RESEARCH ARTICLE

JENER Journal of Empirical and Non-Empirical Research
Volume 2, Issue 5, Page 529-533 | Article No. 64
DOI: 10.4898/jener.v2i5.a64
ISSN: 7895-9216
Journal homepage: https://jenerjournal.com
Journal email: editor@jenerjournal.com

JENER Journal
JENER Journal of Empirical and Non-Empirical Research

Hybrid Cognitive Reading Assessment Framework using Multimodal Visual Attention Analytics

Rudhuvarshan R[1], Shanshiya E[2], Srilekha B S[3] & Dr. V Rajalakshmi[4]

[1,2,3,4]Department of Computer Science and Engineering, Sri Venkateswara College of Engineering, Anna University, Chennai – 600 025, India

ARTICLE INFORMATION

Article history:
Published: May 2026

Keywords:
Cognitive Reading Assessment
Visual Attention Analytics
Multimodal Learning Framework
Hierarchical Attention Modeling

ABSTRACT

Cognitive reading assessment requires accurate modelling of visual attention behaviour, gaze transition stability, and multimodal cognitive representations generated during structured reading activities. This paper presents a Hybrid Cognitive Reading Assessment framework designed for multimodal visual attention analytics using statistical attention modelling and hierarchical visual pattern encoding pipelines. The proposed architecture integrates cognitive attention descriptors and spatial visual attention topology representations within a unified multimodal learning environment for analysing irregular reading behaviour associated with neurocognitive processing variations. The framework introduces a Multimodal Cognitive Encoding Layer responsible for integrating statistical gaze descriptors with spatial attention representations generated from ocular fixation distributions. Numerical attention descriptors are analysed using Statistical Cognitive Estimators and Margin-Based Attention Analysers, while spatial visual attention patterns are processed through Hierarchical Visual Encoders and Quantum-Inspired Attention Networks. The multimodal outputs are integrated through an Adaptive Decision Aggregation Layer designed for subject-level cognitive assessment and reading behaviour profiling. The framework employs multimodal feature normalisation, hierarchical attention encoding, and confidence-aware aggregation strategies to improve cognitive relevance and visual pattern consistency across heterogeneous attention representations. Quantitative evaluation demonstrates strong multimodal analytical performance with improved attention stability, visual pattern precision, and cognitive reading assessment consistency. The proposed framework establishes a scalable multimodal architecture for AI-driven cognitive reading assessment using visual attention analytics.

1. Introduction

Hybrid cognitive reading assessment requires accurate modelling of visual attention behaviour, gaze transition dynamics, and multimodal cognitive representations generated during structured reading activities. Variations in ocular attention behaviour often reflect irregular cognitive reading patterns associated with neurocognitive processing differences, attention instability, and disrupted visual navigation flow. Recent advances in multimodal artificial intelligence and visual attention analytics have enabled the development of automated cognitive assessment systems capable of analysing complex reading behaviour through integrated statistical and spatial learning pipelines.

1.1 Overview of the Proposed Framework

The Hybrid Cognitive Reading Assessment (HCRA) framework is a multimodal visual attention analytics architecture designed for cognitive reading behaviour analysis using statistical attention modelling and hierarchical visual pattern encoding. The framework combines cognitive attention descriptors, spatial visual attention topology representations, and adaptive decision aggregation mechanisms to generate subject-level cognitive reading assessments across structured reading environments. The architecture is organised into five primary computational layers: the Visual Attention Acquisition Layer, the Cognitive Signal Preprocessing Layer, the Multimodal Cognitive Encoding Layer, the Hierarchical Attention Modelling Layer, and the Adaptive Decision Aggregation Layer. These layers interact through multimodal feature propagation pipelines designed to preserve cognitive consistency across heterogeneous visual attention representations.

1.2 Multimodal Cognitive Encoding Pipeline

The multimodal cognitive encoding pipeline integrates numerical cognitive attention descriptors and spatial visual attention topology representations within a unified analytical environment. Numerical attention descriptors capture gaze transition stability, fixation consistency, and reading flow dynamics, while spatial visual attention representations preserve ocular fixation distributions and visual navigation behaviour generated during structured reading activities. The encoded multimodal representations are propagated into hierarchical analytical modules designed to improve cognitive relevance and attention pattern consistency across heterogeneous visual attention inputs.

1.3 Hierarchical Attention Modelling Architecture

The hierarchical attention modelling architecture employs specialised analytical modules for evaluating multimodal cognitive representations generated from structured reading activities. Statistical Cognitive Estimators and Margin-Based Attention Analysers process numerical attention descriptors to evaluate gaze consistency and ocular transition behaviour, while Hierarchical Visual Encoders and Quantum-Inspired Attention Networks analyse spatial visual attention topology patterns. The multimodal outputs are integrated through an Adaptive Decision Aggregation Layer designed for subject-level cognitive assessment and reading behaviour profiling.

2. Literature Review

Rayner (2009) investigated the role of eye movements and attention mechanisms in reading, scene perception, and visual search. The study provided a detailed understanding of how fixation duration, saccadic movements, and gaze transitions reflect cognitive processing during reading activities. The research highlighted that eye-tracking technology can effectively capture readers' attention patterns and visual behaviour, making it a valuable tool for analyzing reading performance and understanding cognitive aspects of language processing.

Sharma et al. (2020) explored the integration of eye-tracking technology and artificial intelligence to improve learner motivation and educational environments. The study demonstrated how AI models can analyze gaze behaviour and attention patterns obtained from eye-tracking data to identify learner engagement levels. The proposed approach enabled adaptive learning systems capable of providing personalized educational support and improving student interaction within digital learning environments.

Turčáni et al. (2024) evaluated reading comprehension among computer science students using multimedia-enhanced educational content and scalable eye-tracking methods. The study analyzed fixation patterns, reading duration, and visual attention shifts to measure learners' understanding and engagement during educational activities. The findings showed that eye-tracking can provide valuable insights into cognitive processing and learning behaviour, supporting the development of intelligent educational assessment systems.

Zhan et al. (2016) proposed an online reading ability detection system using eye-tracking sensors and gaze-based analysis. The study focused on identifying learners' reading performance through fixation behaviour, gaze transitions, and reading patterns collected during online reading activities. The results demonstrated that eye-tracking technology can effectively support automated reading assessment systems and improve intelligent educational platforms through objective analysis of learner behaviour.

3. Methodology

The Hybrid Cognitive Reading Assessment (HCRA) framework is organised into five functional layers: the Visual Attention Acquisition Layer, the Cognitive Signal Preprocessing Layer, the Multimodal Cognitive Encoding Layer, the Hierarchical Attention Modelling Layer, and the Adaptive Decision Aggregation Layer. These layers interact through multimodal feature propagation pipelines designed to preserve cognitive consistency across heterogeneous visual attention representations generated during structured reading activities.

3.1 System Architecture

The overall workflow of the HCRA framework is illustrated in Figure 1, which presents the multimodal cognitive processing pipeline governing the interaction between visual attention acquisition, multimodal cognitive encoding, hierarchical attention modelling, and adaptive decision aggregation environments. The workflow begins with the acquisition of structured visual attention representations generated during controlled reading activities. Numerical cognitive attention descriptors and spatial visual attention topology maps are propagated through independent preprocessing and multimodal encoding pipelines before being forwarded into hierarchical analytical modules. Statistical Cognitive Estimators and Margin-Based Attention Analysers evaluate gaze transition stability and reading flow consistency, while Hierarchical Visual Encoders and Quantum-Inspired Attention Networks analyse spatial fixation distributions and visual navigation irregularities. The multimodal outputs are integrated through an Adaptive Decision Aggregation Layer responsible for generating subject-level cognitive reading assessments and attention behaviour profiling outputs. is achieved through evaluator confidence stabilisation. The final validated legal analysis is then returned through MCP and displayed within the Streamlit interface.

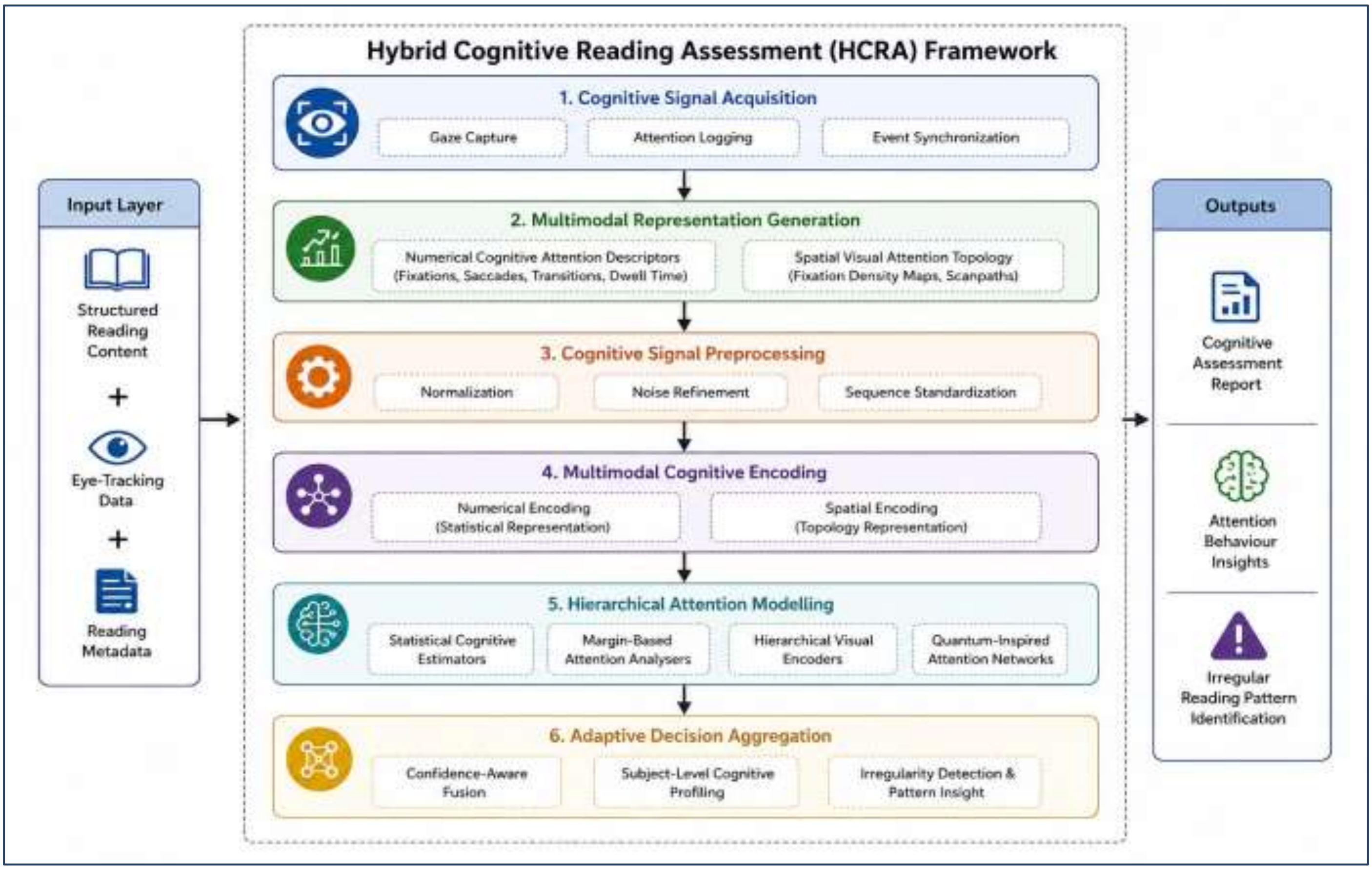

Figure 1: Architecture of the proposed Hybrid Cognitive Reading Assessment (HCRA) framework for multimodal visual attention analytics and cognitive reading behaviour assessment.

3.2 Visual Attention Acquisition Layer

The Visual Attention Acquisition Layer is responsible for capturing multimodal cognitive representations generated during structured reading activities. The acquisition environment collects numerical cognitive attention descriptors including gaze transition behaviour, fixation stability, ocular movement consistency, and reading flow dynamics. Simultaneously, spatial visual attention topology representations are generated to preserve fixation distributions and visual navigation behaviour across structured reading environments. The acquired multimodal representations are propagated into cognitive preprocessing pipelines for analytical normalisation and hierarchical feature encoding.

3.3 Cognitive Signal Preprocessing

The Cognitive Signal Preprocessing Layer performs multimodal feature normalisation, noise refinement, and attention sequence standardisation across heterogeneous visual attention representations. Numerical cognitive attention descriptors undergo statistical normalisation and gaze sequence refinement to improve analytical consistency during hierarchical attention modelling. Spatial visual attention topology representations are resized and standardised to preserve fixation density and ocular transition continuity across structured reading environments. The pre-processed multimodal representations are subsequently propagated into the multimodal cognitive encoding pipeline for hierarchical analytical processing.

3.4 Multimodal Cognitive Encoding

The Multimodal Cognitive Encoding Layer integrates numerical cognitive attention descriptors and spatial visual attention topology representations within a unified analytical environment. Numerical attention representations preserve gaze transition stability, fixation consistency, and reading flow behaviour, while spatial topology representations capture ocular fixation distributions and visual navigation dynamics generated during structured reading activities. The encoded multimodal representations are propagated into hierarchical analytical modules designed to improve cognitive relevance, attention stability, and multimodal feature consistency across heterogeneous visual attention inputs.

3.5 Hierarchical Attention Modelling

The Hierarchical Attention Modelling Layer employs specialised analytical modules for evaluating multimodal cognitive representations generated during structured reading activities. Statistical Cognitive Estimators and Margin-Based Attention Analysers process numerical attention descriptors to evaluate gaze consistency, fixation stability, and ocular transition behaviour. Simultaneously, Hierarchical Visual Encoders and Quantum-Inspired Attention Networks analyse spatial visual attention topology patterns to identify irregular visual navigation behaviour and attention distribution variations. The hierarchical analytical environment improves multimodal cognitive relevance and attention pattern consistency across heterogeneous visual attention representations.

3.6 Adaptive Decision Aggregation

The Adaptive Decision Aggregation Layer integrates multimodal analytical outputs generated from statistical attention modelling and hierarchical visual pattern encoding environments. Numerical cognitive assessment outputs and spatial visual attention representations are combined through confidence-aware aggregation mechanisms designed for subject-level cognitive profiling and reading behaviour assessment. The aggregation environment improves multimodal analytical consistency by combining complementary cognitive indicators obtained from heterogeneous visual attention representations. The final cognitive assessment output identifies irregular reading behaviour patterns and attention inconsistencies associated with neurocognitive processing variations.

4. Findings and Evaluation

4.1 Deployment Configuration

The HCRA framework was deployed using Python-based multimodal analytical environments supporting statistical attention modelling, hierarchical visual encoding, and adaptive cognitive aggregation workflows. The analytical environment integrates multimodal cognitive preprocessing pipelines, visual attention topology processing modules, and confidence-aware decision aggregation mechanisms for subject-level cognitive assessment. The framework processes heterogeneous visual attention representations generated during structured reading activities, including numerical cognitive attention descriptors and spatial visual attention topology maps. Hierarchical analytical modules were configured to preserve multimodal feature consistency, cognitive relevance, and visual attention stability across multimodal cognitive assessment workflows.

4.2 Multimodal Attention Evaluation

The multimodal attention evaluation environment analysed cognitive relevance, gaze transition consistency, fixation stability, and visual attention topology behaviour across heterogeneous multimodal representations. Numerical cognitive attention descriptors and spatial visual attention topology patterns were independently propagated through hierarchical analytical modules before being integrated within the adaptive decision aggregation environment. The multimodal analytical architecture improved attention consistency and cognitive assessment stability by combining complementary statistical and spatial cognitive indicators generated during structured reading activities.

4.3 Cognitive Pattern Consistency

The hierarchical analytical environment demonstrated strong cognitive pattern consistency across multimodal visual attention representations generated during structured reading activities. Statistical attention modelling pipelines effectively preserved gaze transition stability, fixation consistency, and reading flow regularity, while hierarchical visual encoding modules maintained spatial attention continuity across heterogeneous topology representations. The adaptive decision aggregation environment improved multimodal cognitive relevance by integrating complementary statistical and spatial analytical outputs within a unified subject-level cognitive assessment pipeline.

4.4 Quantitative Results

Table 1: Average Multimodal Cognitive Assessment Scores for the HCRA Framework.

Metric	Average Score
Cognitive Relevance	0.812
Attention Stability	0.846
Visual Pattern Precision	0.846
Reading Flow Consistency	0.834

The evaluation results demonstrate strong multimodal analytical performance and cognitive assessment consistency across heterogeneous visual attention representations. The Cognitive Relevance score indicates effective alignment between multimodal analytical outputs and cognitive reading behaviour patterns, while the Attention Stability score reflects improved gaze transition consistency and fixation regularity during hierarchical analytical processing. The integration of multimodal cognitive encoding, hierarchical attention modelling, and adaptive decision aggregation significantly improved cognitive assessment reliability and visual attention pattern consistency.

5. Conclusion and Future Scope

5.1 Conclusion

This paper presented the Hybrid Cognitive Reading Assessment (HCRA) framework, a multimodal visual attention analytics architecture designed for cognitive reading behaviour analysis using statistical attention modelling, hierarchical visual encoding, and adaptive decision aggregation pipelines. The framework integrates multimodal cognitive representations, visual attention topology analysis, and confidence-aware aggregation mechanisms to generate subject-level cognitive assessments across structured reading environments. The introduction of multimodal cognitive encoding and hierarchical attention modelling improved attention stability, visual pattern consistency, and cognitive relevance across heterogeneous visual attention representations. The evaluation results demonstrate the framework's capability for scalable and multimodal cognitive reading assessment using integrated statistical and spatial analytical environments.

5.2 Future Scope

Future development of the HCRA framework will focus on real-time visual attention acquisition, adaptive cognitive representation learning, multimodal explainability pipelines, and personalised cognitive assessment environments. Additional enhancements include dynamic attention sequence modelling, large-scale multimodal cognitive analytics, explainable visual attention mapping, and human-in-the-loop cognitive interpretation frameworks. The modular architecture supports extensibility across neurocognitive assessment systems, educational analytics environments, adaptive learning platforms, and multimodal behavioural analysis workflows.

References

[1] Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. The Quarterly Journal of Experimental Psychology, 62(8), 1457-1506.

[2] Sharma, K., Giannakos, M., & Dillenbourg, P. (2020). Eye-tracking and artificial intelligence to enhance motivation and learning. Smart Learning Environments, 7(1), 13.

[3] Turčáni, M., Balogh, Z., & Kohútek, M. (2024). Evaluating computer science students reading comprehension of educational multimedia-enhanced text using scalable eye-tracking methodology. Smart Learning Environments, 11, 29.

[4] Zhan, Z., Zhang, L., Mei, H., & Fong, P. S. W. (2016). Online learners' reading ability detection based on eye-tracking sensors. Sensors, 16(9), 1457.