CASES: A Cognition-Aware Smart Eyewear System for Understanding Reading

Yingying Zhao
University of Strathclyde
Glasgow, UK
yingying.zhao@strath.ac.uk

Mingzhi Dong
University of Bath
Bath, UK
md2503@bath.ac.uk

Pursuing efficient and effective reading remains an ongoing goal, particularly for dense, knowledge-intensive materials. Academic tasks, such as peer-reviewing manuscripts, exemplify this complexity. Researchers globally spend over 100 million hours annually on peer reviews, with the time valued at nearly $400 million (£312.5M) for UK-based reviewers [1]. These challenges extend beyond academia, impacting professional and everyday reading as well, emphasizing the need to enhance reading efficiency and comprehension. In the broader context, advancing this area of research can provide deeper insights into the cognitive mechanisms of reading and how humans acquire and process knowledge.

A critical step toward improving reading efficiency and effectiveness is to better understand how people read. Reading is a complex, multi-level interaction between eye and mind, with each individual exhibiting a unique reading pattern [2]. These patterns reflect a reader's time-series reading intentions – whether to continue, backtrack, or pause for comprehension or memorization [2]. Recognizing these individualized reading patterns is essential for providing tailored reading support. Another challenge arises when readers lose focus or struggle to locate relevant text sections, especially during (or after) backtracking or cross-referencing. Such disruptions increase cognitive load and interrupt concentration [3], highlighting the need to understand the dynamic progression of reading states. Therefore, understanding the dynamic progression of reading states is critical to mitigating these challenges.

To address these issues, we present the Cognition-Aware Smart Eyewear System (CASES), designed to probe and explain cognitive processes during reading. CASES supports research into reading behavior, human-computer interaction (HCI), and educational applications aimed at improving reading productivity. The system is equipped with two cameras: a forward-facing scene camera capturing the reading material and an inward-facing eye camera recording eye area video for eye tracking. At its core, CASES features a deep neural network, CASES-Net, to extract features related to visual attention and text semantics. These features are fused through a shared convolutional filter mechanism based on temporal convolutional networks to enable accurate reading state estimation at the word and sentence levels. We evaluated CASES in real-world settings through an ablation study with 25 participants, demonstrating that incorporating text semantics improves context awareness and reading state detection. For example, CASES achieves a 20.90% higher accuracy at the sentence level compared to conventional eye-tracking-only methods. Moreover, semantic features enable quantitative explanations of readers' cognitive states. CASES was further integrated into an interactive reading assistant system. An in-field study across 13 participants showed that the system promoted self-awareness and improved reading habits.

Building on these results, our future work includes exploring innovative interactive visualization methods powered by smart eyewear and new computational models, aiming to mitigate the substantial time and cognitive effort required for complex academic reading. By understanding and visualizing the interplay between reader intentions and text structure, we aim to foster metacognitive awareness, helping readers navigate complex, lengthy materials more efficiently and effectively.

ACKNOWLEDGMENTS

We would like to thank our collaborators Xiangyao Qi, Qi Lu, Wentao Pan, Rui Zhu, Yuhu Chang, Qin Lv, Robert P. Dick, Fan Yang, Tun Lu, Ning Gu, and Li Shang for their valuable contributions to this project.

REFERENCES

[1] Balazs Aczel, Barnabas Szaszi, and Alex O Holcombe. 2021. A billion-dollar donation: estimating the cost of researchers' time spent on peer review. *Research Integrity and Peer Review* 6 (2021), 1–8.

[2] Xiangyao Qi, Qi Lu, Wentao Pan, Yingying Zhao, Rui Zhu, Mingzhi Dong, Yuhu Chang, Qin Lv, Robert P. Dick, Fan Yang, Tun Lu, Ning Gu, and Li Shang. 2023. CASES: A Cognition-Aware Smart Eyewear System for Understanding How People Read. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 3, Article 115 (sep 2023), 31 pages. https://doi.org/10.1145/3610910

[3] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. Visar: A human-ai argumentative writing assistant with visual programming and rapid draft prototyping. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–30.