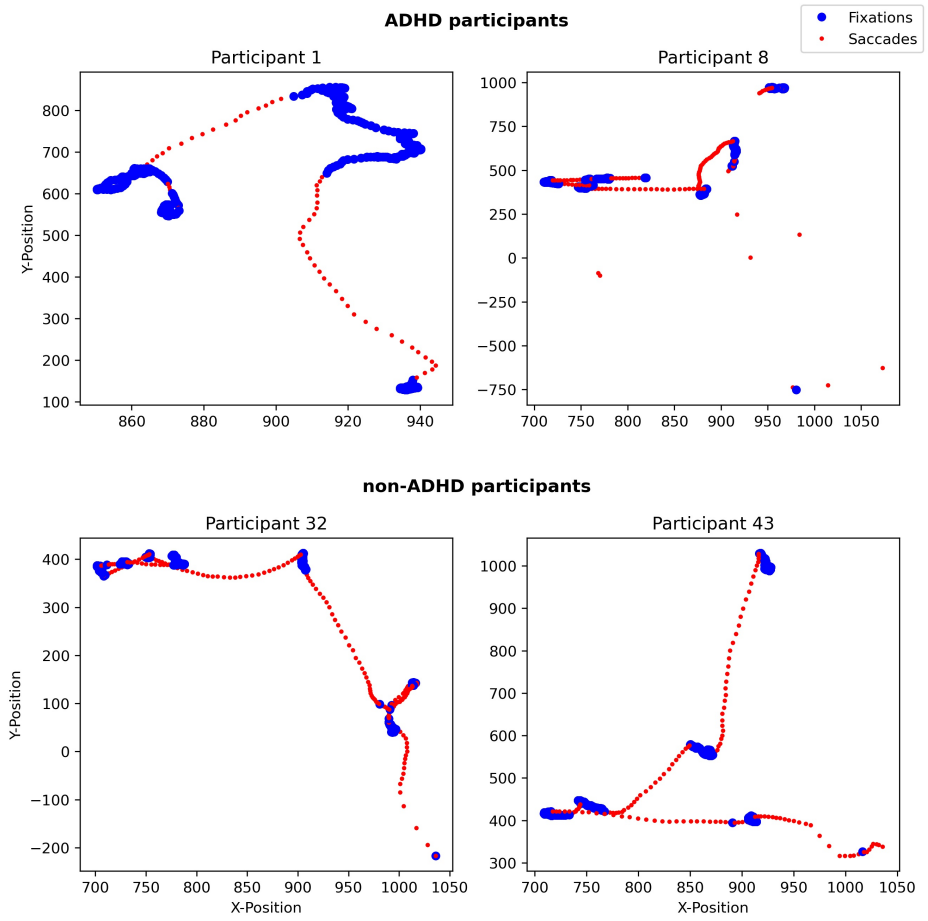

Figure 4.17: Scatter plots of eye movement data, for 1sec duration, for selected ADHD and non-ADHD participants, with fixations and saccades differentiated by color

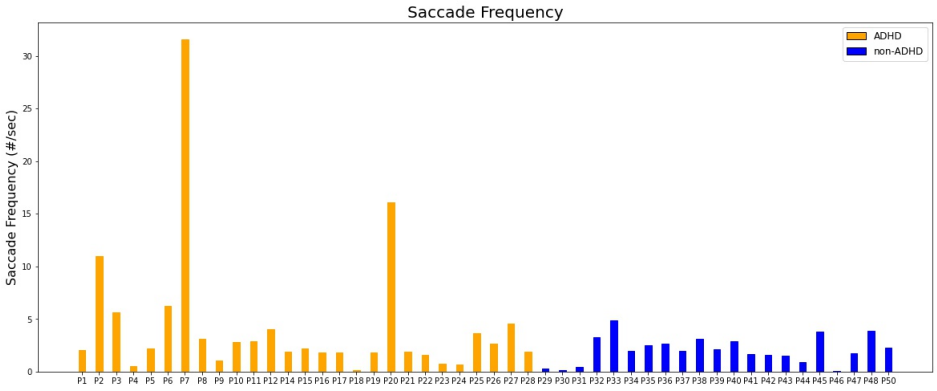

Figure 4.18: Saccade Frequency for all participants.

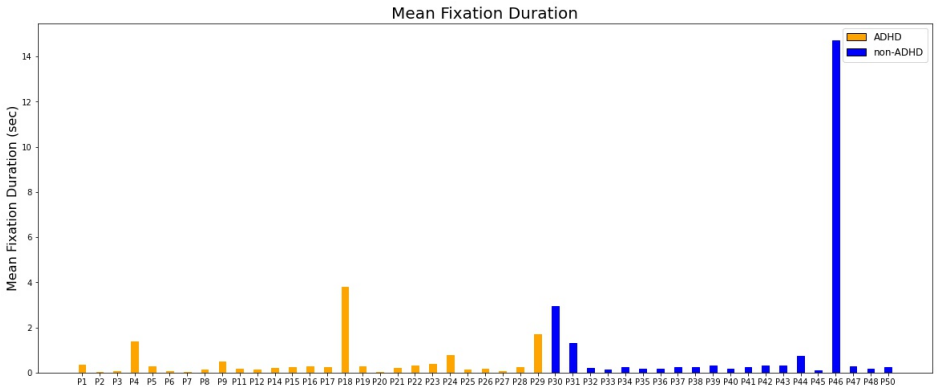

Figure 4.19: Fixation Duration for all participants.

of saccades per unit of time. To calculate this metric for each participant, we first identified the saccades using the participant-specific threshold determined by the FSTR algorithm. We then counted the number of saccades throughtout the complete test period and divided it by the duration of that period, 4.18. The resulting saccade frequency provides an indication of how often a participant's eyes make rapid movements.

Mean fixation duration is another important metric for understanding eye movement patterns, 2.6. To calculate this metric for each participant, we identified the fixations using the participant-specific threshold determined by the FSTR algorithm. We then calculated the duration of each fixation and computed the average duration across all fixations for that participant 4.19. This mean fixation duration offers insights into how long, on average, a participant's eyes remain relatively stable during the eye-tracking task.

However, it should be noted that Participant 10, who has been diagnosed with ADHD, was excluded from the mean fixation duration calculation. This participant's data exhibited an extraordinarily high fixation duration that stood out as an extreme outlier, significantly diverging from the rest of the dataset. This anomaly was likely due to a computational error rather than representing a realistic fixation duration.

Including this data point would skew the overall results, potentially creating a misleading representation of the mean fixation duration for the ADHD group. Consequently, we decided to exclude this participant from this specific analysis to maintain the accuracy and reliability of our findings.

Visually, from the graphs, we do not observe any major difference between the participants with and without ADHD in terms of saccade frequency and mean fixation duration. To further investigate whether there are significant differences between the two groups, we performed independent samples t-tests for both metrics. The t-test results for both Saccade Frequency and Mean Fixation Duration are as follows: Saccade Frequency: t-statistic = 1.58, p-value = 0.1214; Mean Fixation Duration: t-statistic = -1.23, p-value = 0.2255. These results indicate that there is no statistically significant difference between the ADHD and non-ADHD groups in terms of both metrics. This finding suggests that these metrics may not be sufficient to differentiate between individuals with and without ADHD based on their eye-tracking data.

However, similarly with the minimum FSTR value and the Cumulative Sum FSTR value, we will also incorporate these two metrics on our Benchmark model to see if the benchmark accuracy will improve.

4.3.3 Incorporate the Metrics into the Benchmark Model

Now that we have established the four key metrics: the minimum FSTR value, the Cumulative Sum FSTR value, the saccade frequency, and the fixation duration, we will incorporate them into our Benchmark model, to evaluate if its accuracy will be improved.

As it is presented in the Figures 4.20, 4.21, in none of the cases any of the four metrics contributed to an improvement of the benchmark model. On the contrary, by incorporating these metrics, the accuracy of the model dropped by 4pp when using the Cumulative Sum FSTR value, and by 2pp for the other 3 metrics.

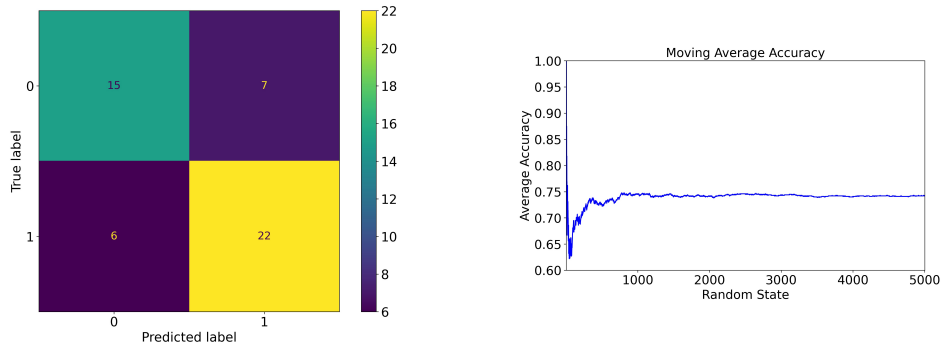

Figure 4.20: Confusion Matrix and Moving Average for the Benchmark Model inclunding either of the 3 metrics (Minimum FSTR Value, Saccade Frequency, Fixation Duration).

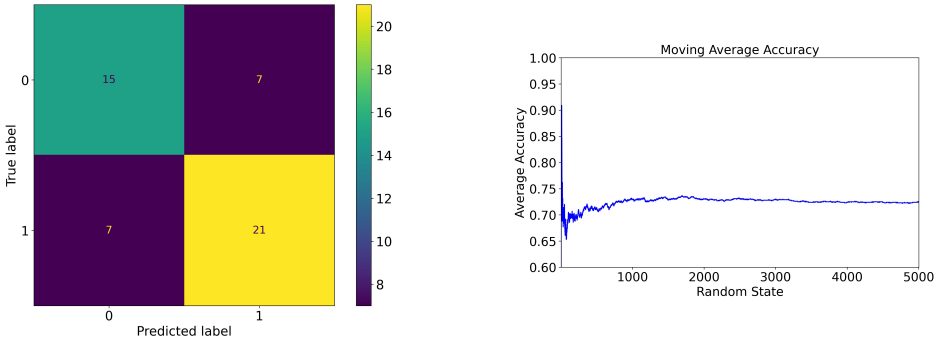

Figure 4.21: Confusion Matrix and Moving Average for the Benchmark Model inclunding the Cumulative Sum FSTR Value.

4.4 Lévy Flight Model: Eye-tracking Metrics

In this section, we present the findings of our investigation into the Lévy flight model as a method for characterizing eye-tracking data, as detailed in the Methodology Section 3.4.2. Our analysis concentrates on the distribution of velocities for individual participants, the scaling exponent, and the model's goodness of fit, with the goal of identifying potential differences in eye-tracking patterns between ADHD and non-ADHD participants.

4.4.1 Lévy Exponent

Following the procedure outlined in the methodology section, we processed the eye-tracking data, computed velocities, and generated histograms for each participant. In Figure 4.22 we present a representative example of 4 participants (2 with ADHD and 2 without ADHD). We then fitted a linear regression model to the log-transformed histograms and estimated the scaling exponent of the power-law distribution. As part of the linear regression model fitting, we calculated the R-squared values (coefficients of determination) for each participant, offering insights into the goodness of fit for the Lévy flight model and the participants' eye-tracking patterns.

To identify the most representative power-law region within the velocity histograms, we employed a factor to define the region's upper limit. This factor is multiplied by the location of the mode (lmode) in the histogram to establish the range where power-law behavior is observed. We iteratively tested various factors, computed the R-squared values for each participant using the linear regression model, and selected the factor that yielded the highest average R-squared value across all participants, as illustrated in Figure 4.23. By opting for the factor that results in the maximum average R-squared value, we aim to capture the optimal power-law behavior in the data for the majority of participants.

Based on the selected factor, the corresponding power-law region is illustrated in Figure 4.24 for the sample of the 4 participants (2 with ADHD

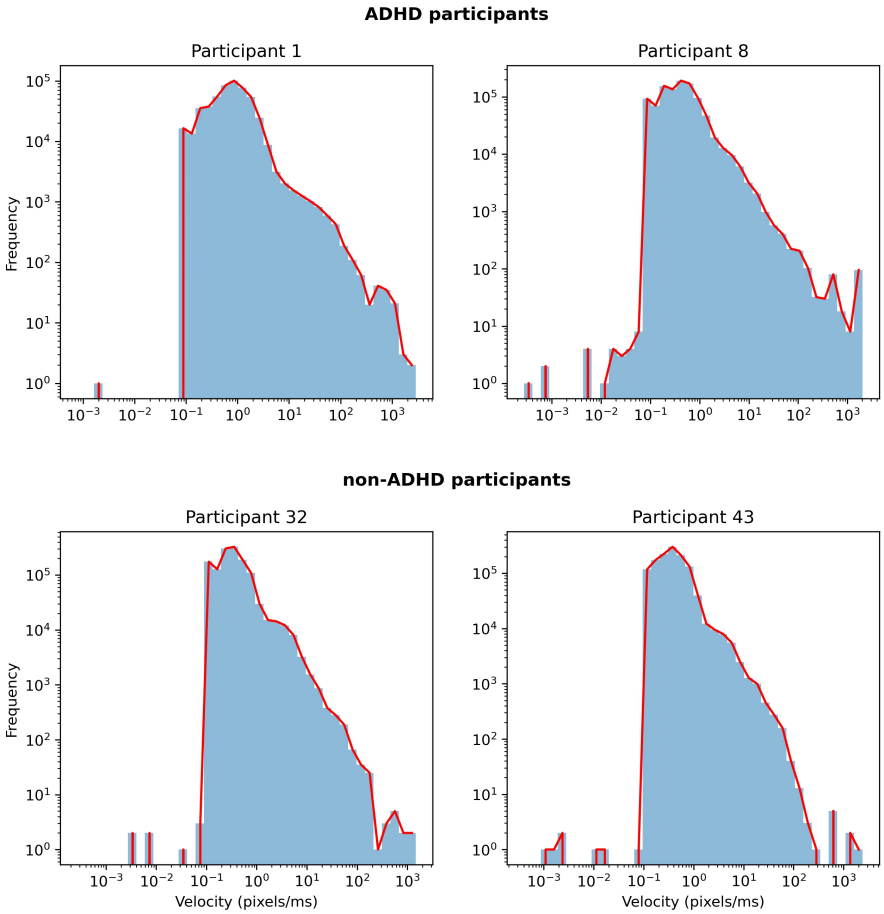

Figure 4.22: Velocity Histograms.

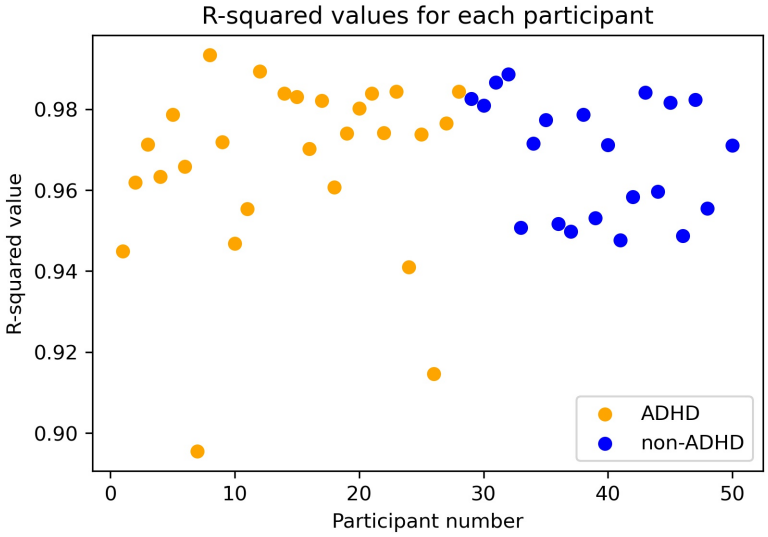

Figure 4.23: R-squared values per participant for the optimal factor.

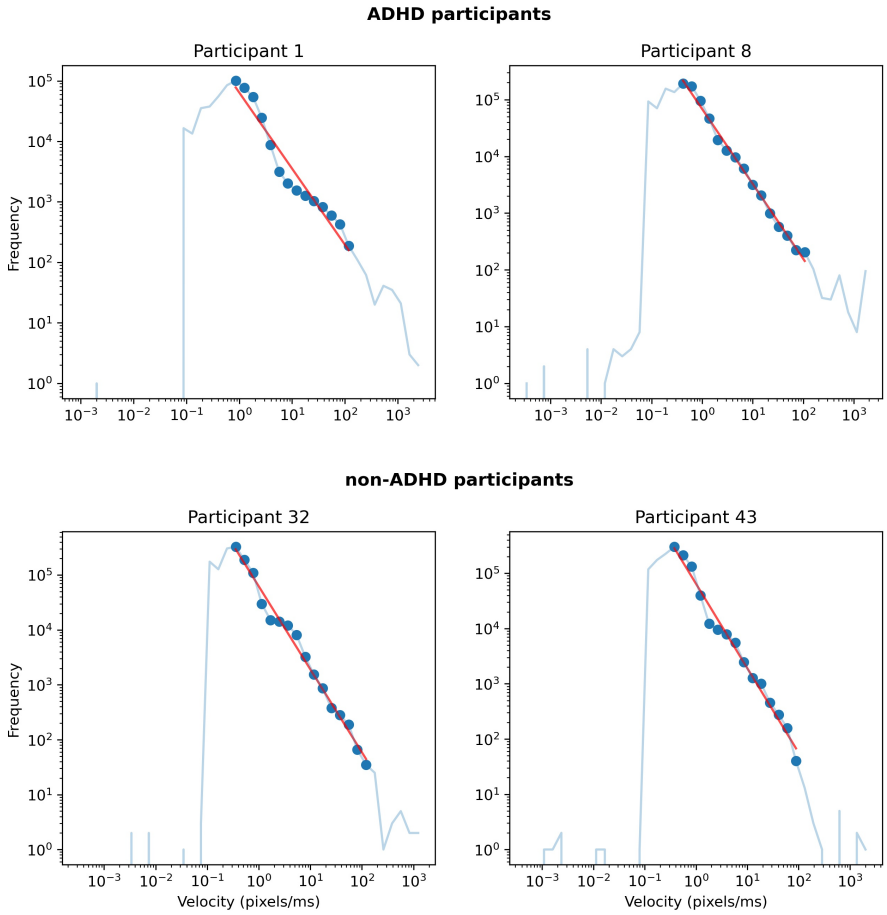

Figure 4.24: Selected Power-Law Region and Fitted Linear Regression Model.

and 2 without ADHD). In this region, as previously mentioned, we fitted a linear regression model (red line) and estimated the scaling exponent of the power-law distribution, as shown in Figure 4.25, and for all the participants in the Figure 4.26.

Average Exponent Values

The average exponent values for the ADHD and non-ADHD groups were calculated based on the Lévy Exponents obtained from the linear regression models. The average exponent values are as follows:

- Mean Value of the exponents of the ADHD group: -1.38
- Mean Value of the exponents of the non-ADHD group: -1.43

These average values indicate that the ADHD group has a slightly lower exponent value than the non-ADHD group, which aligns with our hypothesis that participants with ADHD may have smaller scaling exponents compared to those without ADHD.

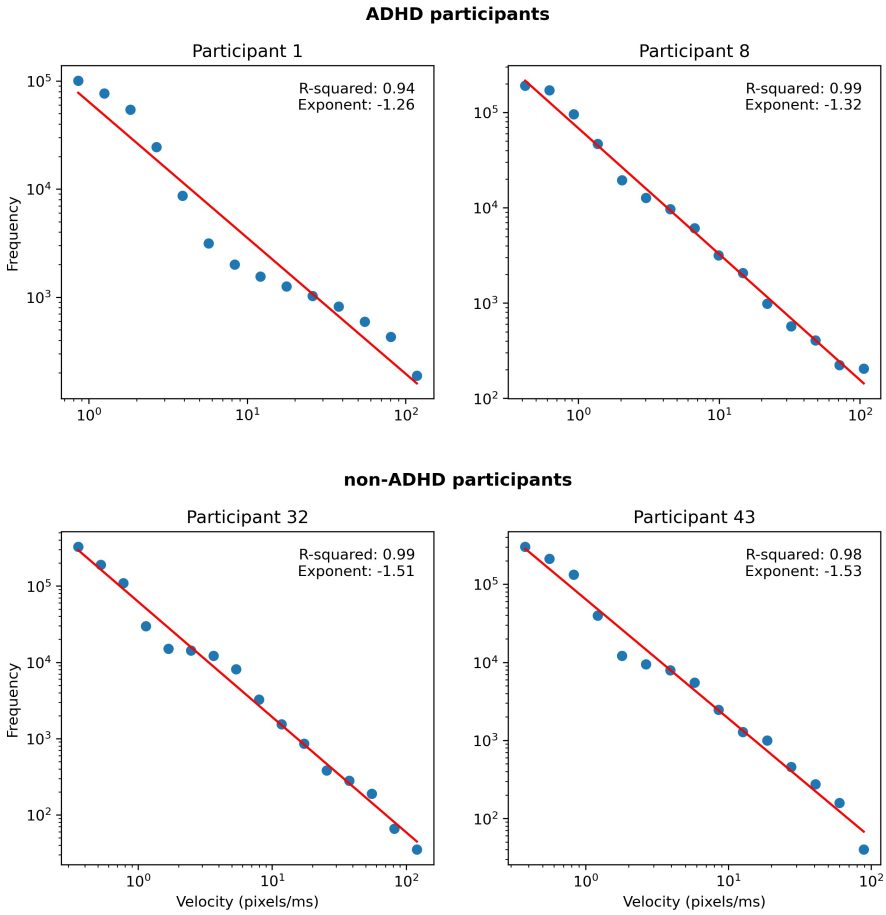

Figure 4.25: Fitted Linear Regression Model, R-squared, Lévy Exponent.

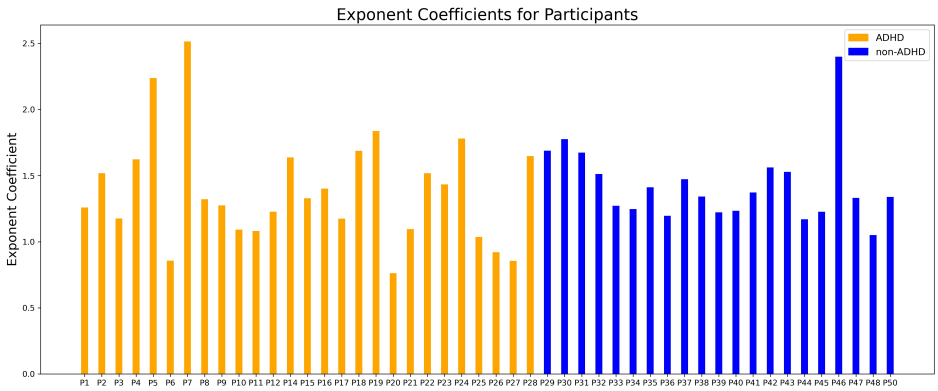

Figure 4.26: Lévy Exponent for all participants.

Comparison to Expected Exponent Values

In the Mathematical Modelling Section 2.2.1 and Methodology Section 3.4.2, we discussed the limited literature on expected exponent values for visual scanpaths. The study by Credidio H. et. al. [40] reported a scaling exponent value close to 2.9, while the general exponent value for foraging

models is approximately 2.

The average exponent values obtained in our study (-1.38 for ADHD and -1.43 for non-ADHD) fall outside the expected range from the literature. This discrepancy may be due to the specific characteristics of our sample or the eye-tracking data. Nevertheless, it is essential to take these differences into account when interpreting the results and considering their implications for the classification of participants with and without ADHD.

T-Test Comparison between the two Groups

To further investigate whether there are significant differences between the two groups, we performed an independent samples t-test. The t-test result is as follows: t-stat = 0.459, p-value = 0.648. This result indicates that there is no statistically significant difference between the ADHD and non-ADHD groups in terms of Lévy Exponent metric. This finding suggests that this metric alone may not be sufficient to differentiate between individuals with and without ADHD based on their eye-tracking data.

Incorporating the Lévy Exponent metric into our benchmark model will provide additional insights into whether this metric can contribute to improving the classification accuracy between individuals with ADHD and those without the condition. By considering the average exponent values and comparing them to the expected values from the literature, we can better understand the eye-tracking patterns of our participants and potentially uncover new insights into the differences between individuals with and without ADHD.

However, the absence of significant differences between the two groups does not necessarily mean that there are no differences between their eye movement patterns. Additional analyses may be necessary to more accurately differentiate between participants with and without ADHD based on their eye-tracking data. For that reason we will incorporate this metric on our Benchmark model to see if the benchmark accuracy will improve.

4.4.2 Incorporate the Metric into the Benchmark Model

Now that we have established the Lévy Exponent metric, we will incorporate it into our Benchmark model, to evaluate if its accuracy will be improved.

As it is presented in the Figure 4.27, the metric did not contribute to an improvement of the benchmark model. Unlike the metrics of the Feed and Fly Model, 4.3, which worsened the Benchmark Model's accuracy, the incorporation of the Lévy Exponent metric maintained the model's accuracy, which could be promising for its future usability.

However, our analysis so far has been conducted on the entire eye-tracking dataset for each participant, without considering the specific events during the memory test. This motivates us to perform an event-based analysis, which might uncover more nuanced differences in the eye movement patterns of ADHD and non-ADHD participants.

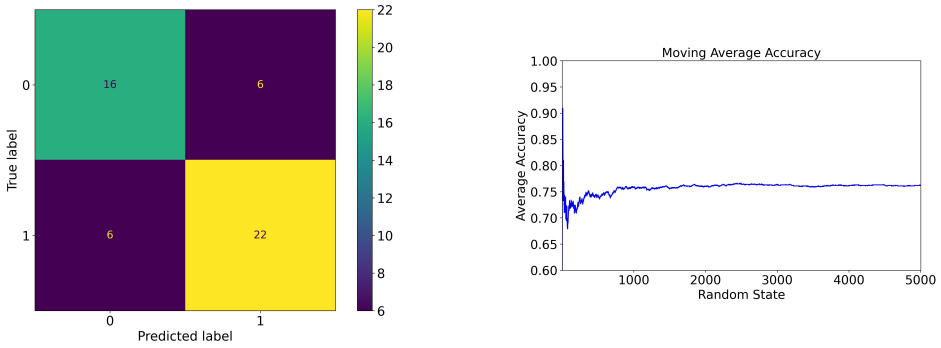

Figure 4.27: Confusion Matrix and Moving Average for the Benchmark Model including the Lévy Exponent.

4.4.3 An Event-Base Analysis

The initial analysis of the Lévy Exponents was conducted using the entire eye-tracking data for each participant, without separating the data based on specific events.

While this approach can offer valuable insights into the overall eye movement patterns of ADHD and non-ADHD participants, it may be interesting to also explore potential differences in eye-tracking patterns during specific events. By analyzing the data on an event-by-event basis, we might be able to uncover more nuanced differences in eye movement patterns between the two groups during various stages of the memory test, such as when focusing on specific memoranda or distractor, as already described in Table 3.1. This additional analysis could potentially provide further insights into the unique eye-tracking characteristics of individuals with and without ADHD, and contribute to a better understanding of the underlying differences between the two groups.

Building upon this idea, we have now performed an event-based analysis of the Lévy Exponents to explore whether analyzing the eye-tracking data on an event-by-event basis would reveal more significant differences between the ADHD and non-ADHD groups. The results of this analysis are presented in Table 4.3.

The event-based analysis reveals statistically significant differences in the Lévy Exponents between the ADHD and non-ADHD groups for Events 1 and 4 (p-values < 0.05). For the other events, the differences are not statistically significant (p-values > 0.05).

These findings suggest that the eye movement patterns of individuals with and without ADHD may differ more significantly during certain stages of the memory test. Specifically, the analysis indicates that the ADHD and non-ADHD groups exhibit different eye-tracking characteristics when viewing one-dot memoranda (Event 1) and empty-grid distractor (Event 4). This insight can be valuable for better understanding the eye movement patterns of individuals with and without ADHD during the memory test and could potentially improve the classification accuracy of the models by incorporating the event-based Lévy Exponents.

Table 4.3: The t-test Results for the Lévy Exponents of the two Groups (ADHD and non-ADHD) per Event

	t-statistic	p-value
Event 1	2.490	0.016
Event 2	1.748	0.087
Event 3	1.591	0.118
Event 4	2.046	0.046
Event 5	1.111	0.272
Event 6	0.666	0.509
Event 7	1.212	0.232
Event 8	1.564	0.125
Event 9	0.780	0.439
Event 10	0.129	0.898

For the purposes of our study we will focus on the Event 1 case, which has the lowest p-value. In terms of data size, the Event 1 case corresponds approximately to a 23% of the complete dataset, ranging from 20% - 25% depending on the participant.

Building on this, we incorporated the Lévy Exponent metric into our Benchmark model to evaluate if the model's accuracy could be improved. In our initial attempt, the accuracy of the benchmark model was maintained at the same level of 76%, as shown in Equation 4.1. In this case, we used the default benchmark model with its four metrics (Age, IQ, FD, and Performance) along with the Lévy Exponent for the Event 1 analysis, as presented in Table 4.3.

However, after utilizing Recursive Feature Elimination with Cross-Validation (RFECV) and GridSearchCV, outlined in Section 3.3.2, the feature selection was refined to include only three metrics: FD, Lévy Exponent, and Performance.

The RFECV process resulted in the exclusion of the Age and IQ metrics. This decision by the RFECV process aligns with previous observations from our benchmark model and findings in the literature. In our initial benchmark model, IQ had a low impact, as indicated by the relatively small coefficient in Equation 4.2. The impact of IQ was less compared to other features such as Performance and FD, suggesting that its role in classifying ADHD might not be as significant in our model. Furthermore, the exclusion of IQ also aligns with the contradictory findings regarding the relationship between IQ and ADHD in the literature. Given these contrasting findings, the exclusion of IQ from our refined model might indeed help to eliminate potential noise and enhance the overall predictive accuracy.

This feature selection refinement, along with hyperparameter tuning, improved the accuracy of our model. As shown in Figure 4.28, the inclusion of the Lévy Exponent metric notably enhanced the performance of the model. Unlike previous metrics and attempts, this addition resulted in an

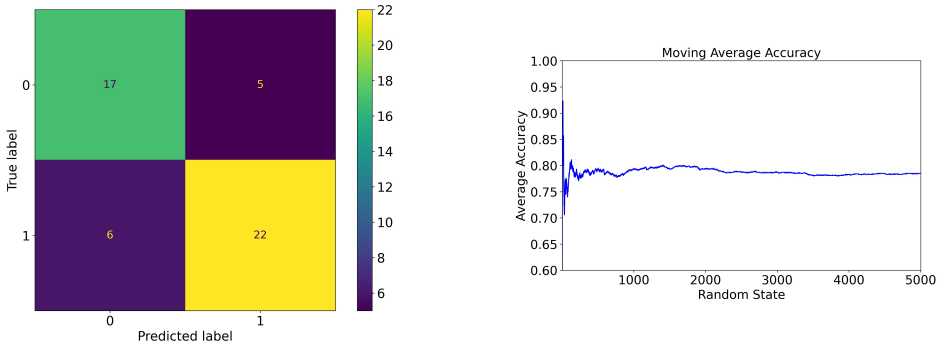

Figure 4.28: Confusion Matrix and Moving Average for the Benchmark Model including the Lévy Exponent of the Event 1.

increase in accuracy by 2 percentage points, reaching an accuracy of 78%, as presented in Equation 4.3.

$$Accuracy = \frac{39}{50} = 0.78 \tag{4.3}$$

In terms of performance metrics, our model achieved a precision of 0.815, a recall of 0.786, and a specificity of 0.773. This demonstrates the model's improved capacity to correctly identify ADHD cases and predict non-ADHD cases accurately. The detailed confusion matrix is presented in Figure 4.28.

Subsequent to this feature refinement and performance improvement, we derived the logistic regression equation for the optimized model, termed $\mathbf{Model_{1/50}}$. The equation is as follows::

$$\text{logit}(p) = 9.7586 - 0.1489 \times FD - 1.1973 \times \textit{Lévy Exp.} - 6.7556 \times \textit{Perf.} \tag{4.4}$$

This equation quantifies the relationship between the logit transformation of the probability of a positive ADHD classification and the selected metrics, facilitating a deeper understanding of the model's predictive capabilities and the influence of each metric on the outcome.

The logistic regression equation for the revised $\mathbf{Model_{1/50}}$ suggests a consistent influence of certain variables compared to the initial model, Equation 4.2.

The Performance metric stands out as a significant predictor of ADHD status in the revised model, which is consistent with its role in the initial model where it also had the highest impact.

Arithmetic and Digital Span (FD), although having a lesser impact, still plays a notable role, mirroring its contribution in the initial model.

The introduction of the Event-based Lévy Exponent in the revised model adds a new dynamic. Its substantial coefficient highlights its importance, suggesting it as a valuable addition to the predictive model.

4.5 Discussion of Results and their Implications

Throughout this investigation, we examined two distinct models - the Feed and Fly Model and the Lévy Flight Model - to analyze and characterize eye-tracking data. We aimed to identify discernible differences in eye-tracking patterns between ADHD and non-ADHD participants.

Our initial exploration centered around the Feed and Fly Model, which incorporated a novel algorithm for dynamic velocity threshold, known as the Fixation-Saccade Transition Ratio (FSTR) Algorithm. This algorithm, designed to offer a more accurate and individualized analysis of eye-tracking data, represents a promising approach to improving the mathematical modeling of eye movements, with potential applications for understanding attentional disorders like ADHD. Despite this, the Feed and Fly Model failed to offer significant differentiation between the ADHD and non-ADHD groups based on the evaluated metrics. Furthermore, incorporating these metrics into our benchmark model led to a decrease in classification accuracy. This suggests that these metrics may not be particularly suitable for predicting ADHD status using our current dataset.

Subsequently, we turned to the Lévy Flight Model, offering a unique approach for characterizing eye-tracking patterns. The Lévy Exponent metric, derived from velocity histograms, hinted at potential variations in eye-tracking behaviors between ADHD and non-ADHD groups. Yet, a preliminary analysis did not reveal any statistically significant differences between the groups.

Despite this, we were driven to further our investigation by conducting an event-based analysis. This examination unveiled statistically significant differences in Lévy Exponents between the groups for specific events. Remarkably, integrating the Lévy Exponent metric from this event-based analysis into our benchmark model led to an enhancement in model performance.

Despite the mixed results from the two models, our research journey underscored the significance of rigorous data analysis and the exploration of varied models and perspectives. Each step, regardless of the outcomes, enriched our understanding of eye-tracking patterns and their potential relation to ADHD. Moreover, this study can serve as a launching pad for future research in this domain, aiming to refine diagnostic tools and our understanding of ADHD.

Interestingly, the three key metrics contributing to the final model each represent different domains of analysis: The Arithmetic and Digital Span (FD) from the WISC test, Performance scores from the memory tests, and the Lévy Exponent derived from eye-tracking analysis. This highlights the necessity for a multi-faceted approach in understanding and diagnosing ADHD, signifying that a single metric or test may not offer a comprehensive picture of the disorder.

Despite the promising findings, this study isn't without limitations. The first limitation is the small dataset size. Restricted to merely 50 subjects, the size of the dataset inevitably limited our analysis. Models trained on such a small dataset could only achieve a maximum accuracy of approximately

70%-80%, and the diminutive size of the test dataset could potentially bias our final results.

The second limitation revolves around the credibility of the experimental setup. There are questions surrounding this, given the conflicting information about the participants' distance from the screen and the lack of details regarding the monitor size.

The third issue pertains to the quality of the eye-tracking data. This type of data is notoriously sensitive and requires meticulous collection processes. Extreme values observed in the velocity skewness and kurtosis raise concerns about the dataset's reliability. Of particular note, the eye-tracking data for two participants - one with ADHD and one without - were incomplete, missing all y-coordinate information. This further compounds the challenges faced in the analysis and potentially limits the generalizability of our findings. Consequently, it is plausible that this dataset may not meet the quality standards necessary for this study.

Lastly, the potential misdiagnosis of ADHD introduces inherent uncertainty. ADHD diagnosis is complex and error-prone, leading to the possibility of misdiagnosis. In a dataset of this size, any misclassified data could significantly impact our analysis and models.

Despite these limitations, this study offers a valuable contribution to the field, providing a stepping stone for future research. We hope that our findings will inspire further exploration into the potential of eye-tracking data in enhancing our understanding and diagnosis of cognitive disorders like ADHD.

Chapter 5

Conclusion

In this thesis, we undertook the ambitious endeavor of enhancing the prediction of Attention-Deficit/Hyperactivity Disorder (ADHD) using eye-tracking data and mathematical modelling. Our objective was to explore the intricate relationship between oculomotor control, neurocognitive traits, and ADHD symptoms, through the lens of eye movement characteristics and mathematical models.

To this end, we adopted a two-fold approach: an initial classification model based on non-eye-tracking metrics, and the subsequent inclusion of eye-tracking metrics derived from two mathematical models, the "Feed and Fly" and Lévy Flight models. Our findings from these investigations have been enlightening, albeit with a mix of results.

The "Feed and Fly" model did not yield significant differentiation between ADHD and non-ADHD groups, yet it added to our understanding of the complex nature of ADHD. On the other hand, the Lévy Flight model, particularly when applied in an event-based analysis, showed promise and led to an enhancement in the performance of our benchmark model.

One of the most striking findings of our study has been the refinement of the feature selection in our model to only three metrics: Arithmetic and Digital Span (FD), Lévy Exponent, and Performance, which represent different aspects of the analysis - WISC test, eye-tracking analysis, and memory test respectively. This highlights that there is no singular approach to diagnosing ADHD, but rather a combination of various metrics from distinct domains that together can improve diagnostic accuracy.

However, it's important to note that our findings do not conclusively confirm our null hypothesis of significant differences in eye movement characteristics between children with and without ADHD. Instead, they offer a glimpse into the potential of eye-tracking data as a diagnostic tool and lay the foundation for future research in this field.

Moving forward, it is important to address the limitations of the present study in future research. With larger and more reliable datasets, the application of the mathematical models explored in this study could be further evaluated. The exploration of additional eye-tracking and non-eye-tracking metrics may also enhance our understanding of ADHD and its diagnosis.

Aside from the direct findings, this research also has broader implications. By demonstrating that eye-tracking data and mathematical modelling can provide useful insights, it encourages their inclusion in future research and clinical practice. Further, it underlines the need for a shift from reliance on single-domain diagnostic tools to a more encompassing approach. Lastly, by striving to develop a more objective and precise diagnostic tool, it aims to improve the quality of care for individuals with ADHD, with potential for enhancing academic achievement, social interactions, and overall quality of life. Future research should continue to explore these promising avenues.

In conclusion, the path to understanding ADHD is complex and nonlinear, filled with challenges and surprises alike. While there is much left to explore, we remain hopeful and committed to this endeavor. We anticipate that the findings from this study will inspire further research, not only in ADHD diagnosis but also in the broader field of cognitive disorders.

Bibliography

[1] I. N. Soria et al. 'Detecting Differences between Clinical Presentations in ADHD through the Cognitive Profile Obtained from WISC-IV'. In: *Universal Journal of Psychology* 21 (2017), pp. 179–186. DOI: 10.13189/ujp.2017.050403.

[2] M. L. Wolraich et al. 'Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents'. In: *Pediatrics* 144 (2019). DOI: 10.1542/peds.2019-2528..

[3] T. Gunnerson. *What's the History of ADHD?* 2022. URL: https://www.webmd.com/add-adhd/adhd-history (visited on 29/12/2022).

[4] Centers for Disease Control and Prevention. *Data and Statistics About ADHD*. 2022. URL: https://www.cdc.gov/ncbddd/adhd/data.html (visited on 29/12/2022).

[5] R. Thomas et al. 'Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis'. In: *Pediatrics* 135 (2015). DOI: 10.1542/peds.2014-3482.

[6] Centers for Disease Control and Prevention. *Mental Health Surveillance Among Children — United States, 2013–2019*. 2022. URL: https://www.cdc.gov/mmwr/volumes/71/su/su7102a1.htm (visited on 29/12/2022).

[7] M. Gaub and C. L. Carlson. 'Gender differences in ADHD: a meta-analysis and critical review'. In: *Journal of the American Academy of Child and Adolescent Psychiatry* 36 (1997), pp. 1036–1045. DOI: 10.1097/00004583-199708000-00011.

[8] M. M. Gottesman. 'Helping parents make sense of ADHD diagnosis and treatment'. In: *Journal of Pediatric Health Care* 17 (2003), pp. 149–153. DOI: 10.1067/mph.2003.55.

[9] B. Kiely and A. Adesman. 'What we do not know about ADHD... yet'. In: *Current Opinion in Pediatrics* 23 (2015), pp. 395–404. DOI: 10.1097/MOP.0000000000000229.

[10] J. M. Swanson et al. 'Etiologic subtypes of attention-deficit/hyperactivity disorder: brain imaging, molecular genetic and environmental factors and the dopamine hypothesis'. In: *Neuropsychology Review* 17 (2007), pp. 39–59. DOI: 10.1007/s11065-007-9019-9.

[11] T. W. Frazier et al. 'ADHD and achievement: Meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students'. In: *Journal of Learning Disabilities* 40 (2007), pp. 49–65. DOI: 10.1177/00222194070400010401.

[12] M. Skounti, A. Philalithis and E. Galanakis. 'Variations in prevalence of attention deficit hyperactivity disorder worldwide'. In: *European Journal of Pediatrics* 166 (2007), pp. 117–123. DOI: 10.1007/s00431-006-0299-5.

[13] A. Crippa et al. 'The utility of a computerized algorithm based on a multi-domain profile of measures for the diagnosis of attention deficit/hyperactivity disorder'. In: *Front Psychiatry* 8 (2017). DOI: 10.3389/fpsyt.2017.00189.

[14] D. Rojas-Líbano et al. 'A pupil size, eye-tracking and neuropsychological dataset from ADHD children during a cognitive task'. In: *Scientific Data 6* 25 (2019). DOI: 10.1038/s41597-019-0037-2.

[15] Lumen. *Anatomy of the Visual System*. 2022. URL: https://courses.lumenlearning.com/waymaker-psychology/chapter/vision/ (visited on 15/01/2023).

[16] BrainFacts/SfN. *Vision: Processing Information*. 2016. URL: https://www.brainfacts.org/thinking-sensing-and-behaving/vision/2012/vision-processing-information (visited on 20/02/2023).

[17] S. Djamasbi. 'Eye Tracking and Web Experience'. In: *AIS Transactions on Human-Computer Interaction* 6 (2014). DOI: 10.17705/1thci.00060.

[18] J. F. Hejtmancik et al. 'Vision'. In: *Conn's Translational Neuroscience* (2017), pp. 399–438. DOI: 10.1016/B978-0-12-802381-5.00031-2.

[19] A. Duchowski. *Eye tracking methodology: Theory and practice*. Springer, 2017.

[20] G. Boccignone and M. Ferraro. 'Feed and fly control of visual scanpaths for foveation image processing'. In: *Ann. Telecommun.* 68 (2013), pp. 201–217. DOI: 10.1007/s12243-012-0316-9.

[21] A. Coutrot, J.H. Hsiao and A.B. Chan. 'Scanpath modeling and classification with hidden Markov models'. In: *Behav Res* 50 (2018), pp. 362–379. DOI: 10.3758/s13428-017-0876-8.

[22] Wikipedia contributors. *Markov chain — Wikipedia, The Free Encyclopedia*. 2023. URL: https://en.wikipedia.org/w/index.php?title=Markov_chain&oldid=1150001760 (visited on 16/04/2023).

[23] A. M. Helmenstine. *An Introduction to Brownian Motion*. 2019. URL: https://www.thoughtco.com/brownian-motion-definition-and-explanation-4134272 (visited on 16/04/2023).

[24] S. Creel et al. 'Hidden Markov Models reveal a clear human footprint on the movements of highly mobile African wild dogs'. In: *Scientific Reports* 10.1 (2020). DOI: 10.1038/s41598-020-74329-w.

[25] A. Edwards et al. 'Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer'. In: *Nature* 449 (2007), pp. 1044–1048. DOI: 10.1038/nature06199.

[26] D. A. Raichlen et al. 'Evidence of Lévy walk foraging patterns in human hunter–gatherers'. In: *Proceedings of the National Academy of Sciences* 111.2 (2014), pp. 728–733. DOI: 10.1073/pnas.1318616111.

[27] H. Yarahmadi and A. A. Saberi. 'A 2D Lévy-flight model for the complex dynamics of real-life financial markets.' In: *Chaos* 32 3 (2022), pp. 033–113. DOI: 10.1063/5.0082926.

[28] R. Metzler and J. Klafter. 'The random walk's guide to anomalous diffusion: a fractional dynamics approach'. In: *Physics Reports* 339.1 (2000), pp. 1–77. DOI: 10.1016/S0370-1573(00)00070-3.

[29] P. Błażejczyk and M. Magdziarz. 'Stochastic modeling of Lévy-like human eye movements'. In: *Chaos* 31 (2021). DOI: 10.1063/5.0036491.

[30] D. Brockmann and T. Geisel. 'The ecology of gaze shifts'. In: *Neurocomputing* 32-33 (2000), pp. 643–650. DOI: 10.1016/S0925-2312(00)00227-7.

[31] Wikipedia contributors. *Lévy flight — Wikipedia, The Free Encyclopedia*. 2023. URL: https://en.wikipedia.org/w/index.php?title=L%C3%A9vy_flight&oldid=1142022673 (visited on 26/03/2023).

[32] B. Dybiec et al. 'Lé vy flights versus Lévy walks in bounded domains'. In: *Physical Review E* 95.5 (2017). DOI: 10.1103/physreve.95.052102.

[33] G. M. Viswanathan et al. 'Lévy flights in random searches'. In: *Physica A: Statistical Mechanics and its Applications* 282 (Jan. 2009), pp. 1–12. DOI: 10.1016/S0378-4371(00)00071-6.

[34] M. F. Shlesinger and J. Klafter. 'Lévy Walks Versus Lévy Flights'. In: *On Growth and Form: Fractal and Non-Fractal Patterns in Physics*. Dordrecht: Springer Netherlands, 1986, pp. 279–283. DOI: 10.1007/978-94-009-5165-5_29.

[35] J. Klafter and I. M. Sokolov. *First Steps in Random Walks: From Tools to Applications*. Oxford University Press, 2011. DOI: 10.1093/acprof:oso/9780199234868.001.0001.

[36] S. Buldyrev et al. 'Generalized Levy-Walk Model for DNA Nucleotide Sequences'. In: *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics* 47 (June 1993), pp. 4514–23. DOI: 10.1103/PhysRevE.47.4514.

[37] O. Bé nichou et al. 'Intermittent search strategies'. In: *Reviews of Modern Physics* 83.1 (2011), pp. 81–129. DOI: 10.1103/revmodphys.83.81.

[38] G. Boccignone and M. Ferraro. 'Modelling gaze shift as a constrained random walk'. In: *Physica A: Statistical Mechanics and its Applications* 331.1 (2004), pp. 207–218. ISSN: 0378-4371. DOI: 10.1016/j.physa.2003.09.011.

[39] T. Rhodes, C. T. Kello and B. Kerster. 'Intrinsic and extrinsic contributions to heavy tails in visual foraging'. In: *Visual Cognition* 22 (2014), pp. 809–842. DOI: 10.1080/13506285.2014.918070.

[40] H. Credidio et al. 'Statistical patterns of visual search for hidden objects'. In: *Sci Rep* 2:920 (2012). DOI: 10.1038/srep00920.

[41] L. R. Rabiner. 'A tutorial on hidden Markov models and selected applications in speech recognition'. In: *Proceedings of the IEEE* 77.2 (1989), pp. 257–286. DOI: 10.1109/5.18626.

[42] V. Levantini et al. 'EYES Are The Window to the Mind: Eye-Tracking Technology as a Novel Approach to Study Clinical Characteristics of ADHD'. In: *Psychiatry Research* 290 (2020), pp. 113–135. DOI: 10.1016/j.psychres.2020.113135.

[43] L. Larsson. *Event Detection in Eye-Tracking Data for Use in Applications with Dynamic Stimuli*. Ph.D. Thesis, Department of Biomedical Engineering, Faculty of Engineering LTH, Lund University. 2016.

[44] B. Birawo and P. Kasprowski. 'Review and Evaluation of Eye Movement Event Detection Algorithms'. In: *Sensors* 22.22 (2022). DOI: 10.3390/s22228810.

[45] B. Mahanama et al. 'Eye Movement and Pupil Measures: A Review'. In: *Frontiers of Computer Science* 3 (2022). DOI: 10.3389/fcomp.2021.733531.

[46] W. Becker and A.F. Fuchs. 'Further properties of the human saccadic system: Eye movements and correction saccades with and without visual fixation points'. In: *Vision Research* 9.10 (1969), pp. 1247–1258. DOI: 10.1016/0042-6989(69)90112-6.

[47] M. Russo et al. 'Oculomotor impairment during chronic partial sleep deprivation'. In: *Clinical Neurophysiology* 114.4 (2003), pp. 723–736. DOI: 10.1016/S1388-2457(03)00008-7.

[48] A. L. Boxer et al. 'Saccade Abnormalities in Autopsy-Confirmed Frontotemporal Lobar Degeneration and Alzheimer Disease'. In: *Archives of Neurology* 69.4 (2012), pp. 509–517. DOI: 10.1001/archneurol.2011.1021.

[49] E. Castello, N. Baroni and E. Pallestrini. 'Neurotological and Auditory Brain Stem Response Findings in Human Immunodeficiency Virus—Positive Patients without Neurologic Manifestations'. In: *Annals of Otology, Rhinology & Laryngology* 107.12 (1998), pp. 1054–1060. DOI: 10.1177/000348949810701210.

[50] H. Hartridge and L. C. Thomson. 'Methods of Investigating Eye Movements'. In: *British Journal of Ophthalmology* 32 (1948), pp. 581–591. DOI: 10.1136/bjo.32.9.581.

[51] R. A. Monty. 'An advanced eye-movement measuring and recording system'. In: *The American psychologist* 30 (1975), pp. 331–335. DOI: 10.1037/0003-066X.30.3.331.

[52] R. Zemblys et al. 'Using machine learning to detect events in eye-tracking data'. In: *Behav Res* 50 (2018), pp. 160–181. DOI: 10.3758/s13428-017-0860-3.

[53] D. D. Salvucci and J. H. Goldberg. 'Identifying Fixations and Saccades in Eye-Tracking Protocols'. In: *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*. New York, NY, USA: Association for Computing Machinery, 2000, pp. 71–78. DOI: 10.1145/355017.355028.

[54] R. Andersson et al. 'One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms'. In: *Behavior research methods* 49 (2017), pp. 616–637. DOI: 10.3758/s13428-016-0738-9.

[55] T. Sen and T. Megaw. 'The Effects of Task Variables and Prolonged Performance on Saccadic Eye Movement Parameters'. In: *Theoretical and Applied Aspects of Eye Movement Research*. Vol. 22. Advances in Psychology. North-Holland, 1984, pp. 103–111. DOI: 10.1016/S0166-4115(08)61824-5.

[56] R. Engbert and K. Mergenthaler. 'Microsaccades are triggered by low retinal image slip'. In: *Proceedings of the National Academy of Sciences* 103.18 (2006), pp. 7192–7197. DOI: 10.1073/pnas.0509557103.

[57] R. Engbert and R. Kliegl. 'Microsaccades uncover the orientation of covert attention'. In: *Vision Research* 43.9 (2003), pp. 1035–1045. ISSN: 0042-6989. DOI: 10.1016/S0042-6989(03)00084-1.

[58] O. V. Komogortsev et al. 'Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors'. In: *IEEE Transactions on Biomedical Engineering* 57.11 (2010), pp. 2635–2645. DOI: 10.1109/TBME.2010.2057429.

[59] G. D. Forney. 'The viterbi algorithm'. In: *Proceedings of the IEEE* 61.3 (1973), pp. 268–278. DOI: 10.1109/PROC.1973.9030.

[60] R. van der Lans, M. Wedel and R. Pieters. 'Defining eye-fixation sequences across individuals and tasks: the Binocular-Individual Threshold (BIT) algorithm'. In: *Behav Res* 43 (2011), pp. 239–257. DOI: 10.3758/s13428-010-0031-2.

[61] M. Nyström and K. Holmqvist. 'An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data'. In: *Behavior Research Methods* 42 (2010), pp. 188–204. DOI: 10.3758/BRM.42.1.188.

[62] L. Larsson, M. Nyström and M. Stridh. 'Detection of Saccades and Postsaccadic Oscillations in the Presence of Smooth Pursuit'. In: *IEEE Transactions on Biomedical Engineering* 60.9 (2013), pp. 2484–2493. DOI: 10.1109/TBME.2013.2258918.

[63] H. Widdel. 'Operational Problems in Analysing Eye Movements'. In: *Theoretical and Applied Aspects of Eye Movement Research*. Vol. 22. Advances in Psychology. North-Holland, 1984, pp. 21–29. DOI: 10.1016/S0166-4115(08)61814-2.

[64] G. Veneri et al. 'Automatic eye fixations identification based on analysis of variance and covariance'. In: *Pattern Recognition Letters* 32.13 (2011), pp. 1588–1593. DOI: https://doi.org/10.1016/j.patrec.2011.06.012.

[65] G. Veneri et al. 'Eye fixations identification based on statistical analysis - Case study'. In: *2010 2nd International Workshop on Cognitive Information Processing*. 2010, pp. 446–451. DOI: 10.1109/CIP.2010.5604221.

[66] F. Shic, B. Scassellati and K. Chawarska. 'The Incomplete Fixation Measure'. In: *Proceedings of the 2008 symposium on Eye tracking research & applications* (Jan. 2008), pp. 111–114. DOI: 10.1145/1344471.1344500.

[67] J. H. Goldberg and J. C. Schryver. 'Eye-gaze-contingent control of the computer interface: Methodology and example for zoom detection'. In: *Behavior Research Methods, Instruments, & Computers* 27 (1995), pp. 338–350. DOI: 10.3758/BF03200428.

[68] R. Den Buurman. 'Eye Movements and the Perceptual Span in Reading'. In: *Reading Research Quarterly* 16 (1981), p. 227. DOI: 10.2307/747557.

[69] A. Lev et al. 'Eye Tracking During a Continuous Performance Test: Utility for Assessing ADHD Patients'. In: *Journal of Attention Disorders* 26 (2020), pp. 245–255. DOI: 10.1177/1087054720972786.

[70] D. Richardson and M. Spivey. 'Eye-Tracking: Characteristics and Methods'. In: *Encyclopedia of Biomaterials and Biomedical Engineering* (2008). DOI: 10.1081/E-EBBE2-120013920.

[71] P. Olsson. *Real-time and Offline Filters for Eye Tracking*. 2007.

[72] A. George. *Image based Eye Gaze Tracking and its Applications*. 2019. arXiv: 1907.04325 [cs.CV].

[73] A. F. Klaib et al. 'Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies'. In: *Expert Systems With Applications* 166 (2021). DOI: 10.1016/j.eswa.2020.114037.

[74] C. Buquet et al. 'Photo-oculography: A new method for eye movements study, interest in ophthalmological and extra pyramidal neurological diseases'. In: *1992 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Vol. 4. 1992, pp. 1555–1556. DOI: 10.1109/IEMBS.1992.5761922.

[75] SR Research Ltd. *EyeLink® 1000 User Manual*. 2009.

[76] SR Research Ltd. *About Eye Tracking*. 2023. URL: https://www.sr-research.com/about-eye-tracking/ (visited on 15/02/2023).

[77] Tobii AB. *Psychology and neuroscience research*. 2023. URL: https://www.tobii.com/solutions/scientific-research/psychology-and-neuroscience (visited on 15/03/2023).

[78] A. Olsen. *The Tobii IVT Fixation Filter Algorithm description*. 2012.

[79] E. S. Dalmaijer. *PyGaze: Open-source toolbox for eye tracking in Python*. 2023. URL: http://www.pygaze.org/ (visited on 15/03/2023).

[80] E. S. Dalmaijer, S. Mathôt and S. Van der Stigchel. 'PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments'. In: *Behav Res* 46 (2014), pp. 913–921. DOI: 10.3758/s13428-013-0422-2.

[81] Z. Hawi, T. Cummins and J. et al. Tong. 'The molecular genetic architecture of attention deficit hyperactivity disorder'. In: *Mol Psychiatry* 20 (2015), pp. 289–297. DOI: 10.1038/mp.2014.183.

[82] J. T. Nigg et al. 'Toward a Revised Nosology for Attention-Deficit/Hyperactivity Disorder Heterogeneity'. In: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 5.8 (2020). Understanding the Nature and Treatment of Psychopathology: Letting the Data Guide the Way, pp. 726–737. DOI: 10.1016/j.bpsc.2020.02.005.

[83] D. N. Maron et al. 'Oculomotor deficits in attention deficit hyperactivity disorder (ADHD): A systematic review and comprehensive meta-analysis'. In: *Neuroscience & Biobehavioral Reviews* 131 (2021), pp. 1198–1213. DOI: 10.1016/j.neubiorev.2021.10.012.

[84] R. G. Ross et al. 'Eye Movement Task Related to Frontal Lobe Functioning in Children with Attention Deficit Disorder'. In: *Journal of the American Academy of Child & Adolescent Psychiatry* 33.6 (1994), pp. 869–874. DOI: 10.1097/00004583-199407000-00013.

[85] F. X. Castellanos et al. 'Executive Function Oculomotor Tasks in Girls With ADHD'. In: *Journal of the American Academy of Child & Adolescent Psychiatry* 39.5 (2000), pp. 644–650. DOI: 10.1097/00004583-200005000-00019.

[86] T. D. Gould et al. 'Altered performance on an ocular fixation task in attention-deficit/hyperactivity disorder'. In: *Biological Psychiatry* 50.8 (2001), pp. 633–635. DOI: 10.1016/S0006-3223(01)01095-2.

[87] G. A. O'Driscoll et al. 'Executive Functions and Methylphenidate Response in Subtypes of Attention-Deficit/Hyperactivity Disorder'. In: *Biological Psychiatry* 57.11 (2005), pp. 1452–1460. DOI: 10.1016/j.biopsych.2005.02.029.

[88] C. Hanisch, R. Radach and K. et al. Holtkamp. 'Oculomotor inhibition in children with and without attention-deficit hyperactivity disorder (ADHD)'. In: *Neural Transm* 113 (2006), pp. 671–684. DOI: 10.1007/s00702-005-0344-y.

[89] N. N. Rommelse et al. 'Deficits in visuo-spatial working memory, inhibition and oculomotor control in boys with ADHD and their non-affected brothers'. In: *Neural Transmission* 115 (2008), pp. 249–260. DOI: 10.1007/s00702-007-0865-7.

[90] I. M. Loe et al. 'Oculomotor performance identifies underlying cognitive deficits in attention-deficit/hyperactivity disorder'. In: *American Academy of Child and Adolescent Psychiatry* 48 (2009), pp. 431–440. DOI: 10.1097/CHI.0b013e31819996da.

[91] E. M. Mahone et al. 'Oculomotor anomalies in attention-deficit/ hyperactivity disorder: evidence for deficits in response preparation and inhibition'. In: *American Academy of Child and Adolescent Psychiatry* 48 (2009), pp. 749–756. DOI: 10.1097/CHI.0b013e3181a565f1.

[92] L. Carr, J. Henderson and J. T. Nigg. 'Cognitive Control and Attentional Selection in Adolescents with ADHD Versus ADD'. In: *Journal of Clinical Child & Adolescent Psychology* 39.6 (2010), pp. 726–740. DOI: 10.1080/15374416.2010.517168.

[93] C. Karatekin, C. Bingham and T. White. 'Oculomotor and Pupillometric Indices of Pro- and Antisaccade Performance in Youth-Onset Psychosis and Attention Deficit/Hyperactivity Disorder'. In: *Schizophrenia Bulletin* 36.6 (2009), pp. 1167–1186. DOI: 10.1093/schbul/sbp035.

[94] Y. Goto et al. 'Saccade eye movements as a quantitative measure of frontostriatal network in children with ADHD'. In: *Brain and Development* 32.5 (2010), pp. 347–355. DOI: 10.1016/j.braindev.2009.04.017.

[95] E. Pishyareh et al. 'Attentional Bias towards Emotional Scenes in Boys with Attention Deficit Hyperactivity Disorder'. In: *Iranian J Psychiatry* 7 (2012), pp. 93–96.

[96] Y. Matsuo et al. 'Gap Effect Abnormalities during a Visually Guided Pro-Saccade Task in Children with Attention Deficit Hyperactivity Disorder'. In: *PLoS ONE* 10 (2015). DOI: 10.1371/journal.pone.0125573.

[97] B. N. Türkan et al. 'Comparison of change detection performance and visual search patterns among children with/without ADHD: Evidence from eye movements'. In: *Research in Developmental Disabilities* 49-50 (2016), pp. 205–215. DOI: 10.1016/j.ridd.2015.12.002.

[98] R. A. Rensink, J. K. O'Regan and J. J. Clark. 'To See or not to See: The Need for Attention to Perceive Changes in Scenes'. In: *Psychological Science* 8.5 (1997), pp. 368–373. DOI: 10.1111/j.1467-9280.1997.tb00427.x.

[99] M. P. Bucci et al. 'Oculomotor Abnormalities in Children with Attention-Deficit/Hyperactivity Disorder Are Improved by Methylphenidate'. In: *Journal of child and adolescent psychopharmacology* 27.3 (2017), pp. 274–280. DOI: 10.1089/cap.2016.0162.

[100] G. Wainstein, D. Rojas-Líbano and N. A. Crossley. 'Pupil Size Tracks Attentional Performance In Attention-Deficit/Hyperactivity Disorder'. In: *Sci Rep 7* 8228 (2017). DOI: 10.1038/s41598-017-08246-w.

[101] V. J. Serrano, J. S. Owens and B. Hallowell. 'Where Children With ADHD Direct Visual Attention During Emotion Knowledge Tasks: Relationships to Accuracy, Response Time, and ADHD Symptoms'. In: *Journal of Attention Disorders* 22.8 (2018), pp. 752–763. DOI: 10.1177/1087054715593632.

[102] S. Caldani et al. 'The Effect of Dual Task on Attentional Performance in Children With ADHD'. In: *Frontiers in integrative neuroscience* 12.67 (2019). DOI: 10.3389/fnint.2018.00067.

[103] J. Fernandez-Ruiz et al. 'Dorsolateral prefrontal cortex hyperactivity during inhibitory control in children with ADHD in the antisaccade task. Brain imaging and behavior'. In: *Frontiers in integrative neuroscience* 14.6 (2020), pp. 2450–2463. DOI: 10.1007/s11682-019-00196-3.

[104] J. H. Huang and Y. S. Chan. 'Saccade eye movement in children with attention deficit hyperactivity disorder'. In: *Nordic Journal of Psychiatry* 74.1 (2020), pp. 16–22. DOI: 10.1080/08039488.2019.1666919.

[105] G. Singh. *ANOVA: Complete guide to Statistical Analysis & Applications*. 2023. URL: https://www.analyticsvidhya.com/blog/2018/01/anova-analysis-of-variance/ (visited on 12/02/2023).

[106] Statistics How To. *ANOVA Test: Definition, Types, Examples, SPSS*. 2023. URL: https://www.statisticshowto.com/probability-and-statistics/hypothesis-testing/anova/ (visited on 12/02/2023).

[107] Wikipedia contributors. *Wechsler Intelligence Scale for Children — Wikipedia, The Free Encyclopedia*. 2023. URL: https://en.wikipedia.org/w/index.php?title=Wechsler_Intelligence_Scale_for_Children&oldid=1139727872 (visited on 18/03/2023).

[108] T. Z. Keith and E. L. Witta. 'Hierarchical and cross-age confirmatory factor analysis of the WISC-III: What does it measure?' In: *School Psychology Quarterly* 12 (1997), pp. 89–107. DOI: 10.1037/h0088950.

[109] O. Moura, P. Costa and M. R. Simões. 'WISC-III Cognitive Profiles in Children with ADHD: Specific Cognitive Impairments and Diagnostic Utility'. In: *The Journal of General Psychology* 146.3 (2019), pp. 258–282. DOI: 10.1080/00221309.2018.1561410.

[110] A. C. Wood et al. 'The relationship between ADHD and key cognitive phenotypes is not mediated by shared familial effects with IQ'. In: *Psychological medicine* 41.4 (2011), pp. 861–71. DOI: 10.1017/S003329171000108X.

[111] S. D. Mayes and S. L. Calhoun. 'WISC-IV and WISC-III Profiles in Children With ADHD'. In: *Journal of Attention Disorders* 9.3 (2006), pp. 486–493. DOI: 10.1177/1087054705283616.

[112] C. Mealer, S. B. Morgan and R. L. Luscomb. 'Cognitive functioning of ADHD and non-ADHD boys on the WISC-III and WRAML: An analysis within a memory model'. In: *Journal of Attention Disorders* 1 (1996), pp. 133–145. DOI: 10.1177/108705479600100302.

[113] R. Lopes et al. 'WISC-III Sensibility in the identification of Attention Deficit Hyperactivity Disorder (ADHD)'. In: *Panamerican Journal of Neuropshychology* 6 (2012), pp. 128–140. DOI: 10.7714/cnps/6.1.208.

[114] IBM. *What is logistic regression?* 2023. URL: https://www.ibm.com/topics/logistic-regression (visited on 12/02/2023).

[115] Wikipedia contributors. *Logistic regression — Wikipedia, The Free Encyclopedia*. 2023. URL: https://en.wikipedia.org/w/index.php?title=Logistic_regression&oldid=1140140318 (visited on 18/03/2023).

[116] IBM. *About SVM*. 2021. URL: https://www.ibm.com/docs/en/spss-modeler/saas?topic=models-about-svm (visited on 18/03/2023).

[117] A. Sasidharan. *Support Vector Machine Algorithm*. 2022. URL: https://www.geeksforgeeks.org/support-vector-machine-algorithm/ (visited on 18/03/2023).

[118] R. Gandhi. *Support Vector Machine — Introduction to Machine Learning Algorithms*. 2018. URL: https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47 (visited on 18/03/2023).

[119] IBM. *What is a Decision Tree?* 2023. URL: https://www.ibm.com/topics/decision-trees (visited on 18/03/2023).

[120] IBM. *What is random forest?* 2023. URL: https://www.ibm.com/topics/random-forest (visited on 18/03/2023).

[121] T. Yiu. *Understanding Random Forest: How the Algorithm Works and Why it Is So Effective*. 2019. URL: https://towardsdatascience.com/understanding-random-forest-58381e0602d2 (visited on 18/03/2023).

[122] IBM. *K-Nearest Neighbors Algorithm*. 2023. URL: https://www.ibm.com/topics/knn (visited on 18/03/2023).

[123] O. Harrison. *Machine Learning Basics with the K-Nearest Neighbors Algorithm*. 2018. URL: https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761 (visited on 18/03/2023).

[124] Analyse Up. *Stratified Kfold Tutorial*. 2023. URL: https://www.analyseup.com/python-machine-learning/stratified-kfold.html (visited on 03/04/2023).

[125] G. I. Webb et al. 'Leave-One-Out Cross-Validation'. In: *Encyclopedia of Machine Learning*. 2011, pp. 600–601. DOI: 10.1007/978-0-387-30164-8_469.

[126] P. Misra and A. S. Yadav. 'Improving the classification accuracy using recursive feature elimination with cross-validation'. In: *Int. J. Emerg. Technol* 11.3 (2020), pp. 659–665.

[127] E. Bisong. 'More supervised machine learning techniques with scikit-learn'. In: *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners* (2019), pp. 287–308. DOI: 10.1007/978-1-4842-4470-8_24.

[128] M. Gusarov. *Do I need to tune logistic regression hyperparameters?* 2022. URL: https://medium.com/codex/do-i-need-to-tune-logistic-regression-hyperparameters-1cb2b81fca69 (visited on 03/04/2023).

[129] R. Pramoditha. *Why do we set a random state in machine learning models?* 2022. URL: https://towardsdatascience.com/why-do-we-set-a-random-state-in-machine-learning-models-bb2dc68d8431 (visited on 03/04/2023).

[130] A. Lekhtman. *Data Science in Medicine — Precision & Recall or Specificity & Sensitivity?* 2019. URL: https://towardsdatascience.com/should-i-look-at-precision-recall-or-specificity-sensitivity-3946158aace1 (visited on 23/03/2023).

Appendix A

Code repository

The codes used in the project are available on GitHub at the following address:

https://github.com/Chris-s366258/Eye_Tracking.

The code files are written in Python format.