Enhancing Learning in Augmented Reality (AR): A Deep Learning Framework for Predicting Memory Retention in AR Environments

Onyeka J. Nwobodo[1] [0000-0002-2706-6830], Godlove Suila Kuaban[2] [0000-0002-3321-4414], Kamil Wereszczyński[1] [0000-0003-1686-472X], and Krzysztof A. Cyran[1] [0000-0003-1789-4939]

[1] Department of Computer Graphics, Vision and Digital Systems, Silesian University of Technology, Akademicka 2A, 44-100 Gliwice, Poland,
{onyeka.nwobodo, kamil.wereszczynski, krzysztof.cyran}@polsl.pl

[2] Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Baltycka 5, 44-100 Gliwice, Poland,
gskuaban@iitis.pl

Abstract. The integration of Artificial Intelligence (AI) with Augmented Reality (AR) has transformed human-computer interaction, offering new opportunities for immersive learning and cognitive assessment. However, the relationship between user engagement in AR environments and memory retention remains underexplored. This study proposes an AI-driven framework for predicting memory retention using behavioural interaction data captured through Microsoft HoloLens 2 sensors. The model estimates the likelihood of object recall in AR-based learning environments by analyzing key interaction metrics such as gaze duration, interaction frequency, revisit counts, and head movement stability. To validate the AI predictions, we compared model-generated retention scores with user-reported recall, demonstrating a strong alignment between predicted and actual memory performance. Our findings align with established cognitive theories, indicating that increased interaction and attentional engagement enhance memory retention. Furthermore, comparisons with prior research on perceptual judgments and spatial memory reinforce the model's effectiveness in capturing real-world cognitive processes. This study introduces a scalable, non-invasive approach to cognitive modeling, bridging AI-driven analytics with AR-based learning. The results have broad implications for education, medical training, AR-based flight simulation training, and workforce development, where optimizing learning efficiency is crucial. By leveraging AI for real-time memory prediction, this research paves the way for more adaptive and personalized AR learning experiences.

Keywords: Augmented reality · Artifical intelligence · Memory retention · Gaze duration · Deep learning · Behavioral interaction.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

1 Introduction

The integration of Artificial Intelligence (AI) and Augmented Reality (AR) has ushered in a new age of human-computer interaction, with significant implications for cognitive science and education. AR integrates digital content with the real world, creating immersive environments that transform how individuals interact with complex information [1]. However, the effect of AR on memory retention, a key factor in learning effectiveness, remains insufficiently explored. This gap in knowledge presents a significant challenge, as memory retention is essential for the long-term success of educational and training interventions.

A study by Chen et al. [2] examined using Microsoft HoloLens for learning anatomy and physiology, comparing AR-based instruction to traditional PowerPoint lectures. Their findings revealed that while AR learning methods improved student engagement and reduced test anxiety, they did not significantly enhance memory recall compared to conventional methods. This highlights the need to refine AR-based learning frameworks to improve long-term retention and recall rather than focusing solely on engagement. Similarly, Gargrish et al. [3] investigated an AR-based Geometry Learning Assistant and found that while AR enhanced student engagement and visualization skills, it did not significantly improve long-term memory retention. Makhataeva et al. [4] developed the ExoMem framework, integrating computer vision and AI-driven spatial localization to augment memory. Their study found that AR significantly reduced cognitive load, improved accuracy, and enhanced performance in object-location memory tasks, with participants making 7.52 times fewer errors and completing tasks 27% faster. While highlighting AR's potential for memory augmentation, the study primarily focused on spatial cognition rather than general memory retention.

Beyond education, AI-driven AR systems have shown significant advancements in manufacturing applications, where AR reduces cognitive load by providing real-time task-related information without disrupting user focus [5]. Traditional AR methods in manufacturing rely on non-AI strategies for detection, tracking, and camera calibration, limiting their adaptability to dynamic environments. AI integration in AR has enhanced real-time adaptability through deep learning, object tracking, and ontology-based knowledge representation, providing a more scalable and practical solution across multiple domains. The potential of AI-enhanced AR systems in adaptive learning environments remains an underexplored area that can benefit from these innovations.

Memory retention is a multifaceted process influenced by cognitive, sensory, and environmental factors [6]. Research in cognitive psychology has established that attentional engagement plays a pivotal role in memory encoding, with increased interaction and focus on stimuli leading to stronger memory traces [7]. Traditional methods for studying memory, such as eye-tracking and electroencephalography (EEG), have provided valuable insights into the relationship between gaze fixation and knowledge retention. However, these methods face significant limitations in dynamic environments, particularly AR settings. For instance, while eye-tracking can identify objects of focus, it cannot capture the underlying cognitive processes or perceptions [8]. Similarly, movement arte-

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

facts often contaminate EEG data, complicating analysis and interpretation [9]. Given these challenges, there is a need for innovative, non-invasive approaches that seamlessly integrate with AR technologies.

Recent studies have explored AI-based models for predicting memory retention, primarily relying on physiological signals such as pupil dilation, heart rate variability, and EEG data [10]. While promising, these approaches rely heavily on biometric data, limiting their scalability and practicality in real-world applications. This research proposes a novel AI-driven memory prediction framework that utilizes behavioural interaction data exclusively captured within AR environments. By focusing on metrics such as gaze duration, interaction frequency, revisit counts, and head movement stability, the framework offers a practical and scalable solution for predicting memory retention.

The key contributions of this research are as follows:

- Development of an AI-driven memory prediction model trained on the Microsoft HoloLens 2 sensor data.
- Design of a computational framework for non-invasive cognitive modeling in AR environments.
- A comparative analysis of AI model predictions against user-reported memory retention demonstrates the model's accuracy and reliability.

This study significantly advances personalized learning and adaptive educational systems by bridging the gap between AI-driven cognitive models and AR-based learning environments. The proposed framework addresses a critical gap in the literature by introducing a scalable, non-invasive solution for real-time memory prediction. This innovation enables the development of more effective and engaging learning experiences. The findings have broad implications across multiple domains, including education, medical training, and workforce development, where optimizing learning efficiency and outcomes is essential.

2 Related Works

Memory retention has been widely studied using eye-tracking and electroencephalography (EEG) techniques to analyze the relationship between gaze fixation and cognitive load [11]. Prior research has demonstrated that gaze duration and pupil dilation correlate with attentional focus, a key factor in memory formation [12]. Kolnes et al. [13] further expanded on this by showing that pupil dilation reflects the breadth of attention, underscoring its utility in assessing cognitive dynamics.

EEG was also used to analyze the relation between brainwave activity and reaction time [14] and mental fatigue [15] in flight simulator sessions. EEG studies highlight theta and gamma oscillations as keys to memory encoding and retrieval. Theta activity in the medial temporal lobe (MTL) and neocortex correlates with memory accuracy and confidence, while theta-gamma phase-amplitude coupling (PAC) supports detailed memory representations [16].

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

Wynn et al. [17] found that parietal theta power correlates with memory confidence, while frontal and parietal gamma oscillations support memory accuracy and decision-making. Increased gamma power during retrieval facilitates pattern completion and neocortical information reinstatement.

Recent research has extended these findings to augmented reality (AR) environments, where physiological measurements, including EEG and eye-tracking, are used to assess cognitive load during AR-based learning and training. However, these methods face challenges in dynamic environments due to signal contamination and practical limitations. Studies suggest that AI-based behavioural analysis offers a scalable alternative for predicting memory retention without requiring intrusive biometric data. Suzuki et al. [18] systematically reviewed physiological methods in AR, identifying EEG and eye-tracking as the most prevalent techniques for assessing cognitive load. Their findings emphasize that a multi-method approach integrating EEG, eye-tracking, and self-rating scales enhances assessment reliability. Vortmann et al. [19] demonstrated that EEG and eye-tracking can distinguish attention between real and virtual objects in AR, achieving 77% accuracy using machine learning. These studies highlight the potential of EEG-based brain-computer interfaces (BCIs) to adapt AR content in real-time, enhancing cognitive training and educational tools.

Gargrish et al. [20] found that AR-based geometry learning significantly improved memory retention compared to interactive simulation (IS) methods. Over two months, AR students showed higher retention scores (12.24 post-learning, 11.76 after one week, 11.32 after two months) versus IS students (9.64, 8.00, 6.44). The immersive, interactive nature of AR enhanced engagement, visualization of abstract concepts, and long-term memory consolidation. Shen et al. [21] further advanced this field by developing a memory augmentation agent using machine learning and natural language encoding: their system, which uses large vision language models, encoded and retrieved egocentric video data from AR headsets. Using the QA-Ego4D dataset, it achieved a BLEU score of 8.3, a metric to evaluate text quality against human references, outperforming previous models (3.4-5.8). A user study showed that the agent enhanced episodic memory recall, surpassing human performance in retrieving spatial and event-based details.

While these advancements demonstrate the potential of AR for memory retention, practical challenges persist in implementing these techniques in real-world applications. Both Suzuki et al. [18] and Vortmann et al. [19] highlight the need for lightweight, non-invasive EEG solutions that can be seamlessly integrated into AR headsets. Building on these studies, our research leverages Microsoft HoloLens 2 sensors to analyze user retention through gaze duration, interaction frequency, revisit counts, and head movement stability metrics. Our work goes beyond traditional AR by using deep learning to model memory retention, an area that has been largely unexplored in prior research. Unlike studies relying on physiological signals, we leverage behavioral data for a more accessible, contact-free approach, making it more practical for real-world education and training.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

3 Material And Methodology

This study presents an AI-based framework for predicting memory retention in augmented reality (AR) environments using deep neural networks (DNN) trained on behavioural interaction metrics. The proposed model is designed to process behavioural data, including gaze duration, interaction frequency, revisit count, and head movement stability, which is captured using Microsoft HoloLens 2 sensors. By leveraging deep learning, this model learns patterns in user interaction behavior and predicts the likelihood of remembering specific objects.

3.1 Experimental Setup

The study was designed to develop and validate an AI-based model for predicting memory retention in augmented reality (AR) environments using behavioural data captured through Microsoft HoloLens 2 sensors. The AR environment was developed using Unity 3D and integrated with the Microsoft HoloLens 2 platform. The environment consists of interactive learning tasks designed to simulate real-world educational scenarios, as illustrated in Figure 1.

Participants: Thirty-six participants with varying augmented reality knowledge were recruited for the study. Before the experiment, participants received a briefing on the study objectives and provided informed consent.

Procedure: The study was conducted in a controlled AR environment using Microsoft HoloLens 2 to evaluate memory retention through interactive object engagement. Each participant was given the HoloLens 2 headset, which was carefully adjusted to ensure proper fit and calibration. Before beginning the experiment, participants received a brief explanation of the study objectives and were given instructions on interacting with the AR environment. Once the experiment started, various virtual objects were instantiated and positioned within the participant's field of view. Participants were instructed to click on each object to reveal its name, which was displayed in a text field within the AR interface. They were also encouraged to revisit and interact with objects multiple times to reinforce memory retention. This process was designed to simulate real-world learning and recall mechanisms.

Throughout the interaction phase, the system continuously recorded the following behavioural data:

Object name: The specific object that the participant interacted with. **Interaction frequency:** The number of times a participant clicked on each object. **Revisit count:** The number of times a participant returned to a previously interacted object. **Head movement stability:** Quantified as the variance of head rotation and position over a fixed time window, used to assess attentional focus and cognitive engagement. The interaction phase lasted for five minutes, during which participants could freely explore the AR environment and interact with the objects as they wished. This ensured that participants had adequate exposure to all objects and could reinforce their memory through repeated interactions.

After the interaction phase, participants received a brief 5-minute pause before completing a free recall form, where they listed objects remembered from

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

the session. No object lists or prompts were provided, ensuring unbiased memory retrieval. No false recalls occurred, though some participants omitted objects, reflecting natural variability. These responses served as ground truth for evaluating the predictions of the AI model.

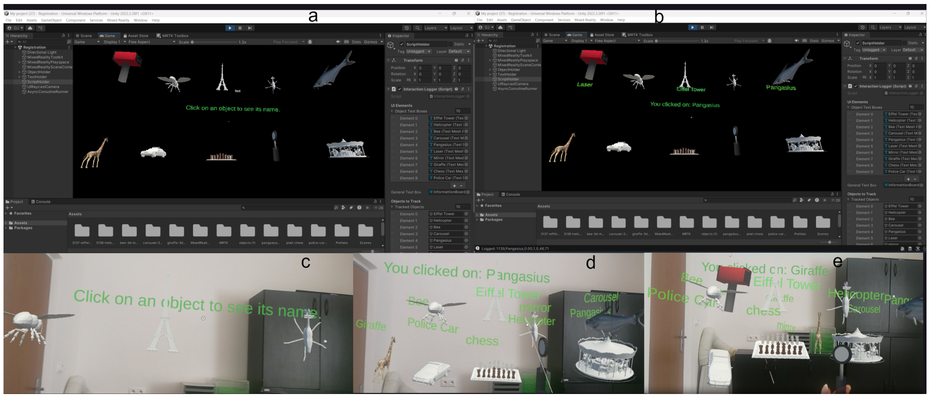

Fig. 1. (a-b) shows the user interface design in Unity for the AR-based memory retention experiment using Microsoft HoloLens 2. (c-e) illustrates the user interactions with objects in the world.

3.2 AI Model Training and Prediction Using Deep Neural Networks (DNN)

This study adopts an efficient approach for predicting memory retention by encoding image identity through category embeddings rather than processing raw image data. The proposed model receives two distinct inputs: behavioural features and object categories. The behavioural features comprise gaze duration, interaction count, revisit frequency, and head movement, which indicate user engagement within the augmented reality environment. The object category is represented as an integer and mapped into a continuous vector space via an embedding layer. This enables the model to learn dense semantic representations of object classes such as “Bee” or “Carousel.”

As illustrated in Figure 2, these two input streams are concatenated and processed through a fully connected neural network of two hidden dense layers with ReLU activation functions, followed by a dense softmax output layer. Dropout layers are incorporated to mitigate overfitting and enhance generalization. The final layer of the model is a softmax classifier that outputs one of four memory states: Strong Recall, Weak Recall, Cognitive Overload, or Lack of Engagement. The detailed configuration of the model layers is presented in Table 1.

To enhance model performance, Bayesian Optimization is employed to fine-tune key hyperparameters, including learning rate, number of units in each dense

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

layer, batch size, and early stopping patience. The model is trained using the Adam optimizer with categorical cross-entropy as the loss function. Early stopping based on validation loss ensures robust generalization.

This approach offers a lightweight yet semantically rich solution by leveraging category embeddings instead of visual feature extraction. It maintains the capacity to capture object identity while ensuring scalability and practicality for deployment in real-world augmented reality learning environments.

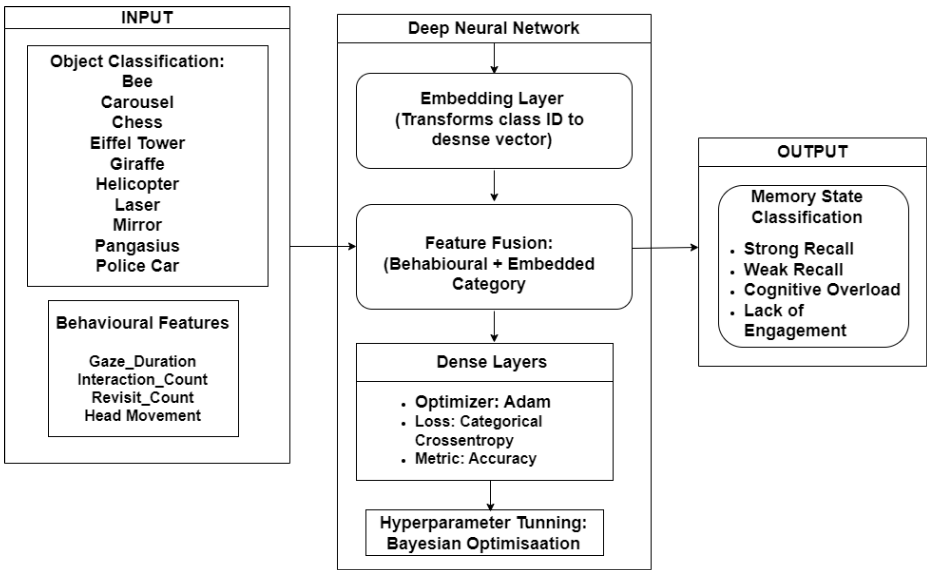

Fig. 2. An overview of the deep learning framework for memory state classification using behavioral features and embedded object categories. Outputs correspond to four memory states, with model parameters optimized via Bayesian optimization.

Table 1. Model architecture summary.

Layer (type)	**Output Shape**	**Param #**	**Connected To**
input_layer_3 (InputLayer)	(None, 1)	0	-
embedding_1 (Embedding)	(None, 1, 4)	40	input_layer_3[0][0]
input_layer_2 (InputLayer)	(None, 4)	0	-
flatten_1 (Flatten)	(None, 4)	0	embedding_1[0][0]
concatenate_1 (Concatenate)	(None, 8)	0	input_layer_2[0][0], flatten_1[0][0]
dense_3 (Dense)	(None, 20)	180	concatenate_1[0][0]
dropout_2 (Dropout)	(None, 20)	0	dense_3[0][0]
dense_4 (Dense)	(None, 30)	630	dropout_2[0][0]
dropout_3 (Dropout)	(None, 30)	0	dense_4[0][0]
dense_5 (Dense)	(None, 4)	124	dropout_3[0][0]

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

Table 2. Categorization of memory states by the AI model based on user interaction metrics.

Predicted Memory State	AI Interpretation (Based on User Interaction Metrics)
Strong Recall	High interaction counts and long gaze duration indicate deep engagement [7]. Frequent revisits reinforce memory encoding, and stable head movement reflects focused attention [12].
Weak Recall	Low interaction counts and short gaze duration suggest brief engagement [12]. Few revisits and stable head movement indicate passive involvement.
Cognitive Overload	Moderate interaction with high revisit counts and frequent gaze shifts implies cognitive strain [13]. Unstable head movement reflects difficulty processing multiple stimuli [18].
Lack of Engagement	Very low interaction, minimal revisits, short gaze duration, and unstable head movement suggest low attention and disengagement [12, 18].

4 Results and Discussions

The proposed deep neural network model was evaluated using the accuracy of training, validation, and loss metrics. As shown in Figure 3, the model demonstrated stable convergence across training epochs. It achieved a final training and validation accuracy of 0.94 and 0.93, respectively, indicating strong predictive performance on the held-out validation data. The accuracy curves rapidly increased during the initial training epochs, followed by convergence beyond epoch 80. This trend suggests that the model effectively learned representations from the behavioural interaction features: gaze duration, interaction count, revisit frequency, and head movement alongside the embedded object categories. Correspondingly, both training and validation loss decreased steadily, stabilizing below 0.2, indicating efficient minimization of prediction error. In addition to classification accuracy and loss, the model achieved a mean absolute error (MAE) of 0.14 and a mean squared error (MSE) of 0.404 on the validation set. These low error values indicate that the predicted memory states closely align with the true labels. The results confirm the model's suitability for generalization and effectiveness in memory state classification using embedded object categories and behavioural interaction features. The observed learning behaviour supports the model's potential application in real-time augmented reality (AR)-based learning environments for personalized memory prediction.

Key interaction metrics: Gaze duration, interaction frequency, revisit count, and head movement stability were analysed for their influence on predicted retention scores (Figure 4). The results show a strong correlation between engagement and memory retention, with higher scores linked to prolonged gaze, frequent interactions, and revisits. Head movement stability appeared most frequently, likely due to its continuous tracking of subtle attentional shifts. Additionally,

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

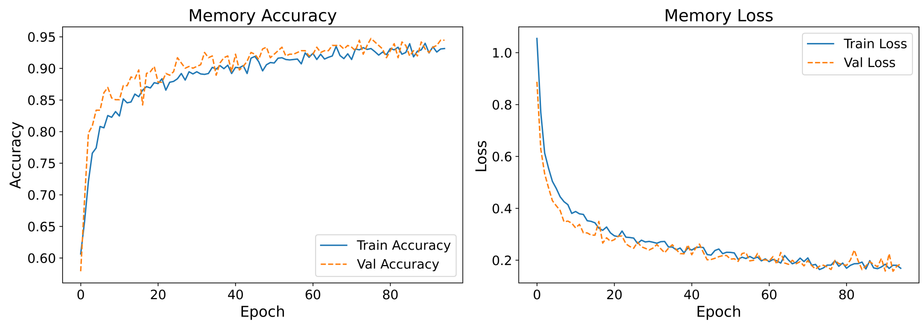

Fig. 3. Training and validation accuracy (left) and loss (right) curves of the deep neural network for memory state classification. The model shows consistent convergence across epochs, indicating effective learning and generalization.

gaze duration may influence head stability, as sustained visual focus tends to reduce head movement [12], highlighting the interplay of these features in the model's interpretation of cognitive engagement.

A detailed breakdown of the predicted and true memory states across different objects is presented in Figure 5. The grouped bar chart reveals that Chess recorded the highest number of predictions and true instances of lack of engagement, followed by Police Car, Pangasius, and Mirror. At the same time, several objects, such as Bee, Laser, and Police Car, also exhibit a considerable number of strong recall cases across both predictions and true states. While this may initially appear contradictory, it reflects the model's instance-based classification, which evaluates each object interaction independently. These outcomes are informed by distinct behavioural features such as gaze duration, interaction frequency, revisit counts, and head movement stability, captured in real-time through the Microsoft HoloLens 2. Thus, depending on their unique engagement profiles during the interaction, the same object may elicit divergent participant memory state outcomes. For example, users who demonstrated sustained attention and multiple revisits were likely associated with strong recall, while others showing minimal gaze or interaction with the same object often aligned with a lack of engagement. The behavioural variability explains the coexistence of predicted and true memory state discrepancies for a single object. In some cases, memory states observed in the true states were not predicted, and vice versa, highlighting the model's sensitivity to subtle behavioural cues that may not always align with labeled outcomes. Additionally, objects like the Carousel and Police Car showed more frequent weak recall. Meanwhile, cognitive overload remained relatively limited, occasionally appearing in objects such as the Giraffe, Eiffel Tower, and Helicopter. The results reinforce the model's ability to capture and differentiate memory outcomes based on users' heterogeneous behavioural interactions.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

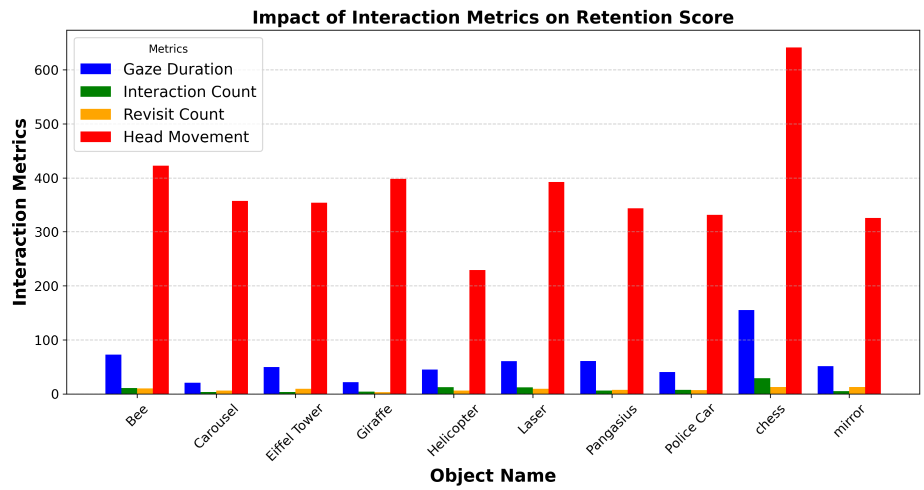

Fig. 4. Impact of interaction metrics: gaze duration, interaction count, revisit count, and head movement on memory retention for different objects.

4.1 Memory State Prediction and User Recall

To evaluate how closely the AI model's memory state predictions reflect actual human memory, we analyzed the predicted memory outcomes for objects that participants recalled after interacting with the AR environment. This comparison provides an empirical basis for assessing the model's cognitive alignment with user recall behaviour.

Figure 6 presents a grouped bar chart showing the AI-predicted memory states: Strong Recall, Weak Recall, Cognitive Overload, and Lack of Engagement for objects listed in the user feedback. Overlaid on the chart is a black line indicating the frequency with which users recalled each object. In Figure 6, Eiffel Tower was the most frequently recalled object, followed by Pangasius and Laser. The AI model also predicted these objects with a high number of Strong Recall and lower Weak Recall classifications, demonstrating a precise alignment between predicted memory strength and participant recall. In contrast, Giraffe, Helicopter, and Mirror had the lowest user recall frequencies. The model accurately captured this in the Giraffe case, which received no Strong Recall predictions, only Weak Recall, Lack of Engagement, and Cognitive Overload classifications. This reflects a strong match between the AI's prediction and actual human memory performance, suggesting that the model could infer the object's low memorability based on user interaction features. Although user recall was similarly low for Helicopter and Mirror, the model still predicted some level of Strong Recall. This indicates a partial misalignment and suggests that while the model may have detected surface-level engagement behaviors (e.g., prolonged viewing or revisits), these interactions alone may not always lead to successful long-term recall [12].

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

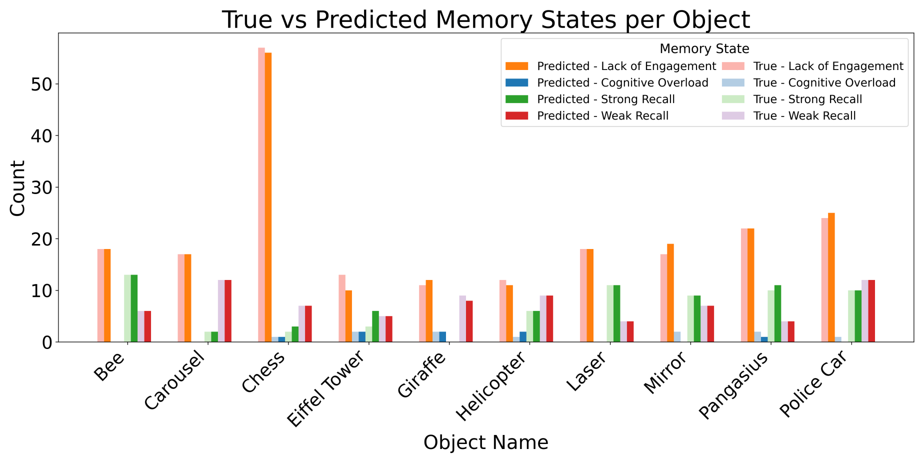

Fig. 5. Grouped bar chart showing AI-predicted and true memory states across objects based on interaction metrics. The distribution reflects object-specific engagement patterns captured during AR experiences.

Our findings align with research by Mynick et al. [22], which suggests that memory is predictive in perceptual judgments, particularly in immersive environments. Their study found that individuals use memory-based expectations to anticipate visual scenes, enhancing perception and recall efficiency. This supports our observation that users also frequently recalled objects with high AI-predicted retention scores, reinforcing that memory-driven expectations influence engagement and recall patterns.

Similarly, research on spatial memory in augmented reality (AR) environments by Maidenbaum et al. [23] further strengthens this relationship. Their study demonstrated that AR-based spatial memory tasks improve recall accuracy and engagement compared to traditional memory tests, emphasizing the role of interactive and immersive experiences. These findings are consistent with our results, where frequently interacted objects, such as the Eiffel Tower, Laser, and Pangasius, were not only the most accurately recalled by users but also closely aligned with the AI model's predictions.

Furthermore, both studies suggest that familiarity and prior exposure to objects enhance recall accuracy. Mynick et al. [22] highlight that learned environments contribute to faster and more precise recall. This aligns with our observation that AI-predicted memory retention closely mirrored user-reported recall patterns. Maidenbaum et al. [23] extend this idea by demonstrating that spatial engagement and movement enhance memory encoding. This supports our finding that increased interaction frequency and prolonged gaze duration correlate with stronger recall probabilities.

These results validate the AI model's ability to predict memory retention by replicating known cognitive processes. The observed alignment between user

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

feedback and AI predictions suggests that AI-driven memory prediction models can effectively simulate real-world memory retention processes, particularly in AR-based learning and interactive environments.

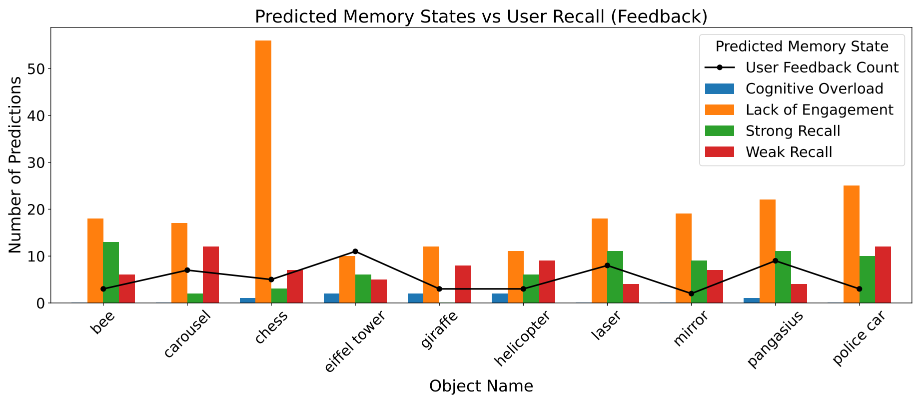

Fig. 6. Comparison of AI-predicted memory states and actual user recall frequencies for recalled objects. The bars represent the predicted memory classification per object, while the black line indicates user recall frequency based on post-interaction feedback.

4.2 Statistical Analysis

We conducted repeated-measures ANOVAs with Tukey HSD post hoc tests at the 5% significance level. Normality was confirmed via Q-Q plots (Figure 7), validating the ANOVA assumptions by showing that residuals approximated a normal distribution. A one-way ANOVA was conducted to examine differences in interaction features across the ten object categories. The dataset consisted of 1,804 object-level interaction samples (N = 1794).

The results indicated statistically significant differences among objects for all four dependent variables. Interaction Count showed a significant effect, $F(9, N) = 54.48, p < 0.001$. Similarly, significant effects were observed for Revisit Count, $F(9, N) = 15.74, p < 0.001$, Head Movement, $F(9, N) = 36.60, p < 0.001$, and Gaze Duration, $F(9, N) = 65.71, p < 0.001$. Given these significant results, a post-hoc Tukey's Honest Significant Difference (HSD) test was performed to determine which objects differed significantly. The analysis included ten distinct objects: Bee, Chess, Pangasius, Mirror, Laser, Eiffel Tower, Helicopter, Giraffe, Carousel, and Police Car.

The Tukey HSD test revealed multiple significant pairwise differences across all dependent variables. Objects with high interaction counts demonstrate significantly different engagement levels, particularly between Chess and Bee ($p < 0.001$). Revisit Count was significantly higher for Mirror and Chess compared to other objects, indicating strong user retention. This finding aligns with the

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

visualization presented in Figure 4, which illustrates the impact of interaction metrics on retention scores, highlighting the increased revisit count for Chess and Mirror. Head movement significantly differed between Pangasius and Helicopter, suggesting varying levels of physical engagement. Gaze Duration showed marked differences, with Bee and Chess receiving more visual attention than other objects.

These findings suggest that object type significantly influences user engagement, memory retention, and interaction behaviour. The results show that specific objects elicit distinct cognitive and physical responses, supporting the hypothesis that object features impact user interaction patterns.

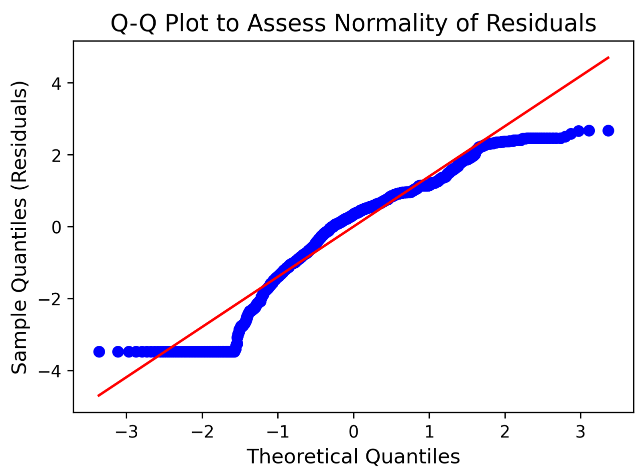

Fig. 7. Q-Q Plot to Assess Normality of Residuals. The points align closely with the theoretical quantiles, indicating approximate normality.

5 Conclusion

This study introduced an AI-driven framework for predicting memory retention in augmented reality (AR) environments using behavioural interaction data captured via Microsoft HoloLens 2 sensors. The proposed model estimated the likelihood of object recall by analyzing key interaction metrics such as gaze duration, interaction frequency, revisit counts, and head movement stability. The results revealed a strong correlation between engagement and memory retention, aligning with cognitive theories and validating the effectiveness of AI-driven approaches in modeling attention and recall. Statistical analysis, supported by ANOVA and Tukey HSD tests, confirmed significant differences in user interaction patterns across objects, highlighting the role of attentional engagement in memory encoding.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

These findings demonstrate the potential of AI-enhanced AR learning systems to improve educational and training outcomes and provide a strong foundation for future work to explore comparative modeling baselines and expand predictive robustness. The results align with prior research on perceptual judgments and spatial memory, supporting that memory-driven expectations influence engagement behaviour in AR. Additionally, the framework offers a more accessible, contact-free alternative for assessing memory processes, bridging the gap between machine learning and human cognition.

This work has important implications for domains such as education, medical training, and workforce development, where optimizing learning efficiency is critical. Future research may build on this by integrating cognitive or physiological signals to complement behavioural data and deploying the framework in real-world settings to evaluate its practical impact. These steps will further validate its potential as a personalized and adaptive learning tool.

Acknowledgement
This publication was supported by the Department of Computer Graphics, Vision and Digital Systems, under statute research project (2025), Silesian University of Technology (Gliwice, Poland).

Conflict of Interest
The authors declare no conflict of interest.

References

1. O. J. Nwobodo, G. S. Kuaban, T. Kukuczka, K. Wereszczyński, and K. Cyran, "Analysis of marker and slam-based tracking for advanced augmented reality (ar)-based flight simulation," in *International Conference on Computational Science*, pp. 208–222, Springer, 2024.
2. L. Zhang, T. Luczak, E. Smith, R. F. Burch, *et al.*, "Using microsoft hololens to improve memory recall in anatomy and physiology: A pilot study to examine the efficacy of using augmented reality in education," *Journal of Educational Technology Development and Exchange (JETDE)*, vol. 12, no. 1, p. 2, 2019.
3. S. Gargrish, D. P. Kaur, A. Mantri, G. Singh, and B. Sharma, "Measuring effectiveness of augmented reality-based geometry learning assistant on memory retention abilities of the students in 3d geometry," *Computer Applications in Engineering Education*, vol. 29, no. 6, pp. 1811–1824, 2021.
4. Z. Makhataeva, T. Akhmetov, and H. A. Varol, "Augmented-reality-based human memory enhancement using artificial intelligence," *IEEE Transactions on Human-Machine Systems*, vol. 53, no. 6, pp. 1048–1060, 2023.
5. C. K. Sahu, C. Young, and R. Rai, "Artificial intelligence (ai) in augmented reality (ar)-assisted manufacturing applications: a review," *International journal of production research*, vol. 59, no. 16, pp. 4903–4959, 2021.
6. M. W. Eysenck and M. T. Keane, *Cognitive psychology: A student's handbook*. Psychology press, 2020.
7. N. Unsworth and A. L. Miller, "An examination of individual differences in levels of processing," *Memory*, pp. 1–12, 2024.
8. Y. F. Li, S. W. Lye, and Y. Rajamanickam, "Assessing attentive monitoring levels in dynamic environments through visual neuro-assisted approach," *Heliyon*, vol. 8, no. 3, 2022.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7

9. D. Nolte, M. Vidal De Palol, A. Keshava, J. Madrid-Carvajal, A. L. Gert, E.-M. von Butler, P. Kömürlüoğlu, and P. König, "Combining eeg and eye-tracking in virtual reality: Obtaining fixation-onset event-related potentials and event-related spectral perturbations," *Attention, Perception, & Psychophysics*, pp. 1–21, 2024.
10. J. Kalafatovich, M. Lee, and S.-W. Lee, "Prediction of memory retrieval performance using ear-eeg signals," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3363–3366, IEEE, 2020.
11. H. Gorin, J. Patel, Q. Qiu, A. Merians, S. Adamovich, and G. Fluet, "A review of the use of gaze and pupil metrics to assess mental workload in gamified and simulated sensorimotor tasks," *Sensors*, vol. 24, no. 6, p. 1759, 2024.
12. N. Binetti, C. Harrison, A. Coutrot, A. Johnston, and I. Mareschal, "Pupil dilation as an index of preferred mutual gaze duration," *Royal Society Open Science*, vol. 3, no. 7, p. 160086, 2016.
13. M. Kolnes, A. Uusberg, and S. Nieuwenhuis, "Broadening of attention dilates the pupil," *Attention, Perception, & Psychophysics*, vol. 86, no. 1, pp. 146–158, 2024.
14. B. Binias, D. Myszor, S. Binias, and K. A. Cyran, "Analysis of relation between brainwave activity and reaction time of short-haul pilots based on eeg data," *Sensors*, vol. 23, no. 14, p. 6470, 2023.
15. B. Binias, D. Myszor, M. Niezabitowski, and K. A. Cyran, "Evaluation of alertness and mental fatigue among participants of simulated flight sessions," in *2016 17th International Carpathian Control Conference (ICCC)*, pp. 76–81, IEEE, 2016.
16. C. Bybee, A. Belsten, and F. T. Sommer, "Cross-frequency coupling increases memory capacity in oscillatory neural networks," *arXiv preprint arXiv:2204.07163*, 2022.
17. S. C. Wynn, C. D. Townsend, and E. Nyhus, "The role of theta and gamma oscillations in item memory, source memory, and memory confidence," *Psychophysiology*, p. e14602, 2024.
18. Y. Suzuki, F. Wild, and E. Scanlon, "Measuring cognitive load in augmented reality with physiological methods: A systematic review," *Journal of Computer Assisted Learning*, vol. 40, no. 2, pp. 375–393, 2024.
19. L.-M. Vortmann, L. Schwenke, and F. Putze, "Using brain activity patterns to differentiate real and virtual attended targets during augmented reality scenarios," *Information*, vol. 12, no. 6, p. 226, 2021.
20. S. Gargrish, A. Mantri, and D. P. Kaur, "Evaluation of memory retention among students using augmented reality based geometry learning assistant," *Education and Information Technologies*, vol. 27, no. 9, pp. 12891–12912, 2022.
21. J. Shen, J. J. Dudley, and P. O. Kristensson, "Encode-store-retrieve: Augmenting human memory through language-encoded egocentric perception," in *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 923–931, IEEE, 2024.
22. A. Mynick, A. Steel, A. Jayaraman, T. L. Botch, A. Burrows, and C. E. Robertson, "Memory-based predictions prime perceptual judgments across head turns in immersive, real-world scenes," *Current Biology*, vol. 35, no. 1, pp. 121–130, 2025.
23. S. Maidenbaum, A. Patel, I. Garlin, and J. Jacobs, "Studying spatial memory in augmented and virtual reality," *bioRxiv*, p. 777946, 2019.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97573-8_7