Classification of Scanpath for Fixations Based on Pupil Measures in Response to Dynamic Visual Stimuli

ADAMU, Muhammad Awwal [1], Fatima Isiaka [2, *]
[1] Department of Computer Engineering, Federal University of Technology, Minna, Nigeria
[2] Computer Science Department, Nasarawa State University, Keffi, Nigeria
Email: [1] muhammad.pg2113215@st.futminna.edu.ng, [2] fatima.isiaka@outlook.com
*Corresponding Author

***Abstract*—As the field of eye-tracking research advances, understanding how visual attention is allocated in dynamic environments has become increasingly relevant. This article explores the classification of scanpaths for fixations based on pupil measures, particularly in response to dynamic visual content. By examining the relationship between pupil dynamics and eye movement patterns, the aim is to enhance our understanding of visual cognition and improve applications, the novel method here is based on attention and performance applying the Euclidean path for shortest distance (EPSD) and a Region Convolutional Network embedded Support vector machine (RCNN-RCNNSVM) for a notable correlation between pupil sizes (Constriction and Dilation) to task performance in relation with scanpaths. The result shows that people who displayed larger pupil diameters during fixations on complex areas of a webpage often also demonstrated slower task completion times and shorter scanpath, suggesting deep contemplation or struggle to process information effectively. These methods applied on scanpaths for fixations through pupil measures provides a deeper understanding of how users interact with dynamic webpages with explicit animated contents. These findings can significantly aid web developers and designers in creating more intuitive and engaging interfaces that accommodate the ways users naturally navigate digital content.**

Keywords—Scanpath, Fixations, Pupil measures, Eye-tracking, Dynamic content, Cognitive processes, Visual attention, Machine learning.

I. INTRODUCTION

The study of eye movements, particularly fixation and saccadic patterns, is crucial for comprehending how individuals interact with visual stimuli. Traditional scanpath analysis focuses on the sequence of fixations but often overlooks physiological measures, such as pupillary response. Recent research indicates that pupil diameter can reveal underlying cognitive processes and emotional states with rapid eye movement (REM) and dilation in sleep and extensive cognitive tasks, and non-rapid eye movement and constriction (NREM), with less cognitive workload (Figure 1). This article presents a method for classifying scanpaths based on both fixation data and pupil measures in the context of dynamic visual stimuli, bridging the gap between eye-tracking metrics and cognitive interpretations using a minimisation problem of the form in Equation 2.

$$\mathrm{w}(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i P_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

where p_i is the predicted probability(output from *cls*), and p_i^* is the ground truth similarly, t_i is predicted bounding box and t_i^* is ground truth bounding box. L_{cls} are classification loss (log loss) and L_{reg} is smooth L_i loss with the minimization problem:

$$\underset{\mathbf{w}}{Minimize} \frac{1}{2} ||\mathbf{w}||^2$$

$$\text{subject to } y_i(\langle \mathbf{w}, \mathbf{x_i} \rangle + b) \geq 1$$

$$\mathbf{y}(\mathbf{xw} + b) \geq 1$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} (\begin{bmatrix} X_{11} & X_{12} & \dots & X_{1d} \\ X_{21} & X_{22} & \dots & X_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nd} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix} + b) \geq 1' \quad (2)$$

b and *w* are the group entries of the dataset of eye movement attributes and the number of fixations or instances collected for each short-term interval.

A. Understanding Scanpaths and Fixation

A scanpath is the trajectory of visual fixation points over time as an observer views a scene. Fixations occur when the eyes are relatively stable, allowing for visual information processing. At the same time, saccades are rapid movements between fixations (Figure 2), and the predicted route for the shortest path can be demonstrated with methods applied in this paper. Researchers often analyse these patterns to discern attentional mechanisms and cognitive processes. However, integrating physiological measures such as pupil diameter can yield deeper insights into cognitive load, arousal, and interest from consecutive fixation points connected by scanpath based on Euclidean space, where:

Received: 22-12-2024
Revised: 28-06-2025
Published: 30-06-2025

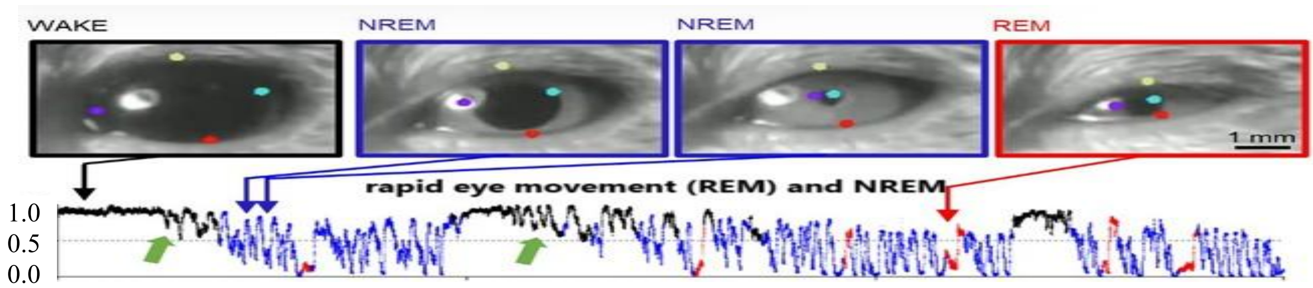
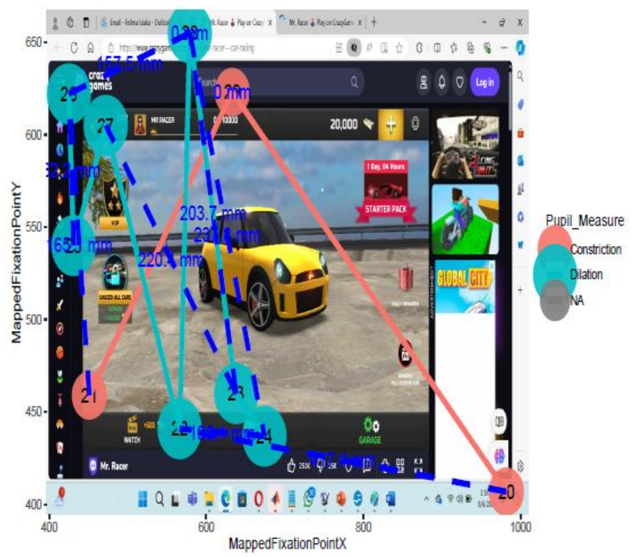
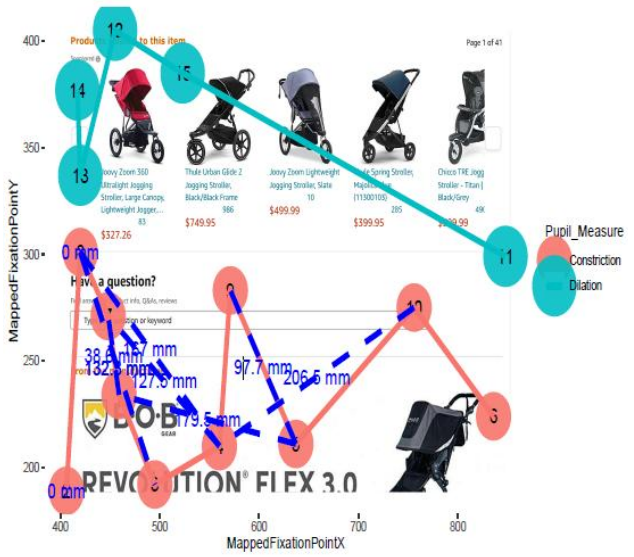


Fig. 1. The pupil diameter can reveal underlying cognitive processes and emotional states with rapid eye movement (REM) and dilation in sleep and extensive cognitive tasks, and non-rapid eye movement and constriction (NREM) with less cognitive workload.



(a) Scanpath and Fixations on First Dynamic webpage



(b) Scanpath and Fixations on Second Dynamic webpage

Fig. 2. Scanpath and fixations made on dynamic webpages

- **p**,**q** = two points in Euclidean n-space.
- $\mathbf{q}_i,\mathbf{p}_i$ = Euclidean vectors, starting from the origin of the space (initial point).
- **n** in Equation 3

$$d(\mathrm{p},\mathrm{q}) = \sqrt{\sum_{i=1}^{n}(\mathrm{q_i}-\mathrm{p_i})^2} \tag{3}$$

B. Pupil Dynamics: A Window into Cognition

Pupil diameter responds dynamically to various stimuli, reflecting cognitive and emotional states. Known as the pupillary light reflex, the pupil dilates in low light but also responds to engagement levels. When viewing dynamic content, changes in pupil size can indicate heightened attention, confusion, or cognitive overload. This physiological measure is critical for interpreting visual engagement during task performance, especially in dynamic environments where distractions abound. The method used here involves a study that employs a series of controlled experiments involving participants viewing dynamic visual content, such as video advertisements or motion graphics. Eye-tracking technology is used to capture both the participant's scanpaths and pupil measures simultaneously. Participants' gaze sequences are recorded, alongside real-time pupil diameter measures, allowing for a hybrid analysis that combines qualitative visual attention patterns with quantitative physiological data. The classification techniques here include the Euclidean path for shortest distance (EPSD) and Support vector machine (RCNNSVM) that classifies the different scanpath types based on distinct pupil size patterns on the dynamic stimuli (four dynamic webpages).

C. Objectives

The main objectives of this paper include:

- Apply Euclidean path for shortest distance (EPSD) and Support vector machine (RCNN-SVM) that classifies the different scanpath types based on distinct pupil size patterns on the dynamic stimuli.
- Present an explicit simulation pattern that describes these phenomena.
- Compare the performance of the classification methods

The preliminary findings suggest a significant correlation between pupil size fluctuations and specific scanpath classifications. For example, fixations associated with high pupil dilation often correspond with critical moments of information processing, suggesting intense cognitive engagement. In contrast, fixations with minimal pupil changes appear to indicate less engagement or familiarity with the content. This classification method implies that pupil measures provide a valuable dimension to standard scanpath analysis. By merging fixation patterns with pupil dynamics, one can gain deeper insights into viewer engagement and cognitive processes. For practical applications, such as in marketing, this information can enhance the targeting of dynamic content to elicit desired viewer reactions or improve educational materials to maintain student engagement. The integration of pupil measures with scanpath analysis offers a promising approach to understanding visual attention in dynamic environments. As technology continues to evolve, future research should focus on refining classification techniques and exploring implications across various domains. Ultimately, this research could lead to more effective designs in advertising, educational tools, and interfaces that cater to the cognitive behaviors of users interacting with dynamic content.

II. LITERATURE REVIEW

In recent years, the field of eye-tracking research has seen significant advances, particularly concerning how visual attention is quantified. One of the most intriguing aspects of this research is the analysis of fixations and the movement patterns known as scanpaths. Scanpaths, which are sequences of fixations and saccades (rapid eye movements between fixations), can yield insights into cognitive processing and attentional mechanisms. This literature review focuses on understanding scanpaths in the context of fixation measurements combined with pupil metrics, illuminating how pupil measures can enhance our comprehension of cognitive processes during visual tasks.

A. Understanding Fixations and Scanpaths

Fixations are intervals during which the eye remains relatively stationary, allowing for information processing from visual stimuli. The duration and location of fixations can indicate where attention is focused, suggesting which elements of a scene are being prioritized. Studies have shown that longer fixation durations correlate with increased cognitive load, indicating that viewers devote more mental resources to understanding complex visual information ([20, 22, 8, 27]).

Scanpaths consist of a sequence of fixations followed by saccades—the rapid, jerky movements of the eyes as they shift from one fixation point to another. Research indicates that scanpaths reveal not only information processing but also the viewer's strategies during visual exploration ([25, 4, 17, 16]).

Differences in scanpaths can reflect individual differences in cognitive style, visual expertise, and task engagement. Pupil dilation is often linked to cognitive and emotional states. McDougal's law, established in the early 20th century, posits that pupil size can be influenced by various factors, including mental effort, emotional responses, and arousal levels ([23, 10, 18, 30]).

As cognitive demands increase, the pupil tends to dilate, providing a secondary metric that enriches our understanding of visual attention and cognitive load.

Recent studies ([12, 24, 26, 6]) increasingly combine pupil diameter measurements with eye-tracking data to create a more comprehensive understanding of visual attention. Researchers have used this integrated approach to assess how cognitive load is distributed throughout scanpaths in various contexts, from reading to complex scene observation. For instance, research by ([13]), and ([21]) emphasized that pupil diameter is a valuable physiological indicator that correlates positively with fixation duration and correlates negatively with the frequency of saccades.

B. Enhancing Cognitive Load Assessment

Several studies ([28, 9, 29]) have explored how the combination of fixations and pupil measures can enhance our understanding of cognitive load. For example, ([19]) and ([7, 15]) showed that pupils dilate significantly during working memory tasks, particularly when the task demands exceeded participants' cognitive resources. Their findings suggest that tracking pupil responses during fixed gaze intervals can indicate underlying cognitive processes, providing additional layers of meaning to scanpath analysis.

Another critical area of research involves investigating how different scanning strategies manifest in fixations and pupil measures. Studies ([5, 11, 3]) have demonstrated that expert readers exhibit shorter but more efficient scanpaths with increased adaptability to varying cognitive loads compared to novices, who tend to exhibit longer, less efficient fixations ([14, 2, 1]). Increased pupil size in novices during reading tasks suggests a higher effort and cognitive load, indicating learning opportunities for educational strategies.

The integration of pupil measures with fixation data has profound implications for various fields, such as marketing, interface design, and education. For instance, in marketing, understanding how consumers' attention is allocated to different visual elements can inform more effective advertising strategies. In educational settings, insights gained from pupil and fixation data can lead to more informed pedagogical approaches that cater to diverse learning styles.

III. METHOD

This section explores novel methodology, significance, and potential applications in the digital landscape. The Scanpaths are a specific sequence of eye movements that the participants make as they explore the visually displayed webpages (four webpages ((Figure 3). These paths consist of the fixations (where the gaze remains relatively stable) and saccades (rapid movements between fixations). Analysing scanpaths allows us to put an authentic definition and understanding of which elements attract attention and how efficiently information is processed in the shortest means of time interval using EPSD and RCNN-SVM.

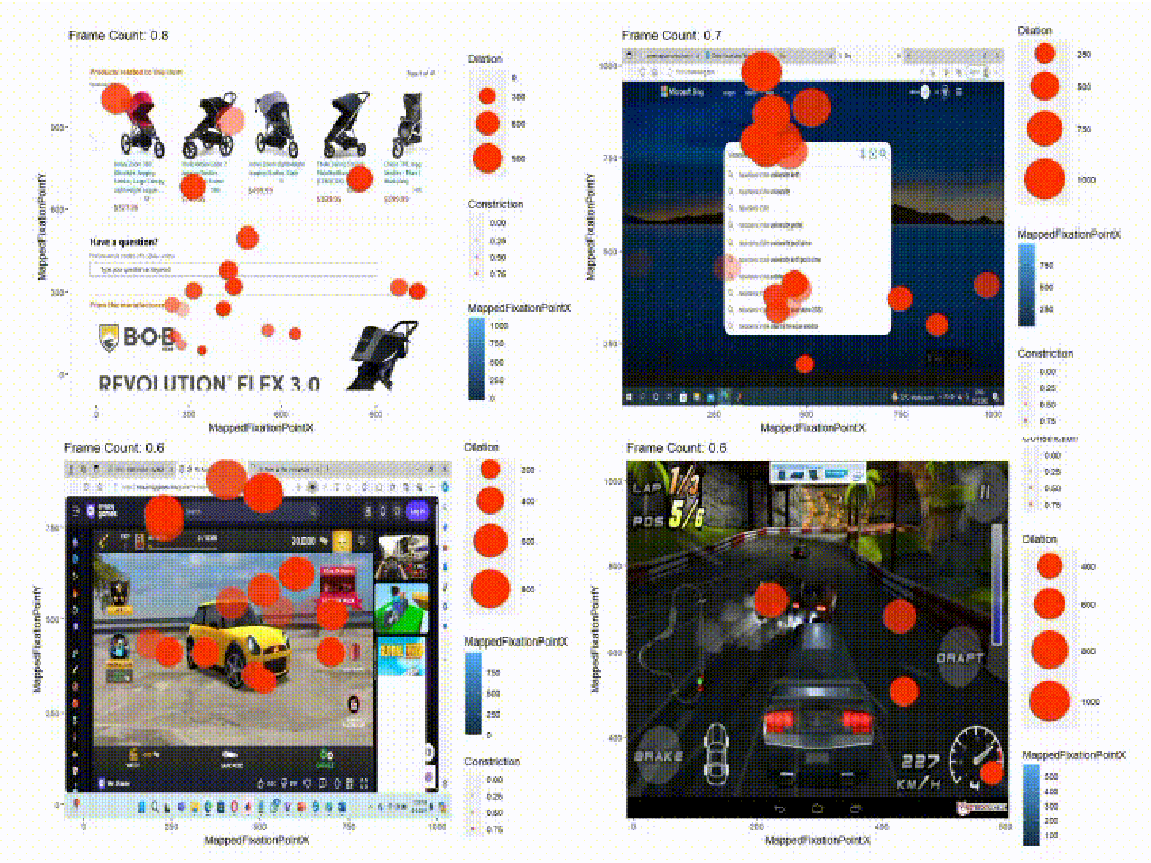


Fig. 3: The different eye movement patterns on four dynamic web pages.

In this new methodology, data is collected simultaneously from eye-tracking devices that monitor pupil size and record scanpaths from pre-determined experiment with four webpages. The main processes includes:

1) Dynamic Webpage Interaction: The participants interact with dynamic webpages while their eye movements and pupil size are monitored. This include scrolling through social media feeds, browsing e-commerce sites, or navigating through multimedia content.
2) Data Aggregation: The collected data is aggregated, noting the timeline of fixations, the duration at each fixation point, pupil Measure fluctuations, and the sequences of saccades (Figure 4).
3) Analysis: Utilizing statistical tools like R and MATLAB to analyze the combined data, looking for correlations between pupil dilation during specific fixations and how this relates to user engagement or cognitive load.
4) Insight Generation: The final stage involved interpreting the data to generate insights about user behaviour, preferences, and potential areas for design optimization useing EPSD and RCNN-SVM by comparing each method of classificaition.

The comparative analysis shows that RCNN-SVM is more reliable than EPSD in interpreting the scanpaths for the first dynamic webpage used in the study and also shows that in overall RCNN-SVM is well suited in interpreting the cognitive response of users to dynamic contents during interaction (Figure 5a and 5b). Also, by incorporating pupil measures into traditional eyetracking processes, these methods applied here provides a richer understanding of how users interact with dynamic content. For instance, a surge in pupil dilation during a particular section of a webpage can indicate high levels of cognitive engagement or emotional activation (Figure 3), signaling that the content was particularly compelling or challenging.

Dynamic webpages often present unique challenges for users, including rapid changes and information overload. This method allows for a nuanced examination of how users process these changes, facilitating the design of more intuitive interfaces that align with cognitive processing capabilities. The proceding section discusses the result obtained from the methods applied.

IV. RESULT

The analysis applied here by employing eye-tracking technology and pupilometry have yielded intriguing insights likee:

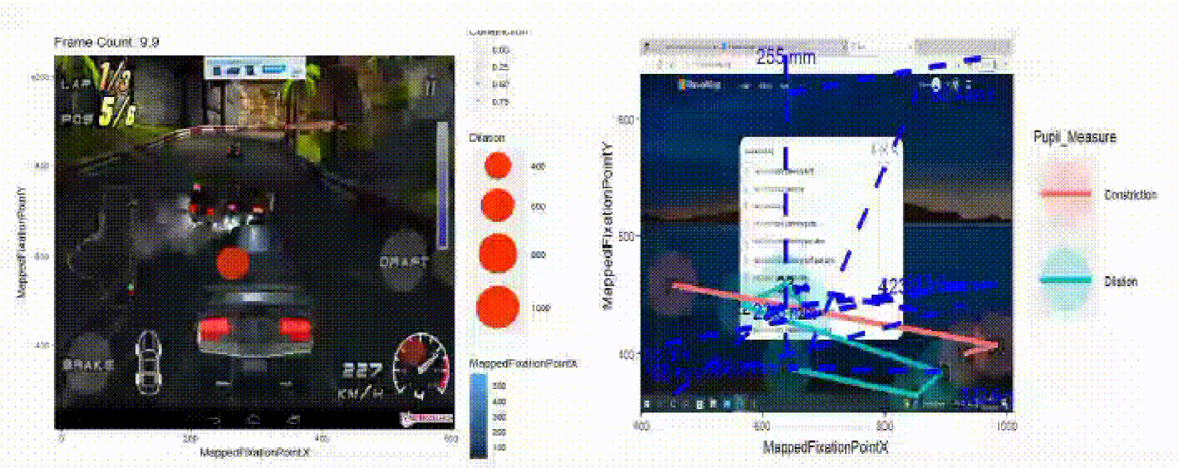


Fig. 4: The varied eye movements of users on content types like users displayed longer fixations on multimedia elements.

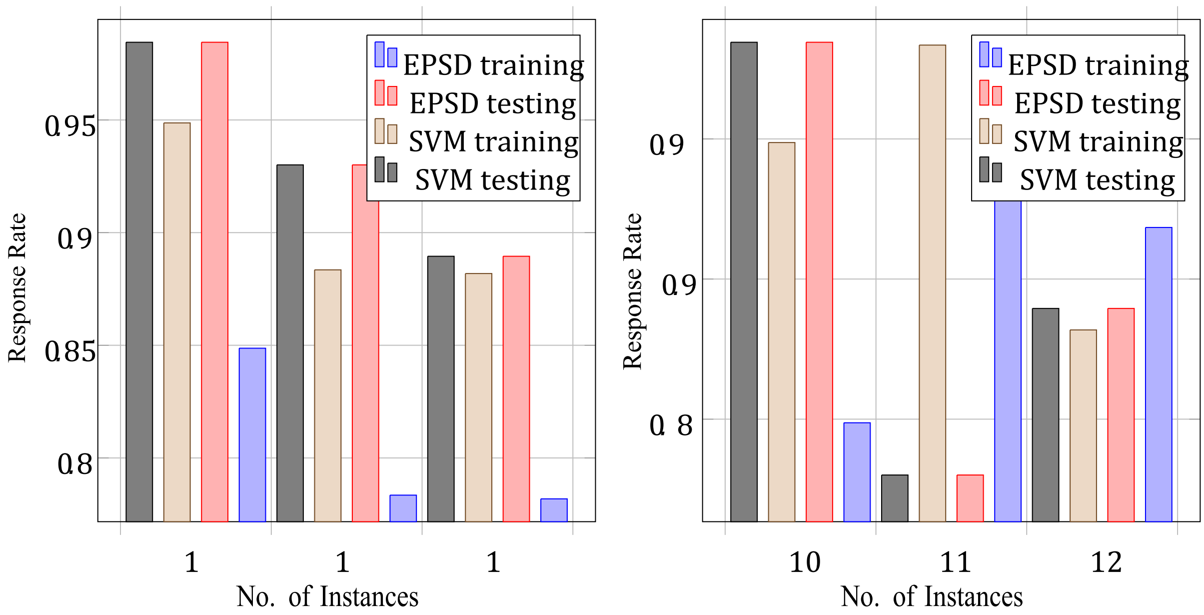
1) Adaptation to Dynamic Content: Users exhibited varying scanpaths and fixations when confronted with rapidly changing information (Figure 3). The participants tends to develop adaptive strategies in response to dynamic updates on the webpages, focusing more on areas that consistently remained relevant while sacrificing attention to shifting elements.
2) Behavioral Patterns: The eye movements varied significantly based on content type. For instance, users displayed longer fixations on games and multimedia elements (Figure 4), such as videos and images, indicating that such formats draw more cognitive engagement than text-heavy sections.
3) Design Implications: Insights from pupil measures on dynamic webpages aid us in creating more user-friendly interfaces by anticipating areas of confusion or overload—allowing for adjustments that foster better navigation and comprehension.

I. Conclusion

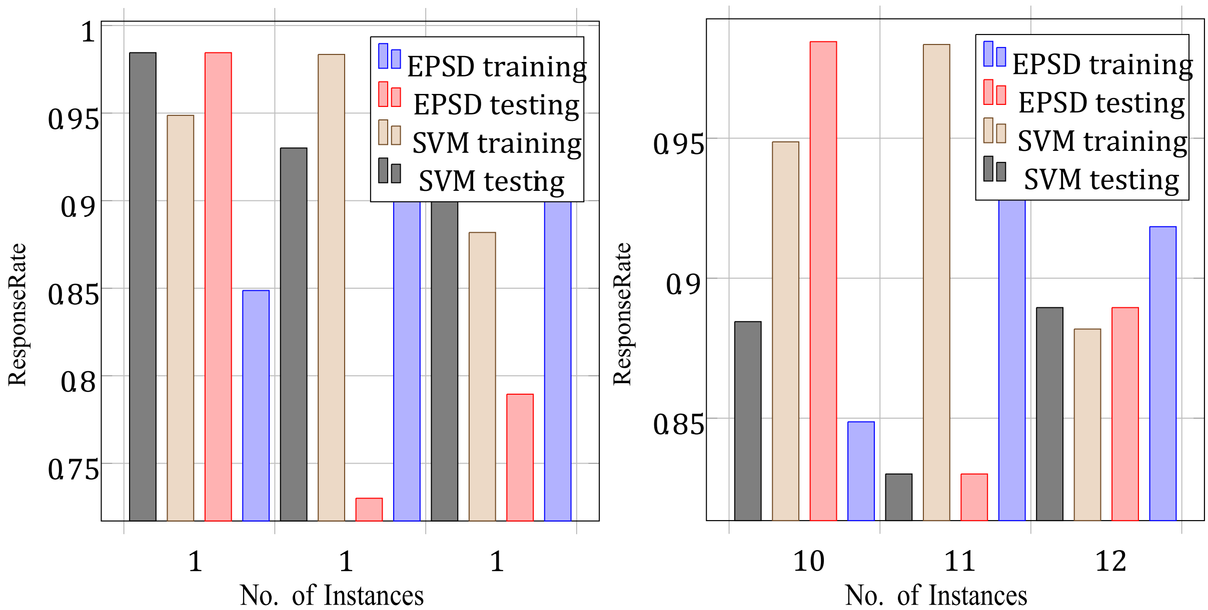
This paper tends to investigate "Classification of Scanpath for Fixations Based on Pupil Measures in Response to Dynamic Visual Stimuli". The study of scanpaths based on fixation and pupil measures presents a promising avenue for advancing our understanding of visual attention and cognitive processes.

As researchers continue to explore the intersection of these metrics, the potential for enhanced insights into how people interact with visual stimuli grows. Further studies are warranted to investigate specific applications and expand the current knowledge base.

By integrating physiological measures like pupil dilation with fixation data, researchers can achieve a more nuanced understanding of human cognition, paving the way for innovative solutions across multiple domains. Investigating scanpaths for fixations through pupil measures provides a deeper understanding of how users interact with dynamic webpages. By correlating eye movement patterns with cognitive engagement indicators like pupil size, researchers uncover the complexities of online behaviour. The findings here shown that the classification of scanpath based on distances can significantly aid web developers and designers in creating more intuitive and engaging interfaces that accommodate the ways users naturally navigate digital content. As digital interactions continue to evolve, the fusion of cognitive psychology and technology will be key to enhancing user experiences in the digital landscape. Future studies, leveraging more advanced tracking technologies, will undoubtedly yield even richer insights into the nuances of user attention and engagement on dynamic platforms.



(a) RCNN-SVM performed better at the testing set than EPSD at first. (b) RCNN-SVM and EPSD performed well in both the training and test pages. set on the second page.



(c) RCNN-SVM and EPSD performed well in both training and test(d) RCNN-SVM and EPSD performed well in both training and test set on third page. set on fourth page.

Fig. 5: Performance metrics between EPSD and RCNN-SVM

References

[1] Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., and Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. Psychological bulletin, 145(10):980.

[2] Castner, N., Kasneci, E., Kubler, T., Scheiter, K., Richter, J., Eder, T., Huttig, F., and Keutel, C. (2018). Scanpath comparison in medical image reading skills of dental students: distinguishing stages of expertise development. In Proceedings of the 2018 ACM Symposium on Eye Tracking Research &S Applications, pages 1–9.

[3] Castner, N. J. et al. (2020). Gaze and visual scanpath features for data-driven expertise recognition in medical image inspection.

[4] Foulsham, T. and Kingstone, A. (2013). Fixation-dependent memory for natural scenes: an experimental test of scanpath theory. Journal of Experimental Psychology: General, 142(1):41.

[5] Gegenfurtner, A., Lehtinen, E., and Saljo, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. Educational psychology review, 23:523–552.

[6] Holmqvist, K. and Andersson, R. (2017). Eye tracking: A comprehensive guide to methods. Paradigms and measures.

[7] Johnson, E. L., Miller Singley, A. T., Peckham, A. D., Johnson, S. L., and Bunge, S. A. (2014). Task-evoked pupillometry provides a window into the development of short-term memory capacity. Frontiers in psychology, 5:218.

[8] Korbach, A., Brunken, R., and Park, B. (2018a). Differentiating different types of cognitive load: A comparison of different measures. Educational Psychology Review, 30(2):503–529.

[9] Korbach, A., Brunken, R., and Park, B. (2018b). Differentiating different types of cognitive load: A comparison of different measures. *Educational Psychology Review*, 30(2):503–529.

[10] Krejtz, K., Zurawska, J., Duchowski, A., and Wichary, S. (2020). Pupillary and microsaccadic responses to cognitive effort and emotional arousal during complex decision making. *Journal of Eye Movement Research*, 13(5).

[11] Krischer, C. and Zangemeister, W. H. (2007). Scanpaths in reading and picture viewing: Computer-assisted optimization of display conditions. *Computers in biology and medicine*, 37(7):947–956.

[12] Liu, H. and Heynderickx, I. (2011). Visual attention in objective image quality assessment: Based on eye-tracking data. *IEEE transactions on Circuits and Systems for Video Technology*, 21(7):971–982.

[13] Mahanama, B., Jayawardana, Y., Rengarajan, S., Jayawardena, G., Chukoskie, L., Snider, J., and Jayarathna, S. (2022). Eye movement and pupil measures: A review. *frontiers in Computer Science*, 3:733531.

[14] Mishra, A., Bhattacharyya, P., Mishra, A., and Bhattacharyya, P. (2018). Scanpath complexity: modeling reading/annotation effort using gaze information. *Cognitively Inspired Natural Language Processing: An Investigation Based on Eye-tracking*, pages 77–98.

[15] Peysakhovich, V., Dehais, F., and Causse, M. (2015). Pupil diameter as a measure of cognitive load during auditoryvisual interference in a simple piloting task. *Procedia Manufacturing*, 3:5199–5205.

[16] Samarin, A., Koltunova, T., Osinov, V., Shaposhnikov, D., and Podladchikova, L. (2015). Scanpaths of complex image viewing: insights from experimental and modeling studies. *Perception*, 44(8-9):1064–1076.

[17] Schwetlick, L., Backhaus, D., and Engbert, R. (2023). A dynamical scan-path model for task-dependence during scene viewing. *Psychological Review*, 130(3):807.

[18] Stanners, R. F., Coulter, M., Sweet, A. W., and Murphy, P. (1979). The pupillary response as an indicator of arousal and cognition. *Motivation and Emotion*, 3:319–340.

[19] Stolte, M., Gollan, B., and Ansorge, U. (2020). Tracking visual search demands and memory load through pupil dilation. *Journal of Vision*, 20(6):21–21.

[20] Stuyven, E., Van der Goten, K., Vandierendonck, A., Claeys, K., and Crevits, L. (2000). The effect of cognitive load on saccadic eye movements. *Acta psychologica*, 104(1):69–85.

[21] Unsworth, N., Robison, M. K., and Miller, A. L. (2019). Individual differences in baseline oculometrics: Examining variation in baseline pupil diameter, spontaneous eye blink rate, and fixation stability. *Cognitive, Affective,* & *Behavioral Neuroscience*, 19:1074–1093.

[22] Valtchanov, D. and Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eyemovements. *Journal of environmental psychology*, 43:184– 195.

[23] Wang, C.-A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., and Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Frontiers in neurology*, 9:1029.

[24] Wang, J. T.-y. (2011). Pupil dilation and eye tracking. A handbook of process tracing methods for decision research: A critical review and user's guide, pages 185–204.

[25] Wang, Y., Bulling, A., et al. (2023). Scanpath prediction on information visualisations. *IEEE Transactions on Visualization and Computer Graphics*.

[26] Wang, Y., Lu, S., and Harter, D. (2021). Multi-sensor eye-tracking systems and tools for capturing student attention and understanding engagement in learning: A review. *IEEE Sensors Journal*, 21(20):22402–22413.

[27] Zagermann, J., Pfeil, U., and Reiterer, H. (2016a). Measuring cognitive load using eye tracking technology in visual computing. In *Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization*, pages 78–85.

[28] Zagermann, J., Pfeil, U., and Reiterer, H. (2016b). Measuring cognitive load using eye tracking technology in visual computing. In *Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization*, pages 78–85.

[29] Zagermann, J., Pfeil, U., and Reiterer, H. (2018). Studying eye movements as a basis for measuring cognitive load. In *Extended Abstracts of the 2018 CHI conference on human factors in computing systems*, pages 1–6.

[30] Zekveld, A. A., Koelewijn, T., and Kramer, S. E. (2018). The pupil dilation response to auditory stimuli: Current state of knowledge. *Trends in hearing*, 22:2331216518777174.