FEDERAL UNIVERSITY OF OURO PRETO
INSTITUTE OF EXACT AND BIOLOGICAL SCIENCE
DEPARTMENT OF COMPUTER SCIENCE

LUCCA ARANTES MARTINS
Supervisor: Prof. Dr. Jadson Castro Gertrudes
Co-supervisor: Prof. Dr. Saul Emanuel Delabrida Silva

USING CLUSTERING TECHNIQUES FOR EXPLORATORY ANALYSIS OF EYE-TRACKING DATA

Ouro Preto, MG
June, 2021

FEDERAL UNIVERSITY OF OURO PRETO
INSTITUTE OF EXACT AND BIOLOGICAL SCIENCE
DEPARTMENT OF COMPUTER SCIENCE

LUCCA ARANTES MARTINS

USING CLUSTERING TECHNIQUES FOR EXPLORATORY ANALYSIS OF EYE-TRACKING DATA

Monography presented to the Computer Science course of the Federal University of Ouro Preto as part of the necessary requirements for obtaining the degree of Bachelor in Computer Science.

Supervisor: Prof. Dr. Jadson Castro Gertrudes
Co-supervisor: Prof. Dr. Saul Emanuel Delabrida Silva

Ouro Preto, MG
June, 2021

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

M386u Martins, Lucca Arantes.
Using Clustering Techniques for Exploratory Analysis of Eye-Tracking Data. [manuscrito] / Lucca Arantes Martins. - 2022.
58 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Jadson Castro Gertrudes.
Coorientador: Prof. Dr. Saul Emanuel Delabrida Silva.
Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Graduação em Ciência da Computação .

1. Machine Learning. 2. Eye-Tracking. 3. Agrupamento de dados. I. Gertrudes, Jadson Castro. II. Silva, Saul Emanuel Delabrida. III. Universidade Federal de Ouro Preto. IV. Título.

CDU 004

Bibliotecário(a) Responsável: Luciana De Oliveira - SIAPE: 1.937.800

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIENCIAS EXATAS E BIOLOGICAS
DEPARTAMENTO DE COMPUTACAO

FOLHA DE APROVAÇÃO

Lucca Arantes Martins

USING CLUSTERING TECHNIQUES FOR EXPLORATORY ANALYSIS OF EYE-TRACKING DATA

Monografia apresentada ao Curso de Ciência da Computação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Ciência da Computação

Aprovada em 22 de Junho de 2022.

Membros da banca

Jadson Castro Gertrudes (Orientador) - Doutor - Universidade Federal de Ouro Preto
Saul Emanuel Delabrida Silva (Coorientador) - Doutor - Universidade Federal de Ouro Preto
Andrea Gomes Campos Bianchi (Examinadora) - Doutora - Universidade Federal de Ouro Preto
Qasim Ali (Examinador) - Mestre - Western Norway University of Applied Sciences

Jadson Castro Gertrudes, Orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 22/06/2022.

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000
Telefone: 3135591692 - www.ufop.br

I would like to dedicate this work to my parents, Paulo Martins da Silva, and Nelsimari Arantes Cabral Silva, and to my brother Fillipe Arantes Martins. You have always been my only certainty.

Acknowledgements

This work marks the end of my graduation, a life cycle that could be finished with a flourish thanks to very special people that crossed my path. I would firstly like to thank my parents and family, for unconditionally supporting me since always. Thanks to Jadson Castro, and Saul Delabrida, for molding my research and supporting me on every doubt and raising insights. My most sincere gratitude to the HVL researchers involved in the projects related to eye-tracking, specially Qasim Ali, Ilona Heldal, Carsten Helgesen, and Chun-Wei Jerry Lin. Thanks to the Computer Science Department (DECOM), and all the professors that instructed me along with the last years. It was a real pleasure to work with you.

"For you the blind who once could see, the bell tolls for thee..."

- Rush

Abstract

Vision impairments in children are harmful to their learning process, cognitive development, social interaction, and scholar performance. In recent years, new technologies have been applied for vision screening tests, sharpening traditional techniques and enabling the early diagnosis of different kinds of debilitation on the visual system functionality. Eye-tracking is a widely used technique applied for different purposes, and when it comes to vision assessment and training, it is greatly suitable. This work aims to apply feature engineering and cluster analysis techniques within eye-tracking data collected from children performing structured visual tasks. Feature engineering creates meaningful attributes for the recordings in terms of performance and data quality, and the exploratory analysis covers different configurations for the clustering methods and their hyper-parameters. Cluster validation metrics evaluate the clustering results' quality, and domain expert acknowledgment is essencial for trustful inferences regarding the children's oculomotor system's health. In order to streamline the exploratory analysis and facilitate the experiments, we also propose a framework for Cluster Analysis of the eye-tracking data.

Keywords: eye tracking, machine learning, cluster analysis, feature engineering, vision impairments, vision screening, vision care.

Resumo

Problemas de visão em crianças são uma ameaça ao processo de aprendizado, desenvolvimento cognitivo, interação social e performance escolar. Nos últimos anos, novas tecnologias têm sido aplicadas à exames de vista, aguçando as técnicas tradicionais e possibilitando o diagnóstico precoce de possíveis debilitações na funcionalidade do sistema visual. O rastreamento ocular (*Eye-Tracking*) é uma técnica amplamente utilizada para diferentes propósitos, e quando se trata de testagem e treinamento de visão, é fortemente adequada. Este trabalho objetiva aplicar engenharia de atributos e técnicas de agrupamento (*Cluster Analysis*) em dados de rastreamento ocular coletados de crianças executando tarefas estruturadas. A engenharia de atributos caracteriza as gravações em termos de performance e qualidade dos dados, e a análise exploratória possibilita diferentes configurações para os métodos de agrupamento de dados e seus hiperparâmetros. Métricas de validação avaliam a qualidade dos agrupamentos resultantes, e a análise dos resultados por parte de especialistas em visão é essencial para se conseguir inferências confiáveis em relação à saúde do sistema oculomotor das crianças. A fim de agilizar a análise exploratória e facilitar os experimentos, também propomos um framework para agrupamento dos dados de rastreamento ocular.

Palavras-chave: rastreamento ocular, aprendizado de máquina, agrupamento de dados, engenharia de atributos, problemas de visão, exames de vista, saúde da visão.

List of Figures

Contents

1 Introduction

Vision is the most fundamental sense for our daily activities, and it plays a massive role in our learning process, cognitive development, environment perception, and social interaction (ZABA; M.A; O.D, 2011; ATOWA; HANSRAJ; WAJUIHIAN, 2019). In children, vision impairments can disturb scholar performance, leading to low cognitive skills and affecting their self-development in different aspects (HOPKINS et al., 2019). According to the World Health Organization, as published in their most recent world report on vision, at least 2.2 billion people have a vision impairment, and at least 1 billion of these impairments could have been prevented with early diagnosis and treatment (WHO, 2019).

Traditional visual tests such as optotype and stereoscopic acuity have been significantly evolved in recent years with the help of computerized vision screening (CHAPLIN et al., 2015). Also, many new tests have been devised, including assessing visual skills and acuity, color and stereo vision, contrast sensitivity, and oculomotor behavior (ALI et al., 2021b). Given that treatment and rehabilitation of oculomotor impairments are facilitated by early diagnosis, children's vision care must be a primary concern. Nevertheless, in many poverty-stricken countries, vision screening is limited by lack of resources and other social factors, including the medical system and the educational level of the population (ALI et al., 2020).

Vision rehabilitation or improvement by training methods has had its efficiency proven by different studies (MATHERS; KEYES; WRIGHT, 2010; LEONG et al., 2014; GALLAWAY; BOAS, 2019; WILHELMSEN; FELDER, 2021). When applied to infants, some difficulties emerge from the fact that even the child is not aware of having a vision impairment, and they usually do not give the proper importance to their vision assessment and care. The use of Serious Games (SG) is a very suitable technique to employ and develop a more entertainer methodology (HELDAL et al., 2021; BORTOLI; GAGGI, 2011). SG can lead children to be more collaborative and assiduous on the training and reach better results by whetting their concentration. Also, another well-suited technology for the training methods is Virtual Reality (VR) (ALI et al., 2021a). In a virtual environment, there is also the possibility for tuning the games to include peripheral vision training (YOUNIS et al., 2017; DAVID; BEITNER; Võ, 2021), which is usually hampered by the screen size limitations.

Eye-tracking (ET) technologies are useful for different research fields, including visual systems, UI/UX, sports and professional performance, psychology, psycholinguistics, computer games, marketing, and consumer research, product design, software engineering, and others (KLAIB et al., 2021). Using an eye-tracker to collect the user's eye features along time (e.g., gaze position, pupil diameter) allows us to analyze their oculomotor system's behavior during a given stimulus with quantitative measures. Moreover, this opens interesting research possibilities on health and education. Many diseases and disorders have been investigated using ET data, such as autism (THAPALIYA; JAYARATHNA; JAIME, 2018; LIU et al., 2015), dyslexia (RELLO; BALLESTEROS, 2015) and major depression (DING et al., 2019). Also, ET fits for assessing scholar skills, such as reading comprehension, learning strategies and cognition development (MEZIERE et al., 2021; CATRYSSE et al., 2018; HESSELS; HOOGE, 2019).

In our work, we analyze ET data collected from children while performing structured tasks in the C&Look software (e.g., following a figure along with the screen in different trajectories). C&Look is a product developed by researchers from Bergen, Norway, in the Western Norway University of Applied Sciences (*Høgskulen på Vestlandet* - HVL) (EIDE et al., 2019), and the database content was collected from children in cities of Tanzania. Since we do not have any labels in the data or previous knowledge about the children's oculomotor system's health, we apply Cluster Analysis within the ET data recordings to find patterns that could breach inferences about children's performances or about the database itself.

Cluster Analysis (CA) is a robust tool considered interdisciplinary because of its vast applicability. The cluster techniques help us understand the structure of unknown data or find characteristics in it that can be useful for solving a specific problem (TAN; STEINBACH; KUMAR, 2005). CA is considered an unsupervised learning methodology (JAIN; DUBES, 1988), and since we do not have any external information about the database objects (ET recordings), it fits our intentions very well.

1.1 Goals

Our work aims to analyze ET data provided by C&Look using Cluster Analysis in order to search for patterns in the recordings that could indicate (dis)similarities in children's oculomotor behavior. We will conduct the specific objectives to ensure the completion of our supreme objective:

- Provide a good overview of the database, counting and filtering the content presented.
- Apply feature engineering in the time series for generating meaningful attributes for the analysis.
- Propose a framework for applying CA within the featured eye-tracking data.
- Explore the best clustering results.

1.2 Document Organization

The document is organized as follows: Chapter 2 presents the main concepts used to build the proposal, including human vision, eye movements, eye tracking technologies, and cluster analysis. In Chapter 3 we cite publications related to our work that make use of eye-tracking data and machine learning methods. In Chapter 4 we present materials and methods used in this research, and describe the pipeline used for exploring the data. Finally, in Chapter 6 we present the conclusion regarding the current accomplishments, cite the study limitations and organize some possible future works for this research.

2 Background

2.1 Human Vision

For gathering visual information from a scene, we change the direction of our eyes so that the light reflected from the point of interest reaches the retina. In the retina, there are two types of photoreceptors: the cons, responsible for our most detailed and colored vision under good light conditions, and the rods, for peripheral and night vision. The layers of the eye's anatomy – shown in Figure 2.1– refract the light to make it fall on the fovea, a small region at the center of the visual field with the highest density of cons, which leads our spatial resolution to decay in peripheral vision (STRASBURGER; RENTSCHLER; JÜTTNER, 2011 apud GOETTKER; GEGENFURTNER, 2021). The ciliary muscle adjusts the shape of the lens according to the distance from where the light reflects, enabling us to change the depth focus (PURVES et al., 2018).

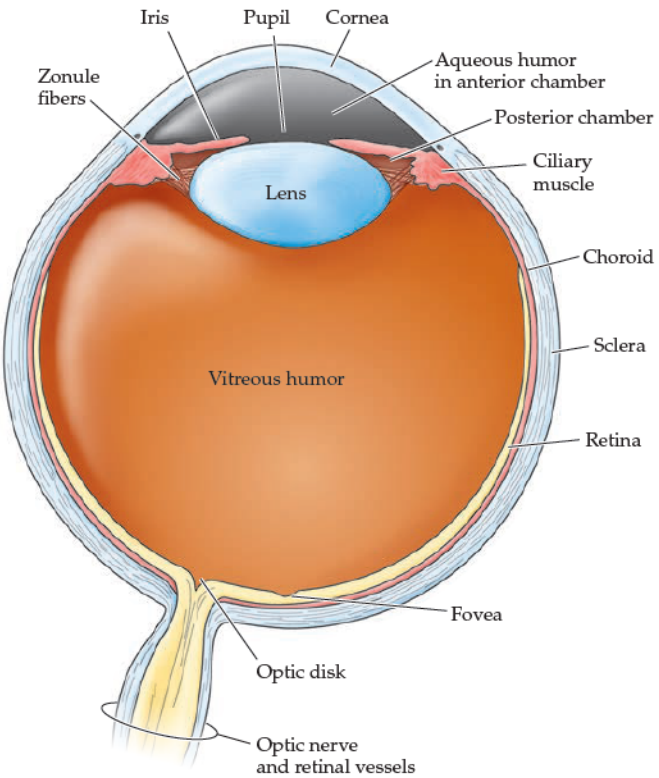

Figure 2.1 – Anatomy of the human eye. Source: Purves et al. (2018).

In short, from Remington (2012), once the light comes in and the visual pathway starts, the photoreceptors convert light energy into neural signals, outputting them through synapses to the bipolar and amacrine cells and then to the ganglion cells. The ganglion cells take the signals from both rods and cones to the optic chiasm through the optic nerve fibers. The signals from each eye cross over so that the right cerebral hemisphere also processes the left visual field information and vice-versa. After this intersection, the optic tract carries the signals to the lateral geniculate nucleus in the thalamus, where every sensory information (except the olfactory) is organized and distributed to the brain area responsible for processing it. In the case of visual signals, they terminate in the occipital lobe, more precisely in the primary visual cortex (also called V1), which is the first brain area to process and respond to visual stimuli. A visualization of the structures in the visual pathway is presented in Figure 2.2.

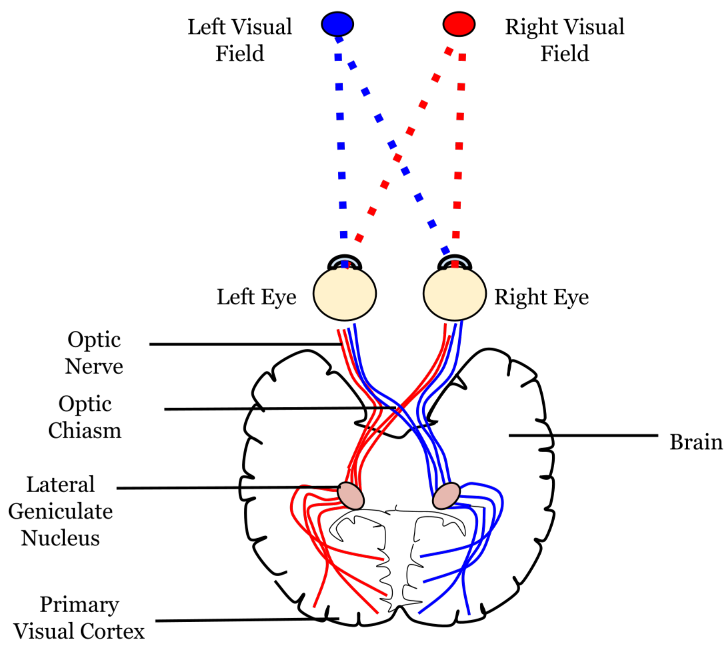

Figure 2.2 – Visual pathway components. Source: Wikimedia (2020).

Six muscles are responsible for the movement of an eye. Each of these muscles is paired with a correspondent one in the other eye to produce equivalent movements on both ocular globes. These muscles respond, essentially, to five types of eye movements for the oculomotor system: **saccades**, **smooth pursuits** and **vergence movements** for shifting the direction of the gaze; and **vestibulo-ocular** and **optokinetic movements** for stabilizing it.

A saccade is every movement that abruptly changes the gaze from one fixation point to another. These movements are ballistic because we are effectively blind during the saccade trajectory and cannot interfere in its direction in the meantime. A smooth pursuit occurs when we follow a moving target aiming to keep it reflecting on the fovea. Curiously, most people cannot perform a smooth pursuit without having a target, ending up doing a sequence of small saccades. The vergence movements are disjunctive (eyes move in different directions) and required when we need to change the focus towards an object that is at a different distance or when we track a target moving closer (eyes converge) or farthest away (eyes diverge).

For stabilizing the gaze, the vestibulo-ocular and optokinetic eye movements operate together, compensating the head movements. The vestibulo-ocular is an involuntary movement and occurs when we fixate on a target and move the head horizontally or vertically. It is empirically easy to note that while we move the head, the eyes go in the opposite direction, making the image stable on the retina. As a complementary feature, the optokinetic movements are small drifts that deal with the visual field changes, e.g., when the subject is rotating its head or body (PURVES et al., 2018; YARBUS, 1967).

In general, when it comes to studies that analyze eye movements concerning a given health condition, researchers focus on three basic events for different purposes: saccades, smooth pursuits, and fixations - some other types of movements are also relevant in some cases - e.g., post-saccades (glissades) (MARDANBEGI et al., 2018), and regressions (WOLFE, 2018). Vision is intimately related to our environment perception and organization of behavior (ZEIL; BOEDDEKER; HEMMI, 2008; RAYNER; POLLATSEK, 1989). This fact opens a wide range of possibilities for the psychological study field, enabling researchers to relate visual performance to many diseases and mental conditions, e.g., Dyslexia, Autism, Alzheimer and Major Depression - see section 3. Also, once vision is fundamental for our cognition development (ASLIN, 2012 apud HESSELS; HOOGE, 2019), many works focus on children's learning process and academic performances. Some examples are Catrysse et al. (2018), which compares students' learning strategies; Meziere et al. (2021) that predicts reading comprehension; and Hessels and Hooge (2019), which focus on developmental cognitive neuroscience in general.

This work is related to the category of studies that aims to evolve the diagnosis of functional impairments in the oculomotor system by using eye-tracking technologies and data analysis. An impaired vision control can lead to a whole range of problems in social interaction, cognition, academic performance, and low productivity (WILHELMSEN; FELDER, 2020; ALVAREZ et al., 2020). Moreover, the natural motivation for focusing on children is that the sooner an impairment is detected, the greater the possibilities of treatment and rehabilitation. Vision training is proven to be effective for improving reading and comprehension capabilities, recovering post-concussion vision disorders, and other problems related to the oculomotor system (BONILLA; OD; ALLISON, 2005; MATHERS; KEYES; WRIGHT, 2010; LEONG et al., 2014; FORTENBACHER et al., 2018; GALLAWAY; BOAS, 2019; WILHELMSEN; FELDER, 2021). In due course, by detecting vision control impairments and correcting them with vision training, the ultimate goal is to guarantee complete oculomotor control in children so that they have a total capacity for using their vision in academic and daily tasks.

2.2 Eye-Tracking

The eye-tracking techniques detect the subject's eye featuress over time, gathering information for different purposes. Even though the eye-tracker devices have been widespread only in recent years, the studies of eye movements began a long time ago, back to the 19th century. The researchers were interested in the reading process, and the first methods did not use any equipment. A specialist observed the eye movements while a subject was reading. Despite the inaccuracy of this method, some first conclusions have been made. In the late 1870s, a French ophthalmologist concluded that reading is not a linear process. In other words, the eyes do not move continuously along a line of text but make small saccades, and fixations (PLUZYCZKA, 2018). Since then, many ET methodologies have been created, and today we can divide them into four categories: The Scleral Search Coil (SSC), the Electro-Oculography (EOG), the Video-Oculography (VOG), and Infrared Oculography (IOG). The figures below show examples of devices for each of these techniques.

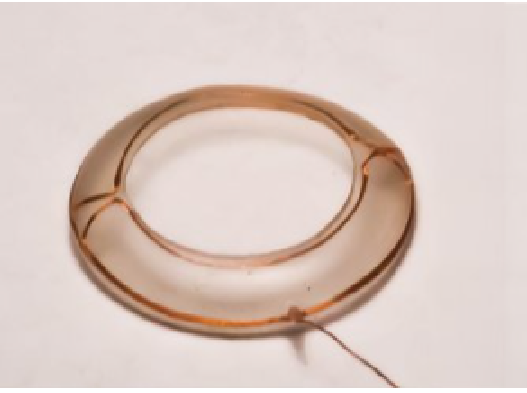

Figure 2.3 – Lens for Scleral Search Coil. Source: Whitmire et al. (2016).

Figure 2.4 – Electro-Oculography. Source: Thankachan (2018).

Figure 2.5 – Video-Oculography. Source: Arar (2017).

Figure 2.6 – Infrared-Oculography. Source: Johns et al. (2007).

Some applications focus on eye movements without regard for the subject's gaze. Those calculate the eye movements in relation to the head. As in our case, other applications aim to identify the subject's gaze in a visual scene, requiring the position of the eye in space (point of regard). For that purpose, techniques applied for tracking the head positions can be combined with the eye positions to find the point of regard, thus finding the subject's focus while scanning a scene (YOUNG; SHEENA, 1975).

The Scleral Search Coil (SSC) is an intriguing technique (ROBINSON, 1963) that uses contact lenses attached to a wire coil moving in a magnetic field. The magnetic field's voltage induced in the coil produces a signal representing the eye position. The contact lenses have mirrors that reflect the infrared waves, and the wavelength ranges are monitored and recorded as their changes represent a movement of the eye. Although this technique has high accuracy, it is not well accepted because of its invasiveness (e.g., reduces visual acuity and increases eye pressure causing discomfort (IRVING et al., 2003)) and expensiveness.

Electro Oculography (EOG) is a simple and inexpensive technique very used in the second half of the 20th century. The basis of this technique is the corneo-retinal electric potential, firstly investigated by Du Bois Reymond in 1849 (HASLWANTER; CLARKE, 2010). The potential range is from 0.4mV to 1.0mV (MALMIVUO; PLONSEY, 1995). When the eyes rotate, they change the electric field causing fluctuations in the skin, which are used by electrodes attached around the eyes to measure the eye movements. Horizontal and vertical movements are recorded separately by different electrodes. Since the electrodes are attached to the head, the tracking is linearly relative to the head position and orientation (KLAIB et al., 2021). EOG is not suitable for daily use, not only because of the use of electrodes but also because the corneo-retinal potential varies during the day and with different light conditions, altering the signal produced (YOUNG; SHEENA, 1975). Nonetheless, it can be handy for medical fields given its simple arrangement and relatively low invasiveness.

A widely studied technique is the Infrared Oculography (IOG) (TOROK; GUILLEMIN; BARNOTHY, 1951), which, similarly to the SSC, uses infrared light to track the eye movements. In this case, the light is pointed to the sclera and the iris. Since the white surface of the sclera has a higher reflection coefficient than the iris, the amount of reflected light changes proportionally with the eye movements. The algorithms for this technique rely on detecting the pupil and the light reflected to yield the eye positions. A reference point in the cornea is used for head movements, called glint (KLAIB et al., 2021). IOG has the advantage of efficiently handling eye blinking and being less invasive, but glasses can be a problem because they will also reflect light.

Video Oculography (VOG) is a technique that uses one or more cameras for recording the subject's head and face. The systems typically use infrared light so they can also work in darkness. The algorithms for this technique aim to detect the eye and pupil position using image processing. The most significant advantage is that it is entirely non-invasive since cameras remotely make the recording. However, although this technique has a simple arrangement, it is expensive and requires more storage and computational power. There are already lots of algorithms for VOG; some examples are (KONG et al., 2018; GONI et al., 2004; KLAIB et al., 2019)

For many ET techniques, machine learning (ML) methods can be applied for training models that increase tracking accuracy. Algorithms including neural networks, regression, naive Bayes classification, support vector machine, and template matching have already been implemented and proven effective for this purpose. The most recently released eye-tracking devices have more sophisticated software programs and ML algorithms.

2.3 Cluster Analysis

In real applications, the data structure or the meaning of its content is often unknown a priori. The elements (objects) do not have any label in these cases. We cannot even know if they have a meaningful distribution or valuable information in their entrails and relationships. Data scientists aim to cluster the objects into different classes composed of the most related ones, based on their (dis) similarities, extracting meaningful information and possible links among the objects in the data. This kind of analysis is called Cluster Analysis (CA), and its applicability is pretty vast, being applied in many disciplines, like medicine, biology, psychology, marketing, image processing, and archaeology (EVERITT et al., 2011). We can use CA to understand or find utilities for the data. The first context is the study of techniques to find potential classes of objects in the data. The second context is the study of techniques to find cluster prototypes (i.e., most representative element – natural or not – of a given cluster).

Since the learning process is not helped by previous external information, CA is generally referred to as an unsupervised classification (TAN; STEINBACH; KUMAR, 2005). Figure 2.7 shows a tree that represents the division of traditional clustering algorithms. The first division of the tree (tree's root node) organizes clustering into hierarchical and partitional algorithms, which have different and straightforward ideas that we discuss in the following sections.

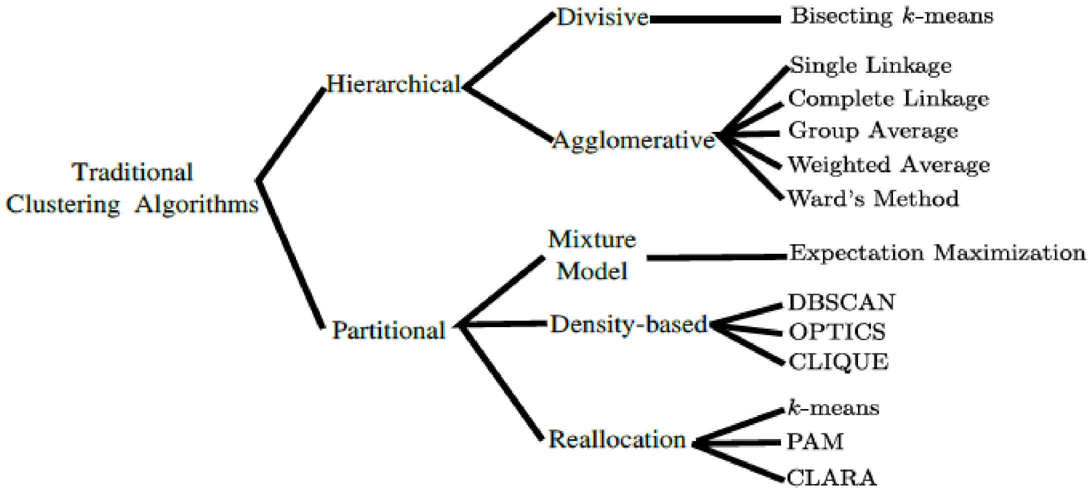

Figure 2.7 – Tree of traditional clustering techniques. Source: (GERTRUDES, 2019).

Cluster Analysis is a powerful tool for human knowledge acquirement and organization. Its resulting clusterings can be the final goal of the desired solution or only an initial step in the process of solving a problem. The clusters can also be used for a future classification problem. The labeled objects can be used for training a neural network that will learn to classify new unknown objects. In our case, since the C&Look recordings are not labeled with any external information, CA fits perfectly for us to find patterns in the data, identifying the performances' characteristics that may reflect an oculomotor impairment or a relationship between subjects, groups or tasks. The recordings' data contain the sequential eye-tracking information extracted at each timestamp along with the task execution, which characterizes it as time series data (MILLS, 2019). The two main branches of CA methods are explained in Sections 2.3.1 and 2.3.2.

2.3.1 Hierarchical Algorithms

This idea describes the relationships between data objects in terms of nested groups that form a hierarchy. This hierarchy is represented as a dendrogram; see Figure 2.8. The **n** leaf nodes of the dendrogram represent the objects (or singleton clusters), and each node represents a cluster (JAIN; DUBES, 1988). The root node represents a cluster that owns all **n** objects. In the basic idea, each level is a solution that has one lesser cluster than the solution of the level right below it.

Consequently, the full dendrogram is a collection of solutions with all possible numbers of clusters for a given clustering configuration. Therefore, by cutting this dendrogram horizontally, we define the number of non-overlapping clusters and their objects (red line in Figure 2.8). A horizontal cut is the simplest way to get a flat partition from the dendrogram, but there are also some more sophisticated methods for the cluster extraction, such as FOSC, that will be presented in Section 2.3.3).

For building the hierarchies, algorithms can use either the **agglomerative** or **divisive** approaches. The first one starts with every object being a singleton cluster, and at each step, two clusters are merged until there is no cluster to nest and a cluster containing all objects in the data is reached. It is more intuitive than the divisive approach, which starts with all objects in the same cluster and splits the most dissimilar object at each step, creating new clusters until reaching the solution with only singleton clusters. For its simplicity, agglomerative algorithms are better known and used. An optional parameter for hierarchical clustering is the minimum cluster size (MCS). If a cluster is smaller than MCS, the algorithm considers that it does not belong to any cluster at that level of the dendrogram, producing a simplified hierarchy.

The most commonly used methods are Single-Linkage (SL), Complete-Linkage (CL), Average-Linkage (AL), Centroid-Linkage (CTL), and Ward's Method (WM), illustrated in Figure 2.9. SL uses the minimum distance between objects in different clusters to define if the clusters should be merged, what can be disturbed by noise between the clusters. CL uses the maximum distance instead, and outliers can distort the greatest distance of two clusters. AL calculates the average of the distances between all possible pairs of elements from different groups. CTL uses the centroid of each cluster (not necessarily a natural object) to calculate its distance. WM firstly pretends to merge the clusters and get the centroid of this new hypothetical cluster. Then it calculates the sum of the squared deviation of every point to this centroid. If this sum in the new cluster is smaller than in the two clusters individually, then the clusters are merged. All linkage methods have their pros and cons, and it is difficult to predict their performance. Therefore, a good practice is to test different methods and analyze their clustering result, trying to find the

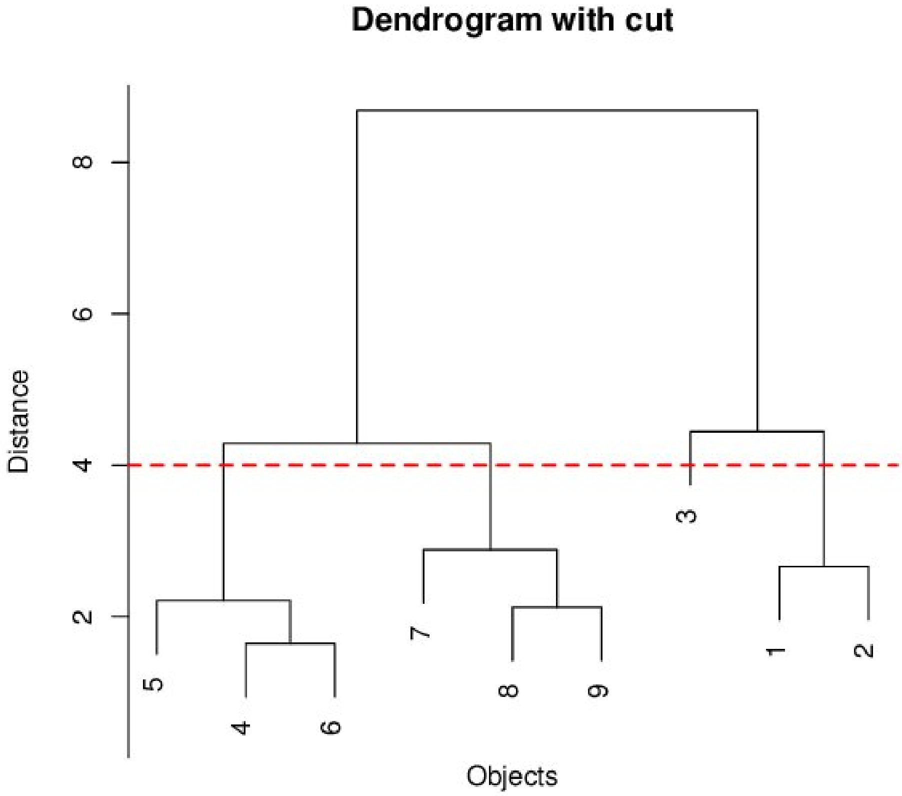

Figure 2.8 – Example of a dendrogram with 9 objects (PEREIRA; MELLO, 2013). The red line in the fourth level cuts the dendrogram extracting four clusters: C1=5, 4, 6; C2=7, 8, 9; C3=3; C4=1, 2.

fittest method for the given data.

2.3.2 Partitional Algorithms

In this approach, the clusters do not overlap. A cluster is a simple division of the data set with its objects, and there are no nested clusters. We can consider the Hierarchical Algorithms (HA) as a sequence of partitional clusterings (TAN; STEINBACH; KUMAR, 2005), as we can take a flat (partitional) solution from a clustering hierarchy by cutting the dendrogram in a given level.

The most famous and widely used partitional algorithm is K-Means (MACQUEEN, 1967). It bases its clustering process on prototypes (centroids) that better represent portions of the data as if each prototype was the center of mass of a group. The first step is to initialize the centroids in the space of data randomly. To reach the solution, at each iteration the algorithm assigns every object to one of the K centroids based on their similarity, which is calculated by the Euclidian distance. Once these assignments are done, each centroid is recalculated as the multivariative mean of the objects assigned to it. In the following iteration, objects are reallocated to their new nearest centroids, and those are updated again. The algorithm stops when the assignments do not

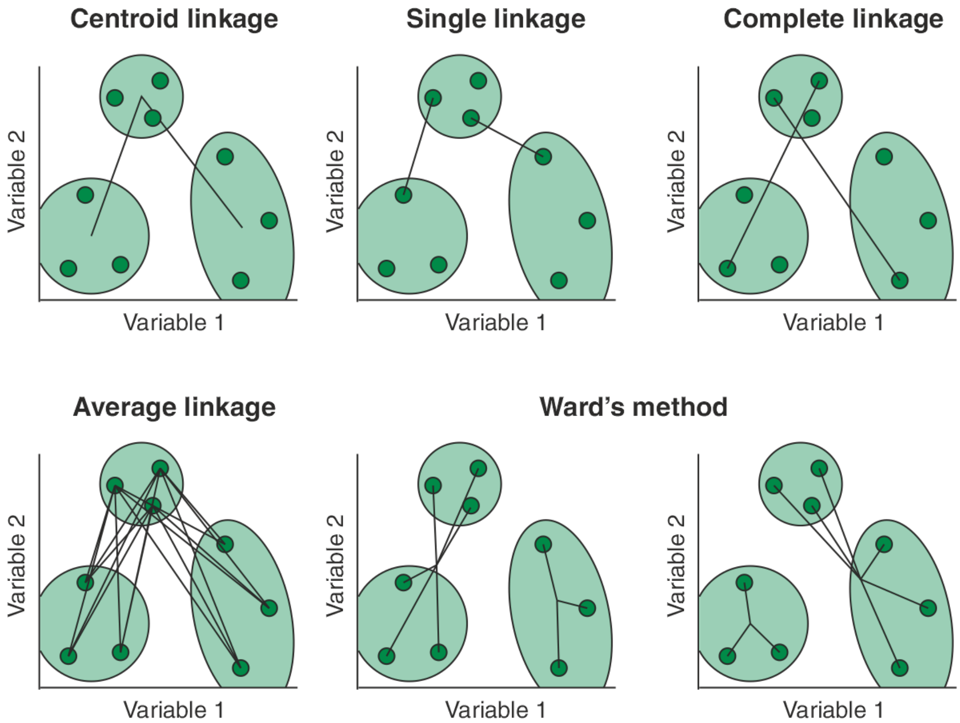

Figure 2.9 – Traditional cluster linkage methods. Available in livebook.

change from the last iteration or when another stop criterion is reached.

By assigning the objects to their nearest centroid, K-Means guarantees to decrease the Sum of Squared Errors (SSE) at each iteration, converging to a solution with a minimum SSE. Note that this solution may not be a global minimum because the results are strongly dependent on the data distribution and the centroids' initialization. A global solution tends to be found by running the algorithm many times with different starting points for the centroids, see (ARTHUR; VASSILVITSKII, 2007). Another critical issue regarding K-Means is the definition of K, the number of clusters to be extracted. Many times in real applications, this number is unknown, so there are also some strategies to find the best number (e.g., Elbow Method (THORNDIKE, 1953) and Average Silhouette Method (KAUFMAN; ROUSSEEUW, 1990), see section 2.3.4).

Other partitional algorithms use the density-based notion of clusters, which is subjectively related to the human intuition for grouping elements by eye. We identify clusters in a bidimensional dataset by simply locating near objects that together form dense regions separated by sparse or empty regions, see Figure 2.10. 2D spaces are used for the purpose of visualization. We cannot do this by ourselves for high dimensional data spaces, but the notion extends to these cases thus can also be applied (KRIEGEL et al., 2011).

DBSCAN is the most famous of these algorithms and stands for "density-based spatial clustering of applications with noise" (ESTER et al., 1996). The essential parameters for this type of algorithm are ε and $MinPts$. The ε-neighborhood of a point are the objects within a radius of size ε from that point. A particular point is considered to be a *core-point* if the number of points in its ε-neighborhood is greater than $MinPts$, a predefined minimum number. If the point is

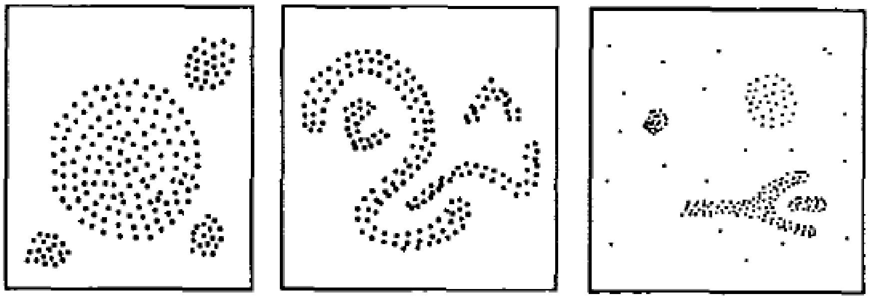

Figure 2.10 – Bidimensional datasets with regions of varying density. Source: Ester et al. (1996).

not a *core-point* but has a *core-point* in its neighborhood, than it is considered a *border-point*. Otherwise, it is considered as noise.

An object **p** is *directly density-reachable* from another object **q** if **p** belongs to the ε-neighborhood of **q**, and **q** is a core point. And an object is *density-reachable* from another object if there is a chain of *directly density-reachable* objects interconnecting them. The clusters are the sets of density-reachable objects distributed along with the data universe. Note that DBSCAN will find a number of clusters that is not previously defined, such as in K-means. Despite this advantage, the technique is very susceptible to its hyper-parameters and may need a large amount of objects to reach consistent results.

2.3.3 FOSC

Hierarchical algorithms provide a lot more information than partitional ones as they produce sequences of partitional solutions (GERTRUDES, 2019). A tricky challenge is finding how to better interpret this whole amount of information in regards to a given research goal. Executing a horizontal cut in a dendrogram's level will give us an optimal solution that is not guaranteed to be globally the best (JAIN; DUBES, 1988). A more sophisticated approach would be to apply local cuts in varying levels of the dendrogram, aiming to extract the best of each flat partition distributed along with the levels. This is what FOSC - which stands for Framework for Optimal Selection of Clusters - does.

Proposed by Campello et al. (2013), FOSC is applied in unsupervised or semi-supervised environments. For the unsupervised context, it uses local measures to quantify the individual quality of each cluster in the dendrogram. These measures can vary depending on the nature of the problem. Once this is done, the goal is to find a flat partition that maximizes the objective function based on a predefined form of aggregation of these individual clusters' qualities. The big difference here is that the clusters can be extracted from different levels of the hierarchy in order to form the globally optimal solution.

See the dendrogram in Figure 2.11. Each node has a value that represents the quality of the corresponding cluster in terms of stability. The measure used by Campello in his study is based on the classic definition of *cluster lifetime*, which is the scale of the hierarchical levels in the dendrogram in which the cluster exists. He made an adjustment: for calculating a cluster's quality, the lifetime of the objects it owns are summed up - the lifetime of an object is the scale until the level where the object is considered as noise. The cluster's quality is related to the concept of Cluster Validation, discussed in 2.3.4.

For the aggregation step, the algorithm uses a bottom-up approach and walks along the hierarchy keeping the best-so-far selection of clusters. For instance, the algorithm would choose the cluster $C5$ instead of its composing clusters $C8$ and $C9$, since $21 > 1.6 + 1.3$. Same thing for choosing $C3$ instead of $C6$ and $C7$ ($36.9 > 1.3 + 1.1$). On the other hand, the cluster $C2$ is not selected because $32.7 < 16.4 + 21.1$.

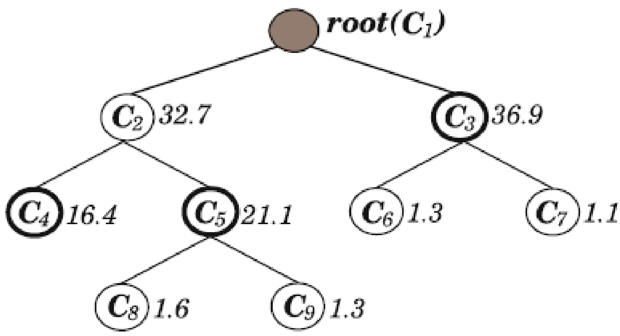

Figure 2.11 – Cluster quality values in terms of stability for a FOSC application example present in Campello et al. (2013)

Two important properties of the cluster stability concept used by FOSC is that it is *local*, since it is calculated for a cluster independently to any other cluster in the dendrogram, and *additive* because it considers all composing objects to calculate the cluster's stability. Due to the *additive* property, the problem FOSC solves is to choose a collection of mutually exclusive clusters that together maximize the sum of their qualities. In the end, it reaches the globally best extraction of clusters from a dendrogram provided by a hierarchical clustering algorithm.

2.3.4 Cluster Validation

Given the usual unsupervised context of CA, a core issue is to evaluate the resulting clusters quantitatively. Since we do not have the labels to tell us whether our classification is right or wrong, statistical methods can provide a solid baseline for inferring if a resulting clustering is a trustful representation of the data. The application of these statistical methods is called Cluster Validation (CV).

Since the definition of a cluster varies depending on the problem, the evaluating methods should also be different for each type of cluster - e.g., SSE can evaluate K-means clusters but would not work well for density-based clusters with non-globular shapes (TAN; STEINBACH; KUMAR, 2005).

It is reasonable to highlight that clustering algorithms will find clusters in the data even when there are no natural clusters, which shows how essential CV is for any CA. Therefore, one of the most important issues for CV is to distinguish non-random structures in the data - in other words, to analyze the **cluster tendency**. Some other issues in the unsupervised context are determining the ideal number of clusters, evaluating the fitness of the results within the data, and deciding between two clusters choosing the best one based on the defined measure.

The two main concepts for unsupervised CV of partitional sets of clusters are: **cohesion** and **separation**. The first concept is related to how similar are the objects in the same clusters. The second one addresses the distinction between the clusters. These measures are called **internal indices** since they do not make use of any external information. The CV can be applied individually to each cluster in the solution or to the solution itself, considering all the composing clusters. The overall validity of a set of clusters can be expressed by the weighted sum of each cluster validity value. This weight also varies according to the measure used (e.g., the size of the cluster or the square root of the cohesion).

2.3.4.1 Silhouette Coefficient

A popular method that combines both concepts described above is the **Silhouette Coefficient** (ROUSSEEUW, 1987). Its process for computing the coefficient for a given point consists of three steps:

1. For the i^{th} object, calculate the average distance to all objects in the same cluster and call it a_i;
2. For the i^{th} and any cluster that it does not belong to, calculate the average distance to all objects in that cluster. Take the minimum value found a call it b_i;
3. Finally, for the i^{th} object, the silhouette coefficient is defined as $s_i = (b_i - a_i)/max(a_i, b_i)$;

Figure 2.12 shows an example of Silhouette Coefficient index calculation for the point P. The values vary between -1 and 1. We want a_i to be as close to 0 as possible and s_i to be positive. Actually, the greater the value of s_i, the better the validation for that given point. By doing that, we analyze how the object within a cluster collaborates to the overall cohesion or separation of the cluster. For validating an entire cluster in terms of silhouette coefficient, we take the average of the silhouette coefficients of each of its composing objects. And for validating the whole clustering solution, we take the average of every cluster's silhouette coefficient index.

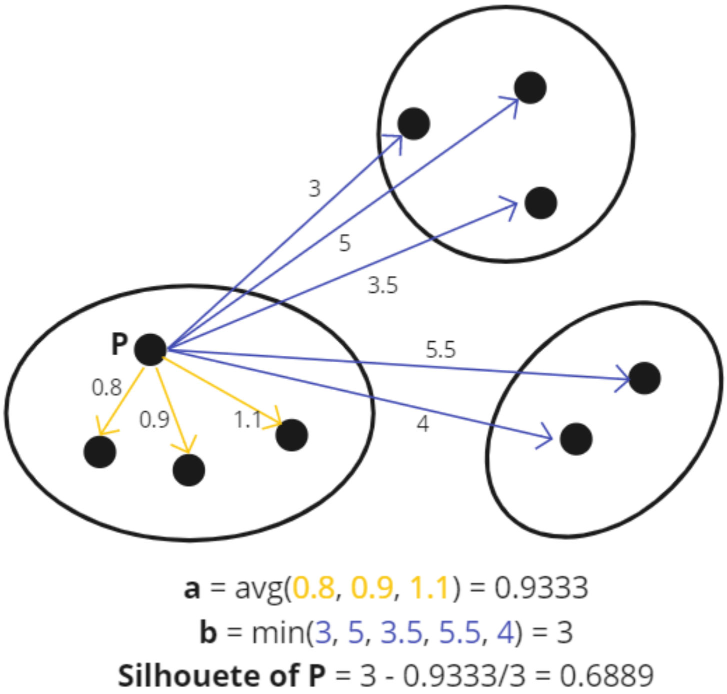

Figure 2.12 – How to calculate Silhouette Coefficient for an specific point P.

2.3.4.2 AUCC

The AUCC (*Area Under Curve for Clustering*) (JASKOWIAK; COSTA; CAMPELLO, 2020) is an adaptation of the AUC - abbreviation for *Area Under the Receiver Operating Characteristics (ROC) Curve* - typically used in supervised contexts to analyze performances of classifiers. To understand the AUCC, we must start with the concept of ROC and AUC.

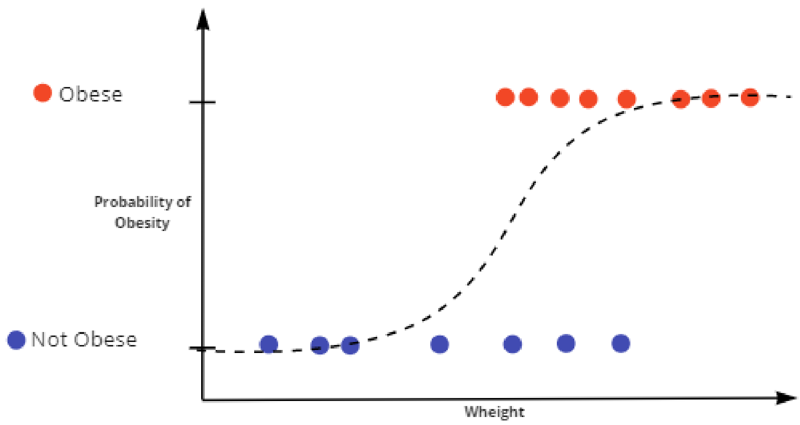

Figure 2.13 – Logistic Regression over 15 data points.

The ROC curve is a way to summarize the predictions' quality of a classifier for every

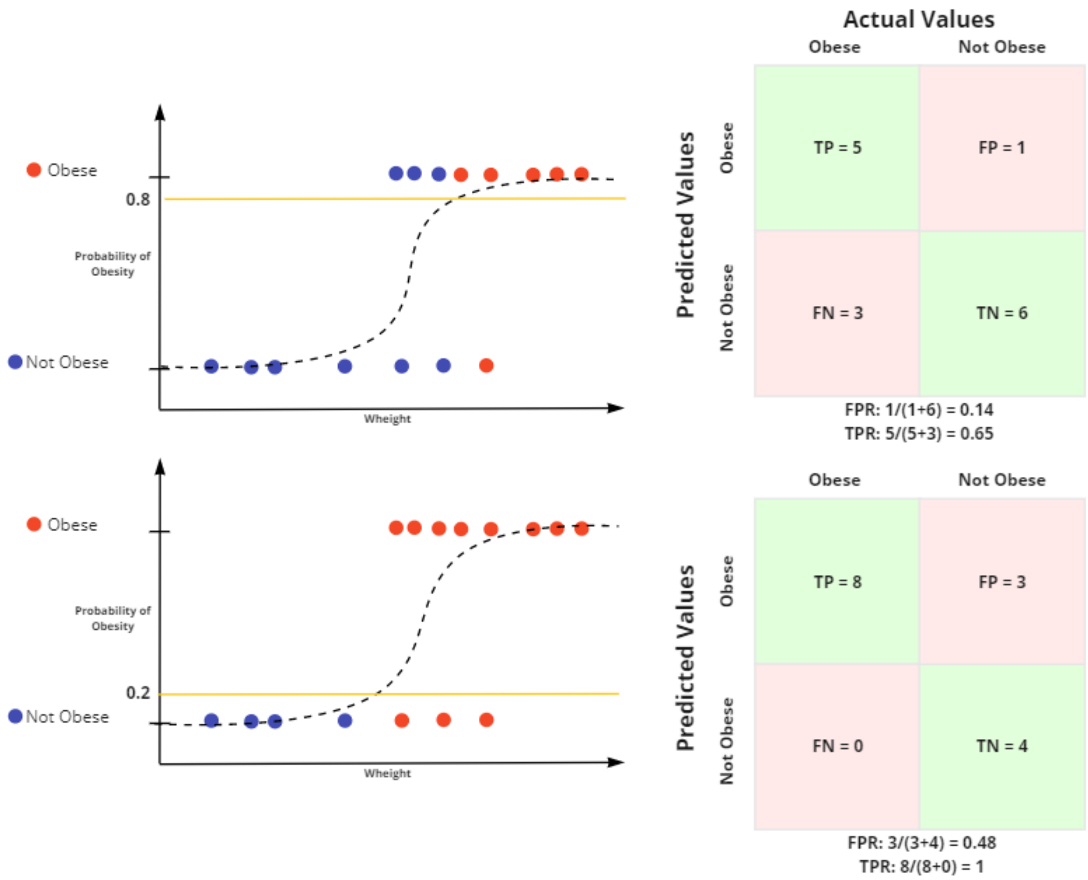

Figure 2.14 – Two probabilty thresholds (0.8 and 0.2), and their respective confusion matrix, with values of FPR and TPR.

possible threshold value based on the actual labels. Consider the problem of classifying if a person is obese based on their weight. Figure 2.13 shows an example of a Logistic Regression that would tell, for each data point, its probability of corresponding to an obese person. Since we want to tell whether the person is obese or not obese, we need to set a threshold to discretize the classifier's output.

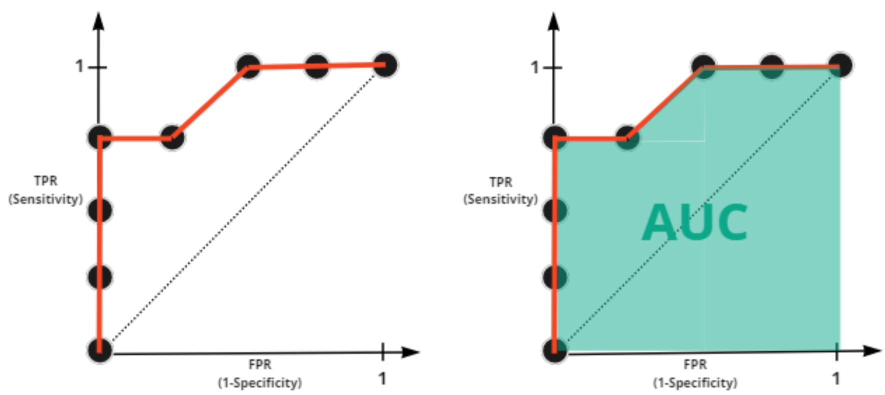

Figure 2.15 – Ilustrative ROC curve and AUC of 0.9.

Each prediction can be labeled as True Positive (TP), False Positive (FP), True Negative (TN), or False Negative (FN), based on the actual classes. A Confusion Matrix presents the number of predictions for each case (Figure 2.14 shows the Confusion Matrices for two threshold examples). By plotting the value of FPR (*FPR = FP / (FP+TN)*) against the value of TPR (*TPR = TP / (TP+FN)*), we can draw a single point on the ROC graph that represent the relation of the FPR and TPR, indicating the quality of the prediction for the current threshold. Figure 2.15 shows an example of a ROC curve for nine different thresholds and the Area Under the ROC, which would have the value of 0.9. The diagonal line represents the solutions where *FPR = TPR*. Therefore, we want the points to be on the left side, meaning that *TPR > FPR*, thus increasing the AUC. A more excellent AUC value indicates a more robust classifier, and the best threshold for a given problem depends on how many False Positives the problem can accept.

As an adaptation for the unsupervised context, if we want to compare clustering methods using the partitions produced, we need to follow these steps:

1. Compute the similarity matrix of all the objects in the clustering population;
2. Obtain an array that indicates the pairwise similarity for each pair of objects;
3. Obtain an array that indicates the pairwise clustering for each pair of objects: 1 if the pair is in the same cluster, 0 if not;
4. Use the two arrays as input to the ROC Analysis. The analogy is: pairwise similarity values correspond to the possible thresholds, whereas the pairwise clustering array corresponds to the true labels;

The output will give the AUC of the clustering partition, which is named AUCC in the unsupervised context. Therefore, when comparing two partitions, the one with the bigger AUCC value can be considered the most robust solution.

3 Related Work

ET data have been used with different approaches. As mentioned, ML algorithms can help with the tracking process, but they are also effectively used to analyze the data collected. In some cases where researchers analyze health conditions related to the eyes, the data is labeled concerning the condition studied. In these cases, supervised classification methods can be used for training diagnostician models, and some examples are shown in the following paragraphs.

In Chen et al. (2018a), they used six different Convolution Neural Network models trained on the public dataset ImageNet for feature extraction of the so-called GaDe (Gaze Deviation) images. These images are generated from gaze data recorded by the Tobii X2-60 eye-tracker while the subjects perform fixations at 9 points on the screen. The CNN outputs a feature vector used by an SVM to detect Strabismus. The SVM is trained with 42 ET records (17 from subjects with Strabismus diagnosed by vision specialists). When compared to other studies, a peculiarity is that the authors do not preprocess the ET data for detecting events (e. g., saccades, fixations). As mentioned by them, these events are not relevant for Strabismus recognition, so the GaDe images are enough for a good classification performance. They have reached the highest performance with 95.2% of accuracy.

Previous psychology research has shown that the lack of eye movements control is a reflection - not a cause - of the difficulty that people with dyslexia have while reading (HYÖNÄ et al., 1995; PIROZZOLO; RAYNER, 1979; RAYNER, 1985). From that, once this lack of eye control is detected, it can be inferred that a subject has a probability of having some level of dyslexia. In (RELLO; BALLESTEROS, 2015) they have used a dataset built with 1135 ET data from participants with and without dyslexia while they were reading Spanish texts. An SVM classifier was used to diagnose. In this case, they detect the eye movement events first to calculate some of the features used by the SVM (e.g., nº of fixations, duration mean of fixations). They have reached an accuracy of 76.38% while not considering the age of the participants (which makes much difference since an old dyslexic subject may have improved their reading capabilities through their years of life).

Besides ET data, other types of signals - e.g. electroencephalogram (EEG) - are also used to detect symptoms of psychological disorders. Different signals can be a good tool for reaching a satisfying diagnosis accuracy. For detecting Autism Spectrum Disorder (ASD), Thapaliya, Jayarathna and Jaime (2018) compares classification performances using many combinations of features, algorithms and data. They compare the performances using EEG data, ET data, and both combined. The results have shown that ET data ended up being more relevant for the algorithms to have better classifications. The dataset used for training contains EEG and ET records of 52 participants (24 with ASD) while watching video clips. They used the Tobii X50 for tracking eye movements, and the fixations were calculated using the Clearview algorithm (from Tobii). Despite using the fixations as features, they also used the score of each individual in a social cognition test. Some of the accuracy performances were pretty good, but as mentioned in the paper, the number of participants was too short to confirm the consistency of the results. Some other studies that have aimed to investigate ASD using ET and ML are Liu et al. (2015), Alie et al. (2011), Jiang and Zhao (2017).

In Ding et al. (2019) they use EEG, ET, and Galvanic Skin Response (GSR) while the 348 participants watch affective and neutral stimuli in 8 short videos. The goal is to diagnose Major Depression Disorder (MDD). Their ET data was recorded using Tobii 4C at 90Hz, and the fixations were defined when the gaze was stable for more than 100ms within a 1º of visual angle. The features used from ET were: percent time attending to each image category, number of fixations per category, mean glance duration, and location of the first fixation. They used Random Forest, Linear Regression, and Support Vector Machine to build the classification model. They have reached the highest f1 score by using EEG data. However, in general, they have found better results when combining the three different measurements.

In Franceschiello et al. (2020) the participants performed 176 trials of a demanding visual search, complicated by distractors on the screen. Since spatial neglected subjects have more difficulty on targets at the left side of the screen, their performance is analyzed considering only the x coordinates. Relevant trajectories' features were extracted by passing these vectors through a CNN. They also made a so-called "crucial" interplay between the ET data-based classification, anatomical markers detected through MRI (Magnetic Resonance Imaging) and DTI (diffusion tensor imaging). The average accuracy on the classification obtained was 74.5%.

On the other hand, in the case of C&Look we do not have labels to tell us whether a subject carries or not an oculomotor dysfunction. Therefore, the ML methods for unsupervised classification are suitable for us. To the best of our knowledge, not much work has been done on clustering ET data with similar goals, but some related examples are highlighted in the following paragraphs.

Murray et al. (2019) aims to examine the reliability of oculomotor metrics to determine normative values through cluster analysis and compare these metrics by age groups. A large sample with 2993 participants performing the RightEye tests was recorded using the SMI 12" 120Hz eye tracker, and the participants were selected only if they were healthy individuals. 5 Right Eye tests were applied, and the normative values were calculated for each one. For data analysis, they used Cronbach's Alpha to evaluate the reliability of RE tests. They also built a clustering model based on Bayesian Information Criterion. Lastly, they examined group differences considering age clusters and gender with MANOVA test, one for each of the 5 tests. The clustering revealed 5 age groups, and the correlations between individuals in each group have shown that age ranges must be considered in an oculomotor test battery.

Göbel and Martin (2018) study the behavior of map readers and how they spent their visual attention. The work evaluates the suitability of unsupervised ML methods for clustering the ET data collected from subjects while they perform everyday comparison tasks on maps. Three different maps with different legend types were used. Similar to our approach, they applied feature extraction on the time series, creating 37 new features. They proposed a framework for clustering the ET data with three different methods: K-means, Spectral Clustering, and DBSCAN. The user can also specify the hyperparameters that the clustering methods will use. Once the CA is done, the resulting clusters are validated with the t-SNE visualization in order to discover their qualities. They tested 29700 combinations of hyperparameters. Although they demonstrated the t-SNE suitability for cluster evaluation, they enhanced the need for domain knowledge to create more elaborated features and detect meaningful results.

Naqshbandi, Gedeon and Abdulla (2016) used K-Means and OPTICS (a density-based cluster technique) within ET data for separating regions of interest in pictures when participants were asked subjective questions to answer verbally; some examples are: "Do you think people in this image are related to one another?", "Why are some people carrying rifles?", "What is this image about?". The fixations were collected from the gaze points in order to detect which parts of the picture call the participants' attention while they gather information to answer the questions. For each task of each subject, the clusters found by K-Means and OPTICS were used in training and decoding tasks with Hidden Markov Models.

Aside from diagnosing various health problems, ET and ML can be used for different aims. The general path of the ET data processing and how it is fed into a classification model can be very similar between studies with different objectives, and that is why any study regarding the use of ET and ML can be useful for this present project. Another exciting application with the fusion of these two technologies is biometric identification. Jäger et al. (2019) developed a deep convolutional neural network for biometric identification focusing on eye micro-movements. The main point of their method is dividing the model into two subnets, one for fast movements and the other for slow movements. They do that because the global normalization would squash the velocities of drift and tremor to near-zero. So what they do is to enhance macro-movements for one subnet and micro-movements for the other one. They have used a database containing 12 reading records from 75 participants. The fast subnet (for microsaccades and saccades) reached an accuracy of 77% after one second. The slow subnet (for tremor and drift) reached an accuracy of 88% after one second. And the full subnet reached 91.4% after one second, 99.77% after 10 seconds, and 99.86% after 40 seconds. As related by the authors, their method overcame all the biometric identification methods known so far.

4 Material and Methods

This Chapter firstly describes in Section 4.1 the database characteristics and the issues to be taken into account. Then, Section 4.2 presents the Event Detection methods considered for identifying the types of eye movements. In Section 4.3, we describe the features used for characterizing the recordings. Finally, Section 4.4 presents the suggested pipeline for applying all the concepts involved.

4.1 Database

The database was collected by the HVL's research group SecEd, which is connected to the project Securing Education for Children in Tanzania. The database is composed of recordings of children performing the C&Look tasks. This software communicates with an eye-tracker collecting the user eyes' information for vision screening (EIDE et al., 2019). Tobii Eye Tracker 4C collected most samples, but the models TX300 and EyeX were also used. The tasks are simple and structured (i.e., the child is not freely observing the screen, it is supposed to be following a figure or reading a text). Once the task is finished, C&Look stores the eye-tracker data with the associated user information in the database for further analysis.

SecEd visited schools in 4 different cities of Tanzania: Moivaro, Tysnes, Austevoll, and Arusha - using C&Look to collect children's performances under specific circumstances. In Moivaro, the experiments were divided into two groups of users, one being composed only of intellectually impaired (I. I.) children. In Arusha, the group realized many visits in order to collect the children's performances before and after vision training sessions - which is also a focus of HVL researchers (ALI et al., 2021a). The valid recordings (i.e., only those collected from children) were divided into nine groups in the database, as shown in Figure 4.1.

Group	Nº of Children	Nº of Records
Moivaro	127	2483
Moivaro I. I.	21	282
Tysnes	15	276
Austevoll	39	699
Arusha	126	1369
Arusha Pre-Training	25	278
Arusha Post-Training	24	278
Arusha Rescreening	21	214
Arusha Control	21	135

Figure 4.1 – Number of recordings and children per group

4.1.1 Issues for CA within C&L database

Some relevant issues about the C&L database and its nature must be considered before applying CA within the recordings. The first issue is the number of valid recordings available for composing a given clustering population. Mainly for unsupervised classification, a proper amount of objects is required so that the models have enough information for training and reaching significant results. The database is composed of 5998 recordings from children. However, they were recorded under different conditions and have different characteristics, leading to the necessity of verifying which recordings can be together in a CA configuration, depending on the research goals. The following subsections will explain the factors considered for filtering the recordings.

4.1.1.1 Device

The performances were recorded with three different eye-trackers from Tobii (4C, EyeX, and TX300). The devices are placed under the screen, tracking the eyes remotely, which is very good when dealing with children. Using recordings of different eye-trackers would unbalance the feature engineering because each eye-tracker has its specific characteristics (e.g., frequency, accuracy). Therefore, we chose to work only with Tobii 4C (see Figure 4.2), because there is a more significant number of recordings available in the database. Also, it is the most recently released of the three ones.

Figure 4.2 – The Tobii Eye-Tracker 4C, released in 2016. Source: https://www.amazon.com/Tobii-Eye-Tracker-Game-changing-Peripheral/dp/B01MAWPMXQ

4.1.1.2 Task

C&Look includes tasks with different stimuli, some in which children have to read a text or sets of syllables, and others in which children have to follow a moving figure along the screen. For the second case, which is our focus, there are tasks in which the figure has different trajectories, thus exploring different functionalities of the oculomotor system (e.g., fixations, vertical/horizontal/diagonal pursuit, big/small saccades).

Since we want to compare children's oculomotor system's behavior, we have to do it using only the very same stimulus. In practical words, the figure has to have the same trajectory and speed (we did not consider the size of the figure or the figure itself). Therefore, we chose to focus on one specific task, in which the figure applies ten fixations over three lines of the screen, moving from the left to the right, starting on the top. Figure 4.3 shows an example of good performance over the chosen task.

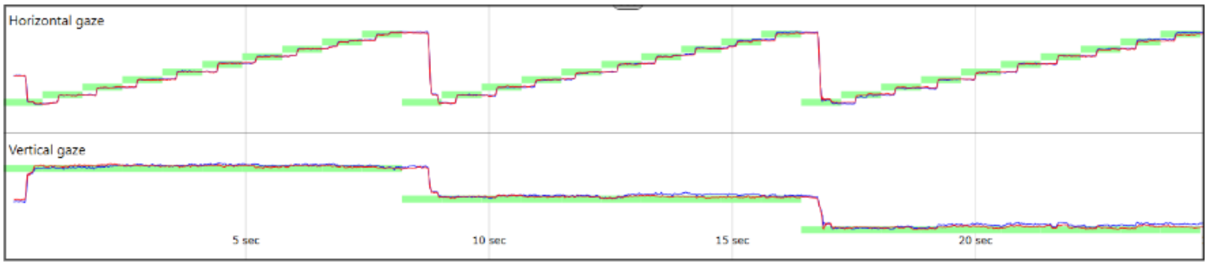

Figure 4.3 – Example of a good performance for the chosen task. Horizontal and Vertical gazes are presented separately. Source: C&Look.

Not every feature extracted from the time series (see Section 4.3) depends on the task presented. However, even so, we decided to realize the experiments only with one task to limit our scope and focus on specific characteristics of the chosen stimulus.

4.1.1.3 Calibration Step

For every eye-tracking application, the calibration step is essential for guaranteeing a good tracking of the eyes. People have very different eye characteristics, which leads to the necessity of calibrating the eye-tracker for each subject. During this procedure, the eye-tracker captures many images of the subject's eye to build an eye model that will provide a more accurate gaze detection. This model has information about the shape, light refraction, and reflection properties of each part of the eye (e.g., cornea, placement of the fovea) (TOBII, 2022). With this information, the device tests different methods (e.g., dark or bright pupil detection) for tracking the eyes in order to choose the most suitable one for the given light conditions. After this procedure, the calibration's quality is illustrated by showing the offset of the gaze points for each calibration dot. Figure 4.4 shows an example of a relatively good calibration.

After the calibration, there is the validation step, which is useful for testing the quality of the new eye model that was built. The validation step uses more targets to evaluate the eye-tracker's performance. Figure 4.5 shows the validation for the calibration on Figure 4.4.

4.1.1.4 Tracking Quality During Performance

Even when the calibration and validation steps seem promising, the tracking quality during the child's performance can occasionally be poor. It probably happens when the child changes its distance or position in relation to the screen right after the calibration process or when there is an abrupt change in light conditions. Figure 4.6 shows an example of a relatively good validation but a poor tracking during the child's performance. Since we want to focus on qualifying performances, the recordings must have consistent tracking of the eyes for not to influence the feature engineering negatively.

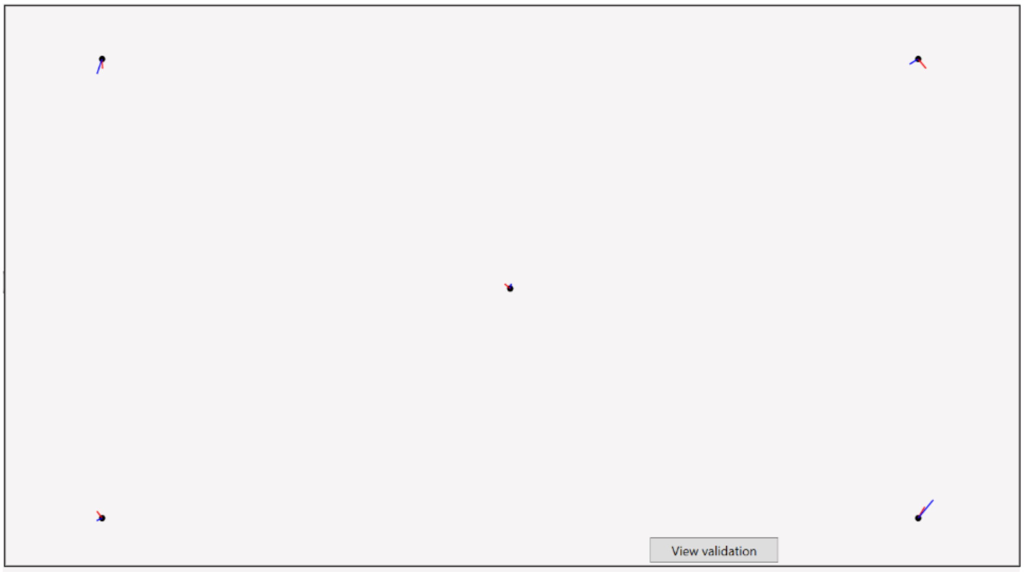

Figure 4.4 – Example of a good calibration step. Source: C&Look.

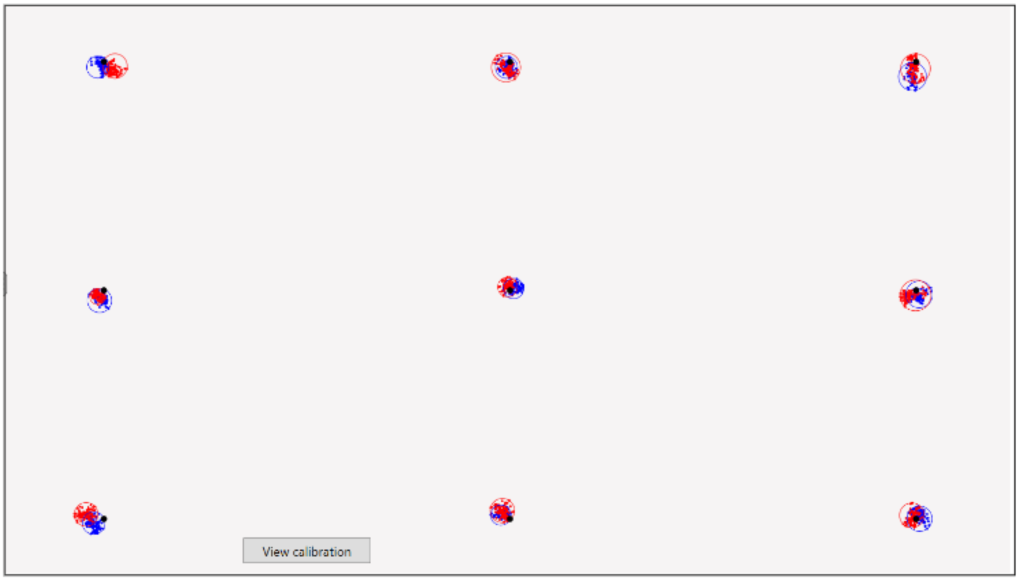

Figure 4.5 – Validation for the calibration in Figure 4.4. Source: C&Look.

The performance is presented with the Horizontal and Vertical plots of the gaze positions, which is a common approach to visualize eye-tracking data. The blue line corresponds to the left eye's gaze, and the red line to the right eye's gaze. The green lines represents the figure's positions.

As in the example of Figure 4.6, there are timestamps when the signals from the eyes are lost. In these moments, the normalized positions of the eyes are set to (0, 0) in the database. Clearly, we must not use these values for building the features because they would reflect false characteristics of the performance. In order to support this issue, recordings with good calibrations but poor tracking during the performance were also filtered.

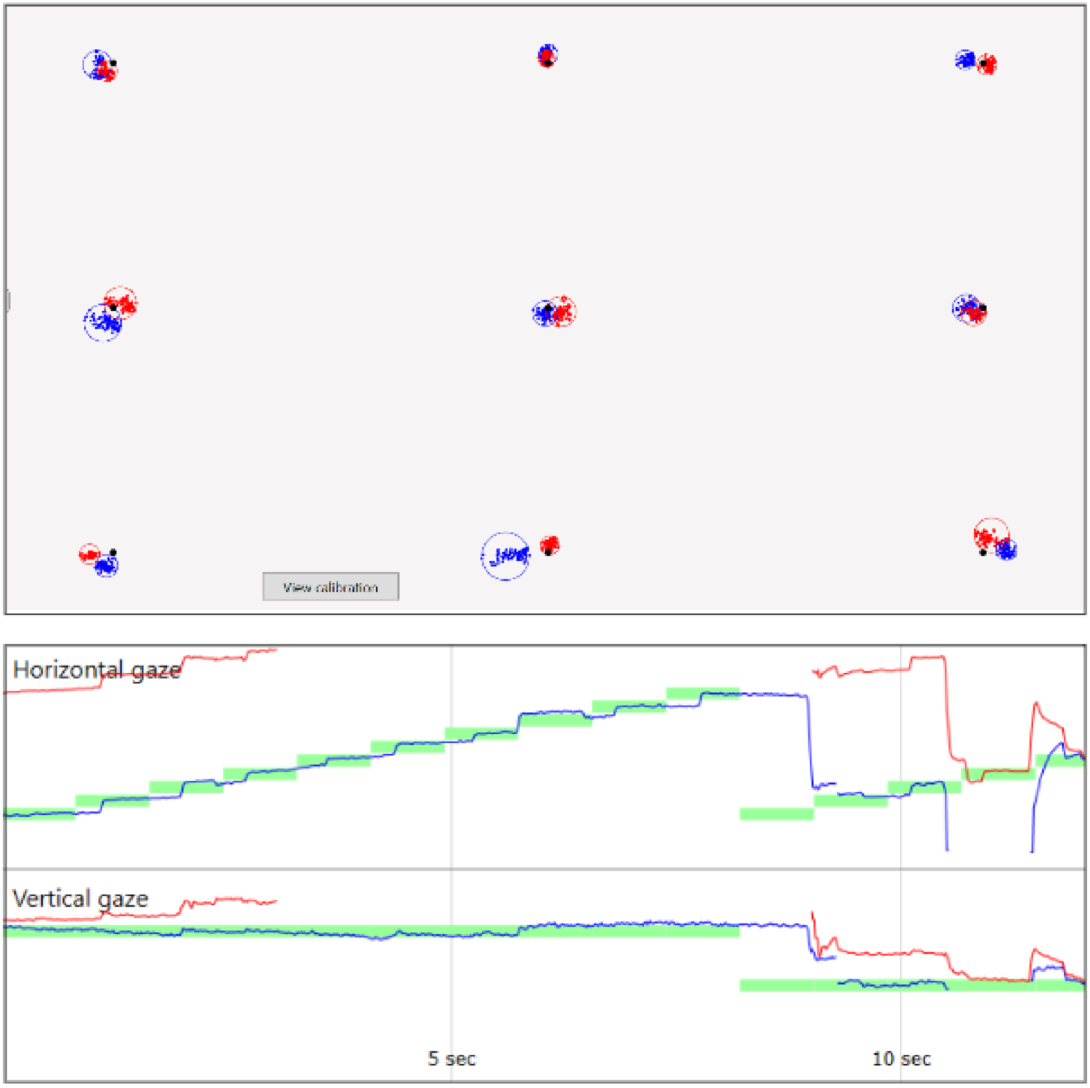

Figure 4.6 – Example of good validation succeeded by a poor tracking during performance. Source: C&Look.

The eye-tracker stores the *tracking status* for each timestamp. The values go from 0 to 3, indicating whether both eyes are tracked or which one is lost:

- 0 - both eyes tracked;
- 1 - both eyes lost;
- 2 - right eye lost;
- 3 - left eye lost;

This values were taken into account during the feature engineering, that will be explored in Section 4.3.

4.2 Event Detection

Identifying eye movements (i.e., fixations and saccades) is a fundamental step in the feature engineering process when working with ET data. The results of this step drastically impact the higher-level analysis. The approaches to use and the events to detect depend on the research goals (WIBIRAMA et al., 2020). Also, equipment variabilities are relevant issues to consider (e.g., quality of the eye-tracker and sample frequency).

There are three classic strands of algorithms for event detection: I-VT, I-DT, and I-AOI (SALVUCCI; GOLDBERG, 2000). I-VT algorithms separate fixations and saccades based on the point-to-point velocities. If the velocity is below a defined threshold, the point belongs to a fixation group of points. Otherwise, it is a saccade point. I-VT is a straightforward and robust strategy, easy to understand and implement. I-DT is based on the dispersion of consecutive points and incorporates a minimum duration threshold of 100-200ms for fixations. It uses a moving window (which width is determined by the duration threshold and sampling frequency) that checks the dispersion of the points inside, D = [max(x) - min(x)] + [max(y) - min(y)]. If the dispersion is below a determined threshold, that sample interval belongs to a fixation. Finally, I-AOI is used in more specific cases when a stimulus is presented (e.g., targets, images, or videos displayed on a screen). This type of algorithm focuses on predefined areas of interest, identifying fixations close to specific parts of the visual field.

A major drawback of these algorithms is manually setting the parameters (i.e., thresholds). A slightly different configuration can end up leading to very different outputs. Another limitation is that they only detect a single event, normally fixations, and cannot classify smooth pursuit, PSO, etc. Given these circumstances, in the last decade, machine learning algorithms for event detection have provided a whole range of new descriptions of eye-tracking data (ZEMBLYS; NIEHORSTER; HOLMQVIST, 2018). However, using these algorithms requires a training dataset with hand-segmented ET samples. For instance, in Hoppe and Bulling (2016) they use a simple CNN to identify eye movements. The model was trained in a dataset introduced by the authors, with 1626 fixations, 2647 saccades, and 1089 pursuit movements. The results outperform existing approaches, classifying smooth pursuits instead of only fixations and saccades.

In our case, event detection is the root of the characterization of the samples, represented as featured objects. Therefore, accurate identification of eye movements is essential for trustful and meaningful clustering performances. There are not many good Python implementations for the described algorithms that would work well for our data, and the moderated sampling rate of Tobii 4c (90Hz) is a limitation for the quality of event detection. We haves used the implementations available in the following public GitHub repository: <https://github.com/ema2159/ET_Project>, choosing the parameters empirically. Even though the event detection algorithms can give different outputs for the same time series, we used only the I-VT implementation, with the same parameters for every sample. Therefore, we must highlight that the event detection was quite a bottleneck for our feature engineering process (see Section 4.3) since we only detect fixations and saccades, not considering blinks, regressions, PSO, or any other event.

4.3 Feature Engineering

Raw data can be numerically represented by features derived from the type of data available. In our case, these features tend to characterize the time series in terms of performances for the chosen task. The number of features to use is a delicate choice to make, because if this number is too small, the model will not have enough information to succeed. And if it is too big, it will be more difficult and costly to train (ZHENG; CASARI, 2018). Therefore, the features used must be meaningful for the ongoing analysis.

Creating features and choosing what to base them on is a tricky job; thus help from specialists in the problem domain is essential. Saiz-Manzanares et al. (2021) used more than 30 features to characterize the eye-tracking data collected for statistical analysis, including counting fixations, saccades, and blinks, such as their frequency, duration, and dispersion. The eye tracker used had a sampling rate of 60Hz. Their work analyzes subjects solving a crossword puzzle task, so not all the features in their study would be relevant for our case.

In our work, we built 15 features to represent the time series of each recording. Their descriptions are in the table in Figure 4.7. After extracting these features, we stored the featured objects in a new database table, so we could use them for the exploratory analysis with different configurations. Although we have to keep straight to our goals, it is relevant to cite that a higher sampling rate (700Hz-1200Hz) would gather more information about the head and eye movements (STEIN et al., 2021), thus opening the way to build more features to represent the time-series.

Initials	Meaning
FC	Fixation count
AFD	Average duration of fixation
AFDisp	Fixation dispersion average
AFDispH	Fixation horizontal dispersion average
AFDispV	Fixation vertical dispersion average
ASD	Average duration of saccade
ASA	Average saccade amplitude
SAMax	Maximum of saccade amplitude
ADTF	Average distance to target during fixation
SPL	Scan Path Length
ADT	Average distance to target
ADB	Average distance between eyes
ADTL	Average distance to figure on the left side of the screen
ADTR	Average distance to figure on the right side of the screen
ADFF	Average distance to first fixations (1°, 11° and 21°)

Figure 4.7 – Features extracted from the time-series. The ones in yellow do not depend on the event detection step.

Most features were based on literature (MURRAY et al., 2019; BARGARY et al., 2017;

SAIZ-MANZANARES et al., 2021), and others were based on the distance between the gaze and the task's figure position. This second type tries to reflect the child's performance in terms of the task goals (i.e., follow the figure).

Since we know that the task's figure itself applies 30 fixations with regular intervals (0.8s), to count how many fixations the subject applies (*FC*) and the duration of each fixation (*AFD*) fits our scenario. The eye gaze data of perfect performance would have exactly 30 fixations (as equal as possible to the figure's fixations); thus, discrepant values may indicate a problem in the child's performance. The *SPL* feature follows the same idea. The dispersion of the fixations (*AFDisp*) is more related to the event detection itself, since it is a important parameter for the algorithms, also studied with horizontal/vertical dispersion (*AFDispH*/*AFDispV*) separated (SALVUCCI; GOLDBERG, 2000).

Since we are detecting only fixations and saccades, counting the number of saccades would not be relevant for the CA because it would always be one less than the *FC*. Nonetheless, information from saccades regarding their duration and amplitude (*ASD*, *ASA*, *SAMax*) can be useful (SAIZ-MANZANARES et al., 2021). *ADT* takes the average distance from the gaze to the figure, and *ADTL*/*ADTR* do it separately, which seems to be an interesting approach since the oculomotor system can have problems with different parts of the visual field (GERSTENBLITH; RABINOWITZ, 2021). Since saccades are ballistic movements - as mentioned in Section 2 - we also built the *ADTF* to collect the distance from gaze to figure only when the subject has control of its eyes. In other words, during fixations (i.e., humans are effectively blind during saccades). *ADFF* was built to focus on assessing how well the child handles the beginning of the task and the line jumps, measuring the average distance for the first fixations of each line. The only feature that involves both gaze data is *ADB*, which takes the average distance of both eye gazes during the recording. This feature can reflect a calibration problem and may also carry valuable information regarding strabismic conditions (CHEN et al., 2018b).

The practical relevance of each feature is complex for us to define since that should be done by a domain expert (i.e.., vision specialist). Nonetheless, we used many subsets of features for each round of Cluster Analysis to try to make the performances be reflected in different ways. Section 4.4 describes how the experiments were conducted.

4.4 Exploratory Analysis

"The path from data to answers is full of false starts and dead ends. What starts out as a promising approach may not pan out. What was originally just a hunch may end up leading to the best solution." - (ZHENG; CASARI, 2018)

The exploratory analysis can cover a wide range of possible clustering configurations. Given the recordings' peculiarities and how subjects and tasks are divided, there are many combinations for us to select the clustering population, the method and hyper-parameters to use, and the features that will characterize the time series. It is impossible to know a priori which settings will produce the best results and overcome our expectations. In order to facilitate the experiments and analysis of the clustering results, we propose a framework for CA based on the pipeline presented in Figure 4.8:

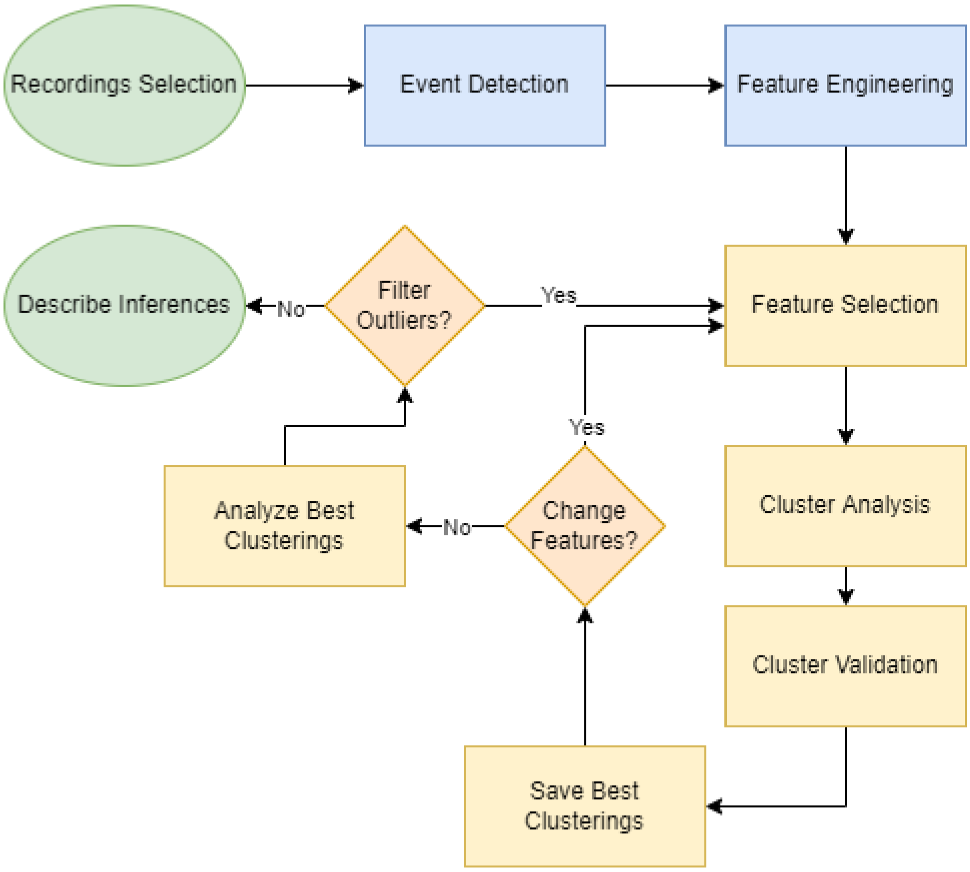

Figure 4.8 – Pipeline for clustering C&Look data

This pipeline is similar to other approaches for knowledge extraction, as KDD, SEMMA, or CRISP-DM (AZEVEDO; SANTOS, 2008). Also, Göbel and Martin (2018) proposed an analogous idea for working with ET data. Nonetheless, the peculiarities of working on C&Look data and building different features may demand a customized pipeline for dealing with our issues.

Each step of the pipeline has the following roles:

- **Recordings Selection**: select the subset of recordings to be turned into featured objects for further cluster analysis. This selection can be based on groups, tasks, or other criteria depending on the objective of the current analysis. In the case of our experiments, we selected the recordings from the task described in Section 4.1.1.2 that were recorded using Tobii Eye-Tracker 4C;
- **Event Detection**: in this step, the detection of eye movements (i.e., fixations and saccades) occurs. These events are used for building the features in the next step. As mentioned,

we have used the I-VT implementation of this freely available Python package: <https://github.com/ema2159/ET_Project>;

- **Feature Engineering**: here, the features are extracted from the events detected in the previous step or from the time series itself. There are 14 features for each eye and 1 feature that involves both of them (ADB - Average Distance Between Eyes, details in Section 4.3);

- **Feature Selection**: in this step, we select the eye data to analyze and the subset of features that will be used in the next round of CA and CV. We can create many different subsets to see which features are more relevant to the cluster separation (again, inferences about the relevance of the features must be supported by a vision specialist).

  To create the subsets, we calculated the Correlation Matrix between all the features to find which features should be used together for clustering the data of each eye (and for both eyes combined). For each feature, we took the 4 least correlated features and built subsets of 2, 3, 4, and 5 features, exploring every combination. Figure 4.9, 4.10 and 4.11 show the correlation matrix for left, right, and both eyes data, respectively, considering all the 464 featured objects.

- **Cluster Analysis**: in this step, we use the clustering methods described in Section 2.3 (K-Means, DBSCAN, and FOSC) within the featured objects. The methods run with varied parameters since we can never know which is the best configuration for that clustering population. Figure 4.12 shows the variations of parameters values for each clustering method;

- **Cluster Validation**: here, we use the Silhouette Coefficient index (described in Section 2.3.4) to assess the quality of the clustering results - AUCC was also calculated;

- **Save Best Clustering**: after the cluster validation, we save the best clusterings based on a minimum threshold for the Silhouette Coefficient index;

- **Change Features?**: here, we check if there are more subsets of features to try. If so, we select the new features and start the cluster analysis again;

- **Analyze Best Clusterings**: we use different visualization tools to analyze the best partitions and the clustering process. We used PCA (Principal Component Analysis (HOTELLING, 1933)) to scatter the data in two dimensions and visualize the partitions created by the best clusterings. Also, we used Decision Trees (BELSON, 1959) that show which features were discriminant during each clustering process. These Decision Trees can be very insightful for vision specialists when analyzing the clusters. The visualizations will be shown in Section 5;

- **Filter Outliers?**: after analyzing the best clusterings with the visualization tools, we can check if the outliers correspond to recordings with poor tracking during the subject's performance. If so, we remove them from the clustering population, because, as mentioned in Sections 4.1.1.3 and 4.1.1.4, they would not reflect a lousy child's performance, but a problem with the tracking itself (which is not the focus of our research);

- **Describe Inferences**: if the outliers do not correspond to bad recordings, there are no more CA configurations to try, and the focus can be on finding inferences about the children's oculomotor system. As mentioned before, these inferences must be confirmed by a vision specialist. Finally, the clusters should not be used to define if a child has or not an oculomotor dysfunction but to indicate which of them should visit a doctor or perform the task again;

The exploratory analysis was conducted based on the presented pipeline, and the results of our experiments will be explored in Section 5.

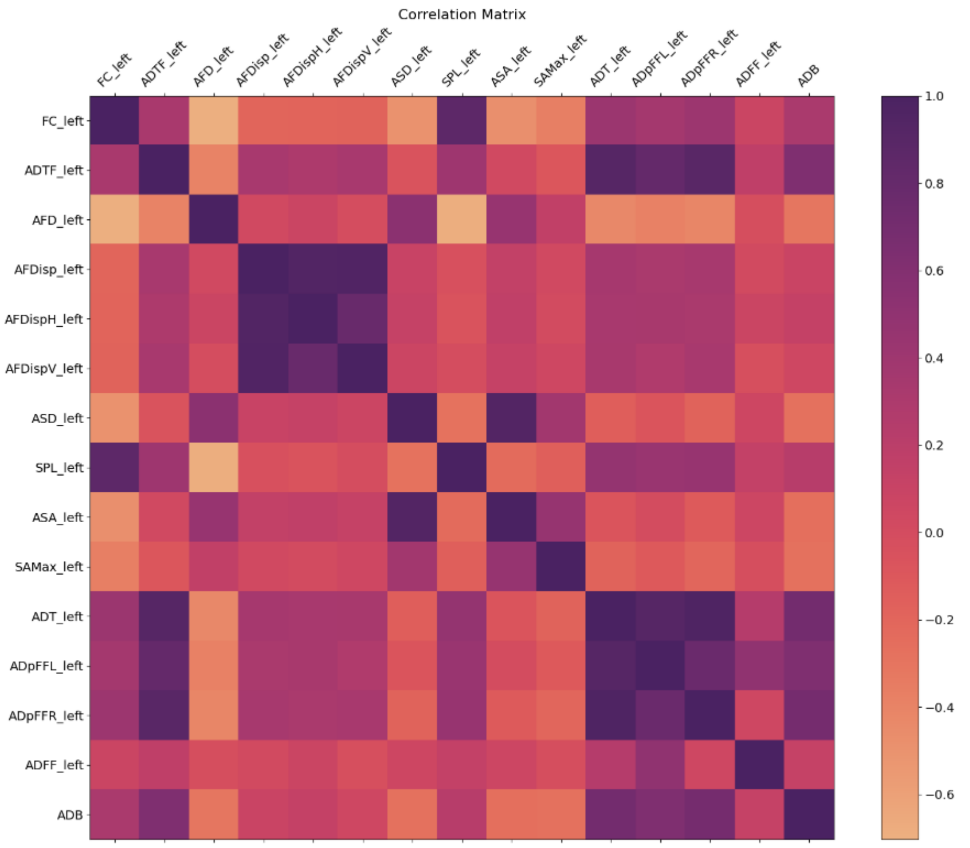

Figure 4.9 – Correlation Matrix between all the features of the left eye data (considering all 464 recordings).

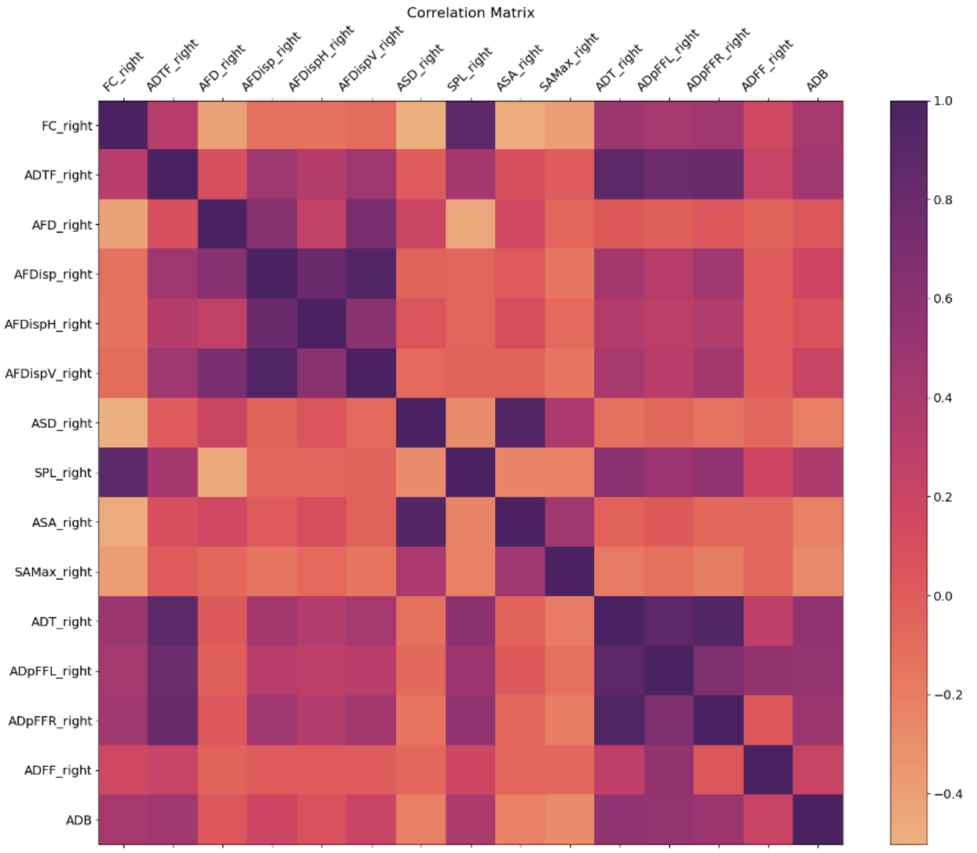

Figure 4.10 – Correlation Matrix between all the features of the right eye data (considering all 464 recordings).

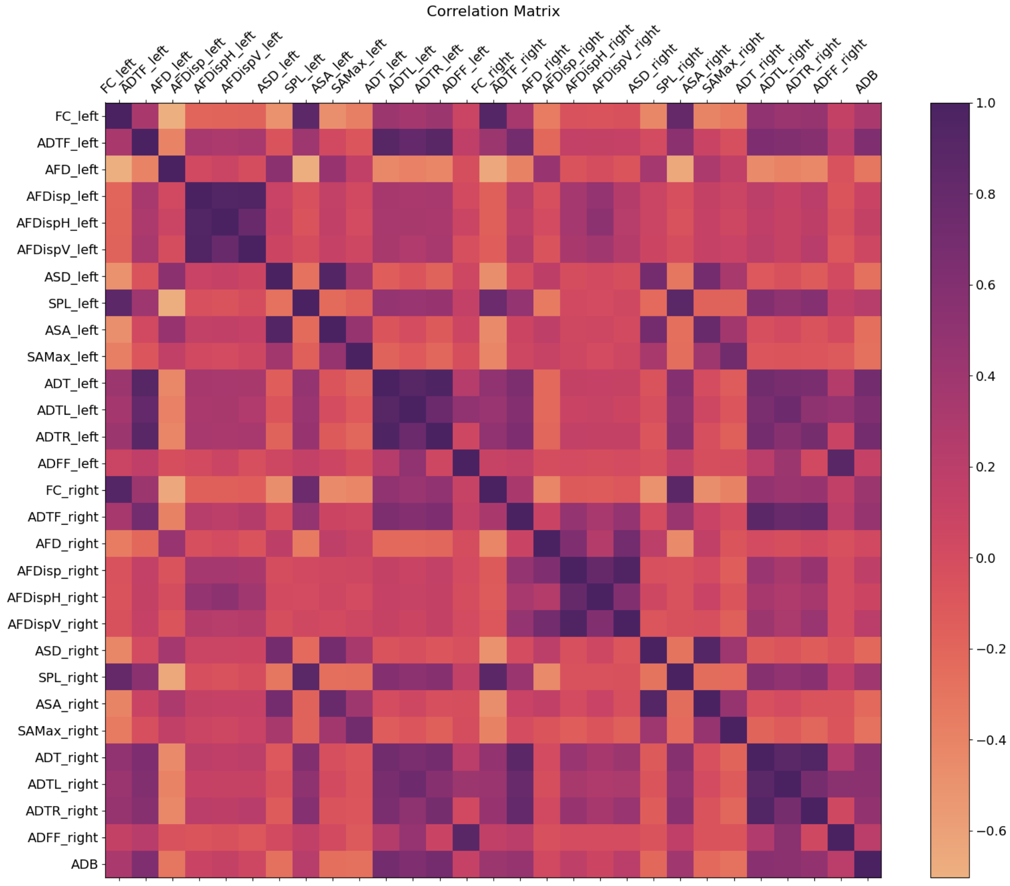

Figure 4.11 – Correlation Matrix between all the features of the both eyes data (considering all 464 recordings).

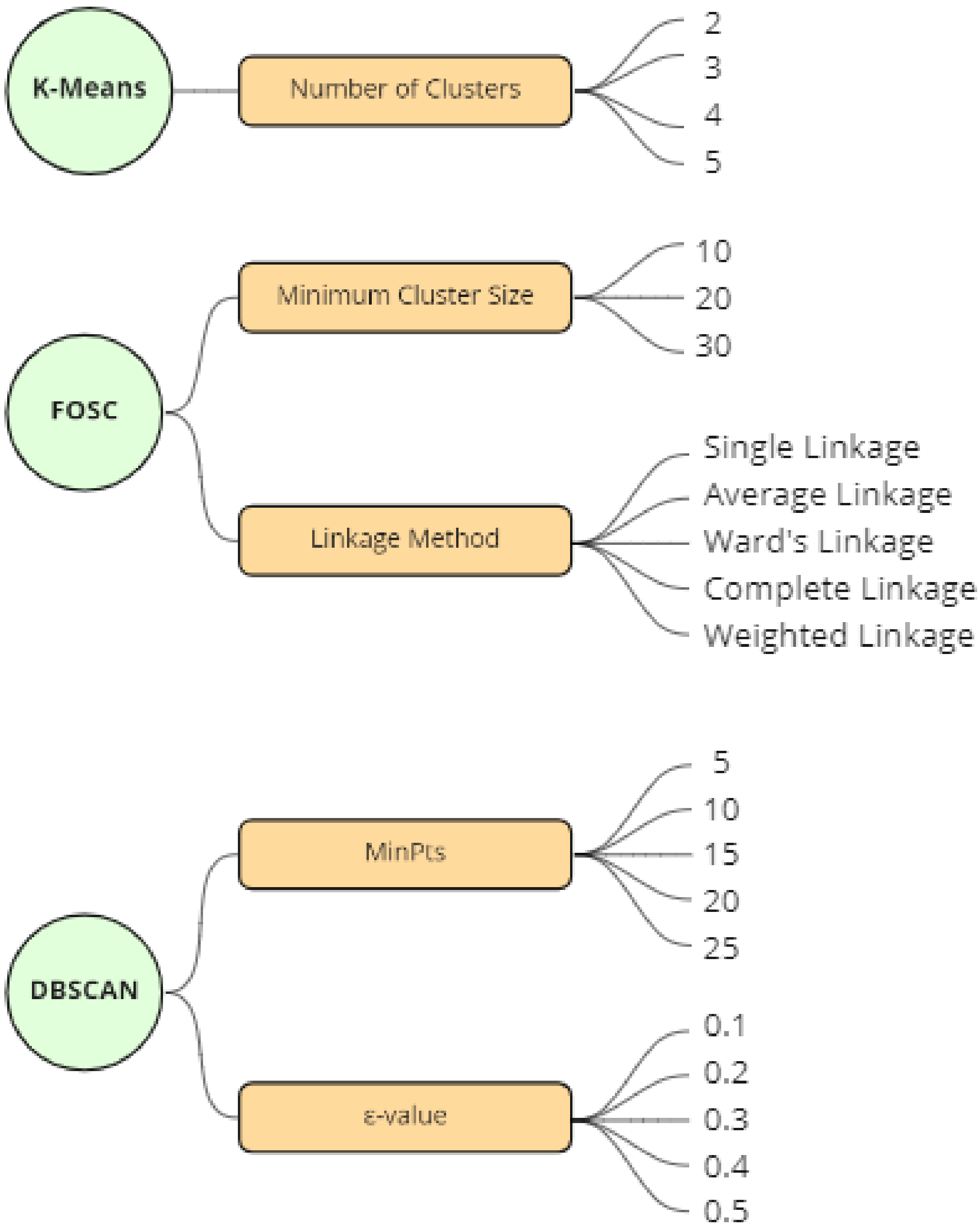

Figure 4.12 – Parameters for each method and the values used.

5 Results

The pipeline presented in Section 4.4 was the base for our experiments. After detecting the eye movements (see Section 4.2) for all the recordings and saving the featured objects in the database (see Section 4.3), we executed 3 rounds of the pipeline loop filtering outliers in order to refine the search for probably bad children's performances. For visualizing the clustering results, we used two visualization tools: PCA (HOTELLING, 1933) for scattering the n-dimensional data in two dimensions in order to see the distribution of the objects for the current subset of features (with size n), and Decision Trees (BELSON, 1959) to see which features were responsible for separating the clustering as presented by the current partition.

The database has 468 recordings from children for the chosen task, but 4 of them were promptly dropped as invalid after having not even a single fixation detected in the event detection step. The replays of these 4 recordings show some really messy time series, leading to a failed event detection, hindering the feature engineering for these recordings. Therefore, the first round included 464 recordings, and the results helped to find the recordings with very bad calibrations or poor tracking during the performance (see Section 4.1.1.3 and 4.1.1.4). In this round, the clusterings reached very good validations for Silhouette Coefficient Index (SCI). Figure 5.1 shows a clustering example with SCI of 0.7966. The objects labeled in orange are recordings with very bad calibrations or clearly poor tracking during the task. Figure 5.2 shows the horizontal and vertical gaze plots of 5 examples of objects belonging to the orange group in 5.1. By analyzing these plots and the replays, we can note that they are really inconsistent recordings regarding the eye-tracking itself.

Figure 5.3 shows another interesting example, with noisy elements. The recordings labeled as noise can be seen in Figure 5.4, and we can note that they all present problems with the eye-tracking, having too many samples with eye gazes not tracked. In the green cluster, we can find recordings that also had problems with the eye-tracker but generated similar time series that are different from the noisy ones, as shown in Figure 5.5. Finally, Figure 5.6 shows examples of the orange cluster aggregating recordings with relatively better performances. We can note that the groups have similar recordings, enabling us to see some of the power of the clustering process.

After analyzing different clustering results with good CV indexes, we filtered those recordings generally labeled as noise or belonging to a smaller and more dispersed cluster. That would indicate that they are different from most of the recordings, and we checked it using the replays on C&Look, confirming the hypothesis. Then, we ended up filtering 89 recordings in the "*Filter Outliers?*" step of the pipeline.

For the second round, we had 375 featured objects, which were mostly better recordings

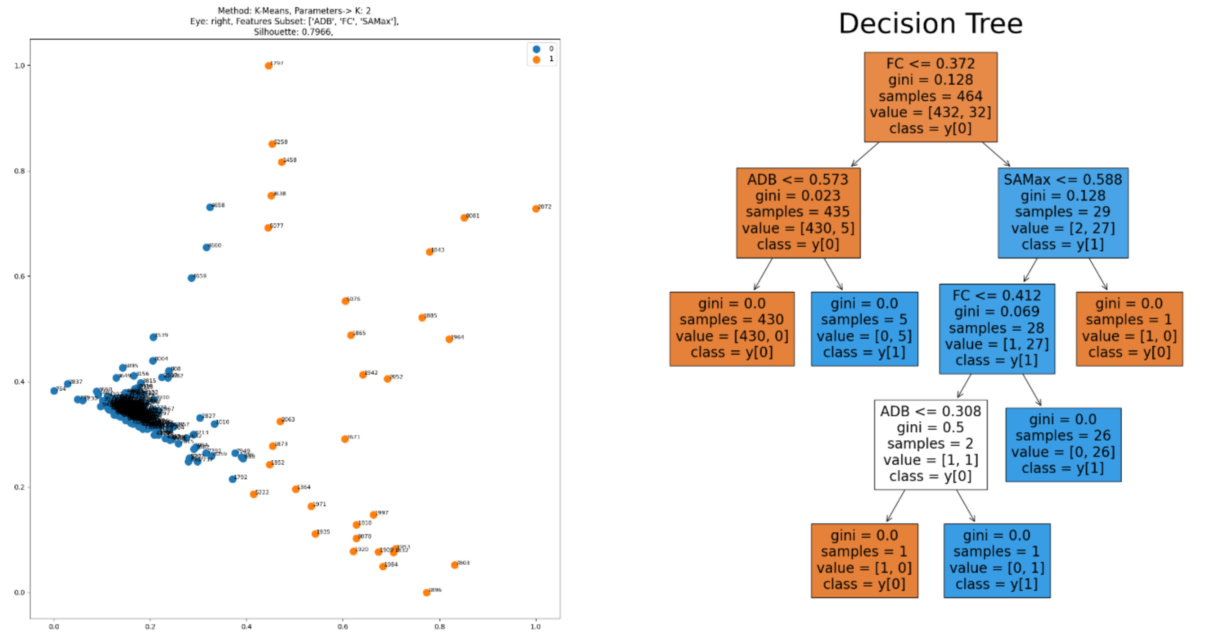

Figure 5.1 – Example of clustering for 464 recordings using data from the right eye. The CA method was K-Means with *k=2*, and the features used were: *ADB (Average Distance Between Eyes)*, *FC (Fixation Count)*, and *SAMax (Maximum Saccade Amplitude)*. The Silhouette Coefficient was 0.7966

in terms of eye-tracking than the recordings dropped in the first round. The feature subsets were generated again based on the correlation between the features of the 375 featured objects. The clustering results did not reach as good indexes as in the first round, but some good partitions also clustered the data in interesting ways. Figure 5.7 shows an example using two features: *ADB (Average Distance Between Eyes)* and *SAMax (Maximum Saccade Amplitude)*. The green cluster calls attention to some featured objects in terms of *ADB*, Figure 5.8 shows that these objects really have time series with higher distances between the eyes. The Decision Tree shows that ADB was the most relevant feature for this clustering process, separating most of the objects.

However, some of these recordings also had problems during the calibration step (see Figure 5.9), so the *ADB* could actually be a problem with the eye-tracker. Therefore, 52 objects were dropped out from the next round, leaving 323 valid recordings.

Now in the third round, there are no recordings with drastically bad eye-tracking, and the best resulting clusters start to show divergences in terms of performance instead of tracking quality. Figure 5.10 shows an example of one of the best clusterings (SCI=0.5082 for FOSC with *MinClSize=20 and Weighted Linkage Method*), that used the left eye data with the features *ADTF (Average Distance to Target During Fixations)*, *ADTR (Average Distance to Target on the Right Side of the Screen)*, and *AFD (Average Fixation Duration)*. In Figure 5.11 we can see the Horizontal and Vertical gaze graphs for recordings belonging to the blue cluster in Figure 5.10. We can note that the recordings have time series with similar segments that may indicate common characteristics in the children's oculomotor behavior. Once confirmed by a vision specialist, these

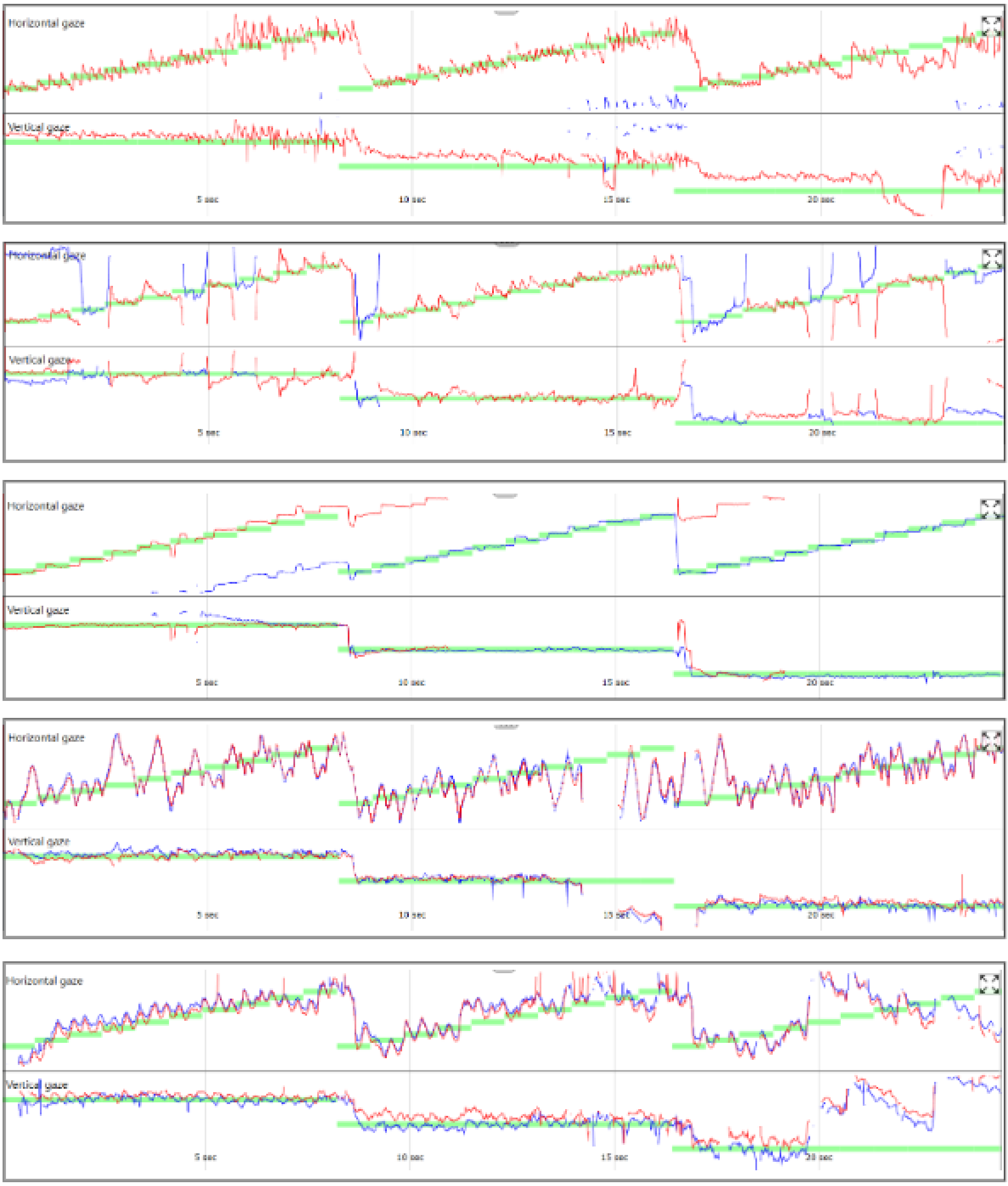

Figure 5.2 – Horizontal and Vertical gaze plots for 5 examples of objects labeled in orange in Figure 5.1.

characteristics could reflect the child's difficulty following the object, fatigue, or loss of focus and attention during the task.

To toughen the calibration quality of the recordings involved, the fourth and last round of our experiments have used only recordings with no apparent problem during calibration, validation, nor tracking during performance. This more rigid filtering left 265 valid recordings.

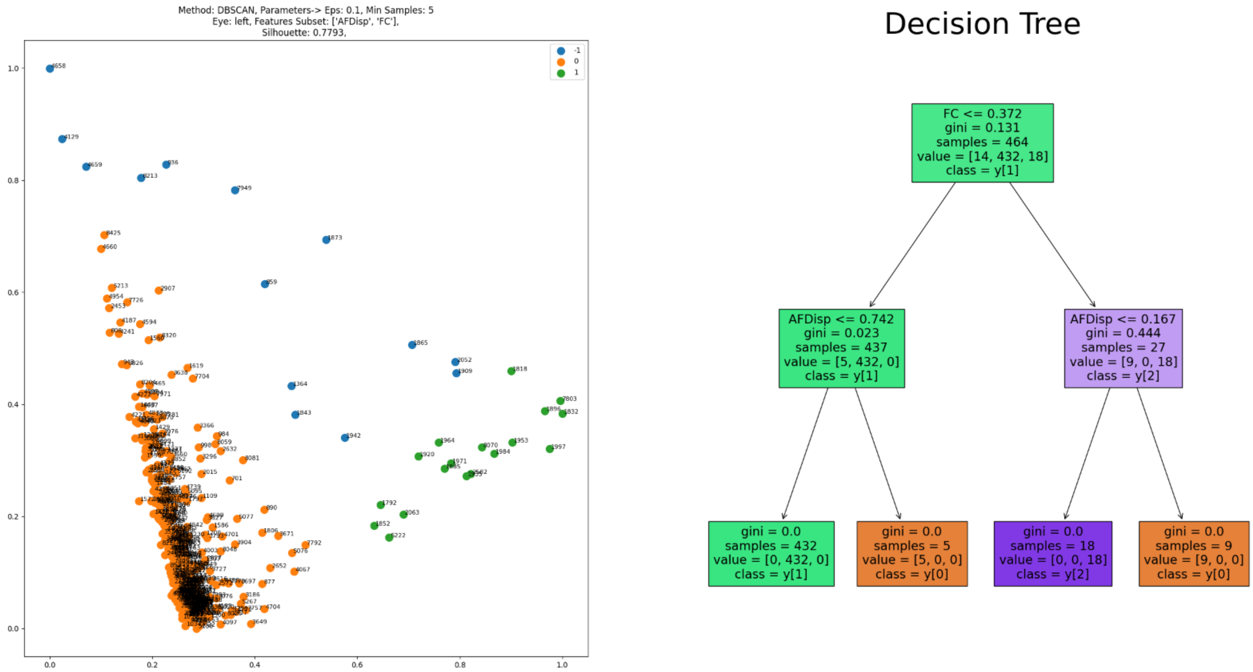

Figure 5.3 – Example of clustering for 464 recordings using data from the left eye. The CA method was DBSCAN with *Eps=0.1* and *MinClSize=5*. The features used were: *AFDisp (Average Fixation Dispersion)* and *FC (Fixation Count)*. The Silhouette Coefficient was 0.7793

Figure 5.12 shows an example of one of the best clustering results for 262 recordings. The features use were *ADB* and *ADT*, and we can note by Figure 5.13 that the recordings clustered together in this round also have similar time series.

Our experiments based on the pipeline illustrated the use of Event Detection, Feature Engineering and all the concepts described in Sections 2 and 4 within the data collected with C&Look. Despite the limitations originated mainly from the low sampling rate of the eye tracker, we could find interesting results that divided the data in seemingly meaningful clusters. The application and improvement of this tool must be supported by a vision specialist in order define the best features to collect, and describe inferences from the resulting clusterings using all the visualization tools: Scattered Data with PCA, Decision Trees, and the C&Look graphs and replays of the recordings.

The Python code and all the best clustering results are available in the GitHub repository <https://github.com/LuccaMartins/CnLook_Clustering>.

Figure 5.4 – Horizontal and Vertical gaze plots for 3 examples of noisy elements in Figure 5.3.

Figure 5.5 – Horizontal and Vertical gaze plots for 3 examples of elements belonging to the green cluster in Figure 5.3.

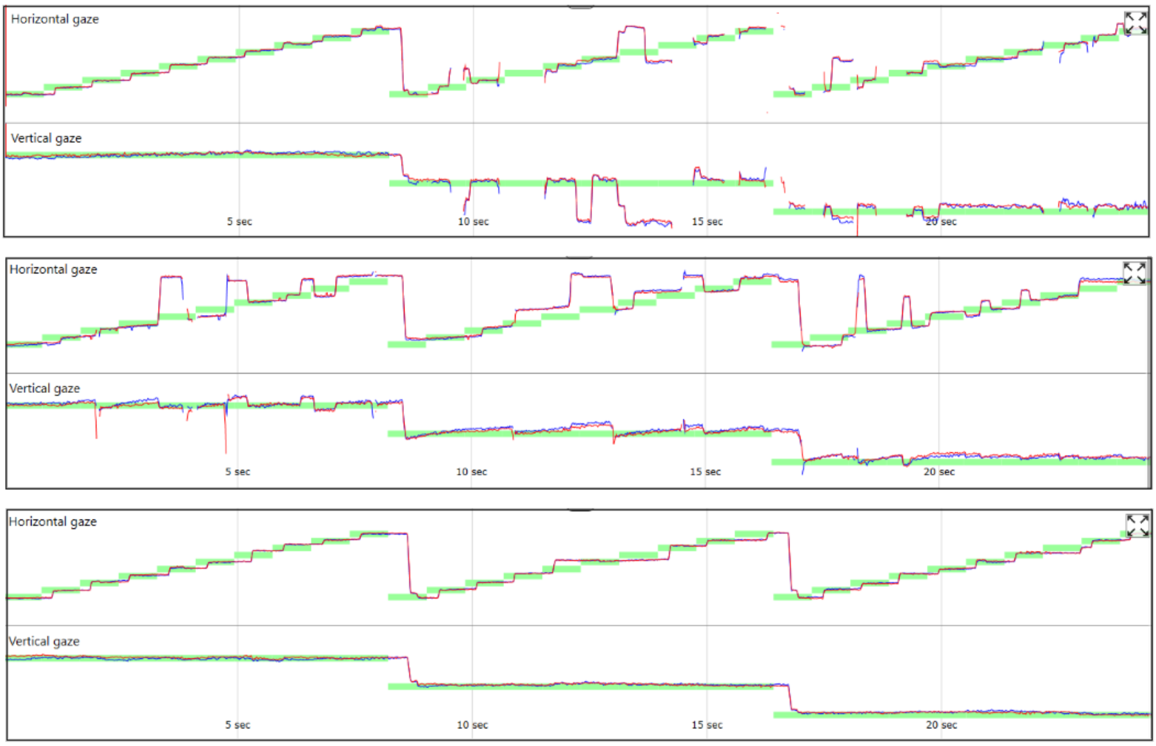

Figure 5.6 – Horizontal and Vertical gaze plots for 3 examples of elements belonging to the orange cluster in Figure 5.3.

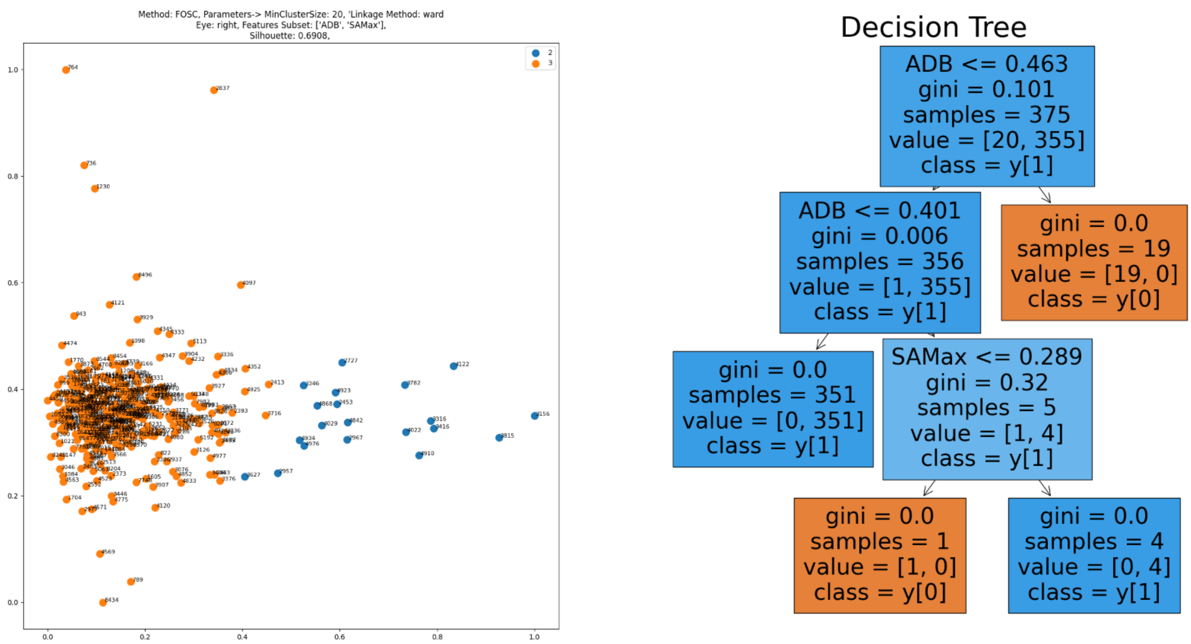

Figure 5.7 – Example of clustering for 375 recordings using data from the right eye. The CA method was FOSC with *MinClSize=20* using the Ward's linkage method. The features used were: *ADB (Average Distances Between Eyes)* and *SAMax (Maximum Saccade Amplitde)*. The Silhouette Coefficient was 0.6980.

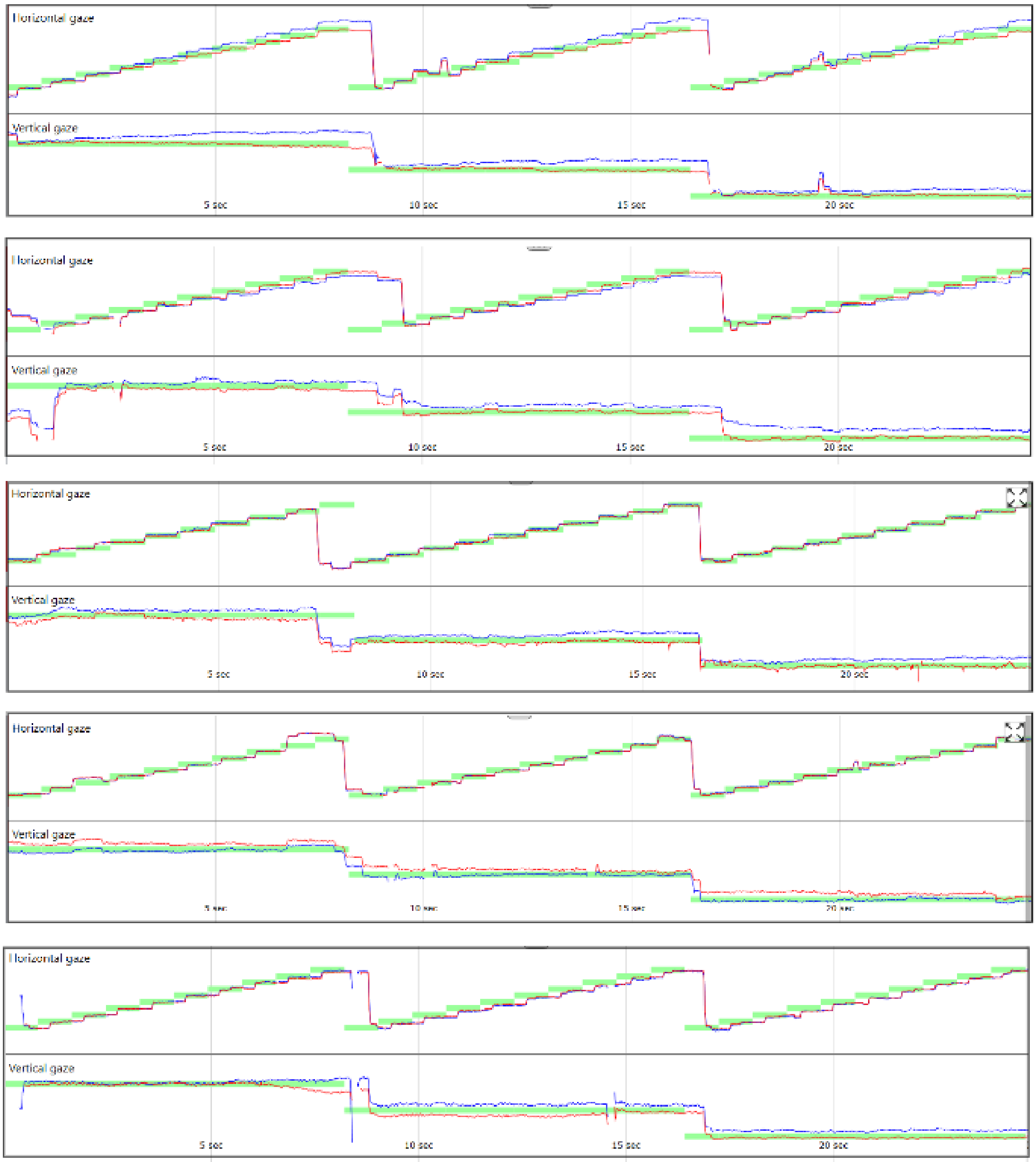

Figure 5.8 – Horizontal and Vertical gaze plots for 5 examples of elements belonging to the blue cluster in Figure 5.7

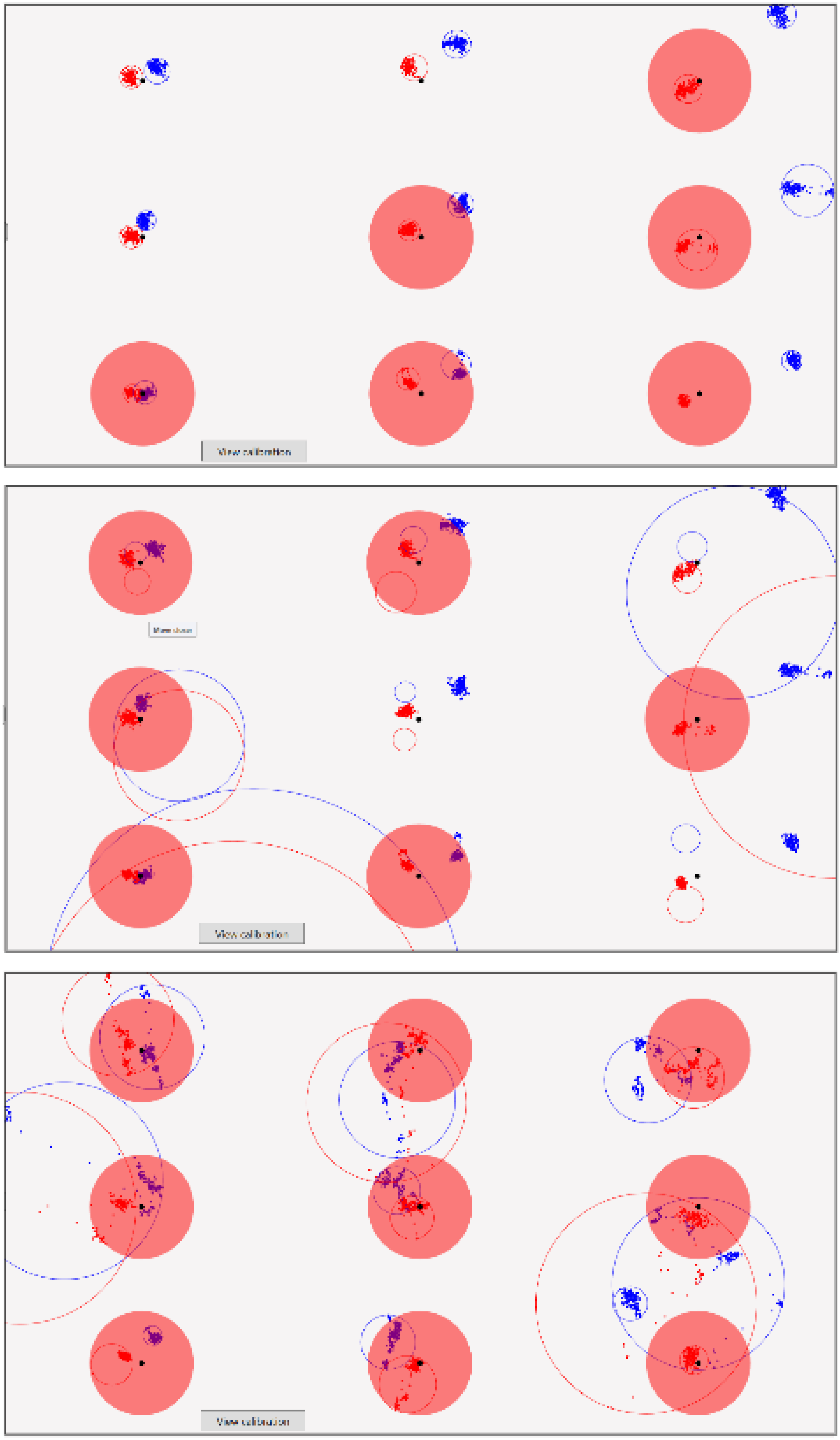

Figure 5.9 – 3 examples of C&Look validation for recordings belonging to the blue cluster in Figure 5.7

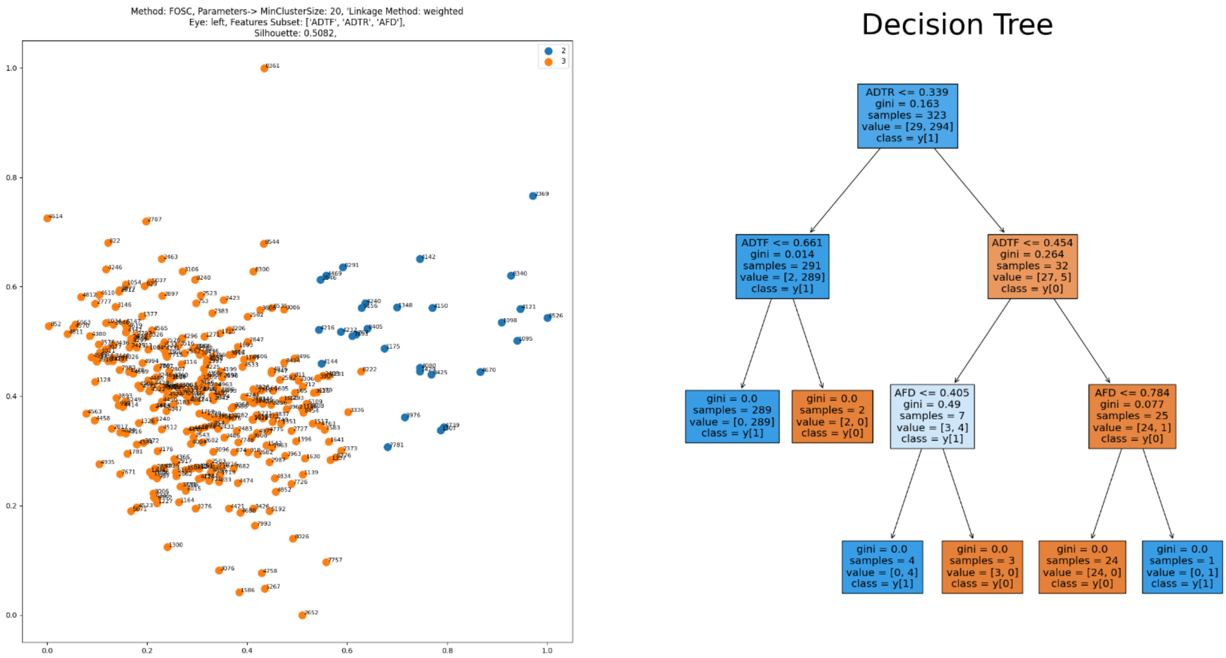

Figure 5.10 – Example of clustering for 323 recordings using data from the left eye. The CA method was FOSC with *MinClSize=20* using the Weighted linkage method. The features used were: *ADTF (Average Distance to Target During Fixations)* and *ADTR (Average Distance to Target on the Right Side of the Screen)*, and *AFD (Average Fixation Duration)*. The Silhouette Coefficient was 0.5082.

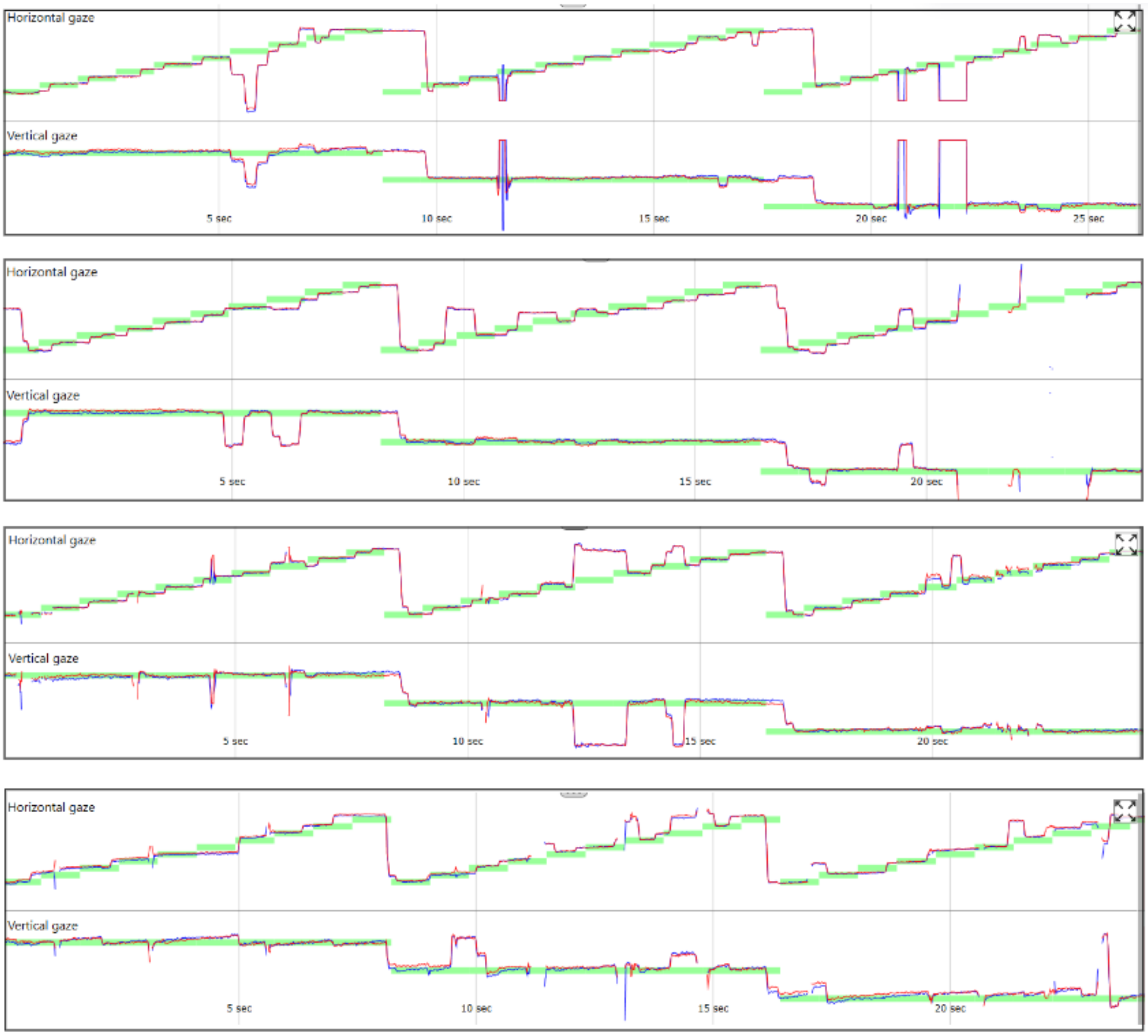

Figure 5.11 – Horizontal and Vertical gaze plots for 4 examples of elements belonging to the blue cluster in Figure 5.10

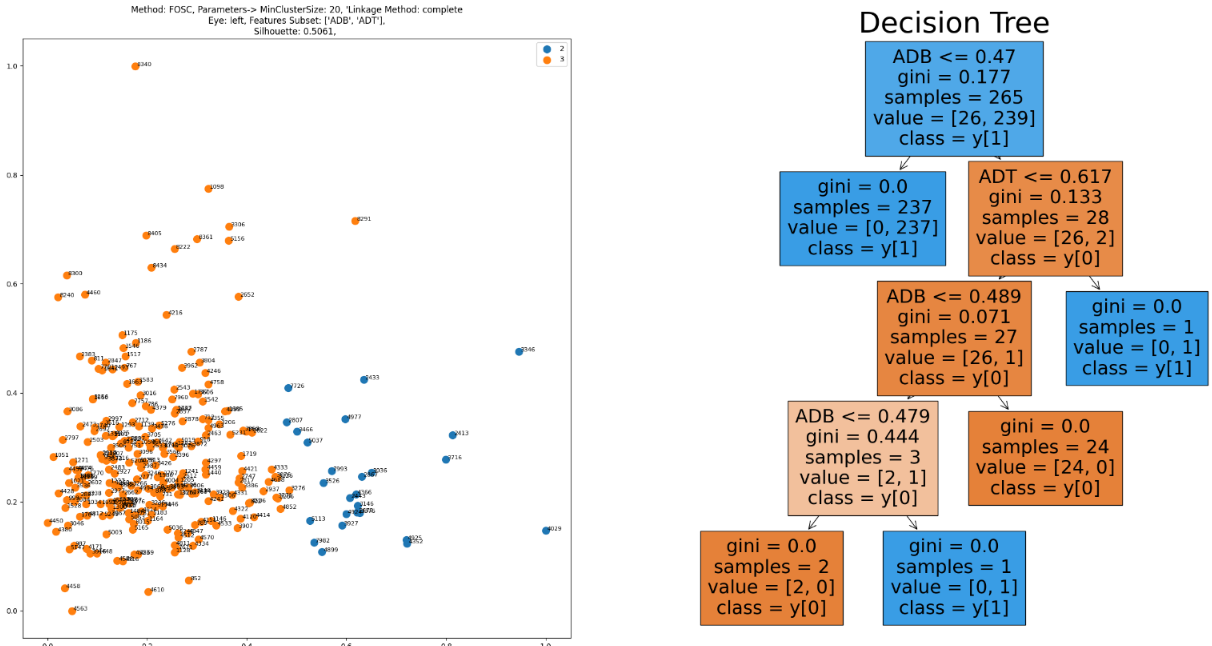

Figure 5.12 – Example of clustering for 265 recordings using data from left eye. The CA method was FOSC with *MinClSize=20* and *Complete Linkage Method*. The features used were: *ADB and ADT*. The Silhouette Coefficient was 0.5061.

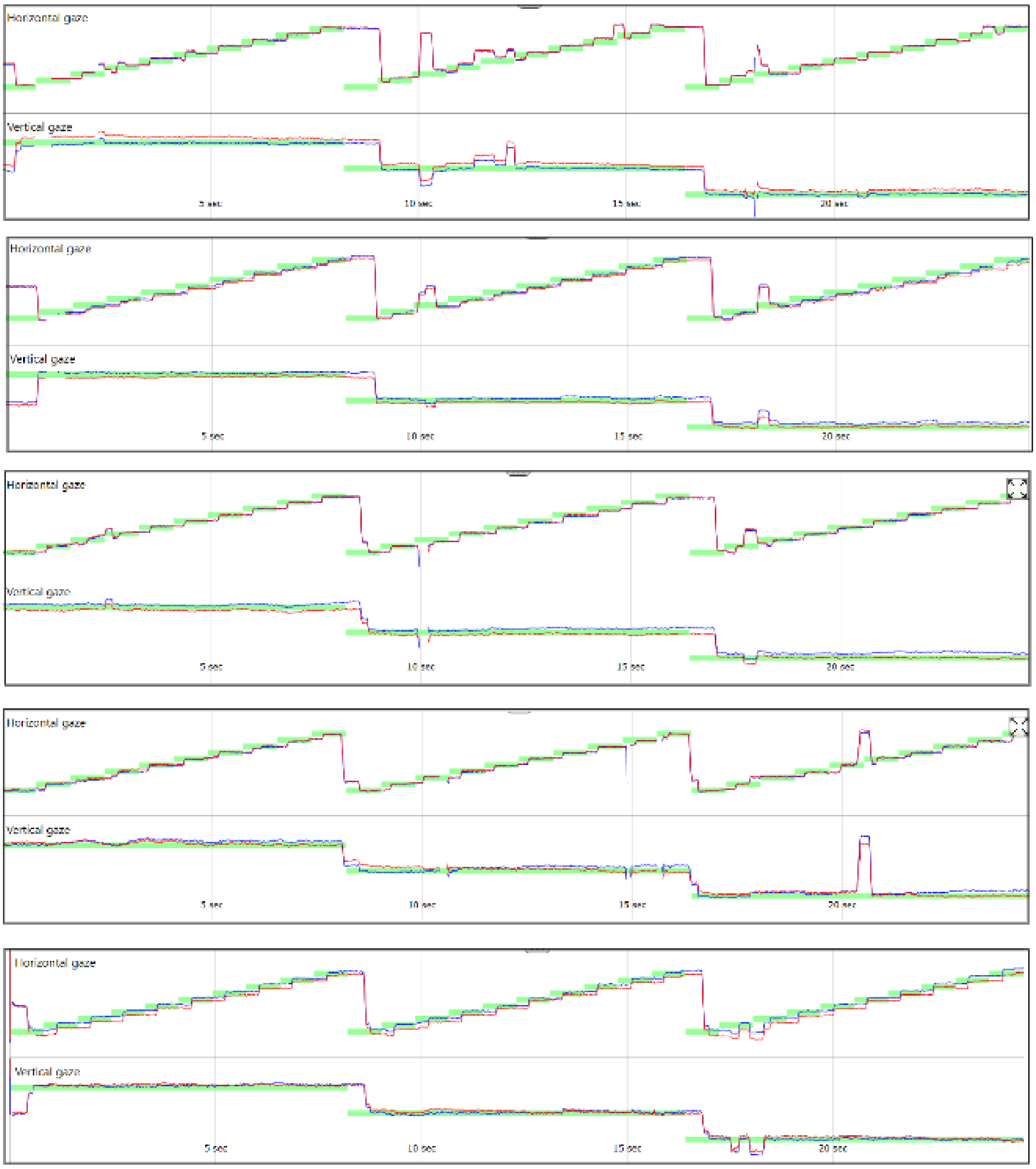

Figure 5.13 – Horizontal and Vertical gaze plots for 5 examples of elements belonging to the blue cluster in Figure 5.12

6 Conclusion

The eyes carry valuable information about a subject's health condition. Assessing the oculomotor behavior can open the way to analysis of psychological aspects, such as cognitive skills, learning process, and brain activity (SAMADANI, 2015; HESSELS; HOOGE, 2019). In children, impaired vision can limit scholar performance and become a hindrance to self-development (HOPKINS et al., 2019). Early diagnosis facilitates rehabilitation and improvement of oculomotor functions (GALLAWAY; BOAS, 2019; WILHELMSEN; FELDER, 2021), so children's vision care must be a primary concern (ZABA; M.A; O.D, 2011). Different stimuli can be used to test a subject's oculomotor system's response, depending on the research goals. When focusing on oculomotor skills in controlling eye movements (i.e. fixations, saccades, and regressions), structured tasks can be a good choice. In our study, the task is to follow a moving figure along with the screen. Therefore, the children's performance are measured with features that aim to reflect their ability to keep the gaze as close as possible to the target.

In order to find patterns in children's performances, we applied Clustering Methods within featured objects that represented the recordings. The resulting clusterings show that Event Detection, Feature Engineering, and Cluster Analysis can be a handy way to separate the data and highlight recordings of doubtful performances. The pipeline presented has enabled the exploratory analysis of the data, providing many different configurations to search for meaningful partitions. Even though we did not always have very good clustering validations, some promising results could be achieved even for the lowest amount of 265 recordings.

Once more, expert domain support is essential for reliable inferences regarding the results. Direct diagnosis of visual impairments is not the goal of this research. The most significant contribution of this work is to indicate eye-tracking recordings that could reflect an impaired oculomotor behavior. A vision specialist's opinion must always be expected to confirm any hypothesis inferred from the results. With the clustering framework presented and the insights that emerged during the study, we hope to have contributed to this branch of the whole eye-tracking research field.

6.1 Limitations and Future Works

Despite the usefulness of the clustering framework provided, it is important to elucidate the limitations of this work.

- **Eye-tracker**: the bottleneck for the quality of the clustering results starts with the device used. The Tobii Eye-Tracker 4C has a sampling rate of 90Hz, which is relatively low compared to the most recently released eye trackers. A low sampling rate limits the accuracy and the details collected by the event detection algorithms, thus undermining feature engineering. Of course, the educational bias of this study requires cheaper materials to enable a vaster use of the C&Look tool. Nonetheless, a better eye-tracker would be a relevant improvement for better detailing the recordings and reaching better clustering results;

- **Event Detection**: the event detection algorithm used (see Section 4.2) also has its limitations. The I-VT method identifies only fixations and saccades, losing much important information that could be gathered if other movements were included (e.g. PSO, Smooth Pursuits). In our case, blinks were also not detected, and they are considered to store valuable information about the subject's oculomotor behavior (COSTELA et al., 2014; YAMADA; KOBAYASHI, 2018). With a more detailed event detection, other features could be created, improving the characterization of the data;
- **Number of Recordings**: This work's limitations also rely on the number of available recordings. A more significant number of recordings for the chosen task would aggregate more information for the clustering methods, thus probably increasing its validation metrics.

Given the limitations of our work and the possible new objectives of this research field, some future works can be done to pursue new achievements. Some of them are cited below:

- Create a GUI for the clustering framework that could be used by non-Python specialists (i.e. vision experts, school teachers). That would be a primary step to popularizing this clustering tool, making it more accessible and easy to use;
- Apply different approaches for feature engineering and cluster analysis. An example would be the use of auto-encoders (BANK; KOENIGSTEIN; GIRYES, 2020) for characterizing the data with features that do not have a physical meaning and can show hidden patterns in the data.
- The event detection is another issue to improve. As stated, the detection quality depends on the quality of the eye-tracker, but different algorithms output different movement detection for the same eye-tracking data (ZEMBLYS, 2016). Therefore, a good improvement would be to apply feature engineering over different event detection approaches and analyze its relevance in the clustering results. Also, designing a specific event detection algorithm for the eye tracker used (considering its sampling rate and accuracy) could also be a worthy contribution;
- Define new features for characterizing the data and use more sophisticated methods for the feature selection, such as Genetic Algorithms (VAFAIE; JONG, 1997). A vision specialist and reliable literature must sustain the choice for new features;

- Improve the Cluster Validation for each algorithm used. As mentioned in Section 2.3.4, validation metrics must be applied over compatible clustering algorithms. Therefore, another good improvement would be to validate the clusterings taking into account the algorithm used (i.e. use DBCV (MOULAVI et al., 2014) for DBSCAN results);

Bibliography

ALI, Q.; HELDAL, I.; HELGESEN, C.; KRUMINA, G.; POP, C.; KÖVÁRI, A.; KATONA, J.; THILL, S. Current challenges supporting school-aged children with vision problems: A rapid review. *Applied Sciences*, v. 11, p. 9673, 10 2021.

ALI, Q.; HELDAL, I.; HELGESEN, C.; KRUMINA, G.; TVEDT, M. Technologies supporting vision screening: a protocol for a scoping review. *BMJ Open*, v. 11, p. 50819, 09 2021.

ALI, Q.; HELDAL, I.; HELGESEN, C. G.; COSTESCU, C.; KÖVÁRI, A.; KATONA, J.; THILL, S. Eye-tracking technologies supporting vision screening in children. In: *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. [S.l.: s.n.], 2020. p. 000471–000478.

ALIE, D.; MAHOOR, M. H.; MATTSON, W. I.; ANDERSON, D. R.; MESSINGER, D. S. Analysis of eye gaze pattern of infants at risk of autism spectrum disorder using markov models. p. 282–287, 2011.

ALVAREZ, C.; TENA, M. A. S.; ANDREU-VÁZQUEZ, C.; COLLAR, C. V. Visual health and academic performance in school-aged children. *International Journal of Environmental Research and Public Health*, v. 17, p. 2346, 03 2020.

ARAR, N. M. Robust eye tracking based on adaptive fusion of multiple cameras. 09 2017.

ARTHUR, D.; VASSILVITSKII, S. K-means++: The advantages of careful seeding. In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. USA: Society for Industrial and Applied Mathematics, 2007. (SODA ’07), p. 1027–1035. ISBN 9780898716245.

ASLIN, R. N. Infant eyes: A window on cognitive development. *Infancy*, v. 17, n. 1, p. 126–140, 2012. Available from Internet: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1532-7078.2011.00097.x>.

ATOWA, U.; HANSRAJ, R.; WAJUIHIAN, S. Visual problems: a review of prevalence studies on visual impairment in school-age children. *International Journal of Ophthalmology*, v. 12, p. 1037–1043, 06 2019.

AZEVEDO, A.; SANTOS, M. Kdd, semma and crisp-dm: A parallel overview. In: . [S.l.: s.n.], 2008. p. 182–185.

BANK, D.; KOENIGSTEIN, N.; GIRYES, R. *Autoencoders*. arXiv, 2020. Available from Internet: <https://arxiv.org/abs/2003.05991>.

BARGARY, G.; BOSTEN, J. M.; GOODBOURN, P. T.; LAWRANCE-OWEN, A. J.; HOGG, R. E.; MOLLON, J. Individual differences in human eye movements: An oculomotor signature? *Vision Research*, v. 141, p. 157–169, 2017. ISSN 0042-6989. Individual differences as a window into the structure and function of the visual system. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0042698917300391>.

BELSON, W. A. Matching and prediction on the principle of biological classification. *Journal of the Royal Statistical Society*, 1959.

BONILLA, N. W.; OD; ALLISON, C. L. A review of the efficacy of oculomotor vision therapy in improving reading skills. 2005.

BORTOLI, A. D.; GAGGI, O. Playwitheyes: A new way to test children eyes. p. 1–4, 2011.

CAMPELLO, R. J. G. B.; MOULAVI, D.; ZIMEK, A.; SANDER, J. A framework for semi-supervised and unsupervised optimal extraction of clusters from hierarchies. *Data Min. Knowl. Discov.*, v. 27, n. 3, p. 344–371, 2013. Available from Internet: <https://doi.org/10.1007/s10618-013-0311-4>.

CATRYSSE, L.; GIJBELS, D.; DONCHE, V.; MAEYER, S. D.; LESTERHUIS, M.; BOSSCHE, P. Van den. How are learning strategies reflected in the eyes? combining results from self-reports and eye-tracking. *British Journal of Educational Psychology*, v. 88, n. 1, p. 118–137, 2018. Available from Internet: <https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/bjep.12181>.

CHAPLIN, P. K. N.; BALDONADO, K.; HUTCHINSON, A.; MOORE, B. Vision and eye health: Moving into the digital age with instrument-based vision screening. *NASN School Nurse*, v. 30, n. 3, p. 154–160, 2015. PMID: 25870098.

CHEN, Z.; LO, W.-L.; CHI, Z.; CHANG, W. Strabismus recognition using eye-tracking data and convolutional neural networks. *Hidawu, Journal of Healthcare Engineering*, 2018.

CHEN, Z. H.; FU, H.; LO, W. L.; CHI, Z.; XU, B. Eye-tracking-aided digital system for strabismus diagnosis. *Healthcare Technology Letters*, v. 5, n. 1, p. 1–6, 2018. Available from Internet: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/htl.2016.0081>.

COSTELA, F.; OTERO-MILLAN, J.; MCCAMY, M.; MACKNIK, S.; TRONCOSO, X.; JAZI, A.; CROOK, S.; MARTINEZ-CONDE, S. Fixational eye movement correction of blink-induced gaze position errors. *PloS one*, v. 9, p. e110889, 10 2014.

DAVID, E.; BEITNER, J.; Võ, M. The importance of peripheral vision when searching 3d real-world scenes: A gaze-contingent study in virtual reality. *Journal of Vision*, v. 21, p. 3, 07 2021.

DING, X.; YUE, X.; ZHENG, R.; BI, C.; LI, D.; YAO, G. Classifying major depression patients and healthy controls using eeg, eye tracking and galvanic skin response data. *Journal of Affective Disorders*, v. 251, p. 156–161, 2019. ISSN 0165-0327. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0165032718330064>.

EIDE, M. G.; HELDAL, I.; HELGESEN, C. G.; WILHELMSEN, G. B.; WATANABE, R.; GEITUNG, A.; SOLEIM, H.; COSTESCU, C. Eye-tracking complementing manual vision screening for detecting oculomotor dysfunction. In: *2019 E-Health and Bioengineering Conference (EHB)*. [S.l.: s.n.], 2019. p. 1–5.

ESTER, M.; KRIEGEL, H.; SANDER, J.; XU, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: SIMOUDIS, E.; HAN, J.; FAYYAD, U. M. (Ed.). *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*. AAAI Press, 1996. p. 226–231. Available from Internet: <http://www.aaai.org/Library/KDD/1996/kdd96-037.php>.

EVERITT, B. S.; LANDAU, S.; LEESE, M.; STAHL, D. *Cluster Analysis: Everitt/Cluster Analysis*. [S.l.: s.n.], 2011.

FORTENBACHER, D. L.; BARTOLINI, A.; DORNBOS, B.; TRAN, T. Vision therapy and virtual reality applications. *Advances in Ophthalmology and Optometry*, v. 3, n. 1, p. 39–59, 2018. ISSN 2452-1760. Advances in Ophthalmology and Optometry. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S2452176018300027>.

FRANCESCHIELLO, B.; BOURGEOIS, A.; MINIER, A.; MURRAY, M. M.; POUGET, P.; BARTOLOMEO, P.; ANSELMI, F. Convolutional neural networks on eye tracking trajectories in patients with spatial neglect. *medRxiv*, Cold Spring Harbor Laboratory Press, 2020. Available from Internet: <https://www.medrxiv.org/content/early/2020/07/05/2020.07.02.20143941>.

GALLAWAY, M.; BOAS, M. Effect of vergence/accommodative therapy on reading in children with convergence insufficiency: A randomized clinical trial. *Optom Vis Sci*, 2019.

GERSTENBLITH, A. T.; RABINOWITZ, M. P. *The Wills eye manual: Office and emergency room diagnosis and treatment of eye disease 8th edition*. [S.l.: s.n.], 2021.

GERTRUDES, J. C. Semi-supervised learning approaches with applications in medicinal chemistry. *Instituto de Ciências Matemáticas e de Computação, USP - São Carlos*, 2019.

GÖBEL, F.; MARTIN, H. Unsupervised clustering of eye tracking data. In: . [S.l.: s.n.], 2018.

GOETTKER, A.; GEGENFURTNER, K. R. A change in perspective: The interaction of saccadic and pursuit eye movements in oculomotor control and perception. *Vision Research*, v. 188, p. 283–296, 2021. ISSN 0042-6989. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0042698921001851>.

GONI, S.; ECHETO, J.; VILLANUEVA, A.; CABEZA, R. Robust algorithm for pupil-glint vector detection in a video-oculography eyetracking system. In: . [S.l.: s.n.], 2004. v. 4, p. 941 – 944 Vol.4.

HASLWANTER, T.; CLARKE, A. H. Chapter 5 - eye movement measurement: electro-oculography and video-oculography. In: *Vertigo and Imbalance: Clinical Neurophysiologyof the Vestibular System*. Elsevier, 2010, (Handbook of Clinical Neurophysiology, v. 9). p. 61–79. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S1567423110090052>.

HELDAL, I.; HELGESEN, C.; ALI, Q.; PATEL, D.; GEITUNG, A. B.; PETTERSEN, H. Supporting school aged children to train their vision by using serious games. *Computers*, v. 10, n. 4, 2021. ISSN 2073-431X. Available from Internet: <https://www.mdpi.com/2073-431X/10/4/53>.

HESSELS, R. S.; HOOGE, I. T. Eye tracking in developmental cognitive neuroscience – the good, the bad and the ugly. *Developmental Cognitive Neuroscience*, v. 40, p. 100710, 2019. ISSN 1878-9293. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S187892931930297X>.

HOPKINS, S.; NARAYANASAMY, S.; VINCENT, S. J.; SAMPSON, G. P.; WOOD, J. M. Do reduced visual acuity and refractive error affect classroom performance? *Clinical and Experimental Optometry*, v. 103, n. 3, p. 278–289, 2019. Available from Internet: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cxo.12953>.

HOPPE, S.; BULLING, A. *End-to-End Eye Movement Detection Using Convolutional Neural Networks*. arXiv, 2016. Available from Internet: <https://arxiv.org/abs/1609.02452>.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 1933.

HYÖNÄ, J.; OLSON, R.; DEFRIES, J.; FULKER, D.; PENNINGTON, B.; SMITH, S. Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, v. 21, p. 1430–1440, 12 1995.

IRVING, E. L.; ZACHER, J. E.; ALLISON, R. S.; CALLENDER, M. G. Effects of Scleral Search Coil Wear on Visual Function. *Investigative Ophthalmology & Visual Science*, v. 44, n. 5, p. 1933–1938, 05 2003. ISSN 1552-5783. Available from Internet: <https://doi.org/10.1167/iovs.01-0926>.

JÄGER, L. A.; MAKOWSKI, S.; PRASSE, P.; LIEHR, S.; SEIDLER, M.; SCHEFFER, T. Deep eyedentification: Biometric identification using micro-movements of the eye. *CoRR*, abs/1906.11889, 2019. Available from Internet: <http://arxiv.org/abs/1906.11889>.

JAIN, A. K.; DUBES, R. C. Algorithms for clustering data. In: . USA: Prentice-Hall, Inc., 1988. ISBN 013022278X.

JASKOWIAK, P. A.; COSTA, I. G.; CAMPELLO, R. J. G. B. *The Area Under the ROC Curve as a Measure of Clustering Quality*. 2020.

JIANG, M.; ZHAO, Q. Learning visual attention to identify people with autism spectrum disorder. p. 3287–3296, 2017.

JOHNS, M.; TUCKER, A.; CHAPMAN, R.; CROWLEY, K.; MICHAEL, N. Monitoring eye and eyelid movements by infrared reflectant oculography to measure drowsiness in drivers. *Somnologie - Schlafforschung und Schlafmedizin*, v. 11, p. 234–242, 12 2007.

KAUFMAN, L.; ROUSSEEUW, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley, 1990. ISBN 978-0-47187876-6. Available from Internet: <https://doi.org/10.1002/9780470316801>.

KLAIB, A.; ALSREHIN, N.; MELHEM, W.; BASHTAWI, H. Iot smart home using eye tracking and voice interfaces for elderly and special needs people. *Journal of Communications*, v. 14, p. 614–621, 07 2019.

KLAIB, A.; ALSREHIN, N.; MELHEM, W.; BASHTAWI, H.; MAGABLEH, A. Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and internet of things technologies. *Expert Systems with Applications*, v. 166, p. 114037, 03 2021.

KONG, Y.; LEE, S.; LEE, J.; NAM, Y. A head-mounted goggle-type video-oculography system for vestibular function testing. *EURASIP Journal on Image and Video Processing*, v. 2018, 04 2018.

KRIEGEL, H.-P.; KRÖGER, P.; SANDER, J.; ZIMEK, A. Density-based clustering. *Wiley Interdisc. Rew.: Data Mining and Knowledge Discovery*, v. 1, p. 231–240, 05 2011.

LEONG, D.; MASTER, C.; MESSNER, L.; PANG, Y.; SMITH, C.; STARLING, A. The effect of saccadic training on early reading fluency. *Clinical Pediatrics*, v. 53, n. 9, p. 858–864, 2014. PMID: 24790022.

LIU, W.; YU, X.; RAJ, B.; YI, L.; ZOU, X.; LI, M. Efficient autism spectrum disorder prediction with eye movement: A machine learning framework. p. 649–655, 2015.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: *In 5-th Berkeley Symposium on Mathematical Statistics and Probability*. [S.l.: s.n.], 1967. p. 281–297.

MALMIVUO, J.; PLONSEY, R. Bioelectromagnetism. appendix b. application of maxwell's equations in bioelectromagnetism. In: ____. [S.l.: s.n.], 1995. p. 455–461. ISBN 978-0195058239.

MARDANBEGI, D.; KILLICK, R.; XIA, B.; WILCOCKSON, T.; GELLERSEN, H.; SAWYER, P.; CRAWFORD, T. J. Effect of aging on post-saccadic oscillations. *Vision Research*, v. 143, p. 1–8, 2018. ISSN 0042-6989. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0042698917302237>.

MATHERS, M.; KEYES, M.; WRIGHT, M. A review of the evidence on the effectiveness of children's vision screening. *Child: care, health and development*, v. 36, p. 756–80, 11 2010.

MEZIERE, D.; YU, L.; REICHLE, E.; MALSBURG, T.; MCARTHUR, G. Using eye-tracking measures to predict reading comprehension. 09 2021.

Front-matter. In: MILLS, T. C. (Ed.). *Applied Time Series Analysis*. Academic Press, 2019. p. i–iii. ISBN 978-0-12-813117-6. Available from Internet: <https://www.sciencedirect.com/science/article/pii/B9780128131176000193>.

MOULAVI, D.; JASKOWIAK, P. A.; CAMPELLO, R.; ZIMEK, A.; SANDER, J. Density-based clustering validation. In: . [S.l.: s.n.], 2014.

MURRAY, N.; KUBITZ, K.; HUNFALVAY, M.; ROBERTS, C.-M.; BOLTE, T.; TYAGI, A. An examination of the oculomotor metrics within a suite of digitized eye tracking tests. v. 5, p. 269–284, 12 2019.

NAQSHBANDI, K.; GEDEON, T.; ABDULLA, U. A. Automatic clustering of eye gaze data for machine learning. In: *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. [S.l.: s.n.], 2016. p. 001239–001244.

PEREIRA, C.; MELLO, R. Common dissimilarity measures are inappropriate for time series clustering. *Revista de Informática Teórica e Aplicada*, v. 20, p. 25, 01 2013.

PIROZZOLO, F. J.; RAYNER, K. The neural control of eye movements in acquired and developmental reading disorders. In: WHITAKER, H.; WHITAKER, H. A. (Ed.). *Studies in Neurolinguistics*. Academic Press, 1979, (Perspectives in Neurolinguistics and Psycholinguistics). p. 97–123. ISBN 978-0-12-746304-9. Available from Internet: <https://www.sciencedirect.com/science/article/pii/B9780127463049500094>.

PLUZYCZKA, M. The first hundred years: a history of eye tracking as a research method. *Applied Linguistics Papers*, v. 4/2018, p. 101–116, 12 2018.

PURVES, D.; AUGUSTINE, G. J.; FITZPATRICK, D.; HALL, W. C.; LAMANTIA, A.-S.; MOONEY, R. D.; PLATT, M. L.; WHITE, L. E. Chapter 11 - vision: The eye, and chapter 12 - central visual pathways. In: ASSOCIATES, S. (Ed.). *Neuroscience)*. Sixth edition. New York: Oxford University Press, 2018. p. 233–280.

RAYNER, K. The role of eye movements in learning to read and reading disability. *Remedial and Special Education*, v. 6, n. 6, p. 53–60, 1985. Available from Internet: <https://doi.org/10.1177/074193258500600609>.

RAYNER, K.; POLLATSEK, A. *The Psychology of Reading*. [S.l.: s.n.], 1989. ISBN 9780203357798.

RELLO, L.; BALLESTEROS, M. Detecting readers with dyslexia using machine learning with eye tracking measures. In: *Proceedings of the 12th International Web for All Conference*. New York, NY, USA: Association for Computing Machinery, 2015. (W4A '15). ISBN 9781450333429. Available from Internet: <https://doi.org/10.1145/2745555.2746644>.

REMINGTON, L. A. Chapter 13 - visual pathway. In: REMINGTON, L. A. (Ed.). *Clinical Anatomy and Physiology of the Visual System (Third Edition)*. Third edition. Saint Louis: Butterworth-Heinemann, 2012. p. 233–252. ISBN 978-1-4377-1926-0. Available from Internet: <https://www.sciencedirect.com/science/article/pii/B978143771926010013X>.

ROBINSON, D. A. A method of measuring eye movemnent using a scieral search coil in a magnetic field. *IEEE Transactions on Bio-medical Electronics*, v. 10, n. 4, p. 137–145, 1963.

ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, p. 53–65, 1987. ISSN 0377-0427. Available from Internet: <https://www.sciencedirect.com/science/article/pii/0377042787901257>.

SAIZ-MANZANARES, M. C.; PÉREZ, I. R.; RODRÍGUEZ, A. A.; ARRIBAS, S. R.; ALMEIDA, L.; MARTIN, C. F. Analysis of the learning process through eye tracking technology and feature selection techniques. *Applied Sciences*, v. 11, n. 13, 2021. ISSN 2076-3417. Available from Internet: <https://www.mdpi.com/2076-3417/11/13/6157>.

SALVUCCI, D. D.; GOLDBERG, J. H. Identifying fixations and saccades in eye-tracking protocols. In: . New York, NY, USA: Association for Computing Machinery, 2000. (ETRA '00), p. 71–78. ISBN 1581132808. Available from Internet: <https://doi.org/10.1145/355017.355028>.

SAMADANI, U. A new tool for monitoring brain function: Eye tracking goes beyond assessing attention to measuring central nervous system physiology. *Neural Regeneration Research*, v. 10, p. 1231, 08 2015.

STEIN, N.; NIEHORSTER, D. C.; WATSON, T.; STEINICKE, F.; RIFAI, K.; WAHL, S.; LAPPE, M. A comparison of eye tracking latencies among several commercial head-mounted displays. *i-Perception*, v. 12, n. 1, p. 2041669520983338, 2021.

STRASBURGER, H.; RENTSCHLER, I.; JÜTTNER, M. Peripheral vision and pattern recognition: A review. *Journal of Vision*, v. 11, n. 5, p. 13–13, 12 2011. Available from Internet: <https://doi.org/10.1167/11.5.13>.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. *Introduction to Data Mining*. Addison Wesley, 2005. ISBN 0321321367. Available from Internet: <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0321321367>.

THANKACHAN, B. Haptic feedback to gaze events. 12 2018.

THAPALIYA, S.; JAYARATHNA, S.; JAIME, M. Evaluating the eeg and eye movements for autism spectrum disorder. p. 2328–2336, 2018.

THORNDIKE, R. Who belong in the family? *Psychometrika*, v. 18, p. 267–276, 02 1953.

TOBII. What happens during the eye tracker calibration? In: . [S.l.: s.n.], 2022.

TOROK, N.; GUILLEMIN, V. J.; BARNOTHY, J. M. Lxxx photoelectric nystagmography. *Annals of Otology, Rhinology & Laryngology*, v. 60, n. 4, p. 917–926, 1951. PMID: 14903813. Available from Internet: <https://doi.org/10.1177/000348945106000402>.

VAFAIE, H.; JONG, K. D. Genetic algorithms as a tool for feature selection in machine learning. 11 1997.

WHITMIRE, E.; TRUTOIU, L.; CAVIN, R.; PEREK, D.; SCALLY, B.; PHILLIPS, J.; PATEL, S. Eyecontact: Scleral coil eye tracking for virtual reality. In: *Proceedings of the 2016 ACM International Symposium on Wearable Computers*. New York, NY, USA: Association for Computing Machinery, 2016. (ISWC '16), p. 184–191. ISBN 9781450344609. Available from Internet: <https://doi.org/10.1145/2971763.2971771>.

WHO. World report on vision. *Geneva: World Health Organization; 2019. Licence: CC BY-NC-SA 3.0 IGO. Available in <https://www.who.int/publications/i/item/9789241516570>*, 2019.

WIBIRAMA, S.; MURNANI, S.; SETIAWAN, N. A.; NURLATIFA, H. A survey of event detection methods for eye movements classification in smooth-pursuit-based interactive applications. In: *2020 International Symposium on Community-centric Systems (CcS)*. [S.l.: s.n.], 2020. p. 1–6.

WIKIMEDIA, C. *File:Neural pathway diagram.svg — Wikimedia Commons, the free media repository*. 2020. [Online; accessed 10-November-2021]. Available from Internet: <https://commons.wikimedia.org/w/index.php?title=File:Neural_pathway_diagram.svg&oldid=508435975>.

WILHELMSEN, G.; FELDER, M. Learning is visual: Why teachers need to know about vision. In: ____. [S.l.: s.n.], 2020. ISBN 978-1-83962-469-8.

WILHELMSEN, G. B.; FELDER, M. Structured visual learning and stimulation in school: An intervention study. *Scientific Research Publishing*, 2021.

WOLFE, P. Using eye movement to study adolescents' comprehension of visual texts. In: ____. [S.l.: s.n.], 2018. p. 235–248. ISBN 9781522528098.

YAMADA, Y.; KOBAYASHI, M. Detecting mental fatigue from eye-tracking data gathered while watching video: Evaluation in younger and older adults. *Artificial Intelligence in Medicine*, v. 91, p. 39–48, 2018. ISSN 0933-3657. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0933365717306140>.

YARBUS, A. L. Chapter 3 to 6. In: ASSOCIATES, S. (Ed.). *Eye Movements and Vision*. Academy of Science of the USSR, Moscow: Springer Science, 1967.

YOUNG, L.; SHEENA, D. Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, v. 7, p. 397–429, 09 1975.

YOUNIS, O.; AL-NUAIMY, W.; AL-TAEE, M.; AL-ATABY, A. Augmented and virtual reality approaches to help with peripheral vision loss. p. 303–307, 03 2017.

ZABA, J. N.; M.A; O.D. Children's vision care in the 21th century & its impact on education literacy, social issues, & the workplace: A call to action. *Journal of Behavorial Optometry*, 2011.

ZEIL, J.; BOEDDEKER, N.; HEMMI, J. M. Vision and the organization of behaviour. *Current Biology*, v. 18, n. 8, p. R320–R323, 2008. ISSN 0960-9822. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0960982208001620>.

ZEMBLYS, R. Eye-movement event detection meets machine learning. In: . [S.l.: s.n.], 2016.

ZEMBLYS, R.; NIEHORSTER, D.; HOLMQVIST, K. gazenet: End-to-end eye-movement event detection with deep neural networks. *Behavior Research Methods*, v. 51, 10 2018.

ZHENG, A.; CASARI, A. *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. 1st. ed. [S.l.]: O'Reilly Media, Inc., 2018. ISBN 1491953241.