Research Square

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

Is mental fatigue a nonlinear accumulation process? Evidence from field EEG-based brain network analysis

Qing Xin
Southwest Petroleum University
Su Hao
steven_haosu@qq.com
Southwest Petroleum University
Wang Xiaoqin
Southwest Petroleum University
Wang Jian
CCDC Drilling & Production Technology Research Institute
Shen Liaoyuan
Southwest Petroleum University

Research Article

Keywords: Mental fatigue, Functional connectivity, Brain network, Transformer, Time-on-task, Cognitive load

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10012393/v1

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

Additional Declarations: No competing interests reported.

Is mental fatigue a nonlinear accumulation process? Evidence from field EEG-based brain network analysis

Qing Xin[1], Su Hao[1,2,*], Wang Xiaoqin[1], Wang Jian[1], Shen Liaoyuan[1]

[1]School of Economics and Management, Southwest Petroleum University, Chengdu 610500, China;

[2]Key Laboratory of Energy Security and Low-carbon Development, Southwest Petroleum University, Chengdu 610500, China

[3]CCDC Drilling & Production Technology Research Institute, No.88, South section 2, Zhongshan Road, Guanghan, 618300, China.

*Corresponding authors: Su Hao, steven_haosu@qq.com;

Abstract

Mental fatigue is a difficult to detect risk state in high-risk industries and is closely associated with sustained cognitive load and prolonged work duration. To overcome the limitations of post-incident management, research has introduced wearable EEG technologies, promoting a shift in safety management from passive response to feedforward warning. However, most existing studies are based on laboratory experiments, treating cognitive load and work duration as independent variables and inducing fatigue over short durations, which limits the characterization of fatigue under extended work periods. To address these issues, this study conducted a field experiment at a drilling site, combining wearable EEG functional connectivity and brain network analyses with a PCA Transformer framework for automated fatigue recognition. The results show that as task duration increased, functional connectivity analysis indicated that mental fatigue generally increased and then decreased; meanwhile, brain network analysis revealed a transition from local strengthening to global integration, accompanied by reduced network stability. The Transformer model demonstrated excellent classification performance, achieving recognition accuracies of 90.7% and 91.0% under low and high load conditions, respectively. Overall, these findings indicate that shift scheduling should not rely solely on cumulative work duration but should instead pay closer attention to changes in risk across different work periods. Based on this, safety management practices may consider reducing the concentration of high cognitive load tasks and critical decision making activities during later stages of work, while machine learning based fatigue state recognition offers feasible technical support for feedforward safety management.

Keywords: Mental fatigue; Functional connectivity; Brain network; Transformer; Time-on-task; Cognitive load

1. Introduction

Mental fatigue among operators has been widely recognized as an important latent factor influencing safety performance in safety-critical work settings. It is generally defined as a psycho-biological state that occurs when individuals engage in prolonged or high-intensity cognitive activity (Yu et al., 2019). When mental fatigue develops, it directly affects operators' cognitive functioning during task execution, leading to reduced attention, slower response speed, impaired judgment, and a markedly higher risk of operational errors and safety incidents (Tran et al., 2020). In high-risk industries, such performance degradations often manifest first as minor errors or near misses rather than immediate severe accidents. Numerous studies have shown that accidents in high-risk industries often conform to the "Heinrich's Law" pyramid model, whereby severe accidents typically evolve from a large number of minor injuries and near misses. This progression highlights the critical role of early-stage risk identification, underscoring the importance of real-time monitoring and proactive prevention of emerging safety hazards.

In line with this, empirical evidence indicates that in high-risk sectors such as oil and gas drilling, human factors constitute the primary cause of major accidents (Wiegmann & Shappell, 2017). According to statistics from the International Association of Oil and Gas Producers (IOGP, 2016), although the overall number of incidents declined between 2007 and 2016, the fatality rate increased from 1.5 in 2015 to 1.7 in 2016, suggesting that while the frequency of accidents decreased, the severity of outcomes worsened. Similar patterns have been observed in other high-risk domains. In nuclear power systems, for instance, operator fatigue—often stemming from excessive workload and insufficient sleep—can markedly impair decision-making and emergency response capabilities, creating serious safety hazards during critical operations (Bell et al., 2023). In the construction industry, workers are frequently exposed to dynamic and complex environments that impose high cognitive demands, rendering them more vulnerable to mental fatigue (Wang et al., 2023).

Additionally, cognitive load is recognized as another critical factor influencing mental fatigue (Staiano et al., 2023). It is commonly defined as "the total amount of mental activity imposed on an individual's cognitive system per unit of time" (Sweller, 1988), reflecting the level of cognitive resource demand during task execution and directly influencing task efficiency and behavioral performance. Accordingly, cognitive load is widely used to characterize individuals' cognitive states (Behrens et al., 2023). In prolonged or high-risk tasks, sustained high cognitive load not only accelerates the depletion of mental resources but also impairs the brain's regulatory and recovery capacities, leading to earlier and more pronounced manifestations of fatigue (Goodman et al., 2025). In contrast, moderate cognitive load may mitigate fatigue accumulation to some extent, as the brain can strategically adjust and redistribute resources to sustain task performance before cognitive capacity becomes saturated, thus delaying the onset of overt fatigue (Harju et al., 2023).

Electroencephalography (EEG), as a direct reflection of neural activity with high temporal resolution and sensitivity to cognitive processes, has been widely recognized as a valuable tool for monitoring mental fatigue (Proost et al., 2022). Compared with traditional self-report and behavioral measures, EEG provides more objective, continuous, and real-time evidence. However, due to its high dimensionality, nonlinear characteristics, and susceptibility to noise, analyses limited to local EEG activity cannot adequately capture the overall dynamics of fatigue evolution (Wu et al., 2023). Functional connectivity and brain network analyses, by contrast, allow for the characterization of dynamic interactions between brain regions. Corresponding graph-theoretical metrics quantify

changes in information transfer and integration, thereby providing novel neurophysiological indicators for individualized fatigue monitoring and intervention. Cross-industry evidence supports the potential of this approach. For instance, in aviation research, integrating EEG with psychological assessments has proven effective in predicting pilots' risk propensity under high-stakes conditions, providing insights for personnel selection and training (Ebrahim et al., 2025). Moreover, graph-theoretical analyses of functional brain networks have elucidated the relationship between network structure and fatigue resilience—for example, under sleep deprivation, variations in mean node strength and characteristic path length have been shown to explain individual differences in fatigue tolerance (Mason et al., 2024). Consistently, findings from long-duration task paradigms suggest that as fatigue accumulates, the brain often compensates for reduced global efficiency by enhancing local efficiency, a mechanism that varies significantly across brain regions and frequency bands (Xu, 2022).

The limited ability of laboratory paradigms to capture the complexity of high-risk operations has led to growing emphasis on ecological validity in mental fatigue research. In recent years, experimental research has increasingly emphasized real-world field measurements, as authentic environments can substantially enhance the external validity and generalizability of research findings (Mudrik et al., 2024). In line with this trend, the rapid advancement of portable EEG technologies has made in-situ neurophysiological research increasingly feasible. Systematic reviews indicate that implementing EEG-based brain activity monitoring in real operational environments is not only practical but also methodologically advantageous, though challenges such as motion artifacts and environmental noise must be addressed through improved hardware and algorithmic solutions (Janssen et al., 2021). In applied contexts, mobile EEG has been successfully employed in cognitive studies involving head-mounted augmented reality and complex outdoor scenarios, confirming the feasibility and scientific value of neural measurement during "unconstrained" tasks (Lee et al., 2022). Recent reviews in occupational safety further suggest that field-deployable EEG monitoring is transitioning from conceptual validation to practical implementation. Its key advantage lies in the integration of real-time responsiveness with operational deployability, and guidelines have been proposed for advancing on-site neuromeasurement in high-risk settings such as construction (Han et al., 2023).

Driven by these practical demands, machine learning has provided a viable technological pathway for EEG-based fatigue monitoring. On one hand, deep neural networks can automatically extract discriminative spatiotemporal and connectivity features from high-dimensional, nonstationary data. On the other hand, data-driven approaches enhance model robustness and contextual adaptability. In line with this, online monitoring studies employing wearable EEG caps have identified pronounced differences in frontal and occipital activity between fatigued and alert states, enabling high-precision classification of driver fatigue (Li et al., 2023). In construction safety research, integrating EEG with immersive virtual reality and employing a two-stage ensemble classification strategy increased multi-class hazard recognition accuracy to 82.3%, demonstrating the feasibility of combining EEG and machine learning under complex conditions (Jeon & Cai, 2022). In the aviation domain, EEG has been incorporated into hybrid Bayesian neural and Bayesian network models to assess situation-awareness reliability by integrating fatigue and emotional states, with empirical validation in real-world experiments (Ding et al., 2025).

Building on the above progress, conducting continuous field experiments in oil and gas drilling—an operational environment characterized by both high workload and high risk—provides

an evidence base with greater explanatory power and practical relevance. Against this background, the present study aims to systematically examine the differential manifestations of mental fatigue under varying work durations and cognitive load conditions, and to introduce machine learning methods to meet the monitoring needs of high-risk industries. The research design comprised several key steps: first, field experiments were conducted at real oil and gas drilling sites to collect EEG data from operators; second, the data were grouped according to work duration and cognitive load intensity; third, raw EEG signals underwent standardized preprocessing, followed by functional connectivity and brain network analyses to compare fatigue-related network metrics across conditions; and finally, principal component analysis (PCA) was applied for dimensionality reduction, Gaussian mixture modeling (GMM) was used to generate pseudo-labels, and a Transformer-based architecture was employed to achieve binary classification between alert and fatigued states, enabling intelligent fatigue recognition. Conducting these experiments in authentic drilling operations ensured high ecological validity of the findings. On this basis, a two-stage "unsupervised pseudo-label–Transformer" framework was proposed to achieve systematic identification of mental fatigue. Theoretically, this study reveals the interactive effects of cognitive load and work duration at the brain network level and captures the nonlinear "second wind" phenomenon, enriching the understanding of mental fatigue dynamics. Practically, the proposed framework provides a feasible technological pathway and data foundation for real-time fatigue monitoring and safety management in high-risk industries.

2. Methodology

2.1. Experimental Set-up

This study was conducted at an oil and gas drilling site, a representative setting of high-risk industries. The working environment is characterized by frequent human–machine interaction, continuous information processing, and uninterrupted tasks, which place sustained demands on cognitive resources and ensure high ecological validity. Eight drillers participated in the experiment, whose professional background closely matched the research objectives and reflected the defining features of high-risk work—high operational hazards, complex tasks, and potentially irreversible consequences in the event of accidents. As key operators responsible for on-site safety, drillers primarily engage in mentally demanding rather than physically exhausting tasks, making them ideal subjects for studying mental fatigue mechanisms. Participants were aged 32–42 years (mean = 36.83 years) with an average of 5.33 years of work experience. All were healthy, right-handed, free of sleep disorders, and passed medical evaluations before the study. Each held a valid operation license and participated voluntarily after providing informed consent. The study protocol and ethical procedures were approved by the Academic Committee of the School of Economics and Management, Southwest Petroleum University.

2.2 EEG experiments and data collection

Field experiments offer the advantage of observing phenomena directly within natural settings. This study was conducted in a real work environment to enhance both ecological validity and

practical applicability. Unlike laboratory experiments, field studies are carried out under natural and largely uncontrollable conditions, allowing causal relationships between variables to emerge more authentically. Their defining features include environmental realism, behavioral naturalness, task complexity, and limited experimental control (Harrison & List, 2004). Accordingly, the oil and gas drilling site was selected as the experimental setting to ensure the authenticity of operations and the external validity of the findings.

The experiment was conducted continuously over 20 days in an active drilling environment to ensure data representativeness and ecological validity. The research team entered the driller’s control cabin at two-hour intervals to record EEG signals while participants performed their actual work. Both day and night shifts included five scheduled recording sessions, each lasting 10 minutes. For statistical analysis, each participant completed five measurements per shift (T1 – T5, as shown in Figure 1A).

Cognitive load levels were defined based on the specific tasks performed during data collection (Figure 1A). Low-load conditions primarily included drilling operations, during which participants mainly monitored a single display, passively observing drilling parameters with minimal manual control or communication, thus requiring relatively low cognitive processing. In contrast, high-load conditions occurred during complex or critical operations such as tripping in and tripping out. During these periods, participants monitored multiple screens simultaneously, processed real-time data from different sources, maintained continuous communication and coordination with other positions, and performed multiple complex manipulations and rapid decisions, resulting in high interaction and operational intensity.

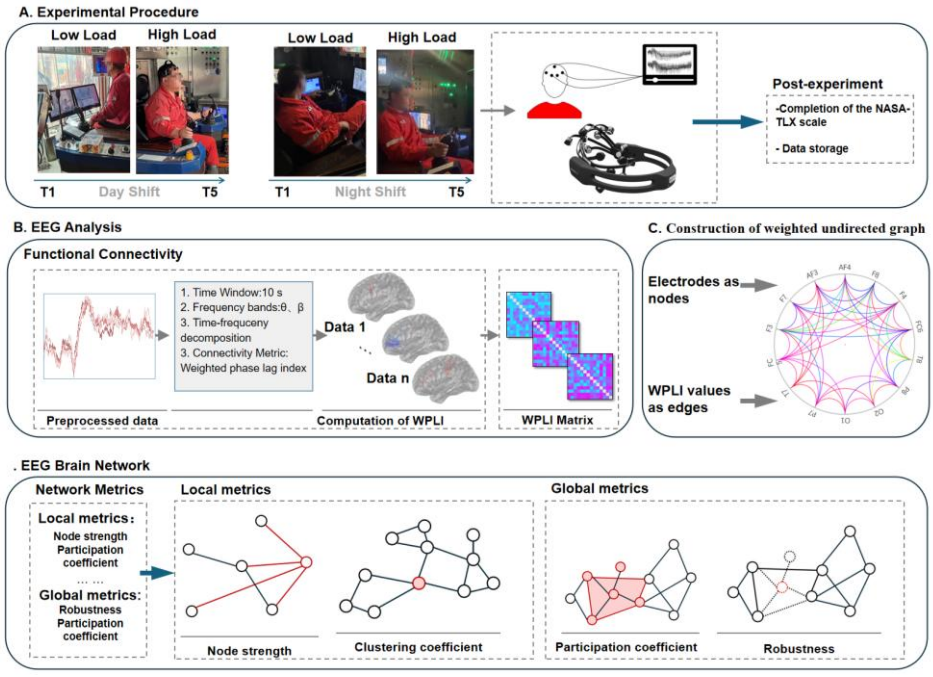


Figure 1. Experimental design and data analysis pipeline. (A) The experiment adopted a 2 × 5 design, combining two cognitive load levels (high vs. low) with five time points (T1 – T5). Upon completion, participants filled in the NASA-TLX scale, and all data were stored for subsequent analysis. (B) Data analysis workflow: raw EEG signals were preprocessed, and functional connectivity matrices were computed in the θ and β frequency bands using the Weighted Phase Lag Index (WPLI). (C) Brain networks were constructed in Python with the

NetworkX and Brain Connectivity Toolbox (BCT) libraries, where electrodes were defined as nodes and WPLI values as edges. (D) Network metrics were then computed at both local and global levels, including node strength, clustering coefficient, participation coefficient, and robustness.

2.3 EEG signal preprocessing

EEG preprocessing was conducted using the open-source platform EEGLAB 2023a (Figure 1B). As the EMOTIV system records differential signals relative to a common-mode sensing electrode, no re-referencing was required. The preprocessing pipeline consisted of several steps. First, channel locations were verified, and a 0.05 – 40 Hz band-pass filter was applied to remove power-line interference and low-frequency drifts while preserving physiological rhythms related to cognitive activity (Hao et al., 2023). Second, independent component analysis (ICA) was used to separate non-neural artifacts, and a 60% rejection threshold was set to manually remove residual artifacts. The data were then segmented into 10-second epochs to ensure each segment contained sufficient valid information. Finally, baseline correction was performed to improve signal accuracy and consistency. Unlike conventional laboratory paradigms that use a fixed pre-stimulus period as the baseline, the drilling environment involved frequent task transitions without clearly defined stimulus events. Therefore, the average EEG recorded during resting states before and after each work period was used as the individual baseline.

Unlike traditional EEG preprocessing, this study optimized the procedure to accommodate the complexity of field environments by implementing a dual-threshold collaborative filtering approach to enhance the robustness of artifact rejection. Conventional single-threshold methods (± 100 μV or ± 80 μV) often lead to excessive data loss under real-world conditions due to head movements, eye blinks, muscle activity, and equipment noise. To address this, a slope threshold was applied to remove high-frequency transient artifacts, while a variance threshold was used to eliminate low-frequency, high-variance noise, thereby suppressing artifacts while retaining valid signals. Furthermore, given the frequent task transitions and lack of fixed stimuli in field settings, a 5-minute resting-state EEG (eyes closed, seated quietly) collected before and after each work shift was used as the individual baseline to ensure consistency and representativeness.

2.4. EEG feature extraction

2.4.1. EEG Functional Connectivity

The interactions among different brain regions were analyzed using phase-based functional connectivity computation (Figure 1B). The instantaneous phases of two EEG channels, $\phi_x(t)$ and $\phi_y(t)$, , yield the following expression for their phase difference:

$$\triangle \phi_{xy}(t) = \left|\phi_x(t) - \phi_y(t)\right| \leq C$$

Where C is a constant boundary. When two channels maintain a stable phase relationship within a specific frequency band, their phase differences remain consistent over time. On this basis, the Phase Lag Index (PLI) was introduced to quantify the consistency of nonzero phase differences between signals, defined as:

$$PLI = \left| \frac{1}{N} \sum_{n=1}^{N} sign(\triangle \phi_{xy}(t_n)) \right|$$

Where N denotes the number of samples, and *sign(·)* is the sign function. PLI minimizes the influence of volume conduction by ignoring zero phase differences but remains relatively sensitive to low signal-to-noise ratios in practice. In order to increase stability, the Weighted Phase Lag Index (WPLI) was employed, which is defined as:

$$WPLI = \frac{|\langle |\Im(S)| sign(\Im(S)) \rangle|}{|\langle \Im(S) \rangle|}$$

In this equation, $\Im(S)$ represents the imaginary part of the cross-spectrum between two signals. Compared with PLI, WPLI not only considers the direction of the phase lag but also introduces amplitude weighting for the imaginary component, thereby improving sensitivity to genuine phase relationships. Considering that field experiments are typically conducted in complex environments characterized by low signal-to-noise ratios and strong volume-conduction effects, WPLI demonstrates greater stability and reliability under such conditions and is thus well suited for functional connectivity analysis of EEG data in this study. The analysis focused on the θ (4–7 Hz) and β (13–30 Hz) frequency bands. The θ band is broadly associated with cognitive control and attentional regulation and is often regarded as a key neural marker reflecting variations in fatigue and cognitive load. The β band is closely related to motor control and higher-order cognitive processing, exhibiting particular sensitivity to network changes under fatigued states.

2.4.2. EEG Brain Network Construction

Based on the functional connectivity matrix, EEG signals were modeled as a directed weighted graph $G = (V, E, W)$, where V represents the set of nodes (EEG electrodes), E represents the set of edges, and $W = [w_{ij}]$ denotes the weighted adjacency matrix, with w_{ij} indicating the connection strength between nodes i and j. Using this representation, both global and local network metrics were systematically evaluated (Figure 1C).

At the local level, closeness centrality (C_c) was used to measure how close a node is to all other nodes in the network, defined as

$$C_c(i) = \frac{N-1}{\sum_{j \neq i} d(i,j)}$$

where $d(i,j)$is the shortest path length between nodes $i$and j, and N is the total number of nodes. A higher $C_c$value indicates that a node occupies a more central position within the network.

Node strength (NS) quantifies the cumulative connection strength between a node and its neighbors, defined as

$$NS = \sum_{j=1}^{N} w_{ij} + \sum_{j=1}^{N} w_{ji}$$

Finally, the participation coefficient (P_i) was used to evaluate the extent to which a node connects to different network modules, defined as:

$$P_i = 1 - \sum_{m=1}^{M} (\frac{k_{im}}{k_i})^2$$

where k_i^mdenotes the sum of connection weights between node $i$and nodes in module m, k_i

represents the total connection strength of node i, and M is the total number of modules. A higher participation coefficient indicates that the node is more evenly connected across different functional modules.

At the global level, six categories of metrics were selected to comprehensively evaluate the overall characteristics of brain networks. Global efficiency ($E_{glob}(G)$) was used to reflect the efficiency of information transfer across the entire network and is defined as

$$E_{glob}(G) = \frac{1}{N(N-1)} \sum_{j \neq i} \frac{1}{d(i,j)}$$

where $d(i,j)$represents the shortest path length between nodes $i$and j. Average closeness centrality quantifies the degree to which each node is close to all other nodes, thereby indicating the overall compactness of the network. Clustering coefficient ($C(G)$) was used to characterize the degree of interconnection among neighboring nodes, with its global form defined as

$$C(G) = \frac{1}{N} \sum_{i=1}^{N} C_i$$

where C_i denotes the local clustering coefficient of node i. Similarly, average node strength represents the mean of all nodes' strengths, describing the overall weighted connectivity level of the network. Network strength ($S(G)$) was used to measure the total amount of connectivity within the network, expressed as

$$S(G) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}$$

This metric captures the overall connection density under varying cognitive load and time conditions. In order to analyze network stability in the face of potential disruptions, robustness was measured by systematically eliminating important nodes and monitoring the retention of the largest connected component or overall efficiency. This allowed evaluation of the network's average resilience to both random failures and targeted attacks.

2.4.3. Fatigue State Recognition and Machine Learning

The identification of mental fatigue continues to pose methodological challenges, particularly due to the lack of an integrated assessment framework. Consequently, most existing studies have adopted one of several commonly used methodological approaches. The first approach relies on physiological indicators derived from experimental manipulation, in which pre-task states are defined as alert and post-task states as fatigued. However, this approach is often influenced by individual baseline variability (Kar et al., 2010; Anwer et al., 2024). The second approach combines subjective rating scales with behavioral tasks, such as the use of the Karolinska Sleepiness Scale in simulated driving paradigms (Gao et al., 2019).

However, both approaches have notable limitations. The first approach may yield distorted judgments when participants are unable to accurately perceive their own fatigue due to limited metacognitive awareness, whereas the second approach overlooks the nonlinear nature of fatigue evolution, for example, alertness at the end of a shift is not necessarily lower than at its beginning. By contrast, the third approach, unsupervised machine learning – based automatic recognition, has shown greater potential in high-risk industrial settings (Li et al., 2024). Building on this approach, the present study does not assume fatigue states to be entirely unstructured; instead, it integrates

unsupervised modeling with temporal prior information to identify latent fatigue patterns from the intrinsic structure of multidimensional EEG features without relying on manual labels. Specifically, the temporal prior is not introduced as a deterministic label but as a gradually evolving state tendency along the work process, such that earlier stages (e.g., T1) are more likely to correspond to alert states, while with increasing time-on-task (T2 – T4), the weighting toward fatigue progressively increases, and temporally consistent pseudo-labels are generated through posterior probabilities. This strategy maximizes data utilization while reducing subjective bias from manual categorization and provides more stable and reliable training signals for subsequent supervised models.

Building on the above rationale, this study further operationalizes unsupervised machine learning – based fatigue recognition by proposing a two-stage modeling framework oriented toward practical application. In the first stage, principal component analysis (PCA) and Gaussian mixture modeling (GMM) are used to perform unsupervised modeling of latent states in the EEG feature space, with temporal priors incorporated to generate high-confidence pseudo-labels; in the second stage, these pseudo-labels are used to train a Transformer-based model to automatically classify alert and fatigued states, thereby achieving mental fatigue identification.

In the first stage, PCA was applied to the EEG features for dimensionality reduction, removing redundancy while retaining the major sources of variance. Let the covariance matrix of the standardized data be $S = \frac{1}{n-1} X^T X = U \Lambda U^T$. The top m principal components were selected such that the cumulative variance contribution reached 90%, yielding the low-dimensional representation $Z = XU_{1:m}$. Subsequently, a GMM was used for clustering in the reduced feature space, modeled as

$$p(z) = \sum_{k=1}^{K} \pi_j \mathcal{N}(z|\mu_k, \sum_k)$$

where π_k, μ_k, and Σ_k denote the weight, mean, and covariance of the k-th Gaussian component, respectively. Model parameters were estimated using the Expectation–Maximization (EM) algorithm, and posterior probabilities were computed as

$$\gamma_{ik} = \frac{\pi_j \mathcal{N}(z|\mu_k, \sum_k)}{\sum_j \pi_j \mathcal{N}(z|\mu_k, \sum_k)}$$

Each sample was then assigned a pseudo-label and a corresponding confidence score. To improve labeling accuracy, semantic mapping and temporal priors were incorporated, and a confidence threshold of 0.70 was applied; samples below this threshold were excluded. Finally, the remaining samples were classified as “alert” or “fatigue” based on posterior probabilities.

In the second stage, the pseudo-labeled data were used to train a Transformer-based classification model. Input sequences were linearly projected into multi-head self-attention layers, where the attention mechanism was computed as

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

Here, Q, K, and V represent the query, key, and value matrices derived from the input feature embeddings, and d_k denotes the dimensionality of each attention head. This mechanism captures long-range dependencies among features and adaptively assigns higher weights to the most discriminative components. The final classification output was obtained through a feed-forward

network following the attention layers. Model performance was evaluated using multiple metrics, including accuracy, recall (sensitivity), and the macro-averaged F1 score (F1-macro).

3. Results

3.1 Functional Connectivity Patterns Across Cognitive Load and Time

Descriptive statistics indicated a general enhancement of functional connectivity across all connectivitys in the load dimension. In the θ-band, all connectivity pairs showed an overall increase with a mean growth of approximately 6.8%. Among them, O2–O1 exhibited the largest increase, rising from (mean ± SD, 0.246 ± 0.193) to (mean ± SD, 0.276 ± 0.210), an approximate change of 12.2%, whereas P7–AF3 showed the smallest increase, from (mean ± SD, 0.263 ± 0.201) to (mean ± SD, 0.270 ± 0.202), representing a 2.7% change. In the β-band, all connectivities also increased, with an average growth rate of 16.1%; the greatest increase was observed for F3–F7 (from 0.137 ± 0.126 to 0.166 ± 0.148, +21.5%), while the smallest occurred for FC5–AF3 (from 0.133 ± 0.123 to 0.145 ± 0.125, +9.3%). Across the entire time series, 91 connectivity pairs were analyzed (Figure 1C). The mean θ-band connectivity ranged between 0.25 and 0.30, whereas the β-band ranged between 0.13 and 0.20. Most connectivitys showed gradual increases or remained stable from T1 to T4 but declined at T5 (Table 1). For instance, in the θ-band, F3–AF3 increased from 0.271 at T1 to 0.275, 0.279, and 0.272 at T2–T4, respectively, before decreasing to 0.258 at T5, showing a distinct turning point. Representative examples of θ and β connectivitys, including F7–AF3, F3–AF3, F3–T8, T7–F3, P7–T7, and O2–F7, are summarized in Table 1.

Two-way statistical analysis revealed that in the load dimension, 89 connectivity pairs in the θ band exhibited significant increases ($p < 0.05$), with only P7–AF3 and P8–P7 remaining non-significant ($p > 0.05$). Among them, 5 pairs fell within $0.01 \leq p < 0.05$, 13 within $0.001 \leq p < 0.01$, and 71 pairs reached $p < 0.001$. In the β band, 90 connectivitys showed significant enhancement ($p < 0.05$), except P7–F3. Further breakdown indicated 2 pairs within $0.01 \leq p < 0.05$, 3 within $0.001 \leq p < 0.01$, and 85 pairs at $p < 0.001$. In the time dimension, 65 connectivitys in the θ band varied significantly across time points ($p < 0.05$), including 21 pairs in $0.01 \leq p < 0.05$, 32 in $0.001 \leq p < 0.01$, and 12 at $p < 0.001$. In the β-band, analysis of variance indicated significant time effects for 68 connectivitys ($p < 0.05$), with 15 pairs in $0.01 \leq p < 0.05$, 12 in $0.001 \leq p < 0.01$, and 41 at $p < 0.001$ (Figure 2A).

All connectivitys were averaged at each time point to derive global indicators for the θ and β bands in order to minimize the influence of random fluctuations in individual connectivitys and better capture the overall trend. For the θ band, mean connectivity values were T1 = 0.2734 (95% CI = 0.2688–0.2779), T2 = 0.2805 (0.2758–0.2851), T3 = 0.2778 (0.2730–0.2825), T4 = 0.2777 (0.2728–0.2827), and T5 = 0.2643 (0.2601–0.2686). One-way ANOVA indicated a significant difference ($F(4, 10781) = 7.339$, $p < 0.001$). Bonferroni post hoc comparisons showed that T5 was significantly lower than T2 ($p < 0.001$), T3 ($p = 0.001$), and T4 ($p = 0.001$). For the β band, mean connectivity values were T1 = 0.1520 (95% CI = 0.1489–0.1551), T2 = 0.1595 (0.1565–0.1626), T3 = 0.1527 (0.1494–0.1560), T4 = 0.1526 (0.1494–0.1557), and T5 = 0.1513 (0.1481–0.1545).

ANOVA results confirmed a significant overall difference ($F(4, 10781) = 4.275$, $p = 0.002$). Bonferroni comparisons further revealed that T2 was significantly higher than T1 ($p = 0.010$), T3 ($p = 0.030$), T4 ($p = 0.024$), and T5 ($p = 0.003$), with no other significant differences among time points (Figure 2B).

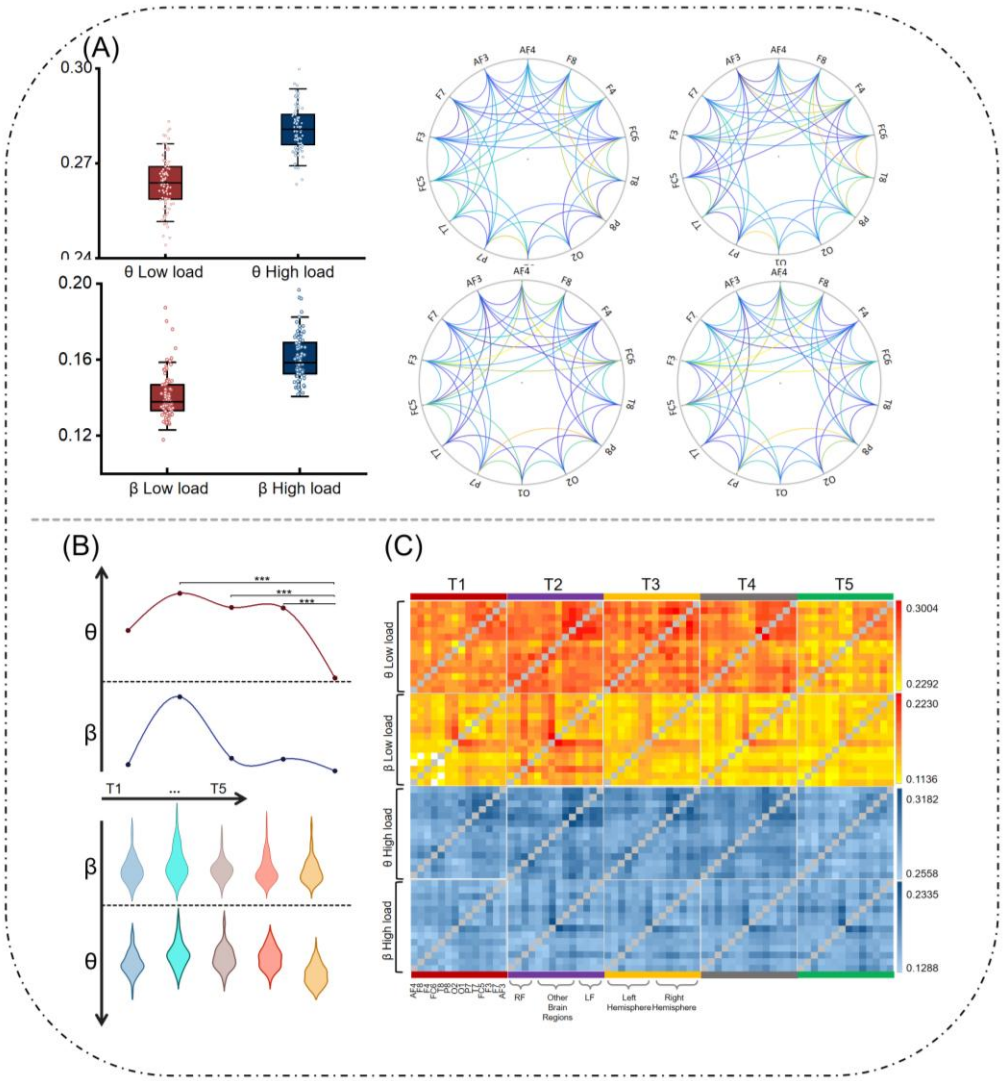


Figure 2. Functional connectivity results. (A) Boxplots illustrate the differences in mean functional connectivity strength between the LCL and HCL groups in the θ and β frequency bands, while chord diagrams depict the connectivity patterns under different load conditions; (B) line plots and violin plots present the temporal variations of mean functional connectivity across the five time points (T1–T5) in both θ and β bands; (C) connectivity matrices compare overall functional connectivity patterns across load levels and time points.

Table 1. Mean changes of selected functional connectivity pairs in θ and β bands over time

Frequency Band	Connectivity Pair	T1	T2	T3	T4	T5
θ	F7 - AF3	0.272	0.275	0.279	0.272	0.259
	F3 - AF3	0.264	0.286	0.274	0.285	0.269

	F3 - T8	0.270	0.288	0.277	0.275	0.262
	T7 - F3	0.269	0.279	0.274	0.278	0.253
	P7 - T7	0.271	0.283	0.280	0.276	0.264
	O2 - F7	0.264	0.268	0.268	0.266	0.252
	F7 - AF3	0.144	0.142	0.139	0.138	0.132
	F3 - AF3	0.144	0.169	0.166	0.167	0.161
β	F3 - T8	0.168	0.196	0.169	0.183	0.174
	T7 - F3	0.142	0.153	0.154	0.151	0.145
	P7 - T7	0.147	0.149	0.150	0.145	0.145
	O2 - F7	0.172	0.177	0.160	0.162	0.158

3.2Alterations in Local Brain Network Metrics

Statistical results for local network metrics focused on closeness centrality, node strength, and participation coefficient, with comparisons conducted across all electrodes. Regarding closeness centrality, the effect of cognitive load was highly significant: in both the θ and β bands, all electrodes showed significant decreases under high-load conditions compared with low-load conditions ($p < 0.001$). Across time, most electrodes in the θ band exhibited significant differences, except F7, P7, and O1 ($p = 0.252$, $p = 0.295$, and $p = 0.114$, respectively). In contrast, all electrodes in the β-band showed highly significant temporal differences ($p < 0.001$).

For node strength, clear distinctions were observed between high and low cognitive load conditions in the θ-band. All electrodes showed significant increases ($p < 0.05$) except P7 ($p = 0.369$) and P8 ($p = 0.219$). The β band revealed even stronger effects: all electrodes except P7 ($p = 0.996$) exhibited highly significant increases ($p < 0.001$). Regarding the time factor, seven electrodes in the θ band showed significant differences—F7 ($p = 0.003$), F3 ($p < 0.001$), P7 ($p = 0.002$), P8 ($p < 0.001$), FC6 ($p = 0.018$), F4 ($p = 0.020$), and AF4 ($p = 0.005$). In the β band, temporal effects were more concentrated, with only FC5 ($p = 0.255$) and FC6 ($p = 0.109$) failing to reach significance, while all other electrodes exhibited significant differences ($p < 0.05$).

For the participation coefficient, differences between load conditions were relatively limited. In the θ band, some electrodes showed increased values (e.g., AF3, F3, P7), while others decreased (e.g., F7, AF4). A similar trend appeared in the β band, with increases in some electrodes (e.g., AF3, F7, O1) and decreases in others (e.g., FC5, P7, O2). In the θ band, only FC6 ($p = 0.010$) and F8 ($p = 0.042$) reached statistical significance; all other electrodes showed nonsignificant differences. In the β band, more electrodes reached significance, including AF3 ($p < 0.001$), F7 ($p = 0.025$), P7 ($p = 0.045$), O1 ($p = 0.019$), and FC6 ($p < 0.001$). From a temporal perspective, significant effects were even fewer. In the θ band, only O2 ($p = 0.005$) and F8 ($p = 0.047$) showed significant temporal differences, whereas in the β band, significant effects were found at AF3 ($p = 0.009$), F7 ($p = 0.019$), P7 ($p = 0.030$), O1 ($p < 0.001$), P8 ($p = 0.026$), and AF4 ($p = 0.018$).

3.3 Alterations in Global Brain Network Metrics

In the analysis of global network metrics, six indicators—global efficiency, average closeness centrality, clustering coefficient, average node strength, network strength, and targeted attack robustness—were selected to systematically evaluate the overall characteristics of brain networks. Results for global efficiency and average closeness centrality revealed differences in information

transfer efficiency and overall network distance (Table 2). In the load dimension, global efficiency in the θ band was significantly higher under high-load conditions compared with low-load conditions (0.574 ± 0.054 vs. 0.571 ± 0.056), $t(10784) = 2.197$, $p = 0.028$. A similar pattern was observed in the β band, with significantly greater efficiency under high load (0.583 ± 0.052 vs. 0.580 ± 0.052), $t(10784) = 3.044$, $p = 0.002$. Across time, no significant differences were found in the θ band, $F(4,10781) = 0.097$, $p = 0.984$, whereas the β band exhibited a significant temporal effect, $F(4,10781) = 8.656$, $p < 0.001$.

Average closeness centrality showed an opposite trend relative to global efficiency. In the θ band, the high-load group was significantly lower than the low-load group (1.189 ± 0.392 vs. 1.252 ± 0.421), $t(10784) = -7.868$, $p < 0.001$; a similar result was found in the β band (2.134 ± 0.809 vs. 2.356 ± 0.838), $t(10784) = -13.638$, $p < 0.001$. Further analysis indicated that average closeness centrality varied significantly across time points in both the θ and β bands (θ: $F(4,10781) = 6.761$, $p < 0.001$; β: $F(4,10781) = 9.472$, $p < 0.001$).

For the clustering coefficient, no significant differences were found between load conditions. In the θ band, the low-load and high-load groups showed values of 0.235 ± 0.114 and 0.238 ± 0.117, respectively, $t(10784) = 1.196$, $p = 0.232$; in the β band, the values were 0.201 ± 0.109 and 0.200 ± 0.111, respectively, $t(10784) = -0.567$, $p = 0.571$. Temporal analysis revealed significant differences across time points in the θ band, $F(4,10781) = 2.587$, $p = 0.035$, but not in the β band, $F(4,10781) = 1.958$, $p = 0.098$.

Average node strength showed highly significant differences across load conditions. In the θ band, the high-load group exhibited greater values than the low-load group (1.616 ± 0.506 vs. 1.536 ± 0.491), $t(10784) = 8.101$, $p < 0.001$. The β band demonstrated a similar pattern, with higher values under high load (0.972 ± 0.413 vs. 0.869 ± 0.370), $t(10784) = 13.126$, $p < 0.001$. Across time, both frequency bands showed significant differences (θ: $F(4,10781) = 6.710$, $p < 0.001$; β: $F(4,10781) = 8.753$, $p < 0.001$).

For network strength, consistent effects were observed: in the θ band, the high-load group was significantly higher than the low-load group (9.697 ± 3.034 vs. 9.215 ± 2.945), $t(10784) = 8.101$, $p < 0.001$; and in the β band, the high-load group similarly exceeded the low-load group (5.833 ± 2.478 vs. 5.212 ± 2.219), $t(10784) = 13.126$, $p < 0.001$. Significant temporal effects were again found for both frequency bands (θ: $F(4,10781) = 6.710$, $p < 0.001$; β: $F(4,10781) = 8.753$, $p < 0.001$).

Regarding targeted attack robustness, the θ band showed significantly lower values under high-load conditions compared with low load (0.528 ± 0.117 vs. 0.533 ± 0.116), $t(10784) = -2.098$, $p = 0.036$. The β band demonstrated the same trend, with reduced robustness under high load (0.520 ± 0.115 vs. 0.527 ± 0.118), $t(10784) = -2.830$, $p = 0.005$. Across time, the θ band did not show significant differences, $F(4,10781) = 1.823$, $p = 0.121$, whereas the β band exhibited significant temporal variation, $F(4,10781) = 5.972$, $p < 0.001$.

Table 2. Differences in global metrics of θ and β bands under different cognitive load conditions

Global Metric	Frequency Band	Low Load (mean ± SD)	High Load (mean ± SD)
Global Efficiency	θ*	0.571±0.056	0.574±0.054
	β**	0.580 ± 0.052	0.583 ± 0.052

	Average Closeness	θ^{***}	1.252 ± 0.421	1.189 ± 0.392
	Centrality	β^{***}	2.356 ± 0.838	2.134 ± 0.809
3.4	Clustering Coefficient	θ	0.235 ± 0.114	0.238 ± 0.117
		β	0.201 ± 0.109	0.200 ± 0.111
	Average Node Strength	θ^{***}	1.536 ± 0.491	1.616 ± 0.506
		β^{***}	0.869 ± 0.370	0.972 ± 0.413
	Network Strength	θ^{***}	9.215 ± 2.945	9.697 ± 3.034
		β^{***}	5.212 ± 2.219	5.833 ± 2.478
	Targeted Attack	θ^{*}	0.533 ± 0.116	0.528 ± 0.117
	Robustness	β^{**}	0.527 ± 0.118	0.520 ± 0.115

Unsupervised Modeling of Fatigue States

During the feature processing stage, all functional connectivity measures and brain network metrics were first standardized and then reduced in dimensionality using PCA, ensuring that 90% of the cumulative variance was retained to preserve the primary information. In the reduced feature space, GMM was applied to cluster the samples into two latent states, and each sample was assigned a cluster label and corresponding confidence value based on posterior probabilities. A soft correction procedure incorporating multi-metric semantic mapping and temporal priors was then performed to refine the results, generating the final pseudo-labels—alert and fatigue. A confidence threshold of 0.70 was applied to exclude samples with insufficient model certainty. After filtering, 4,098 pseudo-labels were generated under the low-load condition, with 4,050 retained; under the high-load condition, 6,688 labels were generated and 6,613 retained. The average confidence across all retained samples reached 0.99 (Figure 3).

Under the low-load condition, unsupervised modeling revealed significant differences between the alert and fatigue states in functional connectivity within the θ-band (all $p < 0.001$, Hedges' $g = 0.55–0.81$), representing medium-to-large effect sizes. The β-band showed a similarly strong pattern (all $p < 0.001$, Hedges' $g = 0.60–1.06$), indicating large effects overall. For global network metrics, no significant differences were observed in overall efficiency between the two states for either the θ or β bands ($p > 0.05$); however, both average node strength and network strength significantly increased across bands (all $p < 0.001$). The clustering coefficient also increased significantly in both frequency bands (all $p < 0.001$), while average closeness centrality decreased markedly (all $p < 0.001$). Targeted attack robustness showed significant differences in both θ and β bands (all $p < 0.001$). At the local level, most nodes exhibited significant increases in clustering coefficient and node strength in both bands (all $p < 0.01$, Hedges' $g = 0.10–0.35$), indicating greater local connectivity and aggregation under fatigue, whereas participation coefficient showed no significant differences across nodes ($p > 0.05$).

Under the high-load condition, results were consistent. The θ band showed highly significant group differences in functional connectivity (all $p < 0.001$, Hedges' $g = 0.49–0.66$), corresponding to medium-to-large effect sizes. The β band demonstrated similarly robust effects (all $p < 0.001$, Hedges' $g = 0.51–0.91$), within the large-effect range. For global metrics, overall efficiency did not differ significantly between the alert and fatigue states in either the θ or β bands ($p > 0.05$). However, both average node strength and network strength increased significantly in both frequency bands (all $p < 0.001$, Hedges' $g = 0.32–0.53$). Likewise, clustering coefficient increased significantly (all $p < 0.001$, Hedges' $g = 0.14–0.20$), average closeness centrality decreased significantly (all $p <$

0.001, Hedges' g = 0.25–0.34), and targeted attack robustness differed significantly across both bands (all $p < 0.001$, Hedges' g = 0.12–0.22). At the local level, most nodes again showed significant increases in clustering coefficient and node strength in both bands (all $p < 0.01$, Hedges' g = 0.10–0.35), reflecting enhanced local connectivity and aggregation during fatigue, while participation coefficient remained statistically unchanged across all nodes ($p > 0.05$).

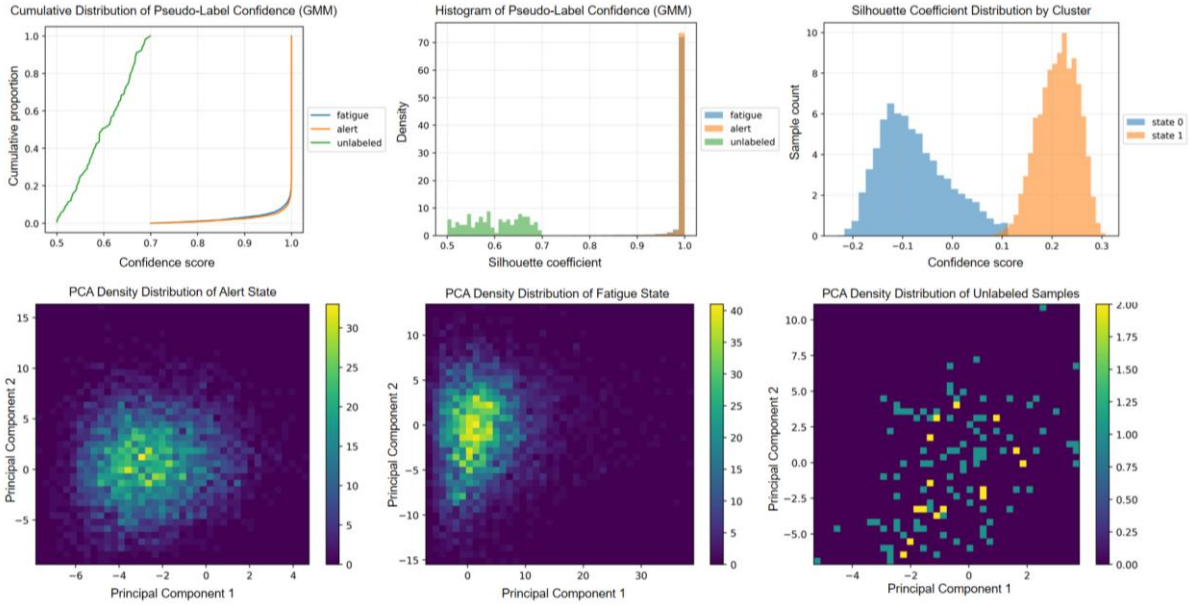


Figure 3. Reliability and distribution patterns of unsupervised clustering. The GMM confidence distribution shows that fatigue and alert labels are mainly concentrated in high-confidence regions, whereas unlabeled samples are located in low-confidence ranges. The silhouette distribution indicates distinct differences in cluster separation and compactness. PCA density maps illustrate label-specific patterns in the reduced space, with the x- and y-axes representing the first and second principal components, i.e., projections of the original features onto the directions of maximum variance. Alert states appear more compact, fatigue states more dispersed, and unlabeled samples more scattered.

3.5 Classification Performance of the Transformer Model

Under the low-load condition, the Transformer model exhibited stable binary classification performance across five-fold cross-validation (Figure 4A). Overall accuracy = 0.899–0.912 (M = 0.907); balanced accuracy = 0.898–0.918. For precision-related metrics, F1-macro = 0.886–0.903; macro-precision > 0.940; macro-recall (sensitivity) = 0.900–0.920; specificity = 0.880–0.940. In terms of discriminative capability, ROC AUC = 0.967–0.978; PR AUC = 0.985–0.989, both indicating high model performance. Matthews correlation coefficient (MCC) = 0.775–0.810, suggesting strong consistency between predicted and true labels. The averaged confusion matrix showed that, per fold, the model correctly identified fatigue samples (TP) = 484 and alert samples (TN) = 218, with false positives (FP) = 23 and false negatives (FN) = 49.

Under the high-load condition, the Transformer model also achieved stable performance (Figure 4B). Overall accuracy = 0.902–0.922 (M = 0.910); balanced accuracy = 0.898–0.921 (M = 0.909); F1-macro = 0.900–0.921 (M = 0.909); macro-precision = 0.899–0.941 (M = 0.924); macro-

recall (sensitivity) = 0.900–0.932 (M = 0.917); specificity = 0.869–0.930 (M = 0.902). Discriminative performance was similarly strong, with ROC AUC = 0.967–0.978 (M = 0.974) and PR AUC = 0.974–0.982 (M = 0.980). MCC = 0.800–0.842 (M = 0.818), again reflecting high agreement between predictions and true labels. The averaged confusion matrix indicated that, per fold, TP = 678, TN = 516, FP = 56, and FN = 61.

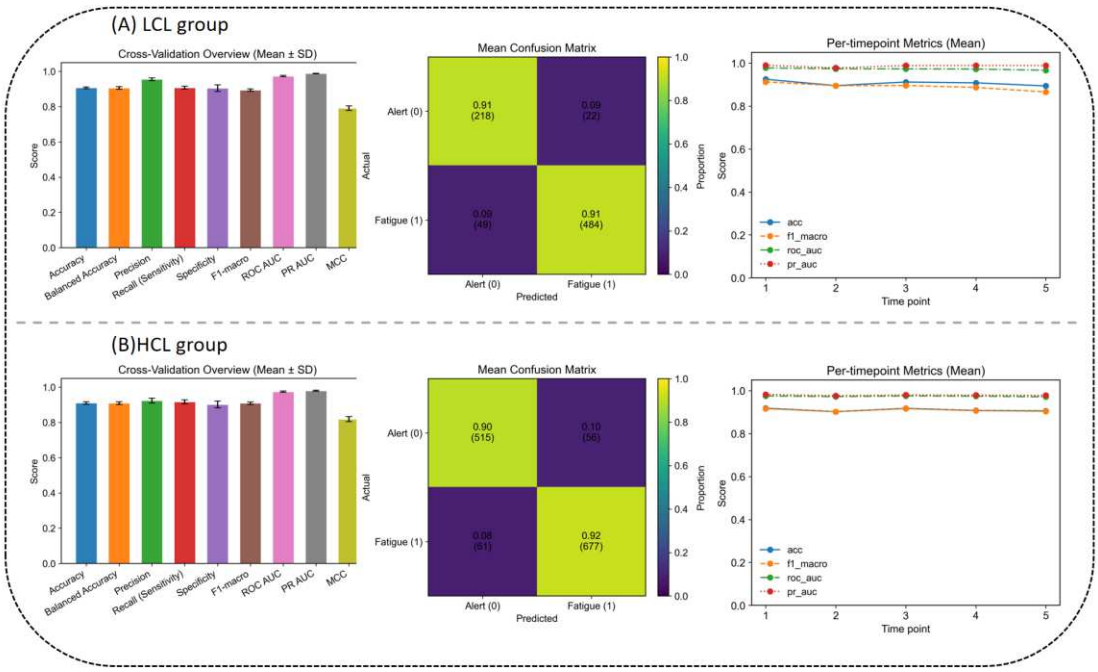


Figure 4. Classification performance of the Transformer architecture. (A) Low-load (LCL) group and (B) high-load (HCL) group models were trained and evaluated using five-fold cross-validation. The left panels show mean performance metrics with standard deviations, the middle panels present the mean confusion matrices, and the right panels illustrate per-time-point metric trajectories. The results indicate that the Transformer achieved stable discrimination between alert and fatigue states under both load conditions.

4. Discussion

4.1 Functional Connectivity Changes During Mental Fatigue Accumulation

Based on a high/low cognitive load × time-on-task experimental design, this study systematically revealed the evolution of functional connectivity during prolonged task engagement under varying cognitive load conditions. Overall, the strength of functional connectivity between most electrode pairs in both the θ and β bands significantly increased over time. This finding aligns with previous research (Xu et al., 2022), indicating that as task duration extends, the demand for interregional information exchange within the brain progressively increases, leading to greater synchronization. Previous research has shown that sustained cognitively demanding tasks, including extended Stroop paradigms, are accompanied by systematic increases in functional connectivity within executive control and attentional networks. Enhanced functional connections along

prefrontal–parietal pathways are often understood as reflecting compensatory coordination that helps operators maintain task performance under increasing cognitive demands (Wu et al., 2023).

It is noteworthy that the magnitude of increase in functional connectivity differed markedly between frequency bands: in the θ-band, the mean increase was approximately 6.8%, while in the β band, it exceeded 16%, with some electrode pairs showing increases greater than 20%. This suggests that β-band connectivity is more sensitive to variations in cognitive load. From a cognitive neuroscience perspective, this difference implies that in high-risk operational settings, drillers must mobilize additional neural resources to maintain vigilance and task performance, resulting in more extensive interregional synchronization. Typically, θ-band activity is closely associated with sustained attention and fatigue accumulation (Pomper & Ansorge, 2021), and its strengthened connectivity reflects enhanced interregional coupling required to maintain stable information processing under cognitive strain (Ueno et al., 2023). In contrast, β-band activity is linked to motor control, sustained attention, and executive demands; thus, increased β-band connectivity often reflects the recruitment of additional neural resources to sustain performance under high task demands (Jing et al., 2020). Accordingly, the marked increase in β-band connectivity observed under high-load conditions in this study can be interpreted as a neural coupling mechanism supporting vigilance and operational stability during extended, cognitively demanding work.

Across the temporal dimension, both θ and β bands showed a gradual rise from T1 to T4, reflecting compensatory enhancement of interregional information transfer under sustained cognitive load. However, at T5, connectivity declined—θ-band markedly and β-band moderately—indicating that fatigue accumulation was not linear. This reversal corresponds to the "second wind" or "wake maintenance zone" (WMZ), a transient recovery of alertness under extreme fatigue (McMahon et al., 2021; Zeeuw et al., 2018; Snipes et al., 2023). Previous studies attribute this effect to the interaction between circadian rhythm and sleep homeostasis, whereby wake-promoting circadian signals temporarily offset sleep pressure (Shekleton et al., 2013; Croce et al., 2018). In light of our results, the T5 decrease in connectivity likely reflects an adaptive neural recalibration—an active reduction of over-synchronization that restores network flexibility. Thus, the "second wind" observed here represents not a random fluctuation but a nonlinear self-regulatory adjustment of the brain to maintain functional stability during prolonged workload.

Integrating the time-series results of both functional connectivity and network metrics reveals a consistent pattern: the first four stages (T1–T4) exhibited progressive enhancement or sustained high levels across multiple indices, while the final stage (T5) presented a clear inflection point characterized by reversal or decline. This turning point represents the neural correlate of the "second wind" phenomenon. Originally a concept in exercise physiology describing temporary recovery following initial fatigue through metabolic and energy redistribution, its cognitive counterpart likely arises from multiple converging mechanisms. First, circadian rhythms exert a strong modulatory influence on alertness and brain function, and the reversal observed at T5 may reflect biological clock effects. Second, during prolonged high-load tasks, cognitive control mechanisms may be actively re-engaged—through motivational drive or attentional reallocation—to sustain task performance. Finally, fatigue adaptation mechanisms may contribute: under persistent fatigue, individuals may develop tolerance or adaptive adjustments that enable partial functional recovery within neural networks. Collectively, the results indicate that mental fatigue does not accumulate linearly but instead fluctuates dynamically under the joint influence of metabolic regulation, circadian rhythm, and cognitive control, with the "second wind" effect representing a salient

manifestation of this complex interplay.

4.2 Brain Network Alterations Associated With Mental Fatigue Induced by Cognitive Load and Time-on-Task

Regarding brain network topology, the results of this study demonstrate that cognitive load exerts consistent and significant effects on the structural characteristics of brain networks. At the local level, node strength across brain regions increased broadly, while closeness centrality decreased overall. This pattern indicates that although individual nodes became less central compared with the alert state, their connection weights significantly strengthened (Qin et al., 2022). The increase in node strength suggests enhanced synchronization and stronger local coupling between brain regions under fatigue, whereas the absence of significant changes in the clustering coefficient implies that local modular structures remained stable across different load conditions (Guo et al., 2025). Similarly, the participation coefficient did not differ significantly, suggesting that fatigue exerted minimal influence on the proportion of intermodular connections. Overall, the brain appeared to maintain stable modular organization while enhancing intra- and inter-module connection strength to facilitate more efficient information flow. Previous studies have reported that as fatigue deepens, global efficiency tends to decrease while local efficiency increases—a compensatory regulatory pattern. Xu et al. (2022), for example, found in a driving task that local efficiency increased with fatigue, reflecting a shift toward stronger local connectivity. Likewise, directional network analyses of driving fatigue have shown longer path lengths and reduced global efficiency, accompanied by increasing local clustering coefficients as fatigue accumulates, suggesting a gradual shift toward local information integration. The present findings are consistent with these reports at the local level: under fatigue, the brain retains its modular architecture but strengthens local connections to sustain core cognitive functions through enhanced local integration (Yan et al., 2024).

At the global level, the current results revealed significantly increased global efficiency under fatigue, decreased average closeness centrality, and unchanged clustering coefficients, alongside elevated mean node strength and network strength in both θ and β bands. These results indicate that during fatigue, the brain enhances interregional integration to improve information transfer efficiency while maintaining stable modular organization, consistent with the absence of significant changes in the participation coefficient. Notably, robustness to targeted attacks declined significantly under fatigue, suggesting that despite increased efficiency and connection strength, the network became more dependent on critical nodes, thereby reducing overall stability. Previous studies have shown that during mild fatigue, brain networks often exhibit shortened characteristic path lengths and increased global efficiency and node degree, temporarily improving information processing efficiency. For instance, Lou et al. (2025) induced cognitive fatigue using a Stroop task and observed increased global efficiency and clustering coefficient along with reduced path length, indicating that the brain transiently adopts a more optimized small-world configuration to counteract fatigue effects. Similarly, Yan et al. (2024) reported concurrent increases in node degree, clustering coefficient, and efficiency during mild fatigue, reflecting enhanced global coordination to meet task demands. Although the increase in global efficiency observed in this study differs from the declining trend reported in some laboratory studies, it may reflect a real-world adaptive mechanism, wherein

the brain compensates for cognitive decline by strengthening cross-regional resource integration under sustained fatigue. This form of global compensatory regulation likely contributes to maintaining cognitive performance in ecologically valid operational settings.

In summary, the results indicate that brain networks under fatigue exhibit a characteristic pattern of compensatory reorganization. Locally, the brain maintains modular stability by enhancing node strength and local coupling; globally, it improves interregional integration and overall efficiency. However, these gains come at the cost of increased dependence on critical nodes, reducing network robustness. Collectively, the brain network under fatigue demonstrates a dynamic pattern of "local reinforcement – global integration – weakened stability," reflecting an adaptive regulatory mechanism that enables the brain to sustain performance under constrained cognitive resources.

4.3 Unsupervised Modeling and Transformer-Based Recognition of Mental Fatigue States

The empirical results of this study demonstrate that mental fatigue exhibits both latent and nonlinear characteristics and is modulated by circadian phase. Consequently, relying solely on subjective ratings or single physiological indicators is prone to distortion. Prior evidence likewise indicates that surrogate measures correlate inconsistently with actual fatigue levels. For example, in physical fatigue contexts, the within-subject correlation between perceived exertion ratings and objective fatigue values can be as low as 0.05 (Ansari et al., 2022). This observation aligns with the temporal–circadian interactions revealed in our results, underscoring that subjective evaluation alone is insufficient for reliably differentiating fatigue states and highlighting the necessity of data-driven and machine learning frameworks (Mehmood et al., 2023).

At the feature level, label scarcity and annotation uncertainty remain major bottlenecks in fatigue recognition. Fatigue thresholds and subjective reports vary substantially across individuals, while objective labels are rarely available in field conditions. To address this, we employed a Gaussian Mixture Model (GMM) to perform unsupervised clustering on physiological features, uncovering latent fatigue-related patterns and generating pseudo-labels to guide model training. This approach is supported by substantial prior evidence: Ansari et al. (2022) applied GMM clustering to drivers' upper-body posture data to automatically identify fatigue- and non-fatigue-related postural patterns, then trained a supervised classifier using the generated pseudo-labels, achieving recognition accuracies of 93% and 90% at the individual level—closely matching EEG-based validation and outperforming fully supervised baselines. In light of our finding of the "second wind" phenomenon, caution is warranted to avoid conflating circadian phase with fatigue intensity during clustering. Accordingly, we propose introducing weak priors or temporal constraints into the GMM (e.g., assigning a higher probability of alertness to T1 and fatigue to T4) to enhance discriminability while preserving the flexibility of unsupervised learning.

The classification results of the deep learning model further demonstrate the practical feasibility of applying deep learning to fatigue recognition in real operational settings. Using pseudo-labeled EEG features derived from functional connectivity and brain network metrics, the Transformer model achieved consistently high classification accuracy under both low- and high-load conditions, indicating stable discrimination between alert and fatigued states. Importantly, this

performance was obtained in a field environment characterized by task variability and environmental noise, suggesting that Transformer-based models can effectively capture fatigue-related patterns without relying on strictly controlled laboratory conditions. By aligning temporal segmentation with actual work periods (T1 – T5), the model is able to learn stage-dependent changes, including the late-stage reversal observed in this study, rather than overfitting short-term fluctuations. Together, these findings support the applicability of deep learning – based approaches for fatigue monitoring in safety-critical operations.

4.4 Practical implications for safety performance under mental fatigue

This study observed, in a real operational context, that mental fatigue does not exhibit a simple linear accumulation pattern as work duration increases, but instead shows a clear stage-dependent reversal toward the end of the work period. Specifically, functional connectivity strength and multiple brain network metrics increased progressively during earlier stages and then declined at the final time point. From a safety management perspective, this stage-dependent reversal does not imply a natural elimination of risk; rather, it is more likely to reflect a compensatory self-regulation of the nervous system under conditions of high workload and pronounced fatigue, accompanied by reduced network stability and increased reliance on critical nodes. This indicates that although operators may display a certain degree of alertness recovery toward the end of a shift, the cognitive system as a whole remains in a structurally vulnerable state. In practical management, assessing late-stage safety risk solely on the basis of cumulative work duration or subjective fatigue perception may therefore lead to an underestimation of this latent risk. In contrast, the nonlinear evolution of functional connectivity and brain network characteristics revealed in this study suggests that safety management practices should integrate circadian rhythm and fatigue stage into task planning for later work periods, for example by avoiding the concentration of high cognitive load tasks or critical decision-making activities toward the end of shifts.

Building on this, the machine learning – based fatigue state recognition framework developed in this study provides a feasible approach for implementing feedforward safety control. By integrating functional connectivity and brain network features and employing an unsupervised pseudo-labeling strategy together with a Transformer architecture, this approach enables stable identification of workers’ fatigue states without reliance on subjective reports or fixed threshold assumptions. This capability is particularly important in high-risk operational settings, where mental fatigue often affects cognitive stability and decision reliability before becoming overtly observable at the behavioral level. Compared with post hoc interventions or passive management approaches based on work-hour rules, machine learning – based identification using EEG network features can capture the emergence and evolution of fatigue risk at an earlier stage, thereby providing safety managers with a basis for initiating interventions before risks translate into operational errors. As a result, fatigue management can gradually shift from dependence on static regulations or experiential judgment toward a feedforward safety control mode supported by real-time state identification, enabling dynamic adjustment of work organization, task reallocation, and the timing of critical decisions.

From a technological implementation perspective, this study demonstrates that functional

connectivity and brain network analyses based on wearable EEG can effectively capture neural signal changes associated with mental fatigue under field operational conditions, thereby supporting the identification of latent risks that are difficult to detect using conventional management approaches. Compared with post-event evaluations based on accident outcomes or behavioral performance, these neural indicators can reflect changes in cognitive system stability before operational errors occur, providing safety management with earlier risk information. On this basis, when combined with machine learning – based fatigue state recognition results, management practices can implement more targeted interventions without altering existing operational workflows, such as dynamically adjusting task allocation, introducing short-term recovery activities during high-risk periods, or temporarily lowering the risk level of critical operations when elevated fatigue risk is identified. By integrating EEG monitoring and intelligent recognition into safety decision-making processes, safety management has the potential to transition from a passive, experience-based, and reactive mode toward an active prevention framework characterized by real-time state awareness and forward-shifted risk control, thereby reducing the accumulation and amplification of mental fatigue – related safety hazards in high-risk operations.

5. Conclusion

Based on continuous EEG data collected from a real oil and gas drilling site, this study systematically examined the interactive effects of cognitive load and work duration on mental fatigue. The results indicate that as work time extended, functional connectivity strength in the θ and β frequency bands, along with multiple brain network metrics, exhibited a nonlinear trajectory characterized by an initial increase followed by a decline, demonstrating that mental fatigue does not accumulate in a linear manner. Under high cognitive load conditions, functional connectivity and network strength continued to increase, whereas network stability decreased accordingly, suggesting that while task performance was maintained, the underlying system became progressively more fragile. Furthermore, results from unsupervised modeling and Transformer-based classification showed that machine learning methods based on functional connectivity and brain network features were able to reliably distinguish between alert and fatigued states in field environments, validating the effectiveness of EEG-based brain network features for mental fatigue recognition.

From a practical perspective, this study highlights the central role of temporal factors and circadian rhythm in fatigue management and, based on experimental evidence, proposes several strategies for improving safety performance. These findings underscore the critical importance of circadian factors in fatigue management. The results have direct implications for safety management in high-risk industries, indicating that reliance on task intensity management alone is insufficient and that effective risk control strategies must integrate physiological time factors. Specifically, this study provides scientific support for the design of circadian-aligned shift schedules, optimization of work duration, and the development of real-time EEG-based monitoring and intervention schemes. Although the “second wind” phenomenon may temporarily enhance alertness, it cannot replace genuine rest. Therefore, during periods characterized by a “second wind,” lighter work tasks should be assigned to workers to mitigate potential safety risks. In addition, this study offers empirical support for the development of wearable EEG-based fatigue monitoring and intervention approaches, enabling managers to implement adjustments before fatigue-related risks become

manifest as operational errors.

The theoretical contribution of this study lies in demonstrating that fatigue does not follow a linear accumulation process. Mental fatigue is not monotonically increasing but instead exhibits complex nonlinear dynamics over time. This finding provides a valuable perspective for understanding the temporal evolution of fatigue. In particular, the observed “second wind” pattern offers initial empirical evidence that fatigue progression may involve fluctuations and stage-dependent characteristics rather than strictly linear deterioration. Moreover, this study is among the first to empirically reveal, within a real high-risk operational environment, the role of circadian rhythm and temporal factors in fatigue formation, showing that cognitive load and task duration jointly shape the development of mental fatigue.

Nevertheless, this study has several limitations. First, due to the constraints of field experimentation and the nature of high-risk operations, the sample size was relatively limited and data were drawn from a single industrial context. Although repeated measurements across multiple days and time points enhanced result stability and internal consistency, future studies should validate these findings using larger samples and across different high-risk industries to improve external generalizability. Second, while wearable EEG technology combined with functional connectivity, brain network analysis, and machine learning enabled a multi-level characterization of mental fatigue, environmental noise and the spatial resolution limitations of wearable devices make it difficult to completely exclude non-neural influences on the recorded signals. Future research could incorporate multimodal physiological measurements or higher-density neuroimaging techniques to further substantiate and refine the neural mechanisms underlying fatigue.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We want to extend our thanks to Chuanqing Drilling Engineering Company Chuanxi Drilling company limited of China National Petroleum Corporation for their invaluable assistance with the design of experimental materials and help with field experiments. The authors gratefully acknowledge the Energy Security and Low-Carbon Development Key Laboratory for providing research facilities and technical support. Some of the research conclusions were derived using the "Near-infrared spectroscopy imager（Fnirs）", "Event-related potential instrument（ERP/EEG）" and "Eye-tracker" equipment from the Energy Security and Low-Carbon Development Key Laboratory.

Funding sources

This work was supported by Sichuan Science and Technology Program (2024NSFSC2047), National Social Science Fund of China (23FJYB039).

Author contributions

Qing Xin, Su Hao contributed equally to this work and are considered co-first authors. **Methodology**: Qing Xin, Su Hao, Wang Xiaoqin, Wang Jian, Shen Liaoyuan; **Software**: Qing Xin, Su Haon; **Experiment Organization**: Qing Xin, Su Hao, Wang Xiaoqin, Wang Jian, Shen Liaoyuan; **Data Curation and Formal Analysis**: Qing Xin, Su Hao, Wang Jian; **Writing – Original Draft**: Qing Xin, Su Hao; All authors have read and approved the final manuscript.

References

[1] Ansari, S., Du, H., Naghdy, F., & Stirling, D. (2022). Automatic driver cognitive fatigue detection based on upper body posture variations. *Expert Systems with Applications*, 203, 117568. DOI: 10.1016/j.eswa.2022.117568.

[2] Anwer, S., Li, H., Antwi-Afari, M. F., Mirza, A. M., Rahman, M. A., Mehmood, I., ... & Wong, A. Y. L. (2024). Evaluation of data processing and artifact removal approaches used for physiological signals captured using wearable sensing devices during construction tasks. *Journal of Construction Engineering and Management*, 150(1), 03123008.

[3] Behrens, M., Gube, M., Chaabene, H., Prieske, O., Zenon, A., Broscheid, K. C., ... & Weippert, M. (2023). Fatigue and human performance: an updated framework. *Sports medicine*, *53*(1), 7-31.

[4] Bell, T., Sprajcer, M., Flenady, T., & Sahay, A. (2023). Fatigue in nurses and medication administration errors: A scoping review. *Journal of clinical nursing*, 32(17-18), 5445-5460.

[5] Chen, K., Chai, S., Xie, T., Liu, Q., & Ma, L. (2024). EEG spatial inter-channel connectivity analysis: A GCN-based dual stream approach to distinguish mental fatigue status. *Artificial Intelligence in Medicine*, 157, 102996.

[6] Ding, S., Hu, L., Pan, X., Zuo, D., & Sun, L. (2025). Assessing human situation awareness reliability considering fatigue and mood using EEG data: A Bayesian neural network-Bayesian network approach. *Reliability Engineering & System Safety*, 260, 110962.

[7] Ebrahim, Y., Molesworth, B., & Molloy, O. (2025). Investigating the predictive validity of subjective and objective measures for general aviation pilots' risk propensity. *Safety Science*, 186, 106818.

[8] Goodman, S. P., Collins, B., Shorter, K., Moreland, A. T., Papic, C., Hamlin, A. S., ... & Marino, F. E. (2025). Approaches to inducing mental fatigue: A systematic review and meta-analysis of (neuro) physiologic indices. *Behavior Research Methods*, 57(4), 102.

[9] Guo, H., Chen, S., Zhou, Y., Xu, T., Zhang, Y., & Ding, H. (2025). A hybrid critical channels and optimal feature subset selection framework for EEG fatigue recognition. *Scientific Reports*, 15(1), 2139.

[10] Han, Y., Huang, J., Yin, Y., & Chen, H. (2023). From brain to worksite: the role of fNIRS in cognitive studies and worker safety. *Frontiers in Public Health*, 11, 1256895.

[11] Harju, L. K., Seppälä, P., & Hakanen, J. J. (2023). Bored and exhausted? Profiles of boredom and exhaustion at work and the role of job stressors. *Journal of Vocational Behavior*, 144, 103898.

[12] Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4), 1009-1055.

[13] Hassan, O. F., Ibrahim, A. F., Gomaa, A., Makhlouf, M. A., & Hafiz, B. (2025). Real-time driver drowsiness detection using transformer architectures: A novel deep learning approach. *Scientific Reports*, 15(1), 17493. DOI: 10.1038/s41598-025-02111-x.

[14] IOGP. Safety performance indicators- 2016 data [Internet]. International Association of

Oil and Gas Producers; 2017 [cited 2017 Jul 6]. Available from:http://www.iogp.org/bookstore/?download_file=5494&order=wc_order_595e260c7e397&email=nwankwo.chizaram%40yahoo.com&key=23a5e5d5ad3e9a956fc64ea322a6429e

[15] Janssen, T. W., Grammer, J. K., Bleichner, M. G., Bulgarelli, C., Davidesco, I., Dikker, S., ... & van Atteveldt, N. (2021). Opportunities and limitations of mobile neuroimaging technologies in educational neuroscience. *Mind, Brain, and Education*, 15(4), 354-370.

[16] Jeon, J., & Cai, H. (2022). Multi-class classification of construction hazards via cognitive states assessment using wearable EEG. *Advanced Engineering Informatics*, 53, 101646.

[17] Jing, D., Liu, D., Zhang, S., & Guo, Z. (2020). Fatigue driving detection method based on EEG analysis in low-voltage and hypoxia plateau environment. *International journal of transportation science and technology*, 9(4), 366-376.

[18] Kar, S., Bhagat, M., & Routray, A. (2010). EEG signal analysis for the assessment and quantification of driver’s fatigue. *Transportation research part F: traffic psychology and behaviour*, 13(5), 297-306.

[19] Kim, M. S., Park, S., Park, U., Kang, S. W., & Kang, S. Y. (2024). Fatigue in Parkinson’s disease is due to decreased efficiency of the frontal network: quantitative EEG analysis. *Journal of Movement Disorders*, 17(3), 304.

[20] Lee, K., Choi, K. M., Park, S., Lee, S. H., & Im, C. H. (2022). Selection of the optimal channel configuration for implementing wearable EEG devices for the diagnosis of mild cognitive impairment. *Alzheimer's research & therapy*, 14(1), 170.

[21] Li, P., Zhang, Y., Liu, S., Lin, L., Zhang, H., Tang, T., & Gao, D. (2023). An EEG-based brain cognitive dynamic recognition network for representations of brain fatigue. *Applied Soft Computing*, 146, 110613.

[22] Lou, Y., Pi, R., Sun, R., Wu, J., Wang, W., Zhu, Z., ... & Gong, W. (2025). Graph theory-based analysis of functional connectivity changes in brain networks underlying cognitive fatigue: An EEG study. *PLoS One*, 20(8), e0329212.

[23] Mason, S. L., Junges, L., Woldman, W., Ftouni, S., Anderson, C., Terry, J. R., & Bagshaw, A. P. (2024). Associating EEG functional networks and the effect of sleep deprivation as measured using psychomotor vigilance tests. *Scientific Reports*, 14(1), 27999.

[24] Mehmood, I., Li, H., Qarout, Y., Umer, W., Anwer, S., Wu, H., ... & Antwi-Afari, M. F. (2023). Deep learning-based construction equipment operators’ mental fatigue classification using wearable EEG sensor data. *Advanced Engineering Informatics*, 56, 101978.

[25] Mudrik, L., Hirschhorn, R., & Korisky, U. (2024). Taking consciousness for real: Increasing the ecological validity of the study of conscious vs. unconscious processes. *Neuron*, 112(10), 1642-1656.

[26] Pomper, U., & Ansorge, U. (2021). Theta-rhythmic oscillation of working memory performance. *Psychological science*, 32(11), 1801-1810.

[27] Prabhakar, S. K., & Won, D. O. (2023). Multiple robust approaches for EEG-based driving fatigue detection and classification. *Array*, 19, 100320.

[28] Proost, M., Habay, J., De Wachter, J., De Pauw, K., Rattray, B., Meeusen, R., ... & Van Cutsem, J. (2022). How to tackle mental fatigue: a systematic review of potential countermeasures and their underlying mechanisms. *Sports Medicine*, 52(9), 2129-2158.

[29] Qin, Y., Hu, Z., Chen, Y., Liu, J., Jiang, L., Che, Y., & Han, C. (2022). Directed brain network analysis for fatigue driving based on EEG source signals. *Entropy*, 24(8), 1093.

[30] Staiano, W., Bonet, L. R. S., Romagnoli, M., & Ring, C. (2023). Mental fatigue: the cost of cognitive loading on weight lifting, resistance training, and cycling performance. *International Journal of Sports Physiology and Performance*, 18(5), 465-473.

[31] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

[32] Tran, Y., Craig, A., Craig, R., Chai, R., & Nguyen, H. (2020). The influence of mental fatigue on brain activity: Evidence from a systematic review with meta‐analyses. *Psychophysiology*, 57(5), e13554.

[33] Ueno, K., Ishii, R., Ueda, M., Yuri, T., Shiroma, C., Hata, M., & Naito, Y. (2023). Frontal midline theta rhythm and gamma activity measured by sheet-type wearable EEG device. *Frontiers in human neuroscience*, 17, 1145282.

[34] Umer, W., Yu, Y., Afari, M. F. A., Anwer, S., & Jamal, A. (2023). Towards automated physical fatigue monitoring and prediction among construction workers using physiological signals: An on-site study. *Safety Science*, 166, 106242. DOI: 10.1016/j.ssci.2023.106242.

[35] Wang, F., Wu, S., Ping, J., Xu, Z., & Chu, H. (2021). EEG driving fatigue detection with PDC-based brain functional network. *IEEE Sensors Journal*, 21(9), 10811-10823.

[36] Wang, Y., Huang, Y., Gu, B., Cao, S., & Fang, D. (2023). Identifying mental fatigue of construction workers using EEG and deep learning. *Automation in Construction*, 151, 104887.

[37] Wiegmann, D. A., & Shappell, S. A. (2017). A human error approach to aviation accident analysis: The human factors analysis and classification system. *Routledge*.

[38] Wu, C. H., De Doncker, W., & Kuppuswamy, A. (2023). Electroencephalography-derived functional connectivity in sensorimotor networks in post stroke fatigue. *Brain Topography*, 36(5), 727-735.

[39] Wu, D., Deng, L., Lu, Q., & Liu, S. (2025). A multidimensional adaptive transformer network for fatigue detection. *Cognitive Neurodynamics*, 19(1), 43. DOI: 10.1007/s11571-025-10224-2.

[40] Wu, Y., Li, W., Zhang, J., Tang, B., Xiang, J., Li, S., & Guo, G. (2023). Driver's hand-foot coordination and global-regional brain functional connectivity under fatigue: Via graph theory and explainable artificial intelligence. IEEE Transactions on Intelligent Vehicles, 9(2), 3493-3508.

[41] Xu, T., Xu, L., Zhang, H., Ji, Z., Li, J., Bezerianos, A., & Wang, H. (2022). Effects of rest-break on mental fatigue recovery based on EEG dynamic functional connectivity. Biomedical Signal Processing and Control, 77, 103806.

[42] Yan, W., He, J., Peng, Y., Ma, H., & Li, C. (2024). Research on brain functional network property analysis and recognition methods targeting brain fatigue. Scientific Reports, 14(1), 22556.

[43] Zeeuw, J. D., Wisniewski, S., Papakonstantinou, A., Bes, F., Wahnschaffe, A., Zaleska, M., ... & Münch, M. (2018). The alerting effect of the wake maintenance zone during 40 hours of sleep deprivation. Scientific reports, 8(1), 11012.

[44] Zhang, Y., Guo, H., Zhou, Y., Xu, C., & Liao, Y. (2023). Recognising drivers' mental

fatigue based on EEG multi-dimensional feature selection and fusion. Biomedical Signal Processing and Control, 79, 104237. DOI: 10.1016/j.bspc.2022.104237.

[45] Zhao, P., Lian, C., Xu, B., & Zeng, Z. (2024). Multiscale global prompt transformer for EEG-based driver fatigue recognition. IEEE Transactions on Automation Science and Engineering, 22, 2700-2711.