GRID 2024
12. INTERNATIONAL SYMPOSIUM ON GRAPHIC ENGINEERING AND DESIGN
14-16. NOVEMBER 2024

12th GRID SYMPOSIUM

AI MODELS FOR PREDICTING VISUAL ATTENTION IN DIGITAL APPLICATIONS: A COMPARATIVE PILOT ANALYSIS WITH EYE TRACKING RESULTS

Neda Milić Keresteš[1], Gala Golubović[1], Sandra Dedijer[1], Živko Pavlović[1], Tamara Janjić
[1] Faculty of Technical Sciences, Graphic Engineering and Design, Novi Sad, Serbia

Introduction

Visual attention encompasses two key mechanisms: exogenous attention, which is reflexively triggered by external stimuli, such as bright objects or unexpected movements, and endogenous attention, which is voluntarily directed by the observer and guided by specific goals and tasks (Corbetta & Shulman, 2002). Understanding these mechanisms of visual attention provides the foundation for developing AI systems that simulate human attention and predict its distribution across visual scenes (Borji & Itti, 2013).

AI-based visual attention prediction relies on data collected through eye-tracking technologies and similar techniqu es that capture how users naturally distribute their attention (Tatler et al., 2011).

The critical components of these systems are saliency and visual priorities. Saliency refers to the characteristics of objects or parts of a scene that make them stand out from their surroundings and attract our attention. These characteristics can be physical (such as contrast, color, and unusual shapes) or semantic (the meaning or relevance of an object in a given context), which together influence the automatic directing of the user's attention (Borji & Itti, 2013).

Problem Description

This study aims to evaluate whether AI can reliably be used during the early stages and iterations of UI/UX design. Specifically, the research focuses on analyzing the performance of a widely used commercial tool *Attention Insight* in a typical UI/UX design scenario.
By conducting a pilot comparative analysis between AI-predicted visual attention maps and actual eye-tracking data obtained from real users, the study seeks to assess the reliability and accuracy of AI-based predictions in the context of UI/UX design.

Methods

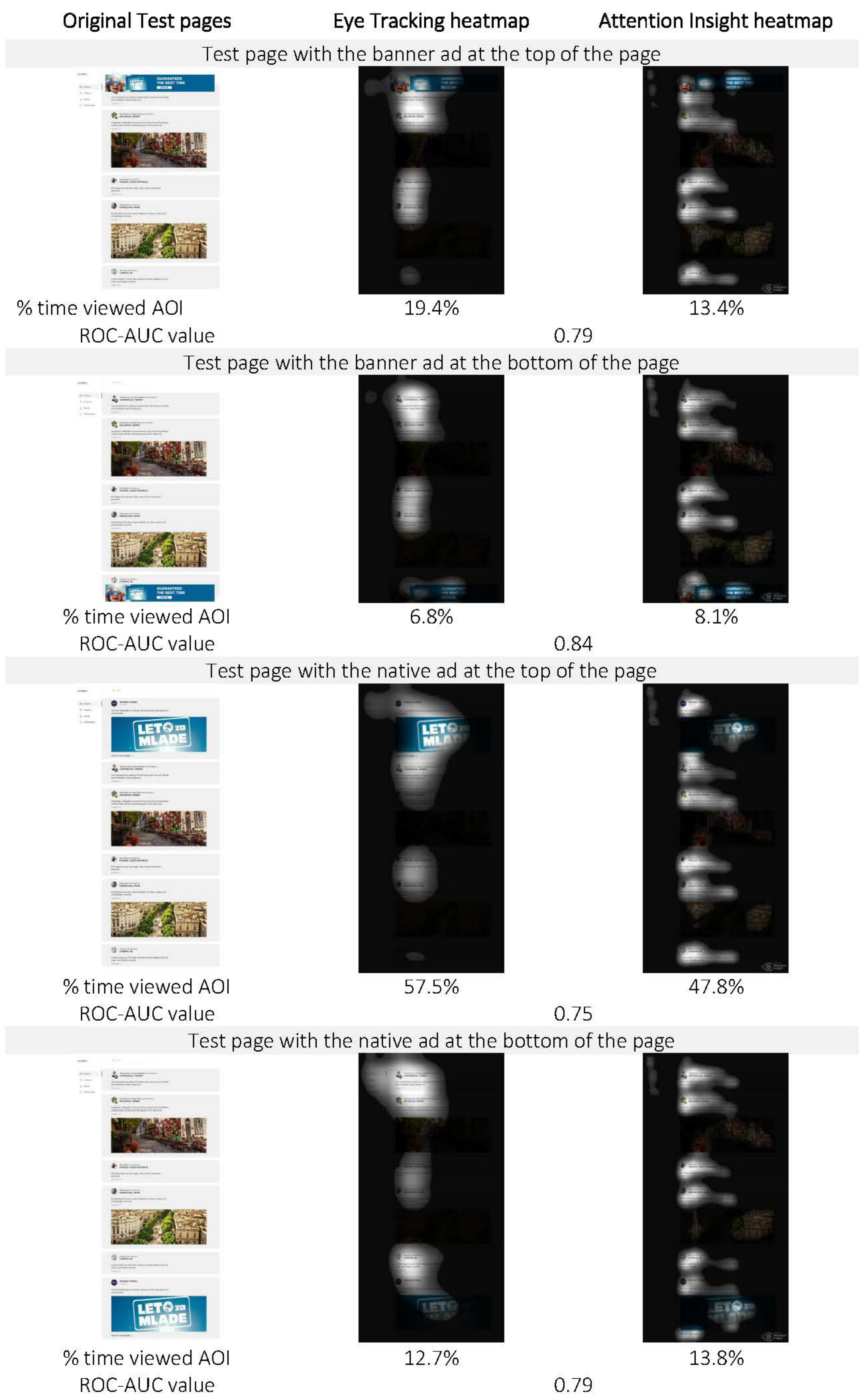
Four versions of a desktop application prototype for sharing travel experiences were created, with two featuring banner ads and two featuring native ads, differing only in the ad's position (top or bottom of the page).
Seventeen participants, with an average age of 26, participated in the eye-tracking study. After eye-tracking data collection, four cumulative heatmaps were generated, summarizing the visual attention of all 17 participants after 5 seconds of viewing each design sample. When comparing an AI-predicted visual attention heatmap to a ground-truth heatmap obtained from eye-tracking data, the Area Under the ROC Curve (AUC-ROC) quantifies the overall ability of the model to discriminate between salient and non-salient areas (Kümmerer et al., 2018):

- AUC = 0.5 suggests random guessing,
- AUC < 0.6 indicates poor performance,
- AUC < 0.8 suggests acceptable performance,
- AUC > 0.8 is considered excellent.

Results

Table 1 provides metrics such as the percentage of attention (Area of Interest, AOI) and ROC-AUC values for each test design.

Table 1 *Focus maps with metrics – AOI % of attention and ROC-AUC*



The comparative analysis between traditional eye-tracking and AI-generated heatmaps revealed a noteworthy alignment in the percentage values of time spent viewing the AOI (ad). In addition, ROC-AUC values, ranging from 0.75 to 0.84, indicate acceptable to excellent model performance, suggesting that the AI model can approximate human visual attention with a reasonable degree of accuracy.
Despite the overall good alignment, discrepancies were observed. For the native ad positioned at the top, the eye-tracking data indicated a significant focus with 57.5% attention time, whereas the AI model estimated a lower value of 47.8%.
Also, the AUC-ROC values are lower compared to their evaluation with the MIT300 dataset (Bylinskii, et al., 2019). This variance may stem from the AI model's training predominantly on images that trigger exogenous attention meaning automatic responses to salient stimuli, whereas UX interfaces often require endogenous attention, which is goal-directed and task-oriented.

Discussion / Conclusion

The findings suggest that AI-based predictions can be a valuable asset in informing design decisions, especially during the early stages of the design process. The acceptable ROC-AUC values indicate that AI tools like Attention Insight can reliably predict areas of user focus, reducing the need for extensive eye-tracking studies, which are time-consuming. However, for crucial design decisions or interfaces requiring a precise understanding of user interactions, supplementing AI predictions with traditional eye-tracking may be beneficial to capture the full spectrum of user attention.

This study's limitations include a relatively small sample size and a focus on desktop applications with specific advertisement types. The AI model's training on images eliciting exogenous attention may limit its applicability in contexts requiring endogenous attention. Future research should involve more extensive eye-tracking data with more diverse test sets.

The pilot analysis demonstrates that AI attention prediction tools like Attention Insight can effectively approximate human visual attention in UX design contexts. Despite being primarily trained on exogenous stimuli, the AI model achieved acceptable ROC-AUC values, indicating its potential utility in predicting user attention without the need for extensive eye-tracking studies. The ability of AI tools to quickly and efficiently predict user attention areas presents a significant advantage in the iterative design process. While traditional eye-tracking provides precise data on user focus and gaze both order and duration, it requires specialized equipment and more time to conduct. AI predictions offer a practical alternative for rapid assessments, enabling designers to make timely and informed decisions to enhance usability and user satisfaction. A combined approach is recommended for comprehensive analyses and a deeper understanding of user behavior. Utilizing AI tools for initial predictions followed by validation through eye-tracking studies can achieve an optimal balance between speed and accuracy.

REFERENCES

Borji, A. & Itti, L. (2013) State-of-the-art in visual attention modeling. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1), 185-207. Available from: doi: 10.1109/TPAMI.2012.89
Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A. & Torralba, A. (2019) MIT saliency benchmark. Available from: https://saliency.tuebingen.ai/results.html [Accessed 5th October 2024].
Corbetta, M. & Shulman, G. L. (2002) Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience, 3(3), 201-215. Available from: doi: 10.1038/nrn755
Kümmerer, M., Wallis, T. S. A. & Bethge, M. (2018) Saliency benchmarking made easy: Separating models, maps and metrics. In: Proceedings of the European Conference on Computer Vision (ECCV), 9-14 September 2018, Munich, Germany. Springer, pp. 630-645.
Tatler, B. W., Hayhoe, M. M., Land, M. F. & Ballard, D. H. (2011) Eye guidance in natural vision: Reinterpreting salience. Journal of Vision, 11(5), 5-5. Available from: doi: 10.1167/11.5.5

ACKNOWLEDGMENTS

This work was supported by the Serbian Ministry of Science and Technological Development, Grant No.:35027 "The development of software model for improvement of knowledge and production in graphic arts industry".