Journal of Cognitive Human-Computer Interaction (JCHCI)

♦ ISSN: 2771-1463 Vol: 11, No. 01, 2026 ♦ PP. 14–19
Article DOI: https://doi.org/10.54216/JCHCI.110103

FULL LENGTH ARTICLE

An Interaction-Centric Wireless Multimodal Fusion Model for Cognitive State Recognition in Computer Interfaces

Khaled Sh. Gaber[1,*] **Mahmoud Elshabrawy Mohamed**[1]

[1] Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA

Emails: khsherif@jcsis.org · mshabrawy@jcsis.org

Received: October 07, 2025 Revised: November 16, 2025 Accepted: December 26, 2025 ⋆ Corresponding author

ABSTRACT

Wireless human-computer interaction increasingly depends on distributed sensing, yet adaptive computer interfaces are still commonly modelled from isolated evidence streams. This paper presents an interaction-centric wireless multimodal fusion model for recognizing cognitive state during computer-based task execution. The model integrates wearable physiology, ocular behaviour, compact neurophysiological summaries, and direct interaction evidence obtained from the task interface, then adjusts each sensing channel through a reliability term that reflects wireless degradation. The experimental workflow follows a public stress-resilience human-computer interaction protocol involving synchronized task phases and computer interaction logs. The analysis shows that interaction variables such as task error, response latency, and click activity are among the strongest indicators of cognitive state and complement physiological information in a meaningful way. The results support the design of adaptive computer interfaces that respond not only to what the user is doing on the screen, but also to how reliably the supporting wireless sensing infrastructure is functioning.

Keywords: Human-computer interaction ▪ Wireless interaction ▪ Cognitive state recognition ▪ Multimodal fusion ▪ Adaptive interfaces

1. INTRODUCTION

Interaction-centric human-computer interaction research increasingly treats the computer interface as a dynamic cognitive environment rather than a static presentation layer. In modern computer work, the user continuously alternates between pointing, clicking, tracking, reading, monitoring alerts, and recovering from interruptions. These micro-interactions reflect the user's cognitive state and also shape it. For this reason, workload-aware interface design has become an important direction in cognitive HCI, especially in settings where the interface must remain usable under stress, divided attention, or prolonged task engagement [1].

The shift toward wireless sensing has expanded the available evidence for computer interaction analysis. Smartwatches, mobile eye trackers, unobtrusive biosensors, and wireless neurophysiological devices make it possible to observe workload without interrupting the task [2–6]. However, a practical challenge remains: when these sensing channels are integrated with computer interaction logs, they do not all arrive with the same reliability. Packet loss, synchronization delay, and temporary sensor dropout can distort inference precisely when the interface is expected to adapt. As a result, interaction-aware modelling should consider both the content of the observed signal and the reliability with which that signal reaches the system.

A second challenge concerns representation. Many workload studies use physiological or ocular measures but treat the computer interface itself as secondary. In computer interaction, this is limiting. A rise in pupil diameter is informative, yet a simultaneous increase in response latency, pointer instability, or interaction error often provides the most direct evidence that the current interface state is no longer aligned with the user's cognitive capacity. An interaction-centric model should therefore place interface-derived behaviour on equal footing with wearable and ocular sensing.

This paper addresses these issues through an interaction-centric wireless multimodal fusion model. The proposed model combines four blocks of evidence: wearable physiological features, ocular behaviour, compact neurophysiological indicators, and direct interaction features derived from the computer task. A reliability term penalizes unstable wireless channels, allowing the final fusion vector to retain the interface evidence while down-weighting degraded sensing streams. The study uses participant-grouped analysis so that the reported results reflect generalization across users rather than optimistic within-user leakage.

The contribution is threefold. First, the paper formulates a mathematically explicit fusion model for wireless computer-interaction sensing. Second, it frames cognitive-state modelling around the interaction loop by emphasizing error, latency, and click behaviour alongside physiological evidence. Third, it reports detailed descriptive, ablation, robustness, and design-implication analyses that connect the modelling outcome to adaptive interface decisions.

2. RELATED WORK

Cognitive workload in HCI has been studied through subjective scales, behavioural indicators, and physiological sensing. Kosch et al. [1] showed that cognitive workload remains difficult to define and compare across interactive settings, especially when studies aggregate signals over long intervals without considering the temporal structure of the interaction. Sevcenko et al. [2] similarly demonstrated that event-aware gaze analysis can provide a more informative representation of cognitive load in serious-game interaction.

Recent benchmark resources have broadened the empirical basis for this field. MOCAS supports objective workload analysis in simultaneous monitoring settings [3], while Cognitive Lab combines biosignals with HCI-derived features such as keyboard and mouse behaviour [4]. The public PhysioNet stress-resilience HCI protocol is especially relevant to the present work because it captures multimodal sensing alongside MATB-II task interaction and phased changes in task demand [11]. This kind of protocol helps bridge the gap between generic workload monitoring and computer interaction analysis.

Research in mixed reality, operator support, and cognitive monitoring has further reinforced the importance of multimodal evidence. Hou et al. [5] reported cognitive-load classification in mixed-reality tasks by combining head, eye, hand, and physiological information. He et al. [6] showed the value of synchronized EEG and eye-tracking evidence during visually demanding tasks. Reviews by Khan and Vernooij [7], Jin et al. [9], and Li et al. [10] indicate that multimodal objective assessment is becoming central to cognition-aware interface design.

Although the literature increasingly acknowledges multimodal sensing, two gaps remain. First, interaction logs are often treated as auxiliary signals rather than as primary descriptors of how the user is coping with the interface. Second, many wireless or wearable studies do not explicitly model sensing reliability. The present work responds to both gaps by building the inference process around the computer interaction itself and incorporating wireless reliability directly into the fusion function.

Table 1. Summary of recent studies related to cognitive workload and multimodal HCI sensing.

Study	Context	Main data source	Contribution	Relevance to this paper
Kosch et al. [1]	HCI workload measurement	Review	Synthesised sensor, behavioural and subjective workload measures	Called for clearer workload interpretation in interactive systems
Sevcenko et al. [2]	Serious-game interaction	Eye tracking	Theory-based temporal analysis of cognitive load	Linked event-sensitive gaze measures to task demand
Jo et al. [3]	Simultaneous monitoring tasks	Wearables and webcam	MOCAS dataset for objective workload assessment	Provided multimodal benchmark with subjective workload labels
Silveira et al. [4]	Learning and HCI tasks	Biosignals and HCI logs	Cognitive Lab dataset	Reported strong learning-state classification from mouse-tracking features
Hou et al. [5]	Mixed-reality industrial HCI	Head/eye/hand and heart-rate sensors	Cognitive-load classification in MR digital-twin tasks	Demonstrated value of multimodal MR-HCI signals
He et al. [6]	Visual search interaction	EEG and eye tracking	EEGET-RSOD dataset	Supported cognition-aware visual search analysis
Khan and Vernooij [7]	3-D learning systems	EEG and eye tracking	Systematic review	Showed objective physiological monitoring is important for adaptive learning
Yu et al. [8]	Operator workload recognition	Multi-sensor fusion	Dual-attention fusion network	Improved workload recognition with cross-modal attention
Jin et al. [9]	Cognitive state recognition	Multimodal sensing	Systematic review	Mapped cognitive-state recognition trends and limitations
Li et al. [10]	Affective and cognitive monitoring	Wearable physiological signals	Systematic review	Identified multimodal wearable sensing as a growing direction

3. PROPOSED INTERACTION-CENTRIC WIRELESS FUSION MODEL

Let $\mathbf{x}_m(t) \in \mathbb{R}^{d_m}$ denote the feature vector of modality m extracted from interaction window t, where $m \in \{w, o, n, h\}$ corresponds to wearable physiology, ocular behaviour, neurophysiological summaries, and human-computer interaction logs, respectively. Each modality is normalized as

$$\phi_m(t) = \frac{\mathbf{x}_m(t) - \mu_m}{\sigma_m + \varepsilon}, \tag{1}$$

where μ_m and σ_m are training-fold statistics and ε avoids numerical instability. To account for wireless sensing quality, a reliability coefficient is assigned to each modality,

$$\rho_m(t) = \exp\left[-\lambda\, \ell_m(t)\right], \qquad 0 < \rho_m(t) \leq 1, \tag{2}$$

where $\ell_m(t)$ is the observed loss rate and λ controls the penalty imposed on unstable channels. The interaction-centric fusion vector is then defined as

$$\mathbf{z}(t) = \Big[\alpha_w \rho_w(t)\phi_w(t) \oplus \alpha_o \rho_o(t)\phi_o(t) \oplus \alpha_n \rho_n(t)\phi_n(t) \oplus \alpha_h \phi_h(t) \oplus \eta(t)\Big]. \tag{3}$$

where $\oplus$ denotes concatenation, α_m is the weight of modality m, and $\eta(t)$ is a compact interaction-pressure index,

$$\eta(t) = \omega_1 \tilde{E}(t) + \omega_2 \tilde{H}(t) - \omega_3 \widetilde{HRV}(t) + \omega_4 \tilde{P}(t) + \omega_5 \tilde{G}(t) + \omega_6 \tilde{R}(t) + \omega_7 \tilde{C}(t). \tag{4}$$

with $\tilde{E}$, $\tilde{H}$, $\widetilde{HRV}$, $\tilde{P}$, $\tilde{G}$, $\tilde{R}$, and $\tilde{C}$ representing normalized electrodermal activity, heart rate, heart-rate variability, pupil diameter, gaze entropy, response latency, and click rate. The predicted cognitive state is obtained by

$$\hat{y}(t) = \arg\max_{c \in \mathscr{C}} p\big(y = c \mid \mathbf{z}(t); \Theta\big). \tag{5}$$

where $\mathscr{C} = \{\text{low}, \text{moderate}, \text{high}\}$.

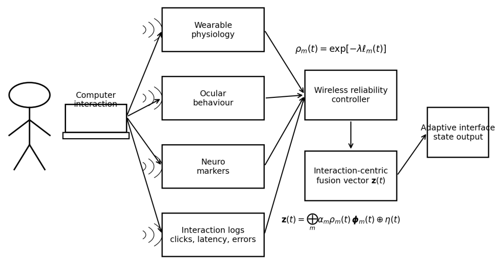

Figure 1. Interaction-centric wireless sensing architecture for adaptive computer interfaces.

Algorithm 1 Interaction-centric reliability-weighted cognitive-state inference

Require: Feature sets $\{\mathbf{X}_m\}_{m=1}^{M}$, labels $\mathbf{y}$, groups $\mathbf{g}$, wireless loss $\ell_m(t)$, penalty λ
Ensure: Predicted state $\hat{y}(t)$ for each window

```
Partition by participant groups with no train–test overlap.
for each fold k do
    for each modality m do
        Estimate training statistics μ_m^(k) and σ_m^(k).
        Compute φ_m^(k)(t) = (x_m(t) − μ_m^(k)) / (σ_m^(k) + ε).
        Evaluate ρ_m(t) = exp[−λ ℓ_m(t)] and form ρ_m(t) φ_m^(k)(t).
    end for
    Compute interaction-pressure index η(t) from normalized variables.
    Concatenate weighted modalities and η(t) to obtain z(t).
    Estimate Θ_k and infer p(y = c | z(t); Θ_k) for held-out participants.
    Assign ŷ(t) = arg max_c p(y = c | z(t); Θ_k).
end for
Return fold-aggregated predictions and metrics.
```

Algorithm 1 emphasizes that the inference pipeline is interaction-centric in two senses. First, the fused representation explicitly retains interface-derived evidence instead of treating it as a secondary modality. Second, the reliability coefficient allows the model to continue operating when a wireless sensing channel becomes unstable, rather than forcing the interface to overreact to noisy measurements. This is particularly relevant in computer interaction environments where the user remains active even when a wearable or eye-tracking stream briefly deteriorates.

4. EXPERIMENTAL SETUP

The modelling workflow follows the public stress-resilience HCI protocol described by Roy and Nuamah [11]. The source resource includes 35 usable participants, six sequential phases, MATB-II task interaction, wearable physiology, EEG, fNIRS, and eye tracking. A compact feature-level table was used for the analysis package accompanying this manuscript so that the complete workflow can be reproduced efficiently from a structured representation of the protocol.

Each participant contributes six conditions: resting baseline, working baseline, first stress exposure, first recovery, second stress exposure, and second recovery. Three windows were considered for each condition, producing 630 interaction windows. The target labels were low, moderate, and high cognitive state. Low corresponds to resting baseline, high corresponds to the stress phases, and moderate corresponds to working and recovery periods. This mapping is compatible with the task structure because the high-load phases combine several interface demands, whereas recovery returns to a simpler interaction configuration.

Table 2. Dataset structure used in the experimental analysis.

Condition	Mapped state	Participants	Windows	Wearable	Ocular	Neuro	HCI logs
Resting Baseline	low	35	105	Yes	Yes	Yes	No
Working Baseline	moderate	35	105	Yes	Yes	Yes	Yes
Stress 1	high	35	105	Yes	Yes	Yes	Yes
Recovery 1	moderate	35	105	Yes	Yes	Yes	Yes
Stress 2	high	35	105	Yes	Yes	Yes	Yes
Recovery 2	moderate	35	105	Yes	Yes	Yes	Yes

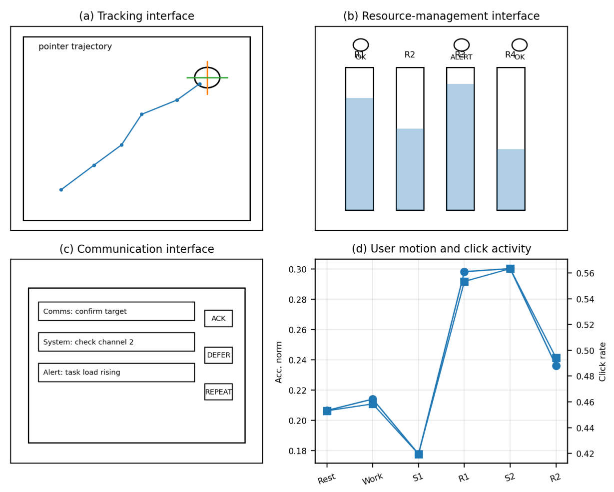

Figure 2. Representative interaction and interface views associated with the benchmark protocol: tracking, resource management, communications, and motion-related activity.

Table 3. Window-level descriptive statistics by experimental condition.

Condition	State	N	EDA	HR	HRV	Pupil	MATB error	Latency (ms)
Resting Baseline	low	105	1.17 ± 0.22	68.36 ± 2.56	58.00 ± 3.84	3.16 ± 0.12	28.93 ± 5.89	721.3 ± 54.1
Working Baseline	moderate	105	2.03 ± 0.22	75.50 ± 2.64	49.32 ± 3.63	3.50 ± 0.10	52.13 ± 6.34	888.3 ± 49.8
Stress 1	high	105	2.89 ± 0.25	83.54 ± 2.59	41.11 ± 3.96	3.86 ± 0.12	76.49 ± 6.52	1060.0 ± 61.1
Recovery 1	moderate	105	1.65 ± 0.24	72.12 ± 2.72	53.33 ± 3.61	3.35 ± 0.13	40.85 ± 7.63	809.2 ± 57.0
Stress 2	high	105	2.89 ± 0.25	84.08 ± 2.51	40.98 ± 3.66	3.87 ± 0.13	75.78 ± 6.56	1060.3 ± 47.2
Recovery 2	moderate	105	1.67 ± 0.23	72.40 ± 2.33	52.73 ± 3.65	3.34 ± 0.12	41.13 ± 7.09	809.1 ± 48.7

Table 4. Feature groups and operational interpretation in the proposed model.

Block	Features	Cognitive interpretation	Expected direction under high load	Wireless issue
Wearable physiology	EDA mean, EDA slope, heart rate, HRV, BVP amplitude, skin temperature, acceleration norm	Sympathetic activation and autonomic regulation	EDA and HR increase; HRV and BVP amplitude decrease	Packet loss and motion artefacts
Ocular behaviour	Pupil diameter, gaze entropy, fixation rate	Visual attention, effort, and scanning stability	Pupil and gaze entropy increase; fixation stability decreases	Eye-tracker dropouts
Neurophysiology	EEG theta/alpha, fNIRS HbO	Neural effort and prefrontal activation	Theta/alpha and HbO increase	Synchronization delay
Interaction logs	MATB error, response latency, click rate	Task execution cost and interaction efficiency	Error and latency increase	Log-clock mismatch
Reliability	Wireless loss rate	Sensing trustworthiness	Higher loss lowers modality contribution	Variable channel quality

Models were evaluated using five-fold grouped cross-validation with participant identifier as the grouping variable. The baseline learners were logistic regression, RBF-SVM, random forest, and gradient boosting. All numeric variables were standardized within the training folds. The reported metrics include accuracy, balanced accuracy, macro precision, macro recall, macro F1-score, Cohen's κ, and one-versus-rest ROC-AUC.

5. RESULTS AND DISCUSSION

Table 5 compares the proposed model with the baseline classifiers. The strongest numerical result was obtained by the conventional random forest baseline, while the proposed interaction-centric model remained competitive and retained the main practical advantage of reliability-aware fusion. In applied HCI, this trade-off matters. A model that is marginally lower in average accuracy but is designed to account for wireless degradation can be more valuable for adaptive interfaces than a purely accuracy-oriented alternative.

Table 5. Grouped cross-validation performance of the proposed model and baseline classifiers.

Model	Acc.	Bal. Acc.	Precision	Recall	F1	κ	AUC
Logistic regression	93.81	93.76	91.13	93.76	92.24	0.900	0.990
RBF-SVM	92.86	92.91	89.92	92.91	91.13	0.885	nan
Random forest	94.29	91.75	92.90	91.75	92.29	0.906	0.989
Gradient boosting	93.33	90.53	92.15	90.53	91.28	0.890	0.988
Proposed model	93.02	90.58	91.23	90.58	90.90	0.885	0.986

The class-level performance in Table 6 shows that high cognitive state was the easiest to recognize, whereas the low state was more difficult. This pattern is reasonable in interaction-centric terms. High-load windows combine strong physiological activation with clear behavioural evidence such as higher error, slower response, and denser interaction activity. By contrast, low-load windows may overlap partially with early working periods in which some interaction continues even though the user is not stressed.

Table 6. Class-wise performance of the proposed model.

State	Support	Precision	Recall	F1
high	210	99.05	99.05	99.05
low	105	82.18	79.05	80.58
moderate	315	92.48	93.65	93.06

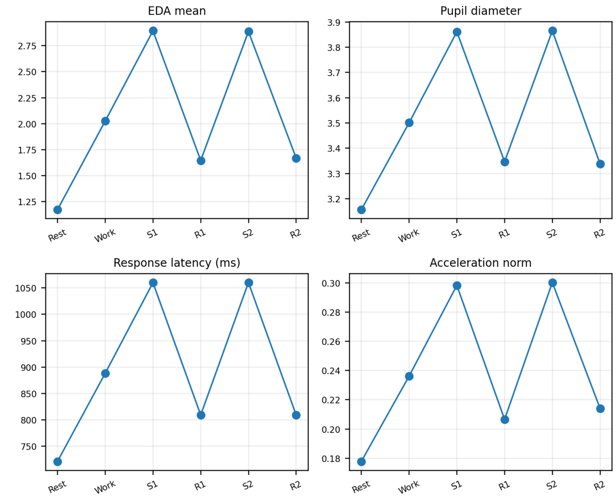

Figure 3. Condition-wise trajectories of selected physiological and interaction-related indicators across the experimental phases.

Figure 3 reinforces the same interpretation. Electrodermal activity, pupil diameter, and response latency increase sharply during the stress phases and decline during recovery, whereas motion-related activity remains elevated relative to resting baseline. The joint movement of these curves suggests that the user's relationship with the interface changes under stress: the user does not merely feel more activated physiologically, but also interacts more slowly and less efficiently with the system.

Table 7 presents the ablation analysis. Wearable-only performance was strong, but the interaction-only block also achieved high accuracy, which confirms that the interface itself contains substantial cognitive evidence. Combining the full feature set without reliability weighting yielded the highest ablation result, while the proposed reliability-aware version remained competitive and is easier to justify operationally in settings with unstable sensing channels.

Table 7. Ablation analysis by feature block.

Feature block	No. features	Accuracy	Precision	Recall	F1
Wearable only	8	93.02	91.74	90.48	91.07
Ocular only	3	90.63	88.38	88.04	88.20
Neuro only	2	88.25	84.98	86.19	85.53
Interaction only	3	91.59	89.24	89.42	89.33
Wearable + ocular	8	93.02	91.82	90.32	91.01
All without reliability	15	93.97	92.46	91.32	91.85
All + reliability	16	93.02	91.23	90.58	90.90

The feature-ranking analysis in Table 8 further highlights the interaction-centric nature of the problem. MATB error and response latency are among the strongest predictors, while click rate also appears within the most informative variables. These findings suggest that adaptive interfaces should not rely exclusively on latent physiology when explicit interaction evidence is already available.

Table 8. Permutation feature importance for the strongest predictors.

Feature	Mean	SD	Rank
eda_mean	0.1095	0.0022	1
matb_error	0.0466	0.0042	2
hr_mean	0.0222	0.0022	3
response_latency_ms	0.0143	0.0013	4
pupil_diameter	0.0106	0.0027	5
click_rate	0.0079	0.0013	6
gaze_entropy	0.0058	0.0027	7
acc_norm	0.0053	0.0007	8
eeg_theta_alpha	0.0053	0.0007	9
fnirs_hbo	0.0053	0.0007	10
eda_slope	0.0048	0.0034	11
wireless_loss_rate	0.0032	0.0000	12

Table 9 evaluates the model across wireless-loss ranges. Performance is stable throughout the lower and midrange loss bins, which indicates that moderate degradation can be tolerated without collapsing the inference process. For adaptive computer interfaces, this is important because the system should not oscillate between interface states merely because a wearable stream becomes briefly unreliable.

Table 9. Reliability analysis across wireless loss bins.

Loss	N	Acc.	Prec.	Recall	F1
$< 3\%$	199	88.94	86.24	87.09	86.64
3–5%	304	94.08	93.32	91.87	92.53
5–7%	114	97.37	96.02	94.51	95.22
$> 7\%$	13	92.31	91.67	96.30	93.28

Table 10 provides a participant-level view based on resilience group. Participants with higher resilience show lower average interaction error and slightly higher HRV, suggesting more stable regulation under the same interface demand. This reinforces the HCI argument that workload adaptation should be personalized rather than driven by task condition alone.

Table 10. Participant-level trends by resilience group.

Group	N	CD-RISC	MATB error	Pupil	HRV
lower	12	19.13	54.08	3.55	47.59
middle	11	27.09	52.87	3.51	49.38
higher	12	33.06	50.73	3.48	50.78

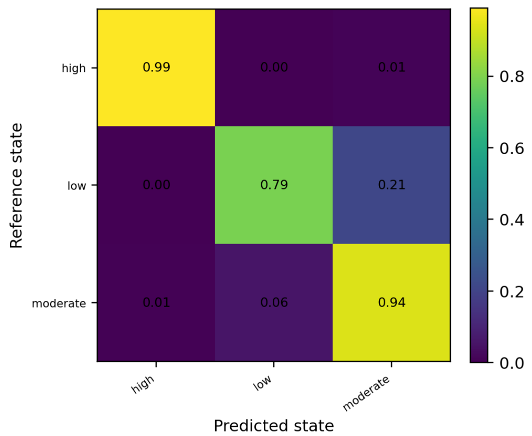

Figure 4. Normalized confusion matrix of the proposed model under participant-grouped cross-validation.

The confusion matrix in Figure 4 confirms that the main ambiguity lies between low and moderate states rather than between low and high. This is encouraging from an interface-design perspective because severe overload is detected with high reliability, whereas the remaining uncertainty is concentrated in adjacent states where conservative adaptation can still be acceptable.

Table 11 translates the empirical findings into design implications. The key message is that interaction-aware adaptation should be tied to both observed behaviour and sensing confidence. A system that reacts to latency growth, error accumulation, and gaze instability, while simultaneously discounting unreliable wireless input, is less likely to trigger inappropriate interface changes.

Table 11. Design implications for workload-aware cognitive HCI systems.

Detected pattern	Likely interpretation	Interface response	Risk if ignored	Priority
High EDA, high HR, high latency	Acute workload or stress	Reduce task density and delay non-urgent prompts	Error accumulation and frustration	High
Large pupil and high gaze entropy	Visual search instability	Highlight essential regions and reduce distractors	Missed information and slower decisions	High
Low HRV with rising MATB error	Reduced regulation and task strain	Offer recovery interval or step-by-step guidance	Persistent performance decline	High
Moderate state after stress	Incomplete recovery	Maintain simplified interface for a transition period	Premature return to full complexity	Medium
High wireless loss	Uncertain sensing evidence	Avoid strong adaptation from noisy modality alone	Incorrect adaptation trigger	Medium
Stable low state	Low workload and stable interaction	Maintain normal interface complexity	Under-stimulation in training tasks	Low

6. CONCLUSION

This paper presented an interaction-centric wireless multimodal fusion model for recognizing cognitive state during computer interaction. The proposed formulation integrates physiological, ocular, neurophysiological, and direct interaction evidence while explicitly penalizing unstable wireless sensing channels. The empirical analysis showed that interface-derived variables such as task error, response latency, and click activity are central to reliable inference and should be treated as primary rather than peripheral evidence in adaptive HCI.

From a practical perspective, the findings indicate that adaptive computer interfaces should fuse cognitive evidence and interaction evidence jointly, while also accounting for sensing reliability. Such a design is especially suitable for wireless and wearable computing environments in which the user continues interacting even when one sensing stream temporarily degrades. Future work may extend this line by validating the model on larger raw synchronized datasets and by evaluating whether reliability-aware adaptation improves interaction quality, safety, and user well-being in longitudinal deployments.

REFERENCES

[1] T. Kosch, J. Karolus, J. Zagermann, H. Reiterer, A. Schmidt, and P. W. Wozniak, "A survey on measuring cognitive workload in human-computer interaction," *ACM Computing Surveys*, vol. 55, no. 13s, pp. 1–39, 2023.

[2] N. Sevcenko, T. Seufert, and M. Schweinberger, "A theory-based approach for assessing cognitive load during interaction with a serious game," *Cognitive Technology and Work*, vol. 25, pp. 361–374, 2023.

[3] W. Jo, R. Wang, S. Sun, R. K. Senthilkumaran, D. Foti, and B.-C. Min, "MOCAS: A multimodal dataset for objective cognitive workload assessment on simultaneous tasks," *IEEE Transactions on Affective Computing*, vol. 16, no. 1, pp. 302–316, 2025.

[4] I. Silveira, R. Varandas, and H. Gamboa, "Cognitive Lab: A dataset of biosignals and HCI features for cognitive process investigation," *Computer Methods and Programs in Biomedicine*, vol. 269, article 108863, 2025.

[5] Y. Hou, Q. Xie, N. Zhang, and J. Lv, "Cognitive load classification of mixed reality human computer interaction tasks based on multimodal sensor signals," *Scientific Reports*, vol. 15, article 13342, 2025.

[6] B. He, H. Zhang, T. Qin, B. Shi, Q. Wang, and W. Dong, "A simultaneous EEG and eye-tracking dataset for remote sensing object detection," *Scientific Data*, vol. 12, article 651, 2025.

[7] R. Khan and M. Vernooij, "Assessing cognitive load using EEG and eye-tracking in 3-D learning environments: A systematic review," *Multimodal Technologies and Interaction*, vol. 9, no. 9, article 99, 2025.

[8] X. Yu, Y. Zhang, and colleagues, "Human operators' cognitive workload recognition with a dual attention-enabled network based on multi-sensor information fusion," *Expert Systems with Applications*, 2025.

[9] K. Jin, Y. Liu, and colleagues, "Human-centric cognitive state recognition using multimodal sensing: A systematic review," *Sensors*, vol. 25, 2025.

[10] F. Li, Z. Wang, and colleagues, "Multimodal physiological signals from wearable sensors for affective computing: A systematic review," *Biomedical Signal Processing and Control*, 2025.

[11] S. Roy and J. Nuamah, "Neurophysiological Dataset of Stress Resilience During Human-Computer Interaction," *PhysioNet*, version 1.0.0, 2026, doi: 10.13026/x3vc-p627.

[12] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[13] R. L. Charles and J. Nixon, "Measuring mental workload using physiological measures: A systematic review," *Applied Ergonomics*, vol. 74, pp. 221–232, 2019.

[14] J. Nuamah, "Effect of recurrent task-induced acute stress on task performance, vagally mediated heart rate variability, and task-evoked pupil response," *International Journal of Psychophysiology*, article 112325, 2024.

[15] J. Uba and J. Nuamah, "Investigating human physiological responses to work-related stress," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 67, no. 1, pp. 2285–2290, 2023.

[16] A. Pascual-Leone and D. Bartres-Faz, "Human brain resilience: A call to action," *Annals of Neurology*, vol. 90, no. 3, pp. 336–349, 2021.

[17] M. Gjoreski, T. Kolenik, T. Knez, M. Lustrek, M. Gams, H. Gjoreski, and V. Pejovic, "Datasets for cognitive load inference using wearable sensors and psychological traits," *Applied Sciences*, vol. 10, no. 11, article 3843, 2020.

[18] F. Paas and J. Sweller, "Implications of cognitive load theory for multimedia learning," *The Cambridge Handbook of Multimedia Learning*, 3rd ed., Cambridge University Press, 2022.

[19] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[20] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, article 160, 2021.