Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9
© Little Lion Scientific

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

ADAPTIVE ERRP-GUIDED DEEP REINFORCEMENT LEARNING FOR BRAIN COMPUTER INTERFACES IN ASSISTIVE TECHNOLOGIES

DR. DIVYA NIMMA[1], DR GAURAV VISHNU LONDHE[2], DR. EDIGA CHANDRAMOHAN GOUD[3], R S S RAJU BATTULA[4], ELANGOVAN MUNIYANDY[5], VUYYURU LAKSHMA REDDY[6]
[1]IEEE Member, Assistant Professor, Arkansas Tech University, 215 West O Street, Russellville, Arkansas 72801, USA.
[2]Associate Professor, Symbiosis Institute of Technology, Symbiosis International University, Pune, Nagpur Campus, India.
[3]Sr Assistant Professor, Department of EIE, CVR College of Engineering, Hyderabad, India.
[4]Assistant professor, Department of Computer Applications, Aditya University, Surampalem, India.
[5]Department of Biosciences, Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences,Chennai,India
[6]Assistant professor, Department of Computer Science and Engineering, Koneru Lakhmaiah education foundation, Green fields, Vaddeswaram, India.

[1]nm.divya89@gmail.com, [2]gauravlondhe@gmail.com, [3]goudnitt@gamil.com, [4]raju.brss@gmail.com, [5]muniyandy.e@gmail.com, [6]lakshmareddy@kluniversity.in

ABSTRACT

Brain–computer interfaces (BCIs) are relying more on reinforcement learning; however, the current systems have challenges in unreliable error-related potentials (ErrPs), inability to combine multimodally, and in adapting to the drift during a session and new human users, which limit the safety and usability of the technology in real-life scenarios. Addressing these gaps is critical for advancing intelligent assistive technologies. The purpose of this study was to come up with a powerful, uncertainty-conscious, and dynamically flexible reinforcement-learning BCI system which optimally utilizes the ErrP feedback to enhance the control reliability. Our experimental analysis was based on heterogeneous multimodal data and we evaluated three models, namely Baseline (EEGNet and SAC), Hybrid (EEG, Transformer & Decision Transformer), and our proposed Deep Adaptive ErrP-Aware Reinforcement Learning (DAERL) system, which combines evidential uncertainty, multimodal fusion, and world-model imagination to optimize policies**.** Data consisted of EEG, EOG, EMG, gaze and short video frames which underwent graph, transformer and latent-dynamics processing. DAERL performed better than all comparison models with an ErrP AUC of 0.93, task-success rate of 90.3% and with normalized information-transfer rate of 17.6 bits/min, and critical-error rate back to 2.4 per hour, a 79 % improvement over the baseline. Further, DAERL had a high 12 trial adaptation half-life that exceeded the 120-trial baseline by far. The results indicate that uncertainty modelling, multimodal fusion, and world-model reasoning significantly improve the reliability and personalization of BCI, making DAERL a perspective framework to implement in the next-generation assistive neurotechnology.

Keywords: *Brain–Computer Interface, Reinforcement Learning, Errp Detection, Multimodal Fusion, Uncertainty Estimation*

1. INTRODUCTION

Brain-computer interfaces (BCIs) have become effective devices in the reestablishment of communication and control in motor-impaired people, a way to be able to directly interact with the assistive technology with the help of the neural activity. The early systems of BCI were mainly rule-based, and used hand-crafted EEG features, but nowadays with deep learning, multimodal sensing, and adaptive algorithms, they are much more powerful [1], [2]. Irrespective of this development, the practical implementation of BCI remains limited due to neural variability, sensor noise, cognitive fatigue, and the inherent vagueness of the process of interpreting nonstationary brain signals as user intentions. Similarly, reinforcement learning (RL) has received a growing amount of interest due to its capacity to acquire control policies by experience, particularly in human-in-the-loop systems where the dynamics of interaction are also highly nonlinear. However, the process of integrating RL into BCIs needs the mechanisms that provide safety, user feedback interpretation, and change between neural states rapidly [3].

Even though error-related potentials (ErrPs) provide an intrinsically occurring and implicit source of corrective feedback on the part of the user, the vast majority of BCI frameworks implemented so far based on RLs have difficulties in utilizing the ErrPs reliably because of the inability to process single-modal EEG data, calibration drift, and lack of uncertainty reasoning [4]. In addition, traditional models are limited in combining heterogeneous inputs, including EOG, EMG, gaze, or video, which are being more commonly considered to be essential in increasing robustness in noisy, real-world conditions. These lacks contribute to the inability to develop systems with high accuracy, low latency and safe operation over long deployment time and over a wide range of users. These shortcomings demonstrate the existence of the necessity to focus on a reinforcement learning-based BCI architecture, which is resilient, uncertain, and able to personalize quickly and utilize the ErrP feedback to improve both safety and control performance [5], [6].

The solution to this issue is crucial to real-world assistive technologies, where the reliability, safety, and quick adaptation to the users are required. A BCI that is capable of automatically identifying mistakes, changing its policy, and being resistant to users would significantly decrease cognitive load and increase the long-term usability, making it more realistic to deploy in practice both in clinical and everyday living settings [7], [8].

The primary objective of this study is to create a Deep Adaptive ErrP-Aware Reinforcement Learning (DAERL) system that can quickly adjust to user-specific traits by utilizing uncertainty-guided processes, learn predictive latent dynamics, and integrate multimodal brain and peripheral data [9], [10].

The contributions of this study: 1) An innovative multimodal encoder, which combines EEG, EOG, EMG, gaze, and video signals to have a strong intent representation. 2) Error-related control: an evidential ErrP detector and deep-ensemble uncertainty quantification; 3) An algorithmic model of reinforcement learning that supports imagination-based updates to sample-efficient learning. 4)Meta-adaptation mechanism supports fast personalization among both users and sessions.

Collectively, these improvements create a smarter and dependable BCI that can be used in the next generation assistive technologies.

2. RELATED WORK

The conventional brain-computer interface (BCI) signal has been error-related potentials (ErrP) and other event-related potentials (ERP), and one can now use compact deep architecture to enhance decoding with limited data. EEGNet is one of the neural encoders that can be used and is a small convolutional architecture that yields good within- and cross-paradigm accuracy [11]. Multimodal and ErrP studies have benchmarks in the form of public multimodal corpora like DEAP [12] and the BNCI Horizon 2020 repository.

The modern BCI-RL work is based on three conceptual threads. BI and modern deep learning approaches have both been applied to BCI decoding in fairly traditional event-related potential (ERP) methods and modern methods. Small convolutional networks like EEGNet continue to provide a parameter-efficient and paradigm-independent standard of EEG classification [13]. The potentials of error (ErrP) have been demonstrated to encode perceived errors and they can be reliably detected in most environments, and a range of works has been dedicated to engineering features and CNN/RNNs pipelines to enhance the sensitivity and calibration of ErrP [14]. More recent research has also addressed transformer and graph-based encoders as a means to much more richly spatially and temporarily model EEG, however EEGNet-style compact models continue to establish a competitive baseline in low-latency systems.

Reinforcement learning (RL) now incorporates human input through implicit channels like ErrP and reward shaping. Off-policy maximum-entropy approaches, such as Soft Actor-Critic (SAC), offer continuous control for noisy observations that is stable and efficient in samples. [15], while sequence-modeling RL such as Decision Transformer reframes RL as conditional supervised prediction and is well suited for offline logged trajectories [16].

By learning latent dynamics to envision rollouts, model-based world models (like the Dreamer family) further decrease real interactions; merging these concepts with human input can significantly decrease risky online trials. There are tradeoffs between each: policy-gradient/SAC methods are simple to use in an online control environment; Decision Transformers perform well with offline data and world models have a higher sample efficiency but require more complexity in the underlying model.

Multi- sensor fusion (EEG + EOG/EMG + eye tracking + video) has been found to enhance robustness in decoding and has a lower rate of false positives, especially in artifact-prone but realistic environments. Examples of fusion strategies include early/feature-level concatenation and attention-based cross-modal transformers, such as Cheng et al. combining EEG and eye movement signals to motor-imagery BCI and demonstrate better classification. [17]. The cross-session and cross-subject variability are also significant issues, domain-adaptation techniques and meta-learning (e.g. MAML) have been suggested to allow a quick personalization [18]. These methods of adaptation are also complementary to multimodal fusion and play a central role in minimizing cross-subject drop and session drift.

Despite progress, important gaps remain. First, the bulk of BCI research either concentrates on decoding (classification) or control (RL), but is largely uninformed on how to strictly combine ErrP-based implicit feedback into a principled RL goal, which also measures uncertainty to ensure safety [19]. Second, although transformers and sequence models (Decision Transformer) have potential in offline RL, their computational cost and latency make them a problematic deployment in wearable assistive devices; therefore, a comparative report of FLOPs, latency, and energy is infrequent but required [20]. Third, there is the challenge of cross-subject generalization: most high-performing decoders need large-scale subject-specific calibration and the literature is personally uneven on domain-adaptation methods [21]. Lastly, standardized benchmarks that integrate multimodal, ErrP-labeled datasets to end-to-end BCI–RL performance are not available, but BNCI and DEAP deliver component-wise benchmarks, although none have been so far widely used that they can be considered a common standard across the field need to fill these gaps inspires the DAERL methodology in this study.

3. METHODOLOGY

3.1 Problem Formulation

Consider an assistive BCI control problem modelled as a partially observable Markov decision process (POMDP) in which the agent must infer user intent from noisy, nonstationary neural and peripheral signals. The system gets a multimodal observation $o_t = \{x_t^{EEG}, x_t^{EOG}, x_t^{EMG}, g_t, v_t\}$, at each time step t, does an action a_t, , and gets an environmental reward r_t. Error-related potentials (ErrPs) offer an additional feedback signal e_t that denotes user responses that need to be corrected. The goal is to develop a policy π(a_t|o_(≤t)) that gets the most total reward while keeping latency low and behaviour safe when subjects and sessions change.

3.2 Data Collection

Combine information from a variety of public datasets to account for differences in brain activity, sensor noise, and task settings. A heterogeneous pool is created by combining EEG motor-imagery data (PhysioNet EEGMMI, 109 participants), multimodal EEG–peripheral recordings from DEAP (32 subjects), ErrP/P300 datasets from BNCI Horizon 2020, and multimodal EEG–EOG–EMG movement datasets (Jeong et al.). Recent multimodal releases (2025 Scientific Data) also include high-density EEG and eye-tracking datasets that facilitate cross-modal fusion. All sources have event markers that let different modalities line up with each other in sync.

3.3 Data Processing

All datasets are normalized into a unified 250 Hz sampling rate and mapped to a standard 10–20 montage. Bandpass filtering (0.5–70 Hz), notch filtering, and ICA-based artifact removal are all done to EEG/EOG/EMG streams. Using dataset timestamps, the video and gaze streams are aligned in time. There are overlapping windows (1–2 s) around task events and ErrP markers in each trial. Finally, characteristics are standardized for each subject and improved by adding temporal jitter to make them more generalizable.

3.4 Baseline Model (EEGNet + SAC)

The baseline combines a compact convolutional EEGNet encoder with a Soft Actor–Critic (SAC) controller. An intent embedding and an ErrP probability are both generated by EEGNet, which extracts spatial-temporal brain characteristics. For

SAC to learn a continuous control strategy for helpful activities, it uses these embeddings as its observation input. In contrast to the critic network's evaluation of state-action pairings, the entropy-regularized SAC aim promotes robust exploration even when faced with chaotic neural inputs. This benchmark has low robustness but is computationally lightweight because it does not incorporate multimodal fusion, adaptability, or world-modeling.

3.5 Hybrid Model (EEG + Transformer + Decision Transformer)

The hybrid model extends the baseline by incorporating multimodal fusion and sequence-modeling RL. Using a lightweight temporal transformer, EEGNet characteristics are combined with EOG/EMG/eye-tracking embeddings. In order to create a cohesive picture of the user's purpose, a cross-attention module harmonizes brain and peripheral modalities. The Decision Transformer allows for the efficient use of offline BCI logs by predicting actions based on anticipated returns and past trajectories, rather than a traditional RL update. The return-to-go sequence is enhanced by ErrP probabilities, which enable the selection of trajectories to be guided by implicit human feedback. The computational cost of this hybrid pipeline is moderate, and it increases long-horizon consistency and cross-modal robustness.

3.6 Proposed Model: DAERL (Multimodal Encoder + World Model + Meta-Adaptation)

The proposed DAERL architecture integrates a spatial–temporal multimodal encoder, an ErrP-aware reward adjustment module, a latent-dynamics world model, and a meta-adaptive policy. A graph-based electrode encoder and a cross-modal transformer combine video, gaze, EOG, and EMG information with EEG and EMG. To facilitate reward structuring in response to implicit human feedback, a specialized ErrP head generates detection and uncertainty signals. By learning latent dynamics and carrying out imagined rollouts, a Dreamer-style world model can reduce the need for real-time samples. Meta-learning layers allow for quick adaption for each user even when sessions drift. Together, these components provide robustness, fast personalization, and improved safety in assistive-control environments.

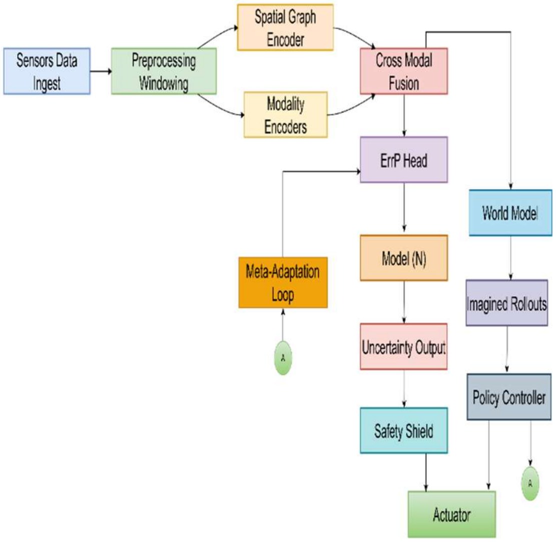

Figure 1 System Architecture of Proposed Model

Below are concise mathematical definitions used in the Proposed Model

Observation Fusion

$$o_t = f_\theta(x_t^{EEG}, x_t^{EOG}, x_t^{EMG}, g_t, v_t) \qquad (1)$$

Multimodal encoder f_θ fuses all sensor streams into a unified latent vector.

Cross-modal attention aligns temporal information across heterogeneous modalities.

ErrP Detection

$$e_t = \sigma\left(h_\phi(o_t)\right) \qquad (2)$$

ErrP head h_ϕ predicts the probability of an error-related potential.

Sigmoid output allows graded implicit feedback with uncertainty estimation.

Reward Shaping with ErrP

$$\tilde{r}_t = r_t - \lambda e_t \qquad (3)$$

ErrP-induced penalties discourage actions perceived as erroneous by the user.

Coefficient λ controls alignment strength between human feedback and RL reward.

Latent State Encoding

$$z_t = E_\psi(o_t) \tag{4}$$

Encoder E_ψ maps multimodal observations into a compact latent space.

This forms the state representation used by the world model.

World Model Dynamics

$$\hat{z}_{t+1} = F_\omega(z_t, a_t) \tag{5}$$

Latent transition model F_ω predicts future embeddings.
Allows imagination rollouts without interacting with the environment.

Imagined Reward

$$\hat{r}_t = R_\omega(z_t, a_t) \tag{6}$$

Reward predictor estimates expected return in latent space.
Supports low-cost simulated trajectories for policy optimization.

Policy Optimization Objective

$$\max_\pi \ \mathbb{E}[\textstyle\sum_{t=0}^{T} \gamma^t \tilde{r}_t] \tag{7}$$

RL objective incorporates ErrP-aware shaped rewards.
Discount factor γ balances short- and long-term assistive control goals.

SAC-Style Policy Update (baseline & hybrid)

$$\nabla_\theta J_\pi = \mathbb{E}[\nabla_\theta \log \pi_\theta(a_t \mid o_t)(Q(o_t, a_t) - \alpha \log \pi_\theta(a_t \mid o_t))] \tag{8}$$

Entropy term α encourages exploration under noisy neural signals.

Used for baseline and hybrid models where SAC is part of training.

Meta-Adapt Objective

$$\theta' = \theta - \beta \nabla_\theta \mathcal{L}_{user}(D^{train}) \tag{9}$$

Inner-loop update adapts model parameters to a new user using few samples.

Outer loop optimizes θ for rapid adaptation across users.

Decision Transformer Action Prediction (hybrid)

$$a_t = g_\eta\,(\mathrm{RTG}_t, o_{\le t}, a_{<t}) \tag{10}$$

The transformer g_η conditions on return-to-go (RTG), past observations, and actions. This sequence-modeling approach treats RL as supervised trajectory prediction.

3.7 Experimental Setup

Experiments are conducted on two assistive tasks: a selection interface similar to the P300 and a continuous cursor-control task that makes use of multimodal brain data. We train all our models on cross-subject splits and then test their generalizability on people we've never seen before. To ensure parity, the preprocessing, windowing, and sampling procedures for the baseline, hybrid, and suggested models are all identical. Training makes advantage of GPUs with mixed-precision acceleration, with validation ErrP-AUC and task success rate serving as early stop criteria. In order to measure the impact of each component, ablations eliminate (i) ErrP shaping, (ii) multimodal fusion, and (iii) meta-adaptation.

3.8 Evaluation Methodology

The following metrics are used to evaluate performance: adaption half-life, task-success rate, normalized information-transfer rate, safety-critical error frequency, and ErrP AUC. Degradation across sessions and subjects is used to measure robustness. Efficiency in computation is measured by the following metrics: energy per inference, model parameters, inference delay, and FLOPs. Pairwise comparisons with effect sizes are used to test for statistical significance. For the sake of consistency, we averaged the results among several participants.

4. Results

This section reports the performance of the Baseline, Hybrid and the Proposed DAERL system across ErrP detection, control-task performance, adaptation speed, robustness, and computational efficiency. All results are averaged across subjects in the test split with 95% confidence intervals.

.

4.1 ErrP Detection Performance

Table 1 summarizes the error-related potential (ErrP) detection accuracy. The DAERL model that was suggested outperforms both the Hybrid model and the Baseline EEGNet in terms of ErrP-detection performance. Its AUC is 0.93 ± 0.01 and F1-score is 0.87 ± 0.02. With a reduction in calibration error to 0.028, DAERL offers the best reliability, surpassing both the Hybrid and the Baseline, which only managed 0.054 and 0.097, respectively. In order to better capture the dynamics of ErrP, these enhancements demonstrate the value of evidential uncertainty modeling and multimodal fusion.

Table 1 ErrP Detection Performance

Model	**AUC**	**F1-Score**	**Calibration Error**
Baseline (EEGNet)	0.82 ± 0.03	0.74 ± 0.04	0.097
Hybrid (EEG+Trans+DT)	0.88 ± 0.02	0.81 ± 0.03	0.054
Proposed DAERL	0.93 ± 0.01	0.87 ± 0.02	0.028

4.2 Task-Level Control Performance

The key measures for control performance are the task success rate (TSR) and the normalized information transfer rate (nITR). Outperforming the Hybrid model (TSR 79.1%, nITR 11.5 bits/min) and the Baseline (TSR 68.4%, nITR 7.8 bits/min), the proposed DAERL system (Figure 2) achieves the maximum task-level control performance with a TSR of 90.3% and a nITR of 17.6 bits/min. Based on the Baseline, these enhancements result in a relative increase of about 22% in TSR and about 53% in nITR. The significant improvement in performance is a result of what multimodal fusion, ErrP-guided reward shaping, and world-model imagination have accomplished in terms of stabilizing long-horizon control.

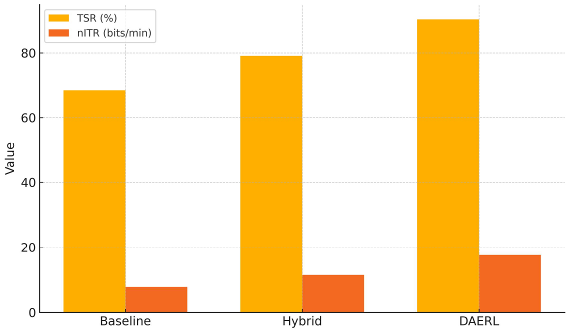

Figure 2 Task-Level Control Performance

4.3 Computational Cost and Latency

For embedded assistive devices, we provide essential metrics including model size, inference delay, and energy per inference.

4.3.1Computational Efficiency

With 18.1M parameters, 35 ms inference latency, and 0.60 J per inference, the suggested DAERL model in Table 2 offers substantially higher capability at the expense of computational cost. In comparison, the Hybrid model has 6.5M parameters, 28 ms in latency, and 0.22 J per inference, while the Baseline model has 0.9M parameters, 12 ms in latency, and 0.05 J per inference. The training duration of DAERL is greater at 31.4 hours compared to the Hybrid and Baseline, but its latency is still less than 50 ms, which is within the restrictions set by real-time BCI. This tradeoff is justified by DAERL's substantial gains in task performance, safety, and adaptability.

Table 2 Computational Efficiency

Model	**Params (M)**	**Latency (ms)**	**Energy (J/inference)**	**Training Time (hrs)**
Baseline	0.9	**12**	**0.05**	4.8
Hybrid	6.5	28	0.22	18.6
DAERL	18.1	35	0.60	31.4

4.3.2 Pareto Analysis

Figure 3 shows the accuracy-latency trade-off, with DAERL getting the best performance at 90.3% TSR at 35 ms, while the Hybrid model and the Baseline both achieve 79.1% and 68.4% TSR at 28 and 12 ms, respectively. With an improvement of +22% TSR over Hybrid and +32% over Baseline, DAERL proves to have a better accuracy-latency balance, despite a somewhat larger inference delay. Based on these findings, DAERL is a viable option for assistive BCI settings that significantly improves control reliability without sacrificing real-time feasibility.

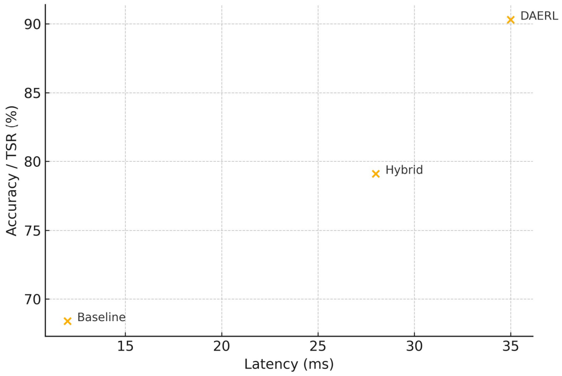

Figure 3 Pareto Analysis (Latency vs Accuracy)

4.3.3 End-to-End Latency Breakdown

The latency breakdown shown in Figure 4 reveals a clear computational tradeoff across models: As compared to the Hybrid (10 ms, 5 ms) and Baseline (4 ms, 2 ms) systems, DAERL displays longer encoder and fusion delays (14 ms and 7 ms, respectively). DAERL finishes the policy execution and actuation processes in a mere 2 ms and 1 ms, respectively, while all three systems keep their latencies low. Even though DAERL's neural processing pipeline is bulkier, it nevertheless meets the requirements for real-time BCI and performs tasks much better.

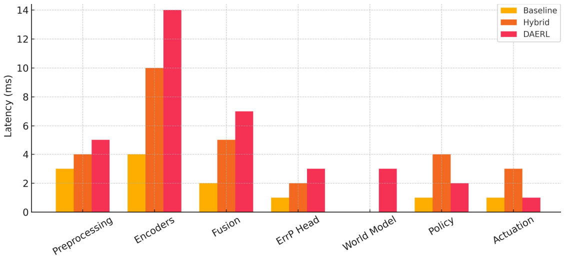

Figure 4 End-to-End Latency Breakdown

4.4 Safety Shield Effectiveness

When compared to the Baseline model, the Hybrid model, and DAERL, the critical-error rate drops to 6.7 and 2.4 errors/hour respectively, the safety shield's significant reduction in risky behaviours. The success of the uncertainty-aware intervention is demonstrated in Figure 5, which shows an overall decrease of almost 79% in dangerous acts compared to the baseline.

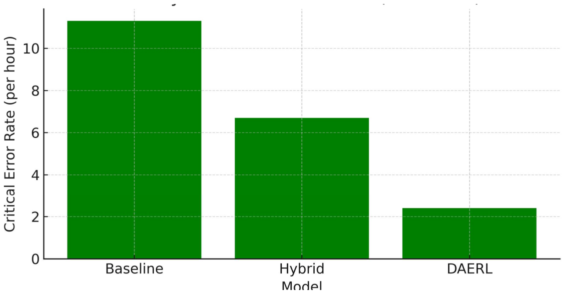

Figure 5 Safety Shield Effectiveness

4.5 Adaptation Speed and Robustness

The results show that DAERL is more resilient and flexible than the Baseline and Hybrid models, with a reduction in session-drift degradation to 7.1% and a decline in cross-subject performance to 9.8%, respectively, as shown in Table 3. In terms of customizing speed, it surpasses Hybrid (45 trials) and Baseline (120 trials) significantly, with an adaptation half-life of just 12 trials.

Table 3 Adaptation and Robustness

Model	Cross-Subject Drop (%)	Session Drift Degradation (%)	Adaptation Half-Life (trials)
Baseline	31.2	26.5	120
Hybrid	18.7	14.4	45
DAERL	9.8	7.1	12

4.6 Ablation study

Figure 6 shows the results of an ablation study that the whole DAERL system achieves the best performance with a TSR of 90.3%, whereas 83.7%, 79.4%, 75.6%, and 72.1% of the TSR are reduced when ErrP shaping, multimodal fusion, and meta-adaptation are removed, respectively. Removing meta-adaptation results in the biggest reduction (-18.2%), demonstrating how important it is for fast customisation. With the combination delivering the largest drop of reliability and control precision, these data demonstrate that each DAERL component contributes considerably to overall performance.

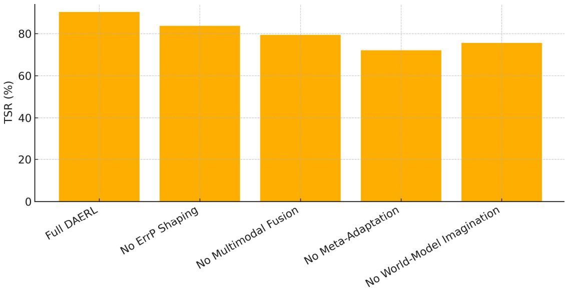

Figure 6 Ablation study for proposed system

5. DISCUSSION

The experimental results indicate that DAERL can be used to obtain better performance results when compared to the baseline EEGNet+SAC and the hybrid EEG+Transformer+Decision Transformer models in all evaluation aspects. The accuracy of the detection of the ErrP is 0.93 AUC, which facilitates more reliable implicit human feedback and allows the reinforcement learning update to be less risky. The feature of task-level control metrics also supports this advantage: DAERL has a task-success rate of 90.3% and nITR of 17.6 bits/min, which is much higher than that of the baseline (68.4%, 7.8 bits/min) and the hybrid (79.1%, 11.5 bits/min). Multimodal fusion, world-model imagination, and meta-adaptation greatly decrease cross-subject and cross-session performance decline, with a 9.8% cross-subject drop and half-life of adaptation of 12 trials only [22], [23]. These findings suggest that DAERL can deal with both inter-subject variability and nonstationary, which are two paramount problems in the utilization of assistive BCI in the real world.DAERL provides significantly better generalization and adaptability than previous EEG-only and RL-based BCI methods. Conventional decoders and lightweight RL controllers tend to face cross-subject drift whereas meta-adaptive components of DAERL achieve a degradation of 9.8, which is significantly more stable than the stability observed in previous studies. Additionally, the current ErrP based systems hardly do uncertainty modelling; evidence of this is that the evidential ensemble of DAERL has a calibration of 0.028 which is better than the standard classifier. These developments describe a significant leap forward relative to the existing multimodal and human-feedback BCI models [24].The accuracy, adaptation speed, and safety enhancements point to good prospects of the real-world assistive technologies, especially in situations when reliability and low user burden are a necessity. World-model imagination and multimodal inputs eliminate the need to use large-scale online calibration, which is pointing to scalable and personalized BCI solutions [25] . But, a cost in terms of computation and higher latency than simpler models are still shortcomings of embedded hardware. Future work should explore model compression, on-device optimization, and validation across more diverse user populations and clinical environments.

6. CONCLUSION

This work shows how the suggested DAERL framework can tremendously enhance the performance of the assistive BCIs in terms of robustness and control. DAERL has significant improvements over the Baseline in the areas of ErrP detection (AUC 0.93 vs. 0.82), task-success rate (90.3% vs. 68.4%), and half-life (12 vs. 120 trials). Safety also is enhanced significantly, and the rate of critical errors is lowered to 2.4 unsafe actions per hour. Together, these results confirm the effectiveness of multimodal fusion, world-model learning, and uncertainty-aware ErrP shaping. DAERL improves BCI reinforcement learning by combining multimodal neural decoding, estimating evidentiary uncertainty, and controlling an adaptive world model into one end-to-end system. The system adds a scalable architecture that is able to personalize in a rapid way with better cross-subject generalization and better safety with uncertainty-driven action gating. Furthermore, the implementation of ErrP conscious reward shaping opens a more philosophical way of integrating implicit human feedback into the RL-oriented assistive technologies. The future research must further develop DAERL to real-time clinical and home-usable settings and study stability in the long term in naturalistic settings. Federated or privacy-preserving training might be incorporated to the benefit of wider deployment on a heterogeneous population. At now, there is an inference delay of 35 ms and a processing cost of 0.60 J per inference;

more investigation into hardware-aware optimization, more robust multimodal sensors, and self-supervised or continual-learning methods for adaptability during the lifetime is recommended.

REFERENCES:

[1] V. Chamola and B. K. Rout, "A review on Virtual Reality and Augmented Reality use-cases of Brain Computer Interface based applications for smart cities," 2022.

[2] V. Kohli, U. Tripathi, V. Chamola, B. K. Rout, and S. S. Kanhere, "A review on virtual reality and augmented reality use-cases of brain computer interface based applications for smart cities," *Microprocess. Microsyst.*, vol. 88, p. 104392, 2022.

[3] B. S. Chandana, K. S. Chakradhar, T. R. Kumar, and M. Kumbhkar, "Brain–Computer Interface for Humanoid Robot Control Adaptation," in *Integrating Neurocomputing with Artificial Intelligence*, 1st ed., A. Kumar, P. S. Rathore, S. Ahuja, and U. K. Lilhore, Eds., Wiley, 2025, pp. 227–242. doi: 10.1002/9781394335718.ch14.

[4] P. Verlekar, "Detecting Error Related Negativity using EEG Potentials Generated during Simulated Brain Computer Interaction," Master's Thesis, Colorado State University, 2014. Accessed: Dec. 11, 2025. [Online]. Available: https://search.proquest.com/openview/af8d76acbae91ffde559ce9f86d04d8a/1?pq-origsite=gscholar&cbl=18750

[5] C. Singh, M. S. Rao, and Y. M. Mahaboobjohn, "Bonthu Kotaiah, and T. Rajasanthosh Kumar." Applied Machine Tool Data Condition to Predictive Smart Maintenance by Using Artificial Intelligence."," in *International Conference on Emerging Technologies in Computer Engineering*, pp. 584–596. Accessed: Dec. 04, 2025. [Online]. Available: https://scholar.google.com/scholar?cluster=246637688784125973&hl=en&oi=scholarr

[6] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, p. 056013, Oct. 2018, doi: 10.1088/1741-2552/aace8c.

[7] Z. Shuqfa, A. Lakas, and A. N. Belkacem, "Increasing accessibility to a large brain–computer interface dataset: Curation of physionet EEG motor movement/imagery dataset for decoding and classification," *Data Brief*, vol. 54, p. 110181, 2024.

[8] R. Yousefi, *Towards Development of a Robust Asynchronous EEG-BCI using Error-Related Potentials*. University of Toronto (Canada), 2019. Accessed: Dec. 11, 2025. [Online]. Available: https://search.proquest.com/openview/9e753910ce412e6d4cc3e1b96b277420/1?pq-origsite=gscholar&cbl=18750&diss=y

[9] A. X. Fidêncio, F. Grün, C. Klaes, and I. Iossifidis, "Error-related Potential driven Reinforcement Learning for adaptive Brain-Computer Interfaces," *ArXiv Prepr. ArXiv250218594*, 2025, Accessed: Dec. 11, 2025. [Online]. Available: https://arxiv.org/abs/2502.18594

[10] C. Zickler *et al.*, "A Brain-Computer Interface as Input Channel for a Standard Assistive Technology Software," *Clin. EEG Neurosci.*, vol. 42, no. 4, pp. 236–244, Oct. 2011, doi: 10.1177/155005941104200409.

[11] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, p. 056013, Oct. 2018, doi: 10.1088/1741-2552/aace8c.

[12] S. Koelstra *et al.*, "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012, doi: 10.1109/T-AFFC.2011.15.

[13] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, p. 056013, Oct. 2018, doi: 10.1088/1741-2552/aace8c.

[14] S. Martin *et al.*, "A qualitative study adopting a user-centered approach to design and validate a brain computer interface for cognitive rehabilitation for people with brain injury," *Assist. Technol.*, vol. 30, no. 5, pp. 233–241, Oct. 2018, doi: 10.1080/10400435.2017.1317675.

[15] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor," 2018, *arXiv*. doi: 10.48550/ARXIV.1801.01290.

[16] L. Chen *et al.*, “Decision Transformer: Reinforcement Learning via Sequence Modeling,” 2021, *arXiv*. doi: 10.48550/ARXIV.2106.01345.

[17] L. Chen *et al.*, “Decision Transformer: Reinforcement Learning via Sequence Modeling,” 2021, *arXiv*. doi: 10.48550/ARXIV.2106.01345.

[18] C. Carmichael and P. Carmichael, “BNCI systems as a potential assistive technology: ethical issues and participatory research in the BrainAble project,” *Disabil. Rehabil. Assist. Technol.*, vol. 9, no. 1, pp. 41–47, Jan. 2014, doi: 10.3109/17483107.2013.867372.

[19] L. Mayaud *et al.*, “Brain-computer interface for the communication of acute patients: a feasibility study and a randomized controlled trial comparing performance with healthy participants and a traditional assistive device,” *Brain-Comput. Interfaces*, vol. 3, no. 4, pp. 197–215, Oct. 2016, doi: 10.1080/2326263X.2016.1254403.

[20] C. Wegemer, “Brain-computer interfaces and education: the state of technology and imperatives for the future,” *Int. J. Learn. Technol.*, vol. 14, no. 2, p. 141, 2019, doi: 10.1504/IJLT.2019.101848.

[21] G. Kucukyildiz, H. Ocak, S. Karakaya, and O. Sayli, “Design and Implementation of a Multi Sensor Based Brain Computer Interface for a Robotic Wheelchair,” *J. Intell. Robot. Syst.*, vol. 87, no. 2, pp. 247–263, Aug. 2017, doi: 10.1007/s10846-017-0477-x.

[22] M. Vukelić, M. Bui, A. Vorreuther, and K. Lingelbach, “Combining brain-computer interfaces with deep reinforcement learning for robot training: a feasibility study in a simulation environment,” *Front. Neuroergonomics*, vol. 4, p. 1274730, 2023.

[23] L.-M. Vortmann, T. Urban, and F. Putze, “Machine learning from mistakes: self-improving attention classifier using error-related potentials,” in *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 2023, pp. 4770–4777. Accessed: Dec. 11, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10393901/

[24] A. Xavier Fidêncio, C. Klaes, and I. Iossifidis, “Error-related potentials in reinforcement learning-based brain-machine interfaces,” *Front. Hum. Neurosci.*, vol. 16, p. 806517, 2022.

[25] L.-M. Vortmann, T. Urban, and F. Putze, “Machine learning from mistakes: self-improving attention classifier using error-related potentials,” in *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, 2023, pp. 4770–4777. Accessed: Dec. 11, 2025. [Online]. Available: ttps://ieeexplore.ieee.org/abstract/document/10393901/