Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Enactive Knowledge Creation in Adaptive Intelligent Systems: Integrating the SECI Model with Multimodal Physiological Computing

Jose Valderlei da Silva

jose.vsilva@unicesumar.edu.br

Centro Universitário Cesumar

Pedro Afonso Cordeiro Sena

Centro Universitário Cesumar

Research Article

Keywords:

Posted Date: June 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9932386/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Intelligent systems that perceive and respond to human physiological states require not only pattern recognition but also the capacity to create, accumulate, and apply knowledge through embodied interaction. Most physiological-computing approaches treat affect recognition as a static classification task, thereby neglecting its situated and enactive nature. We address this gap by integrating Nonaka and Takeuchi's Socialization, Externalization, Combination, and Internalization (SECI) knowledge-creation model with enactivism to design adaptive cognitive systems.

Methods

We developed a hybrid cognitive architecture that combines specialized deep-learning models (CNN-BiLSTM for EEG; bidirectional GRU with attention for ECG/EDA); a multimodal fusion layer with alignment-aware transfer learning; and an adaptive decision policy operating in a closed-loop enactive cycle. Each SECI phase is mapped to a computational mechanism: socialization through sensor coupling, externalization via embedding extraction, combination through multimodal fusion, and internalization via policy learning. The system was evaluated on the WESAD and DREAMER benchmarks using leave-one-subject-out (LOSO) cross-validation and validated through a 30-iteration enactive-cycle simulation.

Results

The bidirectional GRU with attention achieved 61% accuracy on four-class WESAD classification under LOSO, outperforming the LightGBM and XGBoost baselines. The fusion model improved the regression error for valence, arousal, and dominance on DREAMER. The enactive simulation showed progressive physiological regulation toward well-being states (valence + 1.5, arousal − 1.3), with beneficial interventions in 87% of the iterations.

Conclusions

The enactive perception-action-adaptation cycle maps naturally onto the SECI knowledge-creation spiral, yielding a computationally grounded account of how an artificial cognitive system can create, not merely process, knowledge through continuous human-environment interaction. This work bridges enactivism, hybrid AI, and knowledge management as mutually reinforcing foundations for adaptive cognitive systems.

1 Introduction

Intelligent systems capable of perceiving, interpreting, and responding to human physiological states in real-time represent a frontier in knowledge-based systems research [1, 2]. Such systems require not only pattern recognition from sensor data but also the capacity to create, accumulate, and apply knowledge in dynamic, human-centered environments. Despite remarkable advances in deep learning for affective computing [3, 4], most existing approaches treat the problem as a static classification task, neglecting the continuous, embodied, and situated nature of human cognition [5].

The dominant paradigm in physiological computing follows a pipeline architecture: raw signals are preprocessed, features are extracted, and classifier maps feature vectors to discrete emotional labels [6]. Although effective for offline benchmarking, this approach suffers from three fundamental limitations. First, it assumes stationarity–that the relationship between physiological signals and affective states remains constant over time and across contexts. Second, it treats humans as passive data sources rather than active agents whose state is continuously shaped by environmental interactions. Third, it provides no principled mechanism for knowledge accumulation; each classification is independent, with no learning across episodes.

From the perspective of cognitive science, the enactive approach–originally proposed by Varela, Thompson, and Rosch [9]–offers a radically different framework. Enactivism posits that cognition is not the internal manipulation of symbolic representations (as in cognitivism) or the emergent behavior of connectionist networks alone but rather arises from the dynamic coupling between an embodied organism and its environment through perception-guided action. This perspective has been increasingly adopted in human-computer interaction research [7, 8] and provides a richer theoretical basis for designing adaptive systems that respond to the full complexity of human experience.

In parallel, knowledge management (KM) theory–particularly Nonaka and Takeuchi's Socialization, Externalization, Combination, and Internalization (model of organizational knowledge creation [10] model of organizational knowledge creation [10]–provides a structured framework for understanding how knowledge flows between tacit and explicit forms. The SECI cycle (socialization, externalization, combination, and internalization) has been applied in diverse domains, from organizational learning to information systems design [11, 6]. Polanyi's distinction between tacit knowledge ('we know more than we can tell') and explicit knowledge [12] is particularly relevant to physiological computing, where the human body generates rich tacit signals that must be externalized into computable representations. However, the integration of KM theory with AI architectures for physiological computing remains largely unexplored.

This study proposes a novel convergence–an enactive knowledge-based architecture that integrates i) specialized deep-learning models for multimodal physiological data,) a transfer-learning-based fusion layer with alignment loss, (iii) an adaptive decision policy operating in a closed-loop enactive cycle, and (iv) a knowledge-management framework formalized through the SECI model. The architecture was evaluated on two benchmark datasets-WESAD [13] for ECG/EDA-based stress detection and DREAMER [14] for EEG-based emotion recognition-and validated through a simulated enactive interaction cycle.) a

transfer-learning-based fusion layer with alignment loss, (iii) an adaptive decision policy operating in a closed-loop enactive cycle, and (iv) a knowledge-management framework formalized through the SECI model. The architecture was evaluated on two benchmark datasets-WESAD [13] for ECG/EDA-based stress detection and DREAMER [14] for EEG-based emotion recognition-and validated through a simulated enactive interaction cycle.

The contributions of this work are as follows: (1) we formalize the connection between enactivism and knowledge management as a theoretical foundation for adaptive knowledge-based cognitive systems; (2) we design and evaluate a hybrid multimodal architecture with alignment-aware transfer learning; (3) we demonstrate closed-loop adaptive behavior through enactive simulation; (4) we provide a systematic analysis of limitations and a research roadmap; and (5) we deliver a fully reproducible experimental framework with open-source code and public datasets.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the theoretical foundations and the proposed architecture (Methods). Section 4 reports the experimental results. Section 5 discusses their implications. Section 6 analyzes limitations and future directions, and Section 7 concludes.

2 Related Work

The intersection of cognitive science, machine learning, and knowledge management in adaptive systems has received growing but fragmented attention. We position our contribution relative to three streams of related work.

2.1 Enactivism in HCI and adaptive systems

Baranauskas, Duarte, and Valente [8] demonstrated socioenactive interaction in ubiquitous computing through school, museum, and hospital scenarios, providing empirical evidence for the social–physical–digital tripartite coupling that our architecture operationalizes computationally. Their work shows that intersubjective aspects–joint attention, joint action, and coordination–emerge naturally in well-designed socioenactive environments, supporting our claim that the enactive cycle can produce meaningful adaptive behavior. Kaipainen et al. [15] proposed enactive systems in which interaction is controlled by psychophysiological reactions rather than direct interface control, whereas Gambarotto, Potapov, and Sachs [16] recently developed an enactive approach to HCI design based on participatory sense-making and ecological normativity. Our work extends these efforts by providing a concrete computational implementation of the enactive cycle with a quantitative evaluation.

2.2 Physiological computing and multimodal fusion

Schmidt et al. [13] established the WESAD benchmark and reported accuracies between 55% and 80%, depending on the modality combination and class count. Katsigiannis and Ramzan [14] introduced DREAMER for low-cost emotion recognition using EEG and ECG. Picard, Vyzas, and Healey [17] and

Healey and Picard [18] documented the physiological overlap between high-arousal positive and negative states and explained the confusion patterns commonly observed in affect recognition. Our work extends these efforts with a principled integration framework based on transfer learning with an alignment loss, rather than treating each modality and dataset in isolation.

2.3 Knowledge management in AI systems

Although Nonaka and Takeuchi's SECI framework [10] was developed for organizational contexts, its applicability to computational systems has been increasingly explored. Chen and Luo [6] applied KM principles to intelligent information systems, building on foundational work on organizational knowledge [19]. Our contribution is a concrete, computationally grounded instantiation of the SECI cycle within an enactive AI architecture, demonstrating how each phase maps to specific computational mechanisms.

3 Methods

This section presents the theoretical foundations that motivate the design (Section 3.1), the proposed architecture (Section 3.2), and the experimental setup (Section 3.3).

3.1 Theoretical foundations

3.1.1 Cognitive science paradigms: from cognitivism to enactivism

Cognitive science has been shaped by three major paradigms [20, 21]. Cognitivism, which has been dominant since the 1950s, models the mind as a symbol-processing system, analogous to a digital computer, operating on internal representations of the external world. Knowledge is encoded in symbolic structures (rules, frames, schemas), and cognition consists of sequential rule application. While powerful for well-defined domains, cognitivism faces the 'Von Neumann bottleneck' [22]: serial processing of localized symbols limits both scalability and robustness, and the approach has been critiqued for its inability to capture commonsense reasoning [23, 24].

Connectionism emerged as an alternative to modelling cognition as the emergent behavior of large networks of simple, interconnected units. Neural networks have demonstrated remarkable pattern-recognition capabilities; however, they retain the assumption that cognition is essentially an internal computation–the relationship between the agent and environment remains unidirectional (input → processing → output). Moreover, connectionist models depend on either constrained attractor spaces through regularization or backpropagation-based optimization [9], which capture statistical regularities but not the implicit, embodied knowledge that characterizes human understanding.

Enactivism proposes a third path. Two core principles define it [9]: (i) perception consists of perceptually guided action, not passive information reception; and (ii) cognitive structures emerge from recurrent sensorimotor patterns that enable action to be guided by perception. The mind is not 'in the head' but

arises from structural coupling between organisms and the environment. Crucially, meaning is not pre-given in the world or constructed internally; it is enacted through the organism's ongoing interaction with its surroundings [25, 16]. For artificial intelligence, enactivism implies that intelligent behavior cannot be achieved through internal computation alone but requires continuous, closed-loop interaction with the environment.

3.1.2 Knowledge management and the SECI model

Nonaka and Takeuchi's SECI model [10] describes knowledge creation as a spiral process involving four modes of conversion between tacit knowledge (experiential, embodied, difficult to formalize) and explicit knowledge (codified, transmissible, systematic). Socialization (tacit to tacit) involves sharing experiences and creating shared mental models through direct interaction. Externalization (tacit to explicit) articulates tacit knowledge into explicit concepts. Combination (explicit to explicit) systematizes and integrates different bodies of explicit knowledge. Internalization (explicit to tacit) is learning-by-doing, which converts explicit knowledge into operational tacit expertise through practice.

Polanyi's concept of tacit knowledge [12] is foundational here. In physiological computing, the human body continuously generates rich tacit signals (heart-rate variability, electrodermal activity, and brain-wave patterns) that encode affective and cognitive states. These signals are tacit in Polanyi's sense: the person experiencing them typically cannot articulate the precise physiological patterns associated with their emotional state. The challenge for an enactive knowledge-based system is to externalize this tacit physiological knowledge into explicit computational representations, combine them across modalities, and internalize the resulting knowledge through adaptive behavior.

We propose that the enactive AI cycle maps naturally onto the SECI framework (Fig. 1). Socialization occurs when the system shares the human's physiological context through sensor coupling, a form of 'shared experience' between humans and machines. Externalization occurs when deep-learning models extract embedding representations from raw signals, thereby making tacit physiological patterns explicit. Combination occurs in the multimodal fusion layer, where explicit representations from different modalities are integrated. Internalization occurs when the adaptive policy learns from environmental feedback, converting explicit knowledge into operational tacit expertise.

3.1.3 Temporal and spatial reasoning in hybrid AI

Temporal representation and reasoning have been foundational concerns in AI since McCarthy and Hayes introduced situational calculus [24]. Pani and Bhattacharjee [26] reviewed temporal-reasoning frameworks and analyzed the trade-offs between first-order logic (FOL), modal temporal logic (MTL), and reified temporal logic (RTL): FOL treats time as 'just another variable' and benefits from well-known proof theory but is limited for representing concepts like 'now', “” ”since', or 'until'; MTL provides greater expressiveness at a higher computational cost; RTL reifies propositions as terms that can be predicated and quantified over.

Qualitative spatial and temporal reasoning (QSTR) have evolved as a specialized discipline within AI [27, 28]. Allen's interval of algebra, with its 13 mutually exclusive relations between temporal intervals, remains foundational. Bhatt et al. [28] argue for integrating spatial reasoning with action and change formalisms, observing that space and time are inextricably linked and that practical applications require prediction, planning, and explanation beyond mere consistency checking. Sioutis and Long [29] advocated hybrid AI systems that bridge QSTR's symbolic paradigm with machine learning through neuro-symbolic integration. This call resonates with our architecture, which combines temporal deep learning (CNN-BiLSTM for EEG dynamics; GRU for cardiac temporal patterns) with a symbolic decision policy embodying qualitative temporal reasoning about physiological-state evolution. Kasabov's [30] brain-inspired spiking neural networks further show that time–space data modelling requires architectures capable of deep learning in temporal–spatial domains; the NeuCube framework provides a biological precedent for the spatially aware, temporally sensitive processing that physiological computing demands.

3.2 Proposed architecture

The proposed enactive knowledge-based architecture (Fig. 2) consists of four interconnected modules operating in a closed-loop cycle: (i) specialized deep-learning models for each physiological modality; (ii) a multimodal fusion layer with alignment-aware transfer learning; (iii) an adaptive decision policy; and (iv) a knowledge base organized according to the SECI cycle.

3.2.1 Specialised deep-learning models

CNN-BiLSTM for EEG-based emotion recognition. We designed a CNN-BiLSTM architecture for the DREAMER dataset [14], which provides 14-channel EEG recordings during emotional stimulation. A 1D convolutional layer extracts local spectral features from band-power representations across five canonical EEG frequency bands (delta 0.5–4 Hz; theta 4–8 Hz; alpha 8–13 Hz; beta 13–30 Hz; gamma 30–45 Hz). Band powers were estimated using Welch's method [31]. Batch normalization, ReLU activation, and dropout ($p = 0.3$) followed by convolution. The convolutional output feeds a bidirectional long-term memory (LSTM) [3] that processes the sequence in both directions, which is critical because emotional states reflected in EEG may depend on both preceding and subsequent temporal contexts. A temporal-aggregation layer with learned masking handles variable-length sequences, and fully connected layers project the aggregated representation onto valence, arousal, and dominance (continuous 1–5 scales).

Bidirectional GRU with attention for ECG/EDA classification. For the WESAD dataset [13] (ECG, EDA, temperature, and accelerometer signals during baseline, stress, amusement, and meditation), we designed a bidirectional GRU with attention pooling. The GRU [32] was chosen over the LSTM for computational efficiency, achieving comparable performance with fewer parameters. The bidirectional GRU processes resampled physiological windows and captures temporal dynamics in both directions; an attention mechanism weights the most informative temporal segments. The attention weights are

computed as a_t = softmax(w · tanh(W · h_t + b)), where h_t is the GRU hidden state at time t. The attended representation passes through fully connected layers with dropout for four-class classification.

3.2.2 Multimodal fusion with alignment-aware transfer learning

The core innovation of the fusion module is the joint optimization of prediction accuracy and cross-modal alignment. The penultimate-layer activations of each specialized model serve as embedding vectors that encode the learned physiological knowledge. The fusion model comprises i) two linear projection layers that rescale embeddings into a shared latent space, ii) a concatenation layer, iii) fully connected layers with ReLU and dropout, and iv) a prediction head for valence, arousal, and dominance regression. The total loss combines prediction and alignment terms:

$$L_{total} = L_{MSE}(y, \hat{y}) + \lambda_{align} \cdot \| f_{EEG}(e_{EEG}) - f_{ECG}(e_{ECG}) \|_2$$

where L_MSE is the mean squared error, f_EEG and f_ECG are the learned projection functions, e_EEG and e_ECG are the extracted embeddings, and λ_align controls the regularization strength. This dual objective ensures that the fusion model learns a coherent cross-modal latent space in which semantically similar physiological states from different modalities map to nearby points [33].

3.2.3 Adaptive decision policy

The adaptive decision policy translates predicted physiological states into environmental actions. Five actions are defined: increase lighting, decrease lighting, increase background music, decrease background music, and schedule rest periods. Each action has a symbolic effect on the state vector [valence, arousal, dominance], with predefined impact of magnitudes bounded by safety constraints. The current implementation uses a heuristic policy based on qualitative physiological rules, such as “if arousal > threshold_high, reduce auditory stimulation, ” “if valence < threshold_low, increase luminosity,” and 'if dominance < threshold_low and arousal > threshold_mid, schedule rest period, “which embody qualitative temporal reasoning about the expected effects of environmental interventions on physiological states.

3.2.4 The enactive cycle

The four modules operate in a continuous closed-loop cycle: (1) physiological sensors capture the current state (socialization); (2) specialized models encode data into embeddings (externalization); (3) the fusion model predicts the emotional state and updates the knowledge base (combination); (4) the policy selects an environmental action and observes the consequences (internalization); (5) the environment returns a new state and reward signal, and the cycle repeats. In each iteration, the current state is replaced by the resulting state, thus implementing the perception-action-adaptation loop central to enactivism [9].

3.3 Experimental setup

3.3.1 Datasets

WESAD [13]: A multimodal dataset from 15 participants wearing a chest-mounted RespiBAN (700 Hz: ECG, EMG, EDA, temperature, respiration, and 3-axis accelerometer) and a wrist-mounted Empatica E4 (64 Hz: BVP, EDA, temperature, and accelerometer). Participants underwent baseline relaxation (~ 20 min), stress induction via the Trier Social Stress Test [34] (~ 10 min), amusement induction via video clips (~ 10 min), and guided meditation (~ 7 min). After removing invalid labels, four classes were retained: baseline, stress, amusement, and meditation.

DREAMER [14]: EEG (14 channels) and ECG recordings from 23 participants watching 18 emotion-eliciting clips. After each clip, participants rated valence, arousal, and dominance (1–5). Consumer-grade wireless devices make this dataset representative of practical deployment scenarios.

3.3.2 Preprocessing

For WESAD, signals from the chest and wrist sensors were harmonized across different sampling rates (700 Hz vs. 64 Hz) by segmentation into fixed-duration windows and within-window statistical aggregation (mean and standard deviation per channel). Label purity was enforced by discarding windows containing mixed emotional labels and avoiding label noise contamination. The 3-axis accelerometer data were decomposed into separate x, y, and z columns to enable motion-specific features.

For DREAMER, spectral band powers were extracted from raw EEG using Welch's method [31] for the five canonical bands, along with inter-band ratios (e.g., alpha/beta, theta/beta). Two aggregation steps were performed: temporal aggregation (mean and standard deviation of band powers per channel across a video) and intra-channel aggregation (a single representative vector per stimulus), yielding a consolidated feature table per participant per video, suitable for both tabular and deep-learning models.

3.3.3 Training protocol and evaluation

All models were evaluated using leave-one-subject-out (LOSO) cross-validation [35], holding out all data from one subject for testing, which is critical for assessing cross-subject generalization. Tabular baselines (LightGBM [36] and XGBoost [37]) were trained using aggregated feature vectors. Deep models were implemented in PyTorch with the Adam optimizer, early stopping, and learning-rate scheduling. Table 1 summarizes the key hyperparameters.

Table 1
Key hyperparameters for the deep-learning models.

Hyperparameter	Value
Optimizer	Adam
Learning rate	1×10^{-3}
Batch size	32 / 64
Hidden units	128 (BiLSTM) / 64 (GRU)
Dropout	0.3
Epochs	100 + early stopping
Validation	Leave-one-subject-out (LOSO)

4 Results

4.1 Classification performance (WESAD)

LightGBM achieved a global accuracy of 56% (F1 0.55–0.67 for the dominant baseline and stress classes), while XGBoost achieved 58% (F1 0.60–0.71). Both showed strong discrimination for the baseline and stress, which produced clearly distinct physiological signatures (elevated heart rate and EDA during stress); however, they struggled with amusement and meditation, which share subtle, overlapping profiles. These baselines establish the performance floor.

The bidirectional GRU with attention achieved 61% accuracy and a macro F1 of 0.57, outperforming both tabular baselines (Table 2). The confusion matrix showed the largest concentration of correct predictions on the diagonal for the baseline (2,308 correct) and meditation (1,241 correct), while significant confusion occurred between stress and amusement, consistent with the known physiological overlap between high-arousal positive and negative states measured through ECG and EDA alone [17, 18].

Table 2
Classification performance on WESAD (four-class, LOSO).

Model	Accuracy	Macro F1	Precision	Recall
LightGBM	0.56	0.52	0.54	0.53
XGBoost	0.58	0.55	0.57	0.55
GRU-Attention	0.61	0.57	0.59	0.58

These results fall within the range reported for WESAD, typically 55–70% for the four-class problem, reaching 80% only with multiple modalities and subject-dependent splits [13]. Our results, obtained under

the stringent LOSO protocol with ECG/EDA only confirm the difficulty of cross-subject generalization and highlight the value of the attention mechanism in selectively weighting informative temporal segments.

4.2 Regression performance (DREAMER)

The CNN-BiLSTM achieved moderate regression performance. Compared with Fonseca, Guido, and Junior [39] (PCC ≈ 0.6, MAE ≈ 0.18), our model shows lower absolute performance, attributable to using raw EEG with LOSO validation rather than handcrafted features with subject-dependent splits. The fusion model improved MAE across all three dimensions relative to the CNN-BiLSTM alone (Fig. 3), demonstrating the value of multimodal knowledge combination consistent with the SECI combination phase.

4.3 Enactive-cycle simulation

The 30-iteration enactive simulation (Fig. 4) shows that the closed-loop system progressively regulates the simulated physiological state. Starting from [valence 2.5, arousal 3.8, dominance 2.8], the system's interventions produced the following trajectory: valence increased to ~ 4.0 (+ 1.5), arousal decreased to ~ 2.5 (-1.3), and dominance increased to ~ 3.5 (+ 0.7)–consistent with a physiological-regulation pattern that promotes well-being by moving from a stressed, low-valence state toward a relaxed, positive-affect state [38].

At each iteration, the policy selected actions based on the current predicted state: reduction of auditory stimulation when arousal exceeded the threshold, increase in luminosity when valence was low, and initiation of a rest period when both high arousal and low dominance were triggered. The reward signal (magnitude of state improvement) was positive in 87% of the iterations, indicating that the heuristic policy successfully identified beneficial interventions in most cases.

4.4 Knowledge-management mapping

Table 3
maps each architectural component to its corresponding SECI phase, demonstrating that the proposed system implements a complete knowledge-creation cycle. This mapping is not merely metaphorical: each phase corresponds to a concrete computational mechanism that transforms knowledge between tacit and explicit forms.

SECI phase	Conversion	Component	Mechanism
Socialization	Tacit → Tacit	Sensor coupling	Shared physiological context (embodied signals)
Externalization	Tacit → Explicit	CNN-BiLSTM / GRU	Embedding extraction (feature vectors)
Combination	Explicit → Explicit	Fusion layer	Multimodal integration (unified prediction)
Internalization	Explicit → Tacit	Adaptive policy	Feedback learning (operational expertise)

Table 3 Mapping of the SECI knowledge-creation model to the proposed enactive architecture.

5 Discussion

The socialization phase is implemented through continuous coupling between physiological sensors and the human body. Unlike traditional input devices that require conscious user actions, wearable sensors create a form of “shared experience”–the system participates in the human physiological context without requiring explicit articulation. This constitutes a tacit-to-tacit transfer: the embodied, pre-reflective physiological state is shared non-verbally, much as tacit knowledge is transmitted through apprenticeship and shared practice in organizational settings.

Externalization occurs when deep-learning models transform raw physiological signals into embedding vectors. The CNN-BiLSTM and GRU architectures act as ‘articulation engines’ that make implicit patterns explicit by converting the tacit knowledge embodied in heart rate variability, skin conductance, and brain waves into structured computational representations. Combination is implemented in the multimodal fusion layer, where alignment-aware transfer learning ensures that knowledge from different physiological sources is not merely concatenated but genuinely combined, creating new explicit knowledge that transcends any single modality. Internalization occurs when the adaptive policy learns environmental feedback, and each action-observation pair contributes to the system's operational expertise. Although the current heuristic policy provides only shallow internalization, the framework supports deeper learning through reinforcement mechanisms.

Taken together, these results support the central hypothesis: through continuous perception-action-adaptation, the system generates a knowledge spiral in the SECI sense, in which environmental interventions are progressively refined based on accumulated experiential knowledge. This reframes from physiological computing from a static classification pipeline into a knowledge-creating cognitive system and provides a computational reading of enactivism that is testable and reproducible.

6 Limitations and Future Directions

6.1 Limitations

First, the enactive cycle was validated only in a simulation. The symbolic environment, while adequate for demonstrating perception-action-adaptation, does not capture the full complexity of real human environment interactions; the symbolic action effects (e.g., ‘increasing light increases valence by 0.1’) are rough approximations. Validation with real-time sensors and actual actuators is necessary.

Second, the adaptive decision policy is more heuristic than learned. Although the rule-based policy is interpretable and safe, it cannot discover novel intervention strategies or optimize long-term outcomes.

Third, the absolute performance metrics are modest compared with state-of-the-art results obtained with subject-dependent protocols. Our use of LOSO is methodologically more rigorous but yields lower

accuracy, reflecting the genuine difficulty of cross-subject physiological computing.

Fourth, the knowledge-management framework is formalized at the descriptive level. The SECI mapping (Table 3) provides a principled account of knowledge flow but does not yet include quantitative metrics for knowledge-creation efficiency, spiral velocity, or knowledge quality.

Fifth, although the datasets are public and widely benchmarked, they have inherent constraints: WESAD includes only 15 subjects, and DREAMER uses consumer-grade EEG with 14 channels. Sixth, the architecture does not yet incorporate subject-specific adaptation (individual normalization or personalized fine-tuning), which has been shown to improve cross-subject performance [14].

6.2 Future research directions

(1) Real-world deployment in physical smart environments (intelligent offices, classrooms, and healthcare settings) with real-time sensors and actuators; the socioenactive design methodology of Baranauskas et al. [8] provides a methodological template. (2) Learned adaptive policies via reinforcement learning [40], with attention to safe exploration, reward design, and non-stationarity. (3) Brain-inspired architectures: Kasabov's [30] NeuCube and spiking neural networks offer biologically plausible alternatives for temporal–spatial modelling. (4) Quantitative knowledge-management metrics for externalization rate, combination quality, and internalization depth. (5) Personalization through meta-learning [41] to enable rapid adaptation to new users with minimal calibration data. (6) Ethical and privacy considerations: informed consent, data privacy, algorithmic transparency, and prevention of coercive uses of affective technology.

7 Conclusion

This study presents an enactive knowledge-based architecture that integrates deep learning, multimodal fusion, adaptive decision-making, and knowledge management for physiological state estimation. The key theoretical insight is that the enactive cycle of perception-action-adaptation maps naturally onto the SECI knowledge-creation spiral, providing a principled framework for designing adaptive intelligent environments that create, transform, and apply knowledge through ongoing interactions with human users.

Experimentally, a bidirectional GRU with attention achieved 61% accuracy on a four-class WESAD classification under LOSO; multimodal fusion with alignment loss improved the regression error on DREAMER; and a 30-iteration enactive simulation showed progressive regulation toward positive-affect states. Although the absolute metrics are modest, they are consistent with state-of-the-art cross-subject physiological computing and provide a solid foundation for future development. By mapping each architectural component to a specific phase of knowledge conversion, this work demonstrates that the enactive paradigm and knowledge-management theory are not only compatible but mutually reinforcing, and contributes to bridging cognitive science, hybrid AI, and knowledge management.

Declarations

Funding: No funding was received to assist with the preparation of this manuscript.

Competing interests: The authors declare that they have no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Ethics approval: This study used only publicly available, fully anonymized secondary datasets (WESAD and DREAMER). No new human-subject data were collected by the authors.

Data and code availability: WESAD and DREAMER are publicly available from their respective repositories; all source code is available at https://github.com/SenhorAfonso/tcc-socioenativismo

Use of generative AI

During the preparation of this work, the authors used Claude to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

Author contributions (CRediT)

Jose Valderlei da Silva: Supervision, Writing manuscript - review and editing.

Pedro Afonso Cordeiro Sena: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing report - original draft, Visualization.

All authors read and approved the final manuscript.

References

1. van Harmelen F, Lifschitz V, Porter B, editors. Handbook of knowledge representation. Amsterdam: Elsevier; 2007.
2. Davis E, Morgenstern L. Introduction: progress in formal commonsense reasoning. Artif Intell. 2004;153:1-12.
3. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-80.
4. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.
5. Picard RW. Affective computing. Cambridge (MA): MIT Press; 1997.
6. Chen Z, Luo J. A survey on application of knowledge management in AI systems. Knowl Based Syst. 2022;255:109718.
7. Frauenberger C. Entanglement HCI: the next wave? ACM Trans Comput Hum Interact. 2020;27(1):1-27.

8. Baranauskas MCC, Duarte EF, Valente JA. Socioenactive interaction: addressing intersubjectivity in ubiquitous design scenarios. Int J Hum Comput Interact. 2024;40(13):3365-80.
9. Varela FJ, Thompson E, Rosch E. The embodied mind: cognitive science and human experience. Cambridge (MA): MIT Press; 1991.
10. Nonaka I, Takeuchi H. The knowledge-creating company: how Japanese companies create the dynamics of innovation. New York: Oxford University Press; 1995.
11. Alavi M, Leidner DE. Review: knowledge management and knowledge management systems. MIS Q. 2001;25(1):107-36.
12. Polanyi M. The tacit dimension. Chicago: University of Chicago Press; 1966.
13. Schmidt P, Reiss A, Duerichen R, Marberger C, Van Laerhoven K. Introducing WESAD: a multimodal dataset for wearable stress and affect detection. In: Proceedings of the 20th ACM International Conference on Multimodal Interaction. New York: ACM; 2018. p. 400-8.
14. Katsigiannis S, Ramzan N. DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. IEEE J Biomed Health Inform. 2018;22(1):98-107.
15. Kaipainen M, Ravaja N, Tikka P, Vuori R, Pugliese R, Rapino M, et al. Enactive systems and enactive media: embodied human-machine coupling beyond interfaces. Leonardo. 2011;44(5):433-8.
16. Gambarotto A, Potapov K, Sachs CB. Design for autonomy: an enactive approach to human-computer interaction. Adapt Behav. 2025. doi:10.1177/10597123251405703.
17. Picard RW, Vyzas E, Healey J. Toward machine emotional intelligence: analysis of affective physiological state. IEEE Trans Pattern Anal Mach Intell. 2001;23(10):1175-91.
18. Healey JA, Picard RW. Detecting stress during real-world driving tasks using physiological sensors. IEEE Trans Intell Transp Syst. 2005;6(2):156-66.
19. Davenport T, Prusak L. Working knowledge: how organizations manage what they know. Boston: Harvard Business School Press; 1998.
20. Bermudez JL. Cognitive science: an introduction to the science of the mind. 4th ed. Cambridge: Cambridge University Press; 2023.
21. Hayes PJ. The naive physics manifesto. In: Michie D, editor. Expert systems in the micro-electronic age. Edinburgh: Edinburgh University Press; 1979. p. 242-70.
22. Backus J. Can programming be liberated from the von Neumann style? Commun ACM. 1978;21(8):613-41.
23. Merleau-Ponty M. Phenomenology of perception. Landes DA, translator. London: Routledge; 1945.
24. McCarthy J, Hayes PJ. Some philosophical problems from the standpoint of artificial intelligence. Mach Intell. 1969;4:463-502.
25. Di Paolo EA, Rohde M, De Jaegher H. Horizons for the enactive mind. In: Stewart J, Gapenne O, Di Paolo EA, editors. Enaction: toward a new paradigm for cognitive science. Cambridge (MA): MIT Press; 2010. p. 33-87.

26. Pani AK, Bhattacharjee GP. Temporal representation and reasoning in artificial intelligence: a review. Math Comput Model. 2001;34:55-80.
27. Allen JF. Towards a general theory of action and time. Artif Intell. 1984;23:123-54.
28. Bhatt M, Guesgen H, Wolfl S, Hazarika S. Qualitative spatial and temporal reasoning: emerging applications, trends, and directions. Spat Cogn Comput. 2011;11(1):1-14.
29. Sioutis M, Long Z. Hybrid AI systems grounded on qualitative spatio-temporal reasoning. In: Proceedings of the 12th Hellenic Conference on Artificial Intelligence (SETN). New York: ACM; 2022.
30. Kasabov NK. Time-space, spiking neural networks and brain-inspired artificial intelligence. Berlin: Springer; 2019.
31. Welch P. The use of fast Fourier transform for the estimation of power spectra. IEEE Trans Audio Electroacoust. 1967;15(2):70-3.
32. Cho K, van Merrienboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: Proceedings of EMNLP. 2014. p. 1724-34.
33. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell. 2013;35(8):1798-828.
34. Kirschbaum C, Pirke KM, Hellhammer DH. The 'Trier Social Stress Test': a tool for investigating psychobiological stress responses. Neuropsychobiology. 1993;28(1-2):76-81.
35. Esterman M, Tamber-Rosenau BJ, Chiu YC, Yantis S. Avoiding non-independence in fMRI data analysis: leave one subject out. NeuroImage. 2010;50(2):572-6.
36. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems. 2017. p. 3146-54.
37. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD. New York: ACM; 2016. p. 785-94.
38. Russell JA. A circumplex model of affect. J Pers Soc Psychol. 1980;39(6):1161-78.
39. Fonseca EMdS, Guido RC, Junior SB. An emotion recognition approach through EEG signals. Biomed Signal Process Control. 2021;64:102267.
40. Sutton RS, Barto AG. Reinforcement learning: an introduction. 2nd ed. Cambridge (MA): MIT Press; 2018.
41. Hospedales T, Antoniou A, Micaelli P, Storkey A. Meta-learning in neural networks: a survey. IEEE Trans Pattern Anal Mach Intell. 2022;44(9):5149-69.

Figures

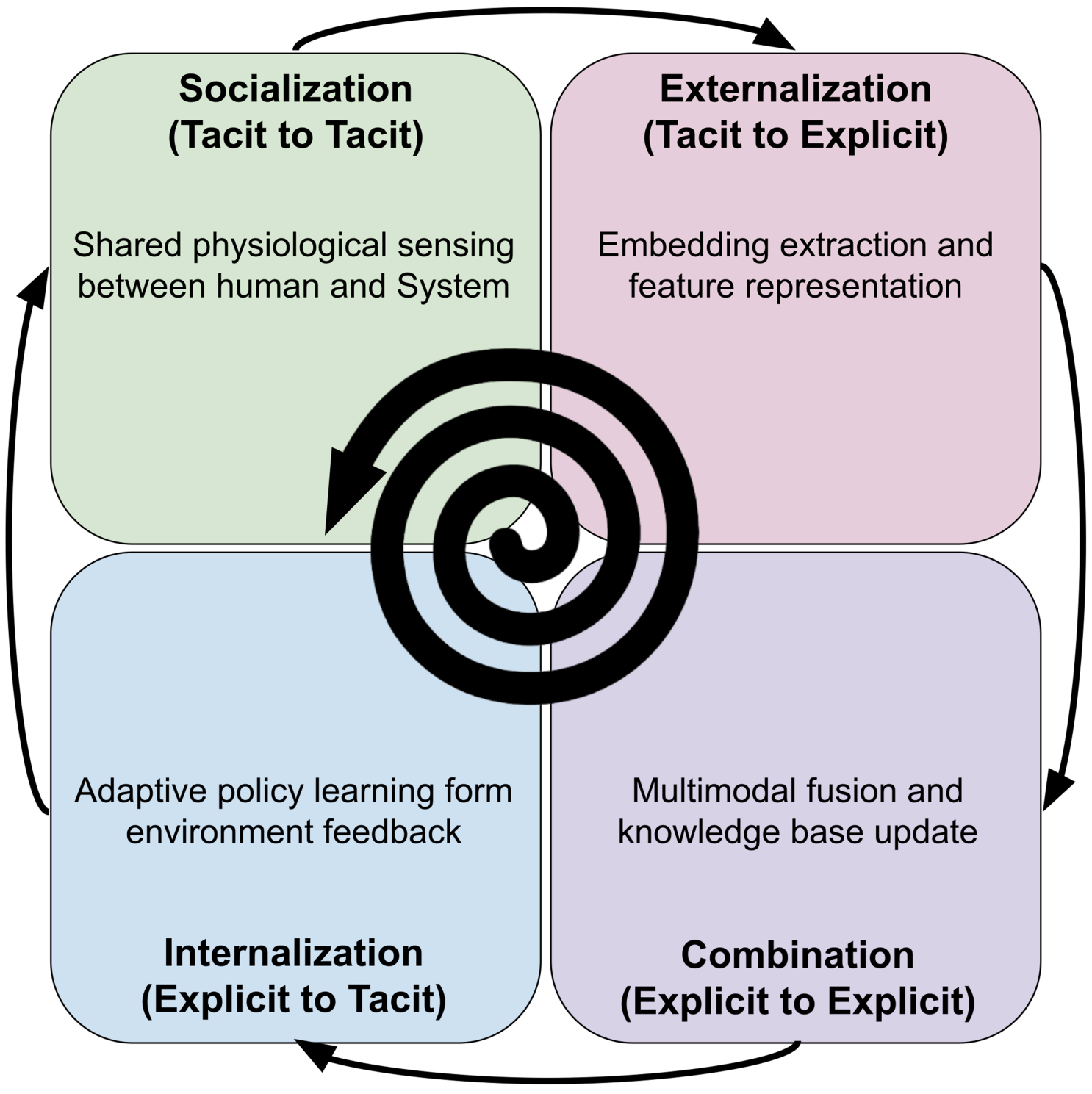


Figure 1

SECI knowledge-creation model adapted for enactive AI systems. Each quadrant maps a SECI phase to its corresponding system component.

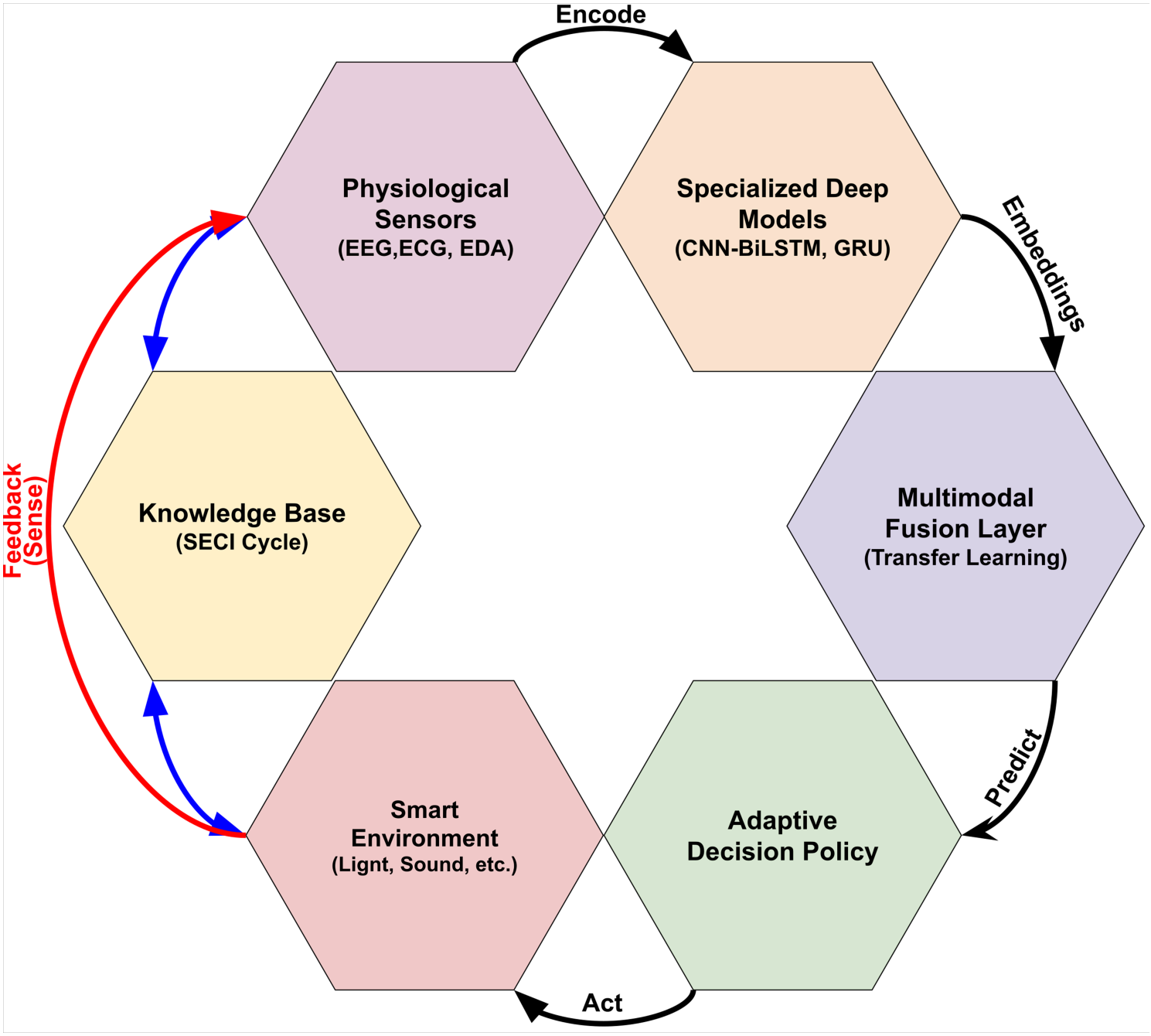


Figure 2

Enactive knowledge-based architecture for physiological-state estimation. Arrows indicate data flow; the feedback loop implements the enactive perception-action-adaptation cycle.

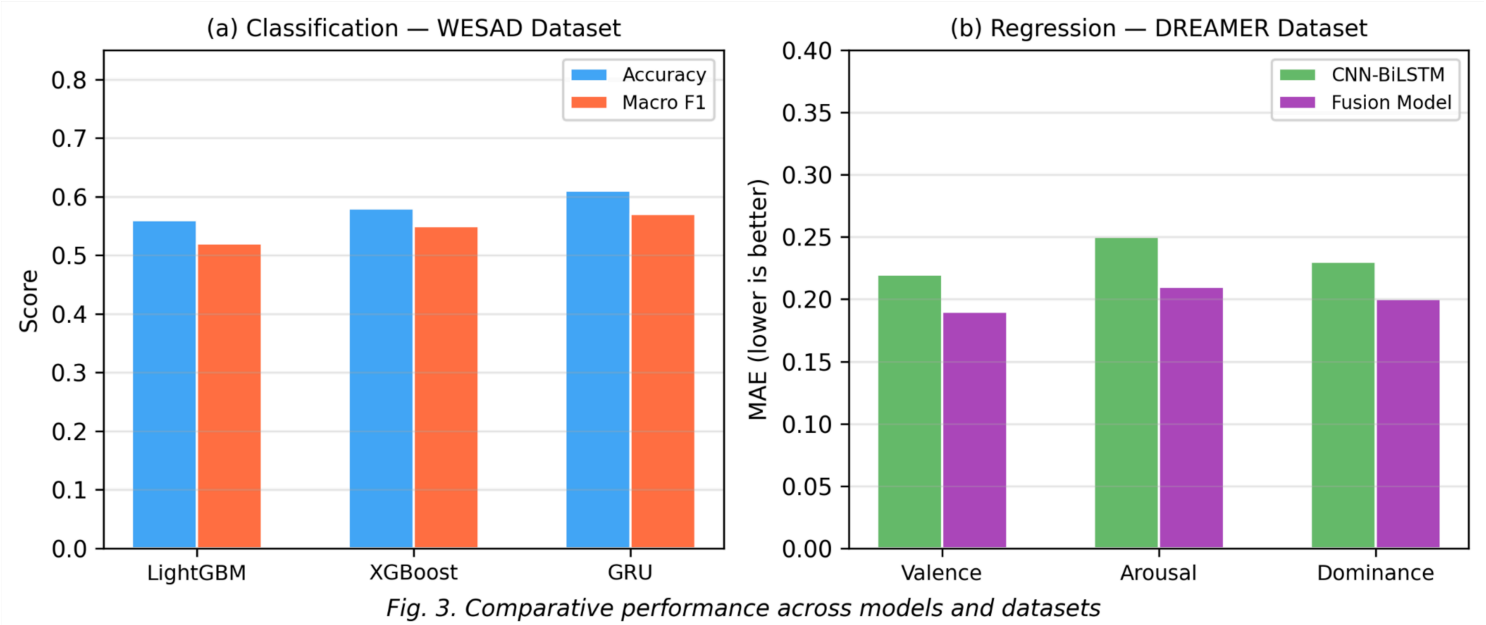


Fig. 3. Comparative performance across models and datasets

Figure 3

Comparative performance: (a) classification on WESAD (four-class); (b) regression MAE on DREAMER by emotional dimension.

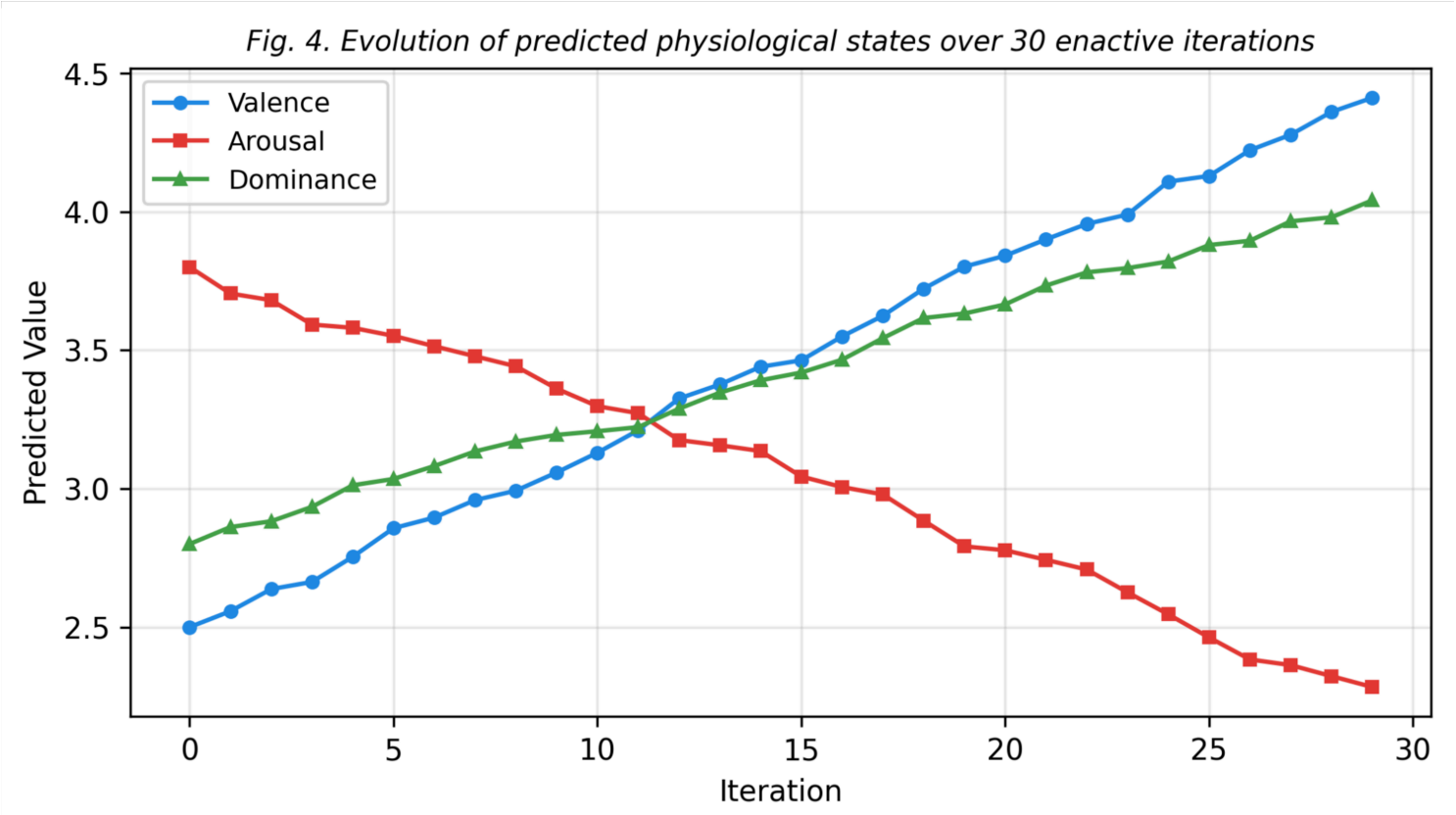


Figure 4

Evolution of predicted physiological states over 30 enactive iterations. The system progressively increases valence and reduces arousal.