Decoding Memory at the Neural Level: wPLI Connectivity and Quantum Tensor Networks for EEG-Based Memorability Prediction

Berkay Bayramoglu[1,*,†], Ismail Erol[1,†], Sohel Akhtar[2,†], Murat Karakus[1,†] and Rukiye Savran Kiziltepe[1,†]

[1]Department of Software Engineering, Ankara University, Ankara, 06830, Türkiye
[2]AI Technology Graduate Program, Ankara University, Ankara, 06830, Türkiye

Abstract

We present the Anlam-Net-M team's solution for the MediaEval 2026 Predicting Media Memorability task (Challenge 1.2): predicting subject-wise movie recall from EEG. To overcome severe inter-subject variability in neural responses, we propose a hybrid architecture processing 32-channel EEG via two parallel streams: a 55-dimensional cognitive functional descriptor and a dynamic weighted Phase Lag Index (wPLI) connectivity graph. These representations are encoded by a P300-aware BiGRU attention mechanism and classified via a 4-qubit variational quantum circuit. Supported by a "Smart Fallback" strategy to recover discriminability when subject-specific training fails, our model achieves a official LOSO-CV (Leave-One-Subject-Out Cross-Validation) AUC of 0.616. This substantially outperforms classical (XGBoost, ElasticNet) and raw-EEG (EEGNet) baselines, demonstrating the vital role of neurocognitive inductive biases and dynamic connectivity in robust memory prediction.

1. Introduction and Related Work

Understanding how the brain encodes complex audiovisual experiences is a core challenge in multimedia research, driving the MediaEval 2026 Predicting Movie and Commercial Memorability task [1]. EEG provides millisecond-scale access to episodic memory gating, complementing visual and semantic features. Previous studies showed that remembered videos elicit stronger right temporal P300 positivity [2], while EEG coherence maps and ERP statistical functionals effectively predict media memorability. Recently, [3] established a functional-descriptor baseline using ElasticNet[4] , achieving a Leave-One-Subject-Out (LOSO) AUC of 0.556 and highlighting the task's strong subject-dependency.

EEG oscillations across canonical bands carry crucial memory and emotion information. Delta drives P300 responses and increases with emotional arousal [5, 6], while theta distribution facilitates multisensory audiovisual integration during encoding [7]. Posterior alpha suppression reflects initial attentional deployment, whereas late alpha and low-beta rebounds protect stored representations [8, 9]. Beyond local power, working memory variations are reliably predicted by spectral entropy [10] and inter-regional phase synchrony. To eliminate volume conduction artifacts, the weighted Phase Lag Index (wPLI) provides robust cross-subject connectivity estimates [11].

MediaEval'26: Multimedia Evaluation Workshop, June 15–16, 2026, Amsterdam, Netherlands and Online
*Corresponding author.
†These authors contributed equally.
24290604@ogrenci.ankara.edu.tr (B. Bayramoglu); 24290049@ogrenci.ankara.edu.tr (I. Erol); sakhtar@ankara.edu.tr (S. Akhtar); mrtkarakus@ankara.edu.tr (M. Karakus); rukiyekiziltepe@ankara.edu.tr (R. S. Kiziltepe)

Our proposed hybrid architecture builds on these findings with a pipeline tailored for Challenge 1.2: (i) a 55-dimensional functional descriptor capturing delta power, theta/alpha wPLI, spectral entropy, and early/late $\theta\alpha\beta$ contrasts; (ii) a dynamic wPLI graph stream (Theta, Alpha, Gamma) processed by a BiGRU with a P300-centered (300–600 ms) Gaussian attention prior; (iii) a 4-qubit tensor network quantum readout layer; and (iv) a "Smart Fallback" strategy that refits on the entire development set if subject-specific training AUC drops below 0.60. This recovers within-subject discriminability, achieving a LOSO-CV AUC of 0.616, substantially outperforming feature-based baselines (ElasticNet AUC 0.594, XGBoost AUC 0.575) and raw-EEG end-to-end models (EEGNet AUC 0.538).

2. Methodology

The proposed framework is a hybrid neural pipeline designed to address the high inter-subject variability inherent in EEG responses to multimedia [3]. It processes 32-channel EEG trials (represented as a tensor $\mathbf{X} \in \mathbb{R}^{32\times5376}$, sampled at 512 Hz over $[-1, 9.5]$ s) through two parallel streams, which are fused and read out by a variational quantum circuit to predict the final recall probability $\hat{p} \in \{0, 1\}$.

Feature Extraction: The first stream computes a 55-dimensional functional descriptor vector. P300 peaks are rescaled to microvolts to prevent vanishing gradients, and gamma-band features are strictly restricted to the early perceptual encoding window $0 - 500$ ms). The complete inventory of these descriptors is detailed in Table 1.

The second stream generates a dynamic connectivity graph. For each 100 ms window (50 ms stride) across θ, α, and γ bands, we compute the weighted Phase Lag Index (wPLI). By utilizing the signed imaginary cross-spectrum, wPLI effectively suppresses zero-lag volume conduction artifacts. Extracting upper-triangular entries and principal eigenvalues yields a dense 1,506-dimensional vector per window.

Table 1
Summary of the 55-dimensional functional descriptor vector, grouped by cognitive domain. Uninformative ERP channels and out-of-window gamma features are excluded.

Domain / Area	Feature Description
Visual & Early Gamma	Early Gamma ($0 - 500$ ms) and Gamma/Beta ratio over occipital areas; α, β, δ power [5, 6]; ERDS.
Auditory	θ power over T7/T8 [7]; long-range θ-band wPLI coherence (e.g., F3-T7, F4-T8) [11].
Semantic	α and θ power over centro-parietal (CP) and Pz [9]; CP δ power; canonical band ratios.
Emotion	Frontal Alpha Asymmetry (F, AF); Frontal-midline θ (Fz) [5, 6]; Temporal β (T7/T8); θ wPLI [11].
Attention	Frontal Engagement index ($\beta/(\alpha + \theta)$); $\alpha + \beta$ wPLI coherence [8, 11]; Individual Alpha Frequency.
Global & Temporal	Spectral Entropy (highlighting FC4) [10]; Global band ratios; Early vs. late $\theta\alpha\beta$ contrasts [9].
ERP (P300)	P300 peak amplitudes rescaled to μV over Pz, Cz, and Fz [2] (vanishing occipital ERPs removed).

Temporal Encoder & Quantum Read-Out: The 1,506-d wPLI sequence is processed by a two-layer BiGRU. Outputs are pooled via scaled dot-product attention, biased by a learnable Gaussian prior $\tilde{a}_t = a_t + s\ \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$ initialized at the expected P300 latency ($\mu = 408$ ms).

This temporal embedding is then concatenated with a projection of the 55-d descriptors and compressed into a 96-dimensional joint representation. To map this classical representation into a quantum feature space, a 4-qubit, 3-layer Variational Quantum Circuit (VQC) consumes this embedding. Recent advancements in Quantum Machine Learning (QML) demonstrate that parameterized quantum circuits act as highly expressive models [12] with intrinsic generalization capabilities [13]. In our architecture, the VQC utilizes data-dependent rotations and softmax entanglement, functioning as a powerful non-linear regularizer and an efficient tensor network

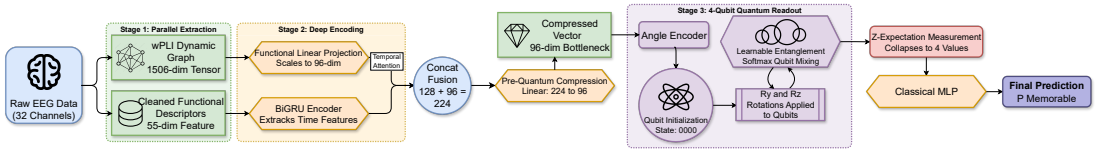
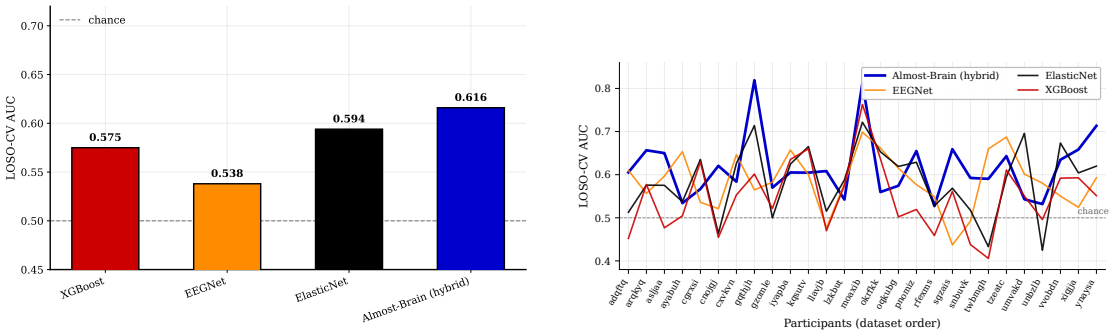


Figure 1: Architecture of the proposed hybrid pipeline.



(a) Official LOSO CV overall AUC comparison.

(b) Parallel coordinate lines illustrating individual AUC trajectories.

Figure 2: Comparative analysis of overall and per-subject model performance using the AUC metric.

read-out for the high-variance EEG data ($\approx 450\,\text{k}$ parameters total). Figure 1 illustrates the complete pipeline: 32-channel raw EEG is processed through two parallel streams to extract wPLI dynamic graphs and functional descriptors, which are deeply encoded, fused, and compressed into a 96-dimensional representation before being passed to a 4-qubit variational quantum circuit paired with a classical MLP for the final memorability prediction.

Smart Fallback Strategy: Models are initially trained subject-specifically using 100% of a subject's training data. However, given the severe inter-subject variance, we introduce a Smart Fallback rule: if a subject's training AUC drops below 0.60, the model automatically discards the subject-specific weights and refits on the entire development population to recover discriminability.

3. Results and Discussion

To evaluate the efficacy of our model, we compared it against three baseline models using LOSO-CV on the development set. The baselines include feature-based classical models (XGBoost [14] and ElasticNet[4]) trained on our 55-d functional descriptors, as well as an end-to-end deep learning model (EEGNet [15]) trained directly on raw EEG signals.

3.1. Overall Performance Comparison

As illustrated in Figure 2(a), the developed pipeline achieves the highest overall performance with an official LOSO-CV test AUC of 0.616. The fact that our hybrid model clearly outperforms both the purely classical feature-based models and the end-to-end EEGNet demonstrates the necessity of integrating hand-crafted cognitive priors (such as P300 amplitude and early-gamma restrictions) with dynamic connectivity graphs (wPLI). Furthermore, EEGNet's relatively lower performance suggests that relying solely on raw EEG without inductive biases struggles to

capture the complex neurocognitive traits of long-term memory encoding.

3.2. Inter-Subject Variability and Inference Strategy

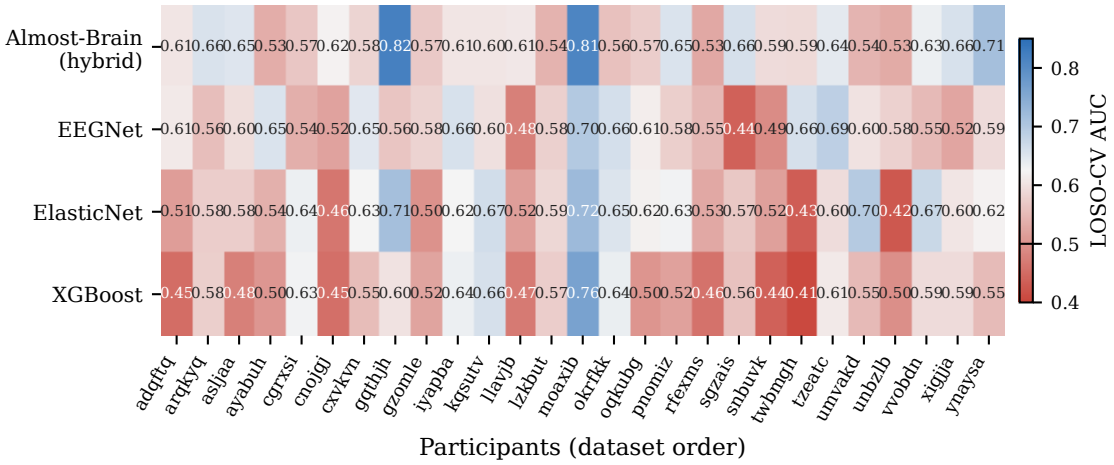


Figure 3: LOSO-CV per-subject AUC heatmap across the 27 participants.

Figures 3 and 2(b) highlight the severe inter-subject variability inherently present in EEG responses. The heatmap and parallel coordinate lines clearly demonstrate that while the models follow a similar performance trend across individuals, our framework consistently maintains superiority. For instance, specific subjects like *gqthjh* achieve exceptionally high predictability (AUC ≈ 0.82), whereas others remain challenging for all models. To address this extreme variance, our Smart Fallback strategy proved vital during testing by dynamically compensating for insufficient within-subject data, allowing the model to recover discriminability when subject-specific training fails.

4. Conclusion and Future Work

This paper presented our proposed hybrid architecture for Challenge 1.2 of the MediaEval 2026 Predicting Movie and Commercial Memorability task. Our results demonstrate that predicting long-term recall from EEG is feasible but severely hindered by inter-subject variability. By integrating hand-crafted cognitive priors—such as P300 amplitude rescaling and early-gamma restrictions—with dynamic wPLI connectivity graphs, our model achieved a LOSO-CV AUC of 0.616. This substantially surpasses purely classical baselines (XGBoost, ElasticNet) and end-to-end deep models (EEGNet). Furthermore, our "Smart Fallback" strategy proved essential in mitigating extreme subject variance. Future work should explore multimodal EEG-visual-textual models and investigate the inner representations of the quantum readout layer to provide deeper neurocognitive insights into memory encoding.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT in order to: Improve writing style, Grammar and spelling check. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] I. Martín-Fernández, A. Ganesh, M. G. Constantin, C.-H. Demarty, M. Gil-Martín, S. Halder, B. Ionescu, A. Matran-Fernandez, R. Savran Kiziltepe, A. García Seco de Herrera, Overview of the mediaeval 2026 predicting movie and commercial memorability task, in: Proceedings of the MediaEval 2026 Workshop, Amsterdam, The Netherlands and Online, 2026.

[2] R. Hamelink, Visual recognition of the memento10k dataset: A p300 component erp study., in: MediaEval, 2022.

[3] I. Martín-Fernández, M. Lobo-Alonso, S. Esteban-Romero, M. Gil-Martín, F. Fernández-Martínez, Exploring movie recall prediction using functional descriptors of the eeg signal (2025).

[4] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, Journal of the Royal Statistical Society Series B: Statistical Methodology 67 (2005) 301–320. doi:10.1111/j.1467-9868.2005.00503.x.

[5] B. Güntekin, E. Başar, A review of brain oscillations in perception of faces and emotional pictures, Neuropsychologia 58 (2014) 33–51. doi:https://doi.org/10.1016/j.neuropsychologia.2014.03.014.

[6] P. Mesa-Gresa, J.-A. Gil-Gómez, J. A. Lozano-Quilis, K. Schoeps, I. Montoya-Castilla, Electrophysiological correlates of the emotional response on brain activity in adolescents, Biomedical Signal Processing and Control 89 (2024) 105754. doi:https://doi.org/10.1016/j.bspc.2023.105754.

[7] Y. Xie, Y. Li, H. Duan, X. Xu, W. Zhang, P. Fang, Theta oscillations and source connectivity during complex audiovisual object encoding in working memory, Frontiers in Human Neuroscience Volume 15 - 2021 (2021). doi:10.3389/fnhum.2021.614950.

[8] J. Sattelberger, H. Haque, J. J. Juvonen, F. Siebenhühner, J. M. Palva, S. Palva, Local and interareal alpha and low-beta band oscillation dynamics underlie the bilateral field advantage in visual working memory, Cerebral Cortex 34 (2024) bhae448. doi:10.1093/cercor/bhae448.

[9] P. A. Packard, T. K. Steiger, L. Fuentemilla, N. Bunzeck, Neural oscillations and event-related potentials reveal how semantic congruence drives long-term memory in both young and older humans, Scientific Reports 10 (2020) 9116.

[10] Y. Tian, H. Zhang, W. Xu, H. Zhang, L. Yang, S. Zheng, Y. Shi, Spectral entropy can predict changes of working memory performance reduced by short-time training in the delayed-match-to-sample task, Frontiers in Human Neuroscience Volume 11 - 2017 (2017). doi:10.3389/fnhum.2017.00437.

[11] M. Vinck, R. Oostenveld, M. van Wingerden, F. Battaglia, C. M. Pennartz, An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias, NeuroImage 55 (2011) 1548–1565.

[12] M. Benedetti, E. Lloyd, S. Sack, M. Fiorentini, Parameterized quantum circuits as machine learning models, Quantum Science and Technology 4 (2019) 043001. doi:10.1088/2058-9565/ab4eb5.

[13] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, P. J. Coles, Generalization in quantum machine learning from few training data, Nature communications 13 (2022) 4919.

[14] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[15] V. J. Lawhern, J. N. Amelia, W. Wu, S. M. Gordon, B. J. Lance, Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces, Journal of Neural Engineering 15 (2018) 056013.