TRANSNAV
http://www.transnav.eu
the International Journal on Marine Navigation and Safety of Sea Transportation
Volume 19
Number 1
March 2025
DOI: 10.12716/1001.19.01.07

Seafarer Mental Workload Assessment Using a Hybrid Deep Learning Model

R. Jiang & S. Fan
The State Key Laboratory of Maritime Technology and Safety, Wuhan, China
Wuhan University of Technology, Wuhan, China

ABSTRACT: Human errors in maritime operations are closely linked to seafarers' mental workload; however, traditional assessment methods lack real-time neurocognitive resolution. This study introduces a novel psychophysiological framework that integrates electroencephalography (EEG) analysis with deep learning to objectively quantify seafarers' mental workload during onboard operations. A high-fidelity bridge simulator was utilized to generate critical maritime scenarios, including ship encounters, narrow channel navigation, poor visibility, and emergency responses. High-density EEG signals were analyzed to extract spectral features (Gamma, Beta, Alpha, Theta, Delta). A hybrid Convolutional Neural Network-Bidirectional Long Short-Term Memory (CNN-BiLSTM) model was proposed to classify workload states of seafarers, combining Convolutional Neural Network (CNN)-extracted frequency patterns with Bidirectional Long Short-Term Memory (Bi-LSTM)-captured temporal dynamics, which achieves 96% accuracy. Furthermore, SHAP interpretability analysis indicated that Theta and Alpha frequencies are key indicators in distinguishing between high and low workloads for seafarers. These results provide a quantitative tool for cognitive assessment of seafarers in maritime training and serve as a guideline for workload allocation in ship bridge teams for shipping companies and maritime authorities.

1 INTRODUCTION

Research in maritime safety has indicated that 75% to 90% of marine accidents are directly associated with human error [1]. In recent years, the maritime industry has witnessed a range of accidents with varying degrees of severity, including but not limited to collisions, groundings, explosions, and man-overboard incidents [2]. Substantiating this trend, the European Maritime Safety Agency (EMSA) reported that 80% of ship collisions and grounding incidents over the past decade were traceable to human-related causes [3]. These findings highlight the urgent need to address human cognitive and behavioral determinants as a way to reduce human error in offshore operations. Human factors such as skill gaps, stress accumulation, excessive mental workloads, and motivational failures can lead to human errors [4].

In recent years, international maritime organizations such as the International Seafarers' Welfare and Assistance Network (ISWAN) and the U.S. Department of Transportation's Maritime Administration (MARAD) have strongly advocated for the mental health and well-being of seafarers [5],[6]. These initiatives emphasize technological innovations and scientific assessments to enhance the work environment and cognitive health of seafarers. The complex marine work environment demands that crew members possess certain non-technical abilities to carry out tasks, including psychological factors, mental workload, and attention, which collectively shape decision-making efficacy. Mental workload is the capacity of available resources to meet task requirements [7]. Ship operations often require seafarers to work around the clock, which can easily result in a heavy workload and poor decision-making. Moreover, adverse weather conditions or high-risk situations at sea can further increase seafarers' mental workload, thereby elevating the risk of accidents in the maritime environment. Consequently, investigating workload dynamics in complex nautical contexts is vital for advancing intelligent monitoring systems that enhance situational awareness and critical emergency risk management.

Traditional workload research methods, mainly relying on subjective evaluation, can identify some failure modes, but it is difficult to capture the fluctuation in seafarers' neurocognitive states during dynamic operations. For example, seafarers' mental workload in high-risk scenarios increases significantly, leading to distraction and delayed decision-making [8]. Existing methods cannot quantify such load changes in real-time, thus limiting the timeliness of accident prevention measures.

The limitations of traditional human factors assessment methods can be addressed using a variety of available biosignals. Among them, electroencephalography (EEG) offers a higher degree of accuracy. EEG equipment utilizes electrodes applied directly to the scalp to non-invasively monitor electrical activity in the human brain. They have highly accurate temporal measurements and can dynamically reflect subtle changes in mental workload [9],[10]. EEG recognition typically involves extracting features like spectral power from raw EEG data and the subsequent application of the derived features to train a classifier. Usually, the raw EEG time-domain signal is transformed into the frequency domain using a transform such as the Fourier transform. Commonly considered frequency bands include the Delta (1-3 Hz), Theta (4-7 Hz), Alpha (8-12 Hz), Beta (13-30 Hz), and Gamma (31-80 Hz) bands. Studies have shown that EEG band characteristics are significantly correlated with workload levels. The power of oscillatory brain activity in the Theta frequency range is positively correlated with workload levels [11], and an increase in Theta power and a suppression of Alpha power often accompany high-workload states [12]. Nevertheless, existing studies lack a targeted analysis of the nautical operating environment, and few have correlated EEG features with specific nautical events, restricting their practical application value in maritime safety. This study proposes to construct an event-driven EEG mental workload assessment framework that realizes real-time quantification of seafarers' cognitive states and operational risk mapping through deep-learning models. The innovations of this study are as follows: 1) It is the first to correlate the temporal-spatial attributes of nautical events with EEG features to construct a scene-specific dataset; 2) A hybrid Convolutional Neural Network-Bidirectional Long Short-Term Memory (CNN-BiLSTM) architecture is proposed to synergistically optimize the frequency-domain and temporal feature characterization; 3) The combination of interpretable analytics and events provides theoretical support and practical guidelines for human factors optimization in intelligent maritime systems.

2 RELATED WORK

2.1 *Human factors in maritime operation*

One of the core challenges in maritime safety is to analyze the complex causal factors of human errors [1],[2]. Among existing methodologies, Human Reliability Analysis (HRA) has gained prominence for its probabilistic approach to error risk assessment through task decomposition [13]. Current HRA variants exhibit unique limitations in addressing the human factors within the complex marine environment. The Technique for Human Error Rate Prediction (THERP) [14] focuses on task-specific error rate prediction but neglects synergistic interactions between personnel competence, organizational culture, and situational stressors. The Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) [15] quantifies factors affecting performance through standardized classification criteria, but it is difficult to capture dynamic marine environmental influences. The Cognitive Reliability and Error Analysis Method (CREAM) [16] contextually models interactions among tasks and the environment, but lacks precision in modeling hierarchical organizational failures. The Human Error Assessment and Reduction Technique (HEART) [17] utilizes error-producing conditions to derive error probabilities, but oversimplifies the dynamics of cognitive load in emergencies. The Human Factors Analysis and Classification System (HFACS) [18] was initially applied to the air transportation System. In recent years, the hierarchical classification framework HFACS has been introduced into the maritime field to reveal the deeper mechanisms of accidents through the causal chain analysis of "organizational management-supervisory failure-precursors of unsafe behaviors". For example, the HFACS-MA model proposed by Chen et al. found that the interaction between communication failure and fatigue accumulation in ship-collision accidents is the critical path leading to operational errors [19]. Soner et al. enhanced HFACS with Fuzzy Cognitive Mapping (FCM) to analyze fire hazards, revealing latent interdependencies between technical failures and team coordination lapses [20]. Meanwhile, virtual maritime simulators have become an important tool for human factors research, which can synchronously collect seafarers' physiological signals and behavioral data in controlled environments through high-precision scenario simulation [10],[21],[22]. Liu et al. demonstrated that simulator training can improve seafarers' decision-making efficiency in emergency collision avoidance tasks. However, the assessment still relies on subjective performance scores and lacks objective means of quantifying cognitive load [10]. The limitations of traditional methods have driven the exploration of neurophysiological indicators. There is an urgent need to integrate objective data, such as EEG signals, to construct an objective and real-time human factors assessment system.

2.2 *EEG signals and mental workload*

Recent advancements in sensor technologies have empowered researchers to leverage physiological metrics for investigating cognitive dimensions of human behavior, particularly mental workload dynamics [23]. Among biosignal modalities, EEG stands out for its millisecond-level temporal precision, non-invasive nature, and validated efficacy in discriminating psychophysiological states, including emotional arousal, cognitive load stratification, and stress patterns [9],[24],[25],[26]. Within this context, mental workload is operationally defined as the cognitive resource allocation required to achieve task-specific goals [27]. To some extent, an increase in workload causes fluctuations in mental workload, which has long been recognized as being strongly associated with human performance. Therefore, it is essential to objectively construct a workload quantification framework.

The association of EEG, a direct physiological representation of neural activity, and its frequency band power characteristics with mental load has been widely validated. The study showed that Alpha and Theta band activity showed higher sensitivity when test subjects increased three task load conditions, associated with escalating acute cognitive demands [28]. All bands are coordinated to amplify in emergency and safety hazard situations [29]. During the extreme fatigue phase caused by increased workload, the local Alpha increases, and the Beta decreases, signaling performance degradation [30]. The peak of Theta-wave power in the frontal region fluctuates significantly with the increased workload, reflecting increased cognitive resource depletion and stress [31]. In conclusion, EEG waveform analysis and its in-band power modulation can be used to measure changes in subjects' workload during a maritime transportation simulation task. Notably, Hogervorst et al. benchmarked EEG against ECG and eye-tracking in N-back task paradigms, establishing EEG’s superior sensitivity to workload-physiology correlations [32]. Christian et al. compared subjects cross-validated with EEG data for a multi-attribute task and 11 computational methods; the results showed that multiple frequencies were required for EEG-based load categorization, and compared with 10 other cognitive state assessment methods, EEG data had a frequency bandwidth advantage [24]. These studies further demonstrate the reliability of EEG-based brain state recognition in assessing subjects' mental workload levels.

3 METHODOLOGY

3.1 *Experimental design and data collection*

The EEG data in this study were obtained from database in Fan et al. [33], which constructed a multi-scenario sailing task to simulate a real sailing environment, as shown in Figure 1. Seafarers must navigate a specific voyage while encountering various events. The key events include ship encounters (simulating multiple ship interactions), emergencies (simulating sudden rudder failures), and poor visibility conditions. All subjects underwent neurological screening to exclude pre-existing conditions and provided written informed consent before participation. The occurrence times, types of events, and operational responses of the participants were recorded. To objectively define mental workload states, EEG data were segmented into two categories: event-driven phases and baseline phases, steady-state navigation intervals without external disturbances, characterized by routine monitoring tasks. This segmentation aligns with prior studies linking dynamic operational demands to neurophysiological responses [10],[29]. 11 male crew members with seafarers' certificates of competency were recruited for the experiment (mean age 31.9 years, mean seafaring experience 7.7 years), with the 30-minute test data from participant 9 selected for analysis. All participants had no history of neurological disorders and signed an informed consent form before the experiment. The experiments utilized a NeuroSky Mindwave single-channel wireless EEG device (sampling rate 512 Hz) to acquire EEG signals from the subjects. The device employed dry-electrode technology to eliminate the interference of conductive gel, ensuring the participants' freedom of movement during dynamic manipulation. To validate the performance of the hybrid model in typical scenarios, the 30-minute data from participant 9 was selected as the main analysis in this study, as it covered the full range of predefined scenarios (ship encounters, poor visibility, and emergencies) and had the best signal quality. Future work will extend to the full dataset to further improve model robustness.

Figure 1. Experimental scenario[33]

3.2 *Feature extraction and preprocessing*

Time-frequency decomposition was performed via an 8-level Discrete Wavelet Transform (Daubechies db8 basis), segmenting raw EEG signals into five relevant bands: Gamma (40-100 Hz), Beta (12-40 Hz), Alpha (8-12 Hz), Theta (4-8 Hz), and Delta (0-4 Hz). Figure 2 shows the 2-second primary EEG data for test subject 9. These five EEG waves are associated with different mental concepts. Gamma-band power reflects anxiety under high-cognitive load, the Beta band is related to active decision-making behavior, the Alpha band characterizes relaxation or distraction, the Theta band indicates emotional fluctuations, and the Delta band is associated with fatigue or deep relaxation. To associate EEG characteristics with nautical events, the time window is dynamically divided based on the type and duration of events. For the vessel-encounter event, 3-minute windows (1.5 min pre/post-event) capture collision-avoidance decision-making. For the poor-visibility event, continuous 3-minute windows during events are used for persistent strategy analysis. For the rudder-failure event, 1-minute windows (30 s pre/post-event) to isolate emergency response dynamics. Longer windows addressed complex behavioral chains in ship encounters/poor visibility, while short windows emphasized abrupt cognitive shifts during failures. Baseline windows represented low-demand navigation. A sliding-window approach (512-sample length, 32-sample step) generated a 26,393×5 temporal feature matrix, ensuring millisecond-level EEG-event synchronization.

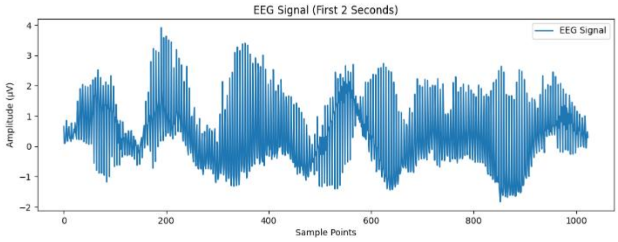

Figure 2. Two-second EEG data of subject 9 was collected by NeuroSky Mindwave headset.

3.3 *Model design*

The mental workload was classified into high and low states using a hybrid CNN-BiLSTM network. The model processes a multi-band EEG power matrix (Gamma, Beta, Alpha, Theta, Delta) combined with temporal dynamics through three core modules. The first module is the CNN layer, which processes the input data in parallel using two sets of one-dimensional convolutional kernels sized 128 and 256, respectively, aiming to capture short- and long-range frequency-domain patterns through different window sizes. After each set of convolutional operations, max-pooling (stride=3) is utilized to compress the feature dimensions and retain key frequency-domain information, effectively extracting the fluctuating features of different time scales in the EEG signals. Following the frequency domain feature extraction in the CNN layer, the data enters the Bi-LSTM layer. A bidirectional LSTM layer with 64 hidden units captures forward and backward cognitive state transitions, augmented by dropout regularization (rate=0.4) and layer normalization to ensure cross-scenario robustness. Fully connected layers splice CNN frequency-domain and LSTM temporal features. Then, a 128-node Swish activation dense layer enables higher-order feature interaction. Finally, the binary classification probability of mental workload (high/low) is output by the Sigmoid function, and the decision boundary is optimized via cross-entropy minimization.

4 RESULTS

4.1 *Model evaluation*

The CNN-BiLSTM hybrid neural network constructed in this study demonstrates excellent performance in mental workload assessment for seafarers. Through the simulation data of typical scenarios such as ship encounters, poor visibility, and rudder malfunction, the model achieves a 96% classification accuracy in high and low workload assessments. The area under the curve (AUC) of the operating characteristics of the subjects reaches 0.986, and the average precision of the precision-recall (PR) curve is 0.99, which indicates that the model has a remarkable ability to differentiate between seafarers' mental workload levels, as shown in Figure 3.

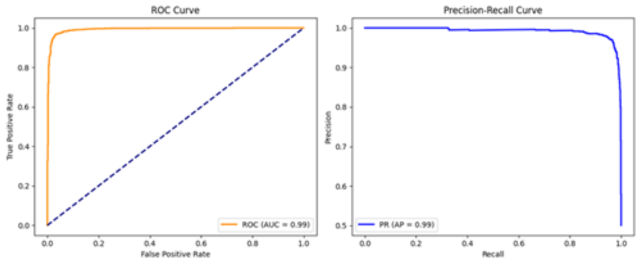

Figure 3. ROC and PR curves of the CNN-BiLSTM model for seafarer mental workload assessment

The confusion matrix shows that the correct recognition rate (recall) of the high load state is 0.941. The correct recognition rate of the low load state is 0.944, with only a few misclassifications, effectively avoiding the omission detection problem, especially in high-risk scenarios, as shown in Figure 4. Notably, the model exhibited superior sensitivity to abrupt workload transitions during emergency scenarios (e.g., rudder failure detection rate exceeding 95%), validating its capability to identify event-induced mental workload states against steady navigation baselines.

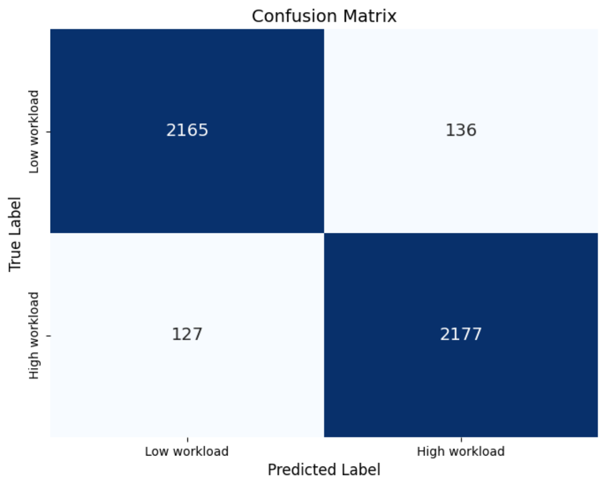

Figure 4. Confusion matrix of high-and low-load forecasts in the CNN-BiLSTM model

4.2 *Interpretability analysis*

To reveal the decision logic of the model, the Shapley Additive Explanations (SHAP) framework is utilized to quantitatively analyze the contribution of each EEG band. As illustrated in Figure 5, the effects of different bands on the classification results vary, with the Theta band (mean |SHAP value| = 0.4346) and the Alpha band (mean |SHAP value| = 0.3609) having the most notable contributions. The positive effect of Theta band indicated that the Theta oscillations in the prefrontal region are enhanced with the increase of mental workload, reflecting the increase in the depletion of cognitive resources and the rise of emotional stress; the negative effect of Alpha band corresponds to the inhibition of attentional resources; the dynamic balance of the two constitutes the core of the neurocognitive load. This result indicates that the Theta band, which reflects the emotional fluctuation state, and the Alpha band, which represents the degree of attentional inhibition, serve as core physiological indicators influencing changes in seafarers' operational performance, thereby providing a key neurocognitive basis for the model output.

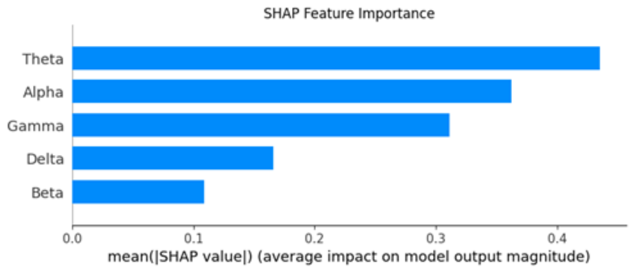

Figure 5. SHAP feature importance analysis of EEG bands in the hybrid CNN-BiLSTM model

5 CONCLUSIONS

This study proposes an event-driven EEG load assessment framework to objectively quantify cognitive states in complex operational scenarios by dynamically correlating nautical events with EEG features. The effectiveness of the CNN-BiLSTM model is demonstrated by its high classification accuracy and feature fusion mechanism, which captures neural dynamics that are difficult to resolve using traditional methods.

The dynamic fluctuations of Theta and Alpha oscillations provide novel neurophysiological insights for maritime human factors research. It reflects both short-term mental workload (Theta power surge) and long-term attention maintenance (Alpha power suppression), can serve as a bridge between neural mechanisms and operational performance. In the future, multimodal data can be further integrated to construct a multidimensional physiological indicator fusion model to enhance the robustness of load assessment in complex scenarios. This study tested the mental workload of seafarers in a more ideal state within a simulated environment and did not consider the factor of fatigue after long shifts. Future research could assess the effect of fatigue accumulation on mental workload within the context of actual seafaring tasks. Additionally, all study subjects were male, necessitating follow-up studies that include female crew members to validate the generalizability of the model. Expanding the sample size and including seafarers with varying levels of experience is also crucial to validating the model's ability to generalize across individual scenarios, making it an important direction for subsequent studies.

ACKNOWLEDGEMENT

The herein study was supported by Wuhan University of Technology's grant 104972025RSCrc0004.

REFERENCES

[1] R. Hanzu-Pazara et al., "Reducing of maritime accidents caused by human factors using simulators in training process," J. Mar. Res., vol. 5, no. 1, Apr. 2008.

[2] V. Ronca et al., "Neurophysiological Assessment of An Innovative Maritime Safety System in Terms of Ship Operators' Mental Workload, Stress, and Attention in the Full Mission Bridge Simulator," Brain Sci., vol. 13, no. 9, Sep. 2023. doi: 10.3390/brainsci13091319.

[3] European Maritime Safety Agency (EMSA), "EMSA Facts & Figures 2022," 2022. [Online]. Available: https://www.emsa.europa.eu/... (Accessed: Apr. 9, 2025).

[4] R. P. E. Gordon, "The contribution of human factors to accidents in the offshore oil industry," Reliab. Eng. Syst. Saf., vol. 61, no. 1, pp. 95–108, Jul. 1998.

[5] International Seafarers' Welfare and Assistance Network (ISWAN), "ISWAN launches allyship guidance to build safer cultures at sea," 2025. [Online]. Available: https://www.iswan.org.uk/news/iswan-launches-allyship-guidance-to-build-safer-cultures-at-sea/ (Accessed: May 6, 2025).

[6] U.S. Department of Transportation, Maritime Administration (MARAD), "Every Mariner Builds A Respectful Culture (EMBARC) Standards Program," 2025. [Online]. Available: https://www.maritime.dot.gov/embarc (Accessed: May 6, 2025).

[7] M. S. Young et al., "State of science: mental workload in ergonomics," ERGONOMICS, vol. 58, no. 1, pp. 1–17, 2015.

[8] M. Arenius, G. Athanassiou, and O. Sträter, "Systemic assessment of the effect of mental stress and strain on performance in a maritime ship-handling simulator," IFAC Proc. Vol., vol. 43, no. 13, pp. 43–46, Jan. 2010.

[9] Y. Liu et al., "EEG-based Mental Workload and Stress Recognition of Crew Members in Maritime Virtual Simulator: A Case Study," in Proc. 2017 Int. Conf. Cyberworlds (CW), Sep. 2017, pp. 64–71.

[10] Y. Liu et al., "Psychophysiological evaluation of seafarers to improve training in maritime virtual simulator," Advanced Engineering Informatics, vol. 44, p. 101048, Apr. 2020.

[11] A. Gevins et al., "Monitoring working memory load during computer-based tasks with EEG pattern recognition methods," Hum. Factors, vol. 40, no. 1, pp. 79–91, Mar. 1998.

[12] D. Dasari, G. Shou, and L. Ding, "ICA-Derived EEG Correlates to Mental Fatigue, Effort, and Workload in a Realistically Simulated Air Traffic Control Task,"Front. Neurosci., vol. 11, p. 297, 2017.

[13] B. Kirwan, A Guide to Practical Human Reliability Assessment, 1st ed. Boca Raton, FL: CRC Press, 1994.

[14] A. D. Swain and H. E. Guttmann, "Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report," Sandia National Lab. (SNL-NM), Albuquerque, NM, USA, Tech. Rep. NUREG/CR-1278, SAND-80-0200, Aug. 1983.

[15] D. I. Gertman et al., "The SPAR H human reliability analysis method," in Proc. 4th Int. Top. Meet. Nuclear Plant Instrum., Control, Human-Machine Interface Technol., pp. 17–24, 2004. [Online]. Available: http://www.scopus.com/... (Accessed: Apr. 17, 2025).

[16] A. Modares et al., "An integrated Cognitive Reliability and Error Analysis Method (CREAM) and optimization for enhancing human reliability in blockchain," Decis. Anal., vol. 12, p. 100506, Sep. 2024.

[17] J. C. Williams, "Heart—A Proposed Method for Achieving High Reliability in Process Operation by Means of Human Factors Engineering Technology," Safety and Reliability, vol. 35, no. 3, pp. 5–25, Dec. 2015.

[18] M. Celik and S. Cebi, "Analytical HFACS for investigating human errors in shipping accidents," Accid. Anal. Prev., vol. 41, no. 1, pp. 66–75, Jan. 2009.

[19] S.-T. Chen et al., "A Human and Organisational Factors (HOFs) analysis method for marine casualties using HFACS-Maritime Accidents (HFACS-MA)," Saf.Sci., vol. 60, pp. 105–114, Dec. 2013.

[20] O. Soner, U. Asan, and M. Celik, "Use of HFACS–FCM in fire prevention modelling on board ships," Saf. Sci., vol. 77, pp. 25–41, Aug. 2015.

[21] Y. Liu et al., "EEG-based Cadets Training and Performance Assessment System in Maritime Virtual Simulator," in Proc. 2018 Int. Conf. Cyberworlds (CW), Oct. 2018, pp. 214–220.
[22] Y. Liu et al., "Human Factor Study for Maritime Simulator-Based Assessment of Cadets," in Proc. ASME 35th Int. Conf. Ocean, Offshore Arctic Eng., vol. 3, New York: ASME, 2016, p. V003T02A073.
[23] M. U. Iqbal et al., "Multi-class classification of control room operators' cognitive workload using the fusion of eye-tracking and electroencephalography," Comput. Chem. Eng., vol. 181, p. 108526, Feb. 2024.
[24] C. A. Kothe and S. Makeig, "Estimation of task workload from EEG data: new and current tools and perspectives," in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2011, pp. 6547–6551.
[25] X. Hou et al., "CogniMeter: EEG-based Emotion, Mental Workload and Stress Visual Monitoring," in Proc. 2015 Int. Conf. Cyberworlds (CW), Oct. 2015, pp. 153–160.
[26] S. Shao et al., "EEG-Based Mental Workload Classification Method Based on Hybrid Deep Learning Model Under IoT," IEEE J. Biomed.Health Inform., vol. 28, no. 5, pp. 2536–2546, May 2024.
[27] C. Berka et al., "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," Aviat. Space Environ. Med., vol. 78, no. 5, Spec. Iss., pp. B231–B244, May 2007.
[28] S. Lei and M. Roetting, "Influence of task combination on EEG spectrum modulation for driver workload estimation," Hum. Factors, vol. 53, no. 2, pp. 168–179, Apr. 2011.
[29] X. Li, L. Yang, and X. Yan, "An exploratory study of drivers' EEG response during emergent collision avoidance," J. Saf. Res., vol. 82, pp. 241–250, Sep. 2022.
[30] B. T. Jap et al., "Using EEG spectral components to assess algorithms for detecting fatigue," Expert Syst. Appl., vol. 36, no. 2, pp. 2352–2359, Mar. 2009.
[31] G. Shou and L. Ding, "Frontal theta EEG dynamics in a real-world air traffic control task," in Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2013, pp. 5594–5597.
[32] M. A. Hogervorst et al., "Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload," Front. Neurosci., vol. 8, p. 322, 2014.
[33] S. Fan et al., "Effects of seafarers' emotion on human performance using bridge simulation," Ocean Eng., vol. 170, pp. 111–119, Dec. 2018.