Integrated Neurosymbolic-CBR Framework for Cognitive State Assessment

Pablo Torija, Javier Bajo, Laura Melgar García[1]

Departamento de Inteligencia Artificial,
Universidad Politécnica de Madrid, Madrid, Spain
[pablo.torijam@upm.es]

Supervisors: Javier Bajo Pérez & Laura Melgar García
Department: Departamento de Inteligencia Artificial, UPM
PhD Start date: October 2024

1 Abstract

Cognitive fatigue, mental overload, and stress are among the leading causes of human error in safety-critical operations such as aviation, air traffic control, and surgical environments. While neurophysiological signals such as EEG, eye-tracking, and ECG have demonstrated potential for objective cognitive state assessment, existing approaches based on traditional machine learning and deep learning models lack two key properties required for operational deployment: the ability to leverage prior operational experience, and the capacity to provide interpretable, domain-grounded explanations of their outputs. This thesis proposes an integrated Neurosymbolic Case-Based Reasoning framework that combines learned signal representations with case retrieval and symbolic domain knowledge. The system uses deep learning models to extract neurophysiological features, a CBR module to classify cognitive states by analogy with previously observed cases, and a neurosymbolic layer that incorporates specific rules to refine classification and provide human-readable explanations. The framework is grounded in three ongoing empirical studies involving multimodal data from real operational environments, including EEG recordings onboard an F-18 with the Spanish Air Force.

This work rests on three guiding hypotheses: (H1) Neurophysiological signals provide objective, continuous markers of cognitive workload that complement subjective and performance-based measures, and combining EEG with eye-tracking and ECG improves classification robustness over single-modality approaches [1], [2]. (H2) Deep learning models achieve high accuracy in this domain but remain largely opaque; existing efforts to explain their outputs in aviation and air-traffic settings rely on post-hoc techniques such as SHAP values or heatmap-based explanations applied on top of already-trained black-box classifiers, rather than on models whose reasoning is interpretable by design [3], [4]. (H3) Case-Based Reasoning is, by its founding principles, an inherently context-sensitive paradigm particularly well suited to small-sample personalized reasoning by analogy, which motivates its use to adapt to new pilots with few available cases [5].

The main objective of this thesis is addressed through five specific sub-objectives that structure the research from signal processing to operational validation. Each sub-objective represents a necessary and independent contribution, while collectively forming the integrated framework proposed.

- **SO1.** To develop and validate a multimodal neurophysiological signal processing pipeline integrating EEG, eye-tracking, and ECG for cognitive state characterisation.

- **SO2.** To design a Case-Based Reasoning module capable of retrieving and adapting prior cognitive state cases using deep learning-based signal representations.

- **SO3.** To develop a neurosymbolic layer incorporating specific domain rules that refines classification outputs and produces human-readable explanations.

- **SO4.** To integrate the three components into a unified framework and evaluate its performance against traditional ML/DL baselines.

- **SO5.** To validate the framework in real operational conditions using data collected in collaboration with the Spanish Air Force.

While each of these components has independent precedent, their combination for this problem is, to the best of our knowledge, unexplored. Recent neurosymbolic-CBR formulations such as Gradual AA-CBR [6] learn argumentation structures for generic classification benchmarks through gradient-based methods, but do not address multimodal physiological data, expert-curated adaptation rules, or cross-context validity. Similarly, retrieval-based explainable diagnosis frameworks such as REMEMBER [7] demonstrate the value of case-based reasoning for few-shot, explainable classification in clinical neuroimaging, but have not been extended to safety-critical, operational human-factors settings. This thesis' original contribution is therefore the integration of (i) deep learning-based multimodal signal representation, (ii) context-filtered case retrieval grounded in established CBR theory, and (iii) expert-curated symbolic rules for adaptation and verification, validated not only in the laboratory but in a military jet simulator and, uniquely, onboard a real F-18 aircraft in collaboration with the Spanish Air Force, an operational validation step largely absent from the neurosymbolic-CBR and XAI literature reviewed above.

2 Methodology and Work Plan

This thesis is being conducted on a full-time basis and is not an industrial doctorate. The estimated defense date is July 2027, pending no significant delays.

2.1 Reasoning Cycle

The framework replicates the classical CBR cycle [5]: Retrieve, Reuse, Review and Retain, illustrated in Fig. 1, extended with an explicit symbolic Rules module that operates alongside, but separately from, the case base. A case will be compounded by a physiological embedding of the neurophysiological signal, together with contextual metadata (laboratory, simulator, real flight).

1. **Retrieve.** Given a new case, the system retrieves the most similar prior cases. Similarity combines a physiological component, computed over the learned embeddings, and a contextual component. In the current version, context is applied as a strict filter rather than a continuously weighted term: retrieval is restricted to cases sharing the same operational context as the new case. This choice is motivated by (i) the limited number of cases available in the early stages of the project, and (ii) the assumption that the relationship between physiological activation and cognitive workload is not directly comparable across contexts.
2. **Reuse.** The solution (workload classification) of the retrieved case(s) is adapted, rather than copied, to the new case through explicit rule-based adaptation: if a relevant variable differs between the retrieved and new case, an adaptation rule adjusts the classification accordingly. This adaptation step is explicit, documented and auditable, in contrast to the implicit reasoning of an end-to-end deep learning classifier.
3. **Review.** The adapted output is verified at two levels: an automatic consistency check against the domain rules stored in Rules (e.g., detecting a physiologically implausible classification given the flight phase), and, when system confidence is low, escalation to a human expert for validation.
4. **Retain.** The validated case, together with its context and final classification, is stored in Cases. In the current version, Retain does not automatically update Rules, as rule revision remains a manual, periodic process.

Failure handling. The cycle distinguishes three failure conditions, each triggering a different corrective action rather than a single generic retry: (a) Reuse $\rightarrow$ Retrieve ("fails"), when adaptation is not possible with the retrieved candidates, prompting a new retrieval; (b) Review $\rightarrow$ Reuse ("invalid adaptation"), when the adapted output fails the consistency check but the retrieved cases could still support a different adaptation; and (c) Review $\rightarrow$ Retrieve ("invalid candidates"), when neither the adaptation nor the underlying retrieved cases prove valid, discarding the attempt entirely. This distinction allows the system to localize at which reasoning stage a failure occurs

Rule elicitation. The Rules module is populated from structured sessions with domain experts and analysis of findings from prior published work relating neurophysiological signal patterns to different cognitive workload levels.

2.2 Work Plan

The research is structured in four phases. The first phase, completed between October 2024 and July 2026, established the empirical foundations of the thesis: a

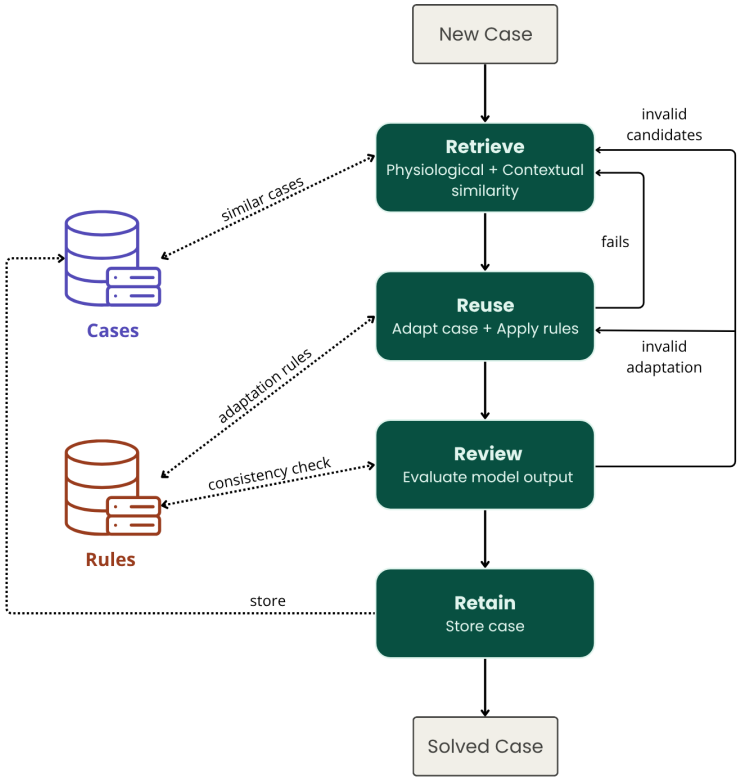


Fig. 1. Integrated Neurosymbolic-CBR framework for cognitive state assessment.

multimodal neurophysiological signal-processing pipeline integrating EEG, eye-tracking, and ECG was developed and validated on laboratory data (F1 ≈ 0.75) from 25 subjects across two sessions each, and two studies applying DL models to cognitive state classification were conducted, currently under review.

The remaining work, spanning September 2026 to July 2027, is organized around five sequential milestones, as seen in Fig. 2. It begins with a state-of-the-art review consolidating the theoretical grounding for the architecture's design choices, followed by the development of the case base and retrieval mechanism, the elicitation and formalization of the symbolic rules module, full integration of the Retrieve - reuse - Review–Retain cycle, and, finally, evaluation against the DL-only baseline. Running throughout this entire period, in parallel to these milestones, are two continuous activities: the rolling acquisition of F-18 simulator data and the ongoing drafting of the thesis manuscript, closing with a final coherence and supervisor-review pass once evaluation results are available.

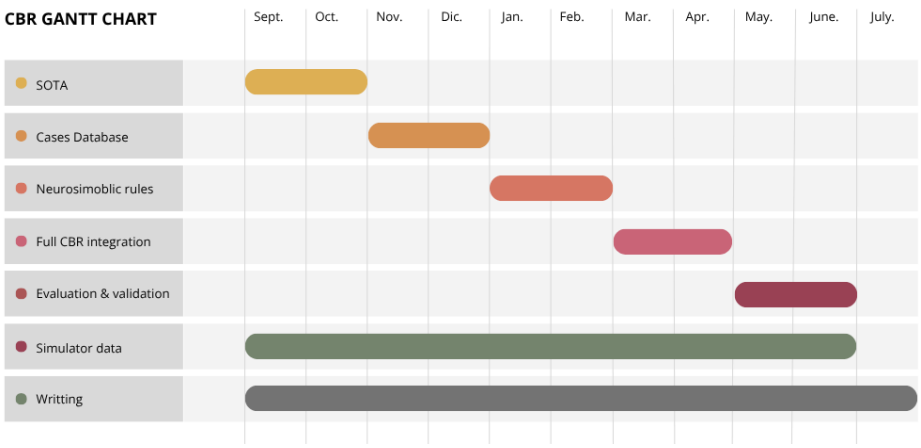


Fig. 2. Gantt Chart of CBR architecture construction

3 Relevance

This research sits at the intersection of three active fields: neurophysiological signal processing, Case-Based Reasoning, and neurosymbolic AI, each with a mature but largely separate body of work. Beyond the specific originality of this proposal, discussed above, the broader scientific relevance of this integration lies in its potential to bridge these communities: it offers CBR and neurosymbolic AI research a demanding, safety-critical testbed grounded in real physiological data, while offering the neurophysiological signal processing community a principled path toward interpretable, experience-driven models, an alternative to the purely data-driven paradigm that currently dominates the field.

Beyond its scientific contribution, this research has direct and near-term transfer potential. Cognitive overload and mental fatigue are responsible for 70–80% of accidents in safety-critical operations, yet no system currently measures the cognitive state of an operator objectively and in real time during the operation itself. The framework developed in this thesis is designed to integrate with flight simulators and operator evaluation systems in defense, aviation, and other high-stakes industries. This potential is already being explored through active funded projects in collaboration with the Spanish Air Force and major aerospace companies, positioning the research not only as an academic contribution but as the foundation of a viable technology transfer pathway.

References

1. E. Debie, R. Fernandez Rojas, J. Fidock, M. Barlow, K. Kasmarik, S. Anavatti, M. Garratt, H. A. Abbass: Multimodal Fusion for Objective Assessment of Cognitive Workload: A Review. IEEE Trans. Cybernetics 51(3), 1542–1555 (2021)

2. M. Rivas-Vidal, A. Calvo Cordoba, C. E. García Cena, F. D. Farfán: Combining EEG and eye-tracking for cognitive and physiological states monitoring: a systematic review. Frontiers in Neuroergonomics (2026). https://doi.org/10.3389/fnrgo.2025.1736672
3. I. Alreshidi, D. Bisandu, I. Moulitsas: Illuminating the Neural Landscape of Pilot Mental States: A CNN Approach with SHAP Interpretability. Sensors 23(22), 9052 (2023)
4. G. Cartocci, A. Veyrié, N. Cavagnetto, C. Hurter, A. Degas, A. Ferreira, M. U. Ahmed, S. Begum, S. Barua, B. M. S. Inguscio, V. Ronca, G. Borghini, G. Di Flumeri, F. Babiloni, P. Aricò: Explainable artificial intelligence in air traffic control: effects of expertise on workload, acceptance, and usage intentions. Brain Informatics 13(1) (2026)
5. A. Aamodt, E. Plaza: Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications 7(1), 39–59 (1994)
6. A. Gould, F. Toni: Neuro-Argumentative Learning with Case-Based Reasoning. arXiv:2505.15742 (2025)
7. D.-C. Can, Q.-H. Tang, H. Ha, B. T. Nguyen, O. Y. Chén: REMEMBER: Retrieval-based Explainable Multimodal Evidence-guided Modeling for Zero- and Few-shot Neurodegenerative Diagnosis. arXiv:2504.09354 (2025)