Dyslexia deep clustering using webcam-based eye tracking

Mohamed Ikermane¹, Abdelkrim El Mouatasim²

¹Laboratory LABSI, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco ²Department of Mathematics and Management, Faculty of Polydisciplinary Ouarzazate (FPO), Ibn Zohr University, Ouarzazate, Morocco

Article Info

Article history:

Received Jan 12, 2023 Revised May 04, 2023 Accepted May 07, 2023

Keywords:

Arabic language Deep clustering Dyslexia K-means Webcam eye tracking

ABSTRACT

Dyslexia is a neurodevelopmental impairment that causes difficulties in reading and can have significant academic, social, and economic impacts. In Morocco, Dyslexia accounts for 37% of children's school failures. Early detection of dyslexia is crucial to help children reach their academic potential and prevent low self-esteem. To address this issue, a dyslexia screening tool using webcam-based eye tracking was developed for the Arabic language. The tool was tested on a dataset of 61 students from three primary schools in southern Morocco, and the results showed that using Arabic dyslexic-friendly typefaces improved reading performance, particularly for those with low reading performance. Deep clustering was used to reduce the dimensionality of the dataset, and the subjects were gathered using unsupervised k-means based on AutoEncoder output. The three clusters produced showed a significant difference in many dyslexia traits, such as the number and duration of fixations, as well as the saccade period. These findings suggest that webcam-based eve-tracking techniques have the potential to be used as an initial dyslexia diagnosis tool to assess if a child exhibits some of the typical symptoms of dyslexia and whether they should seek a professional full dyslexia diagnosis.

This is an open access article under the <u>CC BY-SA</u> license.



Corresponding Author:

Mohamed Ikermane Laboratory LABSI, Faculty of Sciences, Ibn Zohr University Agadir, Morocco Email: ikermane.mohamed@gmail.com

1. INTRODUCTION

Dyslexia is a neurodevelopmental issue that causes difficulty learning to read. It decreases reading performance and fluency and is connected with specific impairments in word immediate recognition [1]. Dyslexia influences a person's reading ability in that he may suffer to decode words, understanding vocabulary, and process proper spelling, which has a huge negative impact on his educational outcomes [2]; therefore, school failure is associated with dyslexia, despite the fact that it has no impact on a child's intellect [3]. In the Organisation for Economic Co-operation and Development OECD's recent International educational survey (PISA) [4], Morocco is ranked in the last fifth position. Students in Morocco scored lower than the OECD average in reading performance. School failure may have many reasons like economic, social problems, psychological and emotional problems and also physical illness and neurodevelopmental disorders reasons.

Dyslexia is regarded in Morocco as one of the most common and most damaging learning disorders in terms of academic success. Thus, instructors must be better trained to assist students who are not progressing; they must also be trained in diagnostic tools, remedial approaches, and accompaniment [5]. Dyslexia accounts for 37% of Moroccan children's school failure [6]. Early dyslexia detection is determining. When diagnosed, children with dyslexia can avoid a high rate of academic failure due to this learning disability, but identifying it is not an easy task. There are several common ways of diagnosing Dyslexia such as WISC-V [7] in which

we solely test individual cognitive processes on which natural reading is based. The result assesses the proportion of right replies and provides an evaluation of success on a specific reading assignment; However, this does not represent the natural reading process [8].

Analyzing approaches based on eye-tracking throughout the reading process provides an assessment of the reading process as it occurs naturally [9]. Many studies in this field have indicated that dyslexics' abnormal eye movements while reading are not only due to their reading difficulties. They are, in fact, rather independent of the reading material, and a comparison of dyslexics and normal readers reveals that eye movement patterns and features can distinguish dyslexics from these groups of readers [10]-[12]. Thus, screening for dyslexia using eye tracking measurements (Blink, Fixation, Saccade...) and machine learning techniques could be extremely beneficial in identifying individuals at risk of dyslexia during the early school years [13]-[15], which can have a significant impact on their academic achievements. This sort of screening employs a device known as an "Eye Tracker," which is an expensive piece of equipment that only works with specified software in some cases. As a result, utilizing eye-tracking technology to identify dyslexia may be costly and complicated. The importance of early dyslexia diagnosis cannot be overstated; a model of dyslexia detection via webcam-based eye tracking may be viewed as a low-cost option that tries to help overcome the expense and limitations of employing an "Eye Tracker" while maintaining high classification accuracy. Such a primary Dyslexia diagnosis technique can be used among students with poor reading performance to identify individuals at high risk of dyslexia. Dyslexic readers experience longer fixations, more fixations, and shorter saccades than non-dyslexic readers [10].

2. ARABIC LANGUAGE AND DYSLEXIA

Arabic text is written in a cursive style, with letters linked together with ligatures. Only 22 of the 28 letters are two-way connectors, whereas the remaining 6 cannot be joined since they are one-way connectors. Furthermore, Arabic is regarded as a bi-directional language since the script is written from left to right while numbers are written from right to left [16]. the likelihood of word migration may have a significant influence on a certain form of dyslexia: Letter Position Dyslexia (LPD), where the Arabic language is written cursively and most of its alphabetical letters - such as /*/./h/- vary their forms depending on where they are in the word: beginning as [17] / \Rightarrow , /middle as/ \Rightarrow , /and final as / \Rightarrow /. To use word lists (without contexts) to test reading accuracy, we must include short vowels known as (zz)" "Harakat" [18], [19] Otherwise, these lists would be nonsensical to Arabic readers, dyslexic or not, since many Arabic words are visually identical and may have distinct pronunciations depending on their meanings. This reveals the importance of short vowels in improving word recognition in the Arabic language.

Identifying eye movements while reading Arabic texts, on the other hand, indicates disparities in the intensity of eye movements and reading patterns, with dyslexics having "longer fixation periods, shorter saccades, and more regressions." Also demonstrates that eye movement patterns mirror the cognitive processes that occur when reading distinct orthographic forms of Arabic texts, vowelized and un-vowelized [19]. Many more contributions [20], [21] for dyslexia in Arabic related gaze metrics to phonological and reading processing issues. demonstrated significant variations in gaze intensity and visual scan patterns between normal and dyslexic subjects. Also confirmed accuracy in recording visual patterns as well as efficiency in assessing for anomalous visual reading patterns, and abnormal intensity on words for struggling readers. The Kuwait Dyslexia Association is still the only Arabic country to undertake a two-year survey of Kuwaiti primary, intermediate, and secondary pupils [22]. It was shown that 6.29% of them have developmental dyslexia. As a result, a rise in this proportion may be fairly predicted in nations with a bigger student population.

3. METHOD

3.1. Stimuli and design

This experiment was carried out through a Web-Based Application. Subjects would be given Arabic normal text materials with various text configurations such as font-sizes, character spacing, and font family to analyze their reading behavior. Participants sat 50 cm away from a 14-inch Lenovo ThinkPad T490 computer with a QHD screen display of 2560x1440px resolution and a refresh rate of 60hz. A Logitech C270 HD WEBCAM 720p/30 fps HD video with diagonal 55° field of view and auto light correction eye tracking their gaze points. The web-based app was written in HTML5/CSS3 and ran on an Apache Server 2.4.51. Owing to the clear documentation and simplicity of integration into a new application, the eye-tracking processes were modified from the eye-tracking WebGazer algorithm [23]. We primarily adjust it to our requirements by employing a clear webcam-based eye-tracking approach. The experimenter must fill out some personal information about the subject before beginning the experiment as illustrated in Figure 1, including the subject's name, age and sex, school grade, and reading level.





Figure 1. Home page of the web-application where we insert the participant's personal information

3.2. Calibration process

Prior to initiating an eye-tracking recording session, a subject is required to undergo a calibration procedure, which involves estimating the geometric characteristics of the subject's eyes in order to compute an accurate gaze point calculation that is fully customized to the subject. The calibration process is made simple by utilizing the WebGazer library. Specifically, as depicted in Figure 2, the subject is instructed to fixate their gaze on the cursor and click on each of the nine points displayed on the screen. The subject must click on each point five times until it turns yellow. Once the calibration process is completed, a window is displayed, which shows the percentage of accuracy measurement, as illustrated in Figure 3. During data collection, accuracy is used to assess the quality and precision of the eye tracker data. WebGazer, with its high accuracy, provides more dependable data as it can accurately describe the location of the subject's gaze on a screen with high precision under favorable experimental conditions. The Ridge Regression model for calibration in WebGazer [23] has demonstrated remarkable accuracy results in a room with good lighting. However, if the lighting conditions were suboptimal, the participant was redirected to the instructional page and requested to recalibrate. Following the successful calibration of the webcam, the actual reading tasks were presented to the participant, preceded by clear and concise instructions.



Figure 2. Calibration processes using 9 dots

Figure 3. WebGazer measuring accuracy

3.3. Reading text Features

The text (العطلة) [24] in the experiment met our requirements. The Arabic text was created based on the "L'Alouette" dyslexia test [25] and has been modified and adapted to reflect the characteristics of the Arabic

language. The content is made up of seven paragraphs and a total of 155 words, with no numerical expressions, acronyms, or foreign words Table 1.

Table 1. Experiment text statistics	
Total word count	155
Total number of characters	884
Number of characters without spaces	728
Unique Words	127
Paragraphs	7

To avoid sequence effects, the text presentation was also controlled. The content and fonts were counterbalanced. As a result, the text paragraphs were divided into two pages, the first of which contained four paragraphs (84 Words) with the most common fonts used on screen and in printed texts [26] Arial, Arial Italic, Times, and Times Italic for each paragraph. All of the texts were center-justified, with a font size of 18 points and with black text color on a white background.

The second page of the paper utilized Arabic dyslexic-friendly typefaces to improve readability for dyslexic readers. Specifically, three different Arabic typefaces were used: Arabic transparent, simplified Arabic, and simplified Arabic fixed fonts with a size of 20 points. These typefaces were formatted using a combination of bold and black color, which is recommended for Arabic web content. Overall, the use of dyslexic-friendly typefaces and formatting can have a significant impact on the readability and comprehension of text for individuals with dyslexia [27].

3.4. Dataset collects

Our dataset is based on eye-tracking data of 61 subjects, who participated in reading 2 texts each. Data were gathered from three primary schools in the south of Morocco: two in Guelmim and one in Sidi Ifni. As illustrated in Figure 4 the individuals ranged in age from 9 to 13 years old, with 28 from 4th grade and 33 from 5th grade. Subjects were divided into three sub-groups depending on their Arabic reading ability, as determined by their teachers (Low, Medium, excellent). Subjects were required to read two distinct texts during Experiments, and their reading eye-tracking data for both were collected independently.



Figure 4. Dataset description of fourth and fifth grade stdeunts

3.5. Deep clustering

Deep clustering [28] is the combination of deep learning with clustering. Learning good feature representations and then running any classical clustering algorithm on the learned representations is one way for deep learning-based clustering. Deep unsupervised learning algorithms can map data points to meaningful low-dimensional representation vectors (Dimensionality reduction) that contain all of the important information about the given data point, allowing clustering on it to provide better results. Deep auto-encoders are a popular method for discovering meaningful representations. As illustrated in Figure 5, the input is processed by a multilayer encoder that produces a low-dimensional code or representation of the input. This code is fed into a decoder, which produces output with the same dimensions as the input. The training goal of the model is to reconstruct the provided input by minimizing the difference between the input and the output. The encoder's learned representations, i.e., the code, must contain all useful information compressed into a low-dimensional vector.



Figure 5. Deep clustering auto-encoder structure

4. **RESULTS AND DISCUSSION**

The findings of eye tracking readings reveal that utilizing Arabic dyslexic-friendly fonts, significant font size, and bold font style may assist in better Arabic text reading, even though the second text is only 16 percent shorter than the first, the majority of participants showed an improvement. The use of Arabic dyslexia-friendly fonts appears to have the greatest influence on subjects with low reading performance, while Excellent readers appear to benefit less. The use of these typefaces and the police styles combination reduces the number of fixations as shown in Figure 6, shortens their duration as shown in Figure 7, and prolongs saccades as shown in Figure 8, which may aid poor reading students and improve their performance.



Figure 6. Number of fixations for each text



Figure 7. Mean duration of a fixation fo each text



Figure 8. Average time of a saccade for each text

4.1. The AutoEncoder

Autoencoders are neural network models that aim to learn a compressed representation of highdimensional data by capturing its essential features. The model consists of two main components, an encoder that transforms the input data into a lower-dimensional code, and a decoder that generates the output data from that code. The goal of the model is to minimize the difference between the input and output data, which ensures that the code contains all the essential information needed to reconstruct the input data. In the dyslexia eyetracking dataset, an artificial neural network (ANN) autoencoder model was used for dimensionality reduction with 3 hidden dense layers and a Code layer of 3 neurons, as shown in Figure 9. By using autoencoders for dimensionality reduction, it is possible to reduce the computational cost of working with high-dimensional data while retaining the essential features needed for accurate analysis.



Figure 9. The auto-encoder architecture

To evaluate the performance of our Autoencoder, we divided our dataset into a training set (with a ratio of 0.75) and a test set (with a ratio of 0.25). From Figure 10, which displays the loss of our Autoencoder over the training and test data, we can infer that the ANN performed well in coding and decoding our dataset inputs. Unsupervised learning involves the challenge of determining the optimal number of clusters, and to tackle this issue, we utilized the elbow method and silhouette analysis techniques in K-means clustering to identify the most suitable number of clusters that best fits our eye-tracking dataset.



Figure 10. The autoencoder model loss over training and test data

4.2. Elbow method and silhouette analysis

The elbow method is a popular technique used in unsupervised machine learning to determine the optimal number of clusters in a dataset. It works by plotting the sum of squared distance (SSE) values against the number of clusters, which measures how well each data point is assigned to its respective cluster centroid. As the number of clusters increases, the SSE generally decreases, but at a decreasing rate. The Elbow method looks for the point on the SSE plot where the rate of decrease starts to slow down and form an elbow, which suggests that additional clusters do not significantly improve the clustering performance. Figure 11 is a

common graphical representation of the elbow method, showing how the SSE values change with the number of clusters and highlighting the elbow point as a candidate for the optimal number of clusters.

The term "silhouette" refers to a way of interpreting and validating consistency inside data clusters, it can also assist in determining the appropriate k value. used to investigate the separation distance between the clusters that resulted. The silhouette plot shows how near each point in one cluster is to points in neighboring clusters, the range of this metric is [-1, 1]. Results showed that 3 is the best K value of our dataset with a silhouette score of 0,65. Analyzing the elbow curve and the silhouette scores we can determine that K=3 is the number of clusters that fit best our dataset. When evaluating our three clusters, the first one is represented by shorter saccades, longer duration of fixation, and more fixations compared to the other two clusters, also this group of participants benefited the most from using Arabic dyslexic-friendly typefaces when comparing the eye-tracking reading outcomes for the two texts. The mean duration of fixation has decreased for all members of this cluster, ranging between 24% and 60%, and saccades have also increased, with an average of 63% to 81%. This cluster was identified as "High Risk of Dyslexia". The second cluster was less affected by Dyslexia traits with fewer fixations, long saccades, and shorter fixations. we categorized these cluster members as "Typical Readers". We labeled the third cluster as "Risk of Dyslexia", participants also showed longer fixation durations but fewer fixations and saccades than the first group. Figure 12 illustrates our eye-tracking dataset clustering using K-means and Deep clustering technics.



Figure 11. Elbow method for evaluating the number of clusters



Figure 12. K-means eye-tracking clusters

5. CONCLUSION

Dyslexic people's eye motions differ from those of normal readers. People with dyslexia tend to have longer fixations, longer reading periods, and more fixations than typical readers. In this research, we demonstrated how these traits, in conjunction with machine learning and webcam-based eye-tracking may be used to identify dyslexia. We got promising results by demonstrating that the use of Arabic Dyslexia friendly typefaces such as Arabic transparent, simplified Arabic, and simplified Arabic fixed fonts with significant font size could improve reading performance, and that poor readers benefited the most from the use of these typefaces by reducing fixation number and duration and expanding saccades Furthermore, the results confirmed that webcam-based eye-tracking technics can be used as an assessing tool for investigating the student's reading performance and the potential risk of dyslexia. To summarize, we utilized a JavaScript-based eye-tracking framework and consumer-grade webcams to record participants' eye movements and gaze positions during reading Arabic text using a web Application and Deep Clustering technics to extract the best features and reduce the dimensionality of our eye-tracking data collection. then we applied the k-means clustering method to estimate Dyslexia prevalence among students.

ACKNOWLEDGEMENTS

The authors thank Primary school teachers and directors for their cooperation.

REFERENCES

- I. C. J. Adubasim and J. T. Nganji, "Dyslexia-A Learning Difference," *Autism Opne Access*, vol. 7, no. 1, pp. 1–4, 2017, [Online]. Available: https://www.omicsgroup.org/journals/the-rapidly-changing-landscape-of-multiple-sclerosisimmunomodulatorytherapy-a-retrospective-chart-review-in-the-unitedarab-emirat-2376-0389-1000203.php?aid=90534.
- M. J. Snowling, "1. Does dyslexia exist?," Dyslexia: A Very Short Introduction, pp. 1–11, 2019, doi: 10.1093/actrade/9780198818304.003.0001.
- [3] S. Gustafson and S. Samuelsson, "Intelligence and dyslexia: Implications for diagnosis and intervention," Scandinavian Journal of Psychology, vol. 40, no. 2, pp. 127–134, 1999, doi: 10.1111/1467-9450.00109.
- [4] OECD, 21st-Century Readers: Developing literacy skills in a digital world. OECD, 2021.
- [5] S. Ihbour, "Dyslexia in Morocco: An overview of recent research," IOSR Journal of Research & Method in Education (IOSR-JRME), vol. 9, no. 1, pp. 24–26, 2019, doi: 10.9790/7388-0901012426.
- [6] M. Abdellah, "Invisible disabilities between education and health."
- [7] D. Study Citation Wechsler, "Technical and interpretive manual supplement," Pearson, p. 84, 2014.
- [8] M. N. Benfatto, G. Ö. Seimyr, J. Ygge, T. Pansell, A. Rydberg, and C. Jacobson, "Screening for dyslexia using eye tracking during reading," *PLoS ONE*, vol. 11, no. 12, 2016, doi: 10.1371/journal.pone.0165508.
- [9] J. Everatt, M. F. Bradshaw, and P. B. Hibbard, "Individual differences in reading and eye movement control," *Eye Guidance in Reading and Scene Perception*, pp. 223–242, 1998, doi: 10.1016/b978-008043361-5/50011-0.
- [10] G. F. Eden, J. F. Stein, H. M. Wood, and F. B. Wood, "Differences in eye movements and reading problems in dyslexic and normal children," *Vision Research*, vol. 34, no. 10, pp. 1345–1358, 1994, doi: 10.1016/0042-6989(94)90209-7.
- [11] J. Hyönä and R. K. Olson, "Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 21, no. 6, pp. 1430–1440, 1995, doi: 10.1037/0278-7393.21.6.1430.
- [12] R. K. Olson, "Eye movements in dyslexic and normal readers," Journal of Experimental Psychology: Human Perception and Performance, vol. 9, no. 5, pp. 816–825, 1983.
- [13] L. Rello and M. Ballesteros, "Detecting readers with dyslexia using machine learning with eye tracking measures," W4A 2015 -12th Web for All Conference, 2015, doi: 10.1145/2745555.2746644.
- [14] A. Jothi Prabha and R. Bhargavi, "Prediction of dyslexia from eye movements using machine learning," *IETE Journal of Research*, vol. 68, no. 2, pp. 814–823, 2022, doi: 10.1080/03772063.2019.1622461.
- [15] T. Asvestopoulou *et al.*, "DysLexML: Screening tool for dyslexia using machine learning," 2019, [Online]. Available: http://arxiv.org/abs/1903.06274.
- [16] F. AlRowais, M. Wald, and G. B. Wills, "An Arabic framework for dyslexia training tools," *1st International Conference on Technology for Helping People with Special Needs (ICTHP-2013) (19/02/13 20/02/13)*, pp. 63–68, 2013, [Online]. Available: http://eprints.soton.ac.uk/id/eprint/353261.
- [17] N. Friedmann and M. Haddad-Hanna, "Letter position dyslexia in Arabic: From form to position," *Behavioural Neurology*, vol. 25, no. 3, pp. 193–203, 2012, doi: 10.1155/2012/296974.
- [18] S. A. Rabia and A. Salfeety, "Reading in Arabic orthography: The influence of short vowels on reading accuracy and comprehension of poor and normal Arabic readers," *Journal of Advances in Linguistics*, vol. 5, no. 2, pp. 723–735, 2015, doi: 10.24297/jal.v5i2.2848.
- [19] S. Abu-Rabia, "Reading Arabic texts: Effects of text type, reader type and vowelization," *Reading and Writing*, vol. 10, no. 2, pp. 105–119, 1998, doi: 10.1023/A:1007906222227.
- [20] A. Al-Edaily, A. Al-Wabil, and Y. Al-Ohali, "Dyslexia explorer: A screening system for learning difficulties in the arabic language using eye tracking," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), vol. 7946 LNCS, pp. 831–834, 2013, doi: 10.1007/978-3-642-39062-3_63.
- [21] A. Al-Edaily, A. Al-Wabil, and Y. Al-Ohali, "Interactive screening for learning difficulties: Analyzing visual patterns of reading arabic scripts with eye tracking," *Communications in Computer and Information Science*, vol. 374, no. PART II, pp. 3–7, 2013, doi: 10.1007/978-3-642-39476-8_1.
- [22] T. M. Al-ghizzi, "Dyslexia in the Arab world," International Journal of Language and Literature, vol. 3, no. 2, 2015, doi: 10.15640/ijll.v3n2a7.

- [23] A. Papoutsaki, "Scalable webcam eye tracking by learning from user interactions," Conference on Human Factors in Computing Systems - Proceedings, vol. 18, pp. 219–222, 2015, doi: 10.1145/2702613.2702627.
- [24] G. Q. Salahia, "Constructing a test of reading ability in Arabic," University of Algiers, 1998.
- [25] P. Lefavrais, "Description, définition et mesure de la dyslexie. Utilisation du test 'l'Alouette," Revue de Psychologie Appliquée, vol. 15, pp. 33–44, 1965.
- [26] L. Rello and R. Baeza-Yates, "Good fonts for dyslexia," Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2013, 2013, doi: 10.1145/2513383.2513447.
- [27] F. Benmarrakchi and J. El Kafi, "Investigating reading experience of dyslexic children through dyslexia-friendly online learning environment," *International Journal of Information and Communication Technology Education*, vol. 17, no. 1, pp. 105–119, 2021, doi: 10.4018/IJICTE.2021010107.
- [28] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep clustering for unsupervised learning of visual features," Jul. 2018, doi: 10.1007/978-3-030-01264-99.

BIOGRAPHIES OF AUTHORS



Mohamed Ikermane b K s is a PhD student member of Engineering Science Laboratory at Faculty of Sciences, Ibn Zohr University, Morocco. With a master in distributed systems computing at Faculty of Science, Ibn Zohr University, Agadir (2014). He obtained Bachelor Degree in computer science at Polydisciplinary Faculty of Errachidia. He works as IT teacher at regional center for education and training professions in Guelmim, Morocco. His research interests on autism detection using deep learning techniques. He can be contacted at email: ikermane.mohamed@gmail.com



Abdelkrim El Mouatasim **(i)** S **(S)** is a full professor (PES) Faculty Polydisciplinary Ouarzazate, Ibn Zohr University, Morocco coordinator of mathematical and computer sciences (SMI) President of the Moroccan Association of artificial intelligence representative of UNESCO's member states (Morocco) for recommendation on the ethics of AI. His research interest is numerical analyses and optimization, operation research, mathematical modeling and statistical modeling, stochastic processes, algorithms, programming and computer simulation, image processing, data mining, data science and big data, machine learning and deep learning. He can be contacted at email: a.elmouatasim@uiz.ac.ma.