

Machine Learning-Based Fatigue Detection from EEG Signals: A Methodological Pilot Framework with Potential for Nursing Applications

● Serife Cetin,¹ ● Eda Dagdevir²

¹Department of Medical Services and Techniques, Incesu Vocational School of Health Services, Kayseri University, Kayseri, Türkiye

²Department of Electronic and Automation, Vocational School of Technical Science, Kayseri University, Kayseri, Türkiye

ABSTRACT

Objective: Fatigue is a significant occupational health issue that adversely affects attention and decision-making. This study aimed to objectively predict fatigue levels using EEG signals and machine learning models and to develop a physiologically based assessment framework with potential future applications in nursing.

Materials and Methods: Time-domain, frequency-domain, entropy, Hjorth, and wavelet features were extracted from EEG data collected from 12 participants aged 21–40 years, with fatigue levels labeled using various scales.

Results: Classification performance varied across validation strategies. Under 5-fold cross-validation, Random Forest achieved the highest performance (accuracy=1.00, macro-F1=1.00), while KNN and NB also demonstrated strong performance. However, under LOOCV, sensitivity, precision, and macro-F1 values decreased despite relatively high accuracy. LOSO validation yielded more balanced results, with Random Forest maintaining the best performance (accuracy=0.97, macro-F1=0.91).

Conclusion: EEG-derived features provide valuable information for fatigue assessment. Although high performance was achieved under certain validation strategies, subject-independent evaluation revealed more realistic performance levels. The proposed framework offers a practical approach to fatigue monitoring. Future studies should focus on larger, domain-specific datasets and multimodal integration to enhance generalizability.

Keywords: EEG, fatigue prediction, machine learning, nursing.



Cite this article as:

Cetin S, Dagdevir E. Machine Learning-Based Fatigue Detection from EEG Signals: A Methodological Pilot Framework with Potential for Nursing Applications. J Clin Pract Res 2026;48(5):530–539.

Address for correspondence:

Eda Dagdevir.
Department of Electronic and Automation, Vocational School of Technical Science, Kayseri University, Kayseri, Türkiye
Phone: +90 549 607 00 50
E-mail:
edadagdevir@kayseri.edu.tr

Submitted: 05.02.2026

Revised: 04.06.2026

Accepted: 09.09.2026

Available Online: 30.09.2026

Erciyes University Faculty of
Medicine Publications -
Available online at www.jcpres.com

INTRODUCTION

Fatigue is a complex phenomenon involving both central and peripheral components. Central fatigue is linked to dysfunction within the central nervous system, resulting in reduced motivation, attention, and cognitive capacity.^{1–3} Neuroimaging studies reveal that structural and functional changes, along with neurotransmitter imbalances in the prefrontal cortex and basal ganglia, play a critical role in this process.⁴ Peripheral fatigue is associated with physiological changes at the

Copyright © Author(s)
This work is licensed under
a Creative Commons
Attribution-NonCommercial
4.0 International License.



muscle level; force production decreases as a result of reduced muscle contractility, irregularities in calcium release, and disruptions in energy metabolism.⁵

The subjective and self-reported nature of fatigue presents significant challenges for its measurement. The existence of more than 250 assessment scales reflects its complexity and marked interindividual variability.^{6,7} Current evidence indicates that fatigue is a multidimensional phenomenon involving interacting central and psychological processes, leading to declines in both physical and mental performance.⁸

Prolonged shift durations and insufficient rest periods are major contributors to nurse fatigue, particularly in critical care settings. Fatigue negatively affects nurses' physical and mental health, increasing the risk of errors and compromising patient safety.⁹ Nurse fatigue has been linked to higher rates of medication errors, patient falls, and other safety incidents.^{10–13} Notably, Scott et al.¹² reported a 28% increase in patient safety incidents as shift duration increased. Consequently, machine learning–based approaches integrating subjective and objective fatigue measures offer promising strategies for understanding and managing fatigue in nursing and other professional groups.

In recent years, physiologically based objective measures have been increasingly used to address the limitations of subjective fatigue scales. Wearable technologies enabling automatic monitoring of indicators such as heart rate variability and sleep quality have advanced fatigue detection models.¹⁴ Electroencephalography (EEG) has emerged as a key indicator in fatigue detection because it directly reflects changes in mental state.¹⁵ Xu et al.¹⁶ showed that deep learning models can effectively capture complex EEG patterns, achieving 94.51% accuracy with a convolutional neural network (CNN) applied to multichannel EEG data. Yang et al.¹⁷ developed a hybrid fatigue detection model using publicly available EEG data from athletes. The model achieved 97.1% accuracy in classifying fatigue-related EEG patterns, demonstrating strong performance in early mental fatigue detection and performance optimization. Several studies have demonstrated the effectiveness of EEG-based AI models for fatigue detection, particularly in driving applications; reported accuracy rates using both traditional machine learning and deep learning approaches generally exceed 90%.^{18–22} These findings highlight the high accuracy and reliability of EEG-based AI models in driver fatigue detection and their potential to enhance driving safety. Chen et al.²³ assessed mental fatigue in healthcare personnel using multiscale entropy features extracted from single-channel EEG signals. Fatigue classification using Random Forest, XGBoost, and backpropagation neural network models achieved area under the curve (AUC) values of up to 95%, supporting EEG as a robust biomarker for mental fatigue detection.

KEY MESSAGES

- EEG-based machine learning models can effectively estimate fatigue using multidomain signal features.
- A data-driven hybrid labeling approach enables a more reliable representation of fatigue than single-scale methods.
- The proposed framework provides a practical and scalable approach for physiologically based fatigue monitoring in clinical settings.

While Chen et al.²³ presented a high-performance classification approach, the present study offers a more comprehensive and novel methodological pilot framework by integrating multiple fatigue scales and multidimensional EEG features. Given the critical importance of fatigue for patient safety, the limited evidence on objective fatigue assessment approaches specific to nursing represents a notable gap in the literature. In this context, rather than proposing a nursing-specific model, the present study provides a physiologically based, general methodological pilot framework with strong potential for future application in nursing by considering factors such as cognitive load and shift work conditions. This approach aims to serve as a foundation for future research focused on supporting employee well-being, enhancing patient safety, and improving the quality of care.

Aims and Objectives

This study introduces a novel framework for EEG-based fatigue detection by integrating multiple self-reported fatigue scales into a single composite label and combining this labeling strategy with a multidomain feature extraction and evaluation pipeline. In this respect, rather than proposing a model specifically tailored to nursing, the study presents a promising, physiologically based general methodological pilot framework with strong potential for future application in the nursing field. Unlike traditional approaches that rely on a single subjective scale, the proposed method uses a data-driven weighting scheme to combine different scale scores and create a composite fatigue index. To the best of our knowledge, such a hybrid label-generation strategy has not previously been reported in the literature.

In addition, the study employs a comprehensive feature representation that captures complementary information from the time, frequency, and time–frequency domains. Furthermore, the classification framework is evaluated using multiple validation strategies, such as 5-fold cross-validation, leave-one-out cross-validation (LOOCV), and leave-one-subject-out (LOSO) validation, enabling both sample-level and subject-independent performance assessments.

The results demonstrate that the proposed framework is more effective at capturing broad fatigue-related differences while also highlighting the challenges of subject-independent generalization. Overall, this study contributes to the literature by proposing a data-driven labeling strategy, integrating multidomain EEG features, and providing a robust evaluation framework for fatigue detection.

MATERIALS AND METHODS

Study Design

This study was designed as an investigation integrating both descriptive and methodological components. The descriptive aspect aims to reveal the relationships among variables in the existing dataset, while the methodological aspect seeks to develop a preliminary methodological framework for physiologically based models applicable to nurses using EEG-based machine learning approaches. This framework includes the development of a fatigue assessment model based on objective biophysiological indicators and a systematic evaluation of its potential applicability in nursing practice.

Participants and Dataset

The dataset, collected in 2021 as part of a study conducted by Kalanadhabhatta and colleagues, included a total of 12 participants (mean age=30.75±5.78 years), comprising 9 males and 3 females, aged 21–40 years. One participant had mild asthma, another had a history of seasonal asthma, and none of the other participants had any current or previous health issues. All participants provided informed consent before the study.²⁴

Experimental Procedure

The original study included three sessions conducted on separate days, with a maximum interval of 19 days, all held at the same time of day to control for circadian effects. Demographic data were collected before the first session. Personality traits (BFI-10) and chronotype (MCTQ) were assessed at baseline. Daily fatigue (FSS) and physical fitness (IFIS) were also evaluated. Before each session, sleepiness (SSS) and energy and mood (GVAS) were measured. Participants were equipped with four wearable devices to monitor physiological signals during each session. EEG data were recorded using the Muse device from an open-access dataset. Participants completed three sessions involving baseline EEG recording, physical exercise, cognitive tasks, and repeated fatigue assessments while physiological signals were continuously recorded.^{25–29} Detailed information regarding the scales used in this study is provided in Supplementary File 1.

EEG Recording and Preprocessing

The flowchart for this study is shown in Figure 1. The process consists of five main stages: labeling, data preprocessing,

feature extraction, classification, and performance evaluation. This flowchart provides a comprehensive overview of the process from EEG data to fatigue estimation.

Questionnaire-Based Labeling

In this study, participants' fatigue was assessed using scores from four self-report instruments: the SSS, GVAS, FSS, and IFIS. According to the dataset providers, some participants had missing FSS values. These missing values were estimated using SSS and GVAS data through a logistic regression model (fitglm), thereby minimizing information loss due to missing data.

Because each scale measures fatigue using different metrics, the scale scores were normalized and combined using the weighted-average method described in Equation 1. The weights were determined based on the variance values of the scales.

$$M_w = \frac{a.SSS + b.FSS + GVAS + d.IFIS}{a + b + c + d} \quad (1)$$

where a, b, c, and d denote the weighting coefficients corresponding to the SSS, FSS, GVAS, and IFIS, respectively, and M_w is the weighted mean score.

The variance was computed across all samples as defined in Equation 2. The corresponding weights (a, b, c, and d) were derived from these variance values, allowing each scale to contribute proportionally to its variability within the dataset.

$$w_i = \frac{var(x_i)}{\sum var(x_i)} \quad (2)$$

where var(x_i) represents the variance of the ith scale, and w_i represents its corresponding weight. This normalization ensures that the sum of all weights equals one.

In this context, scales with higher variance contribute more to the overall fatigue score because they provide greater discriminatory power for distinguishing fatigue levels among participants.

To the best of our knowledge, previous studies have not explicitly reported the integration of multiple fatigue-related scales into a single composite label using a data-driven weighting strategy.

The overall fatigue scores were divided into two categories using a data-driven threshold. Specifically, samples with fatigue scores less than or equal to the median value were labeled as low fatigue (Class 0), whereas those above the median were labeled as high fatigue (Class 1).

Feature Extraction and Selection

EEG signals were recorded from participants using the Muse device at a sampling frequency of 256 Hz from the TP9, AF7,

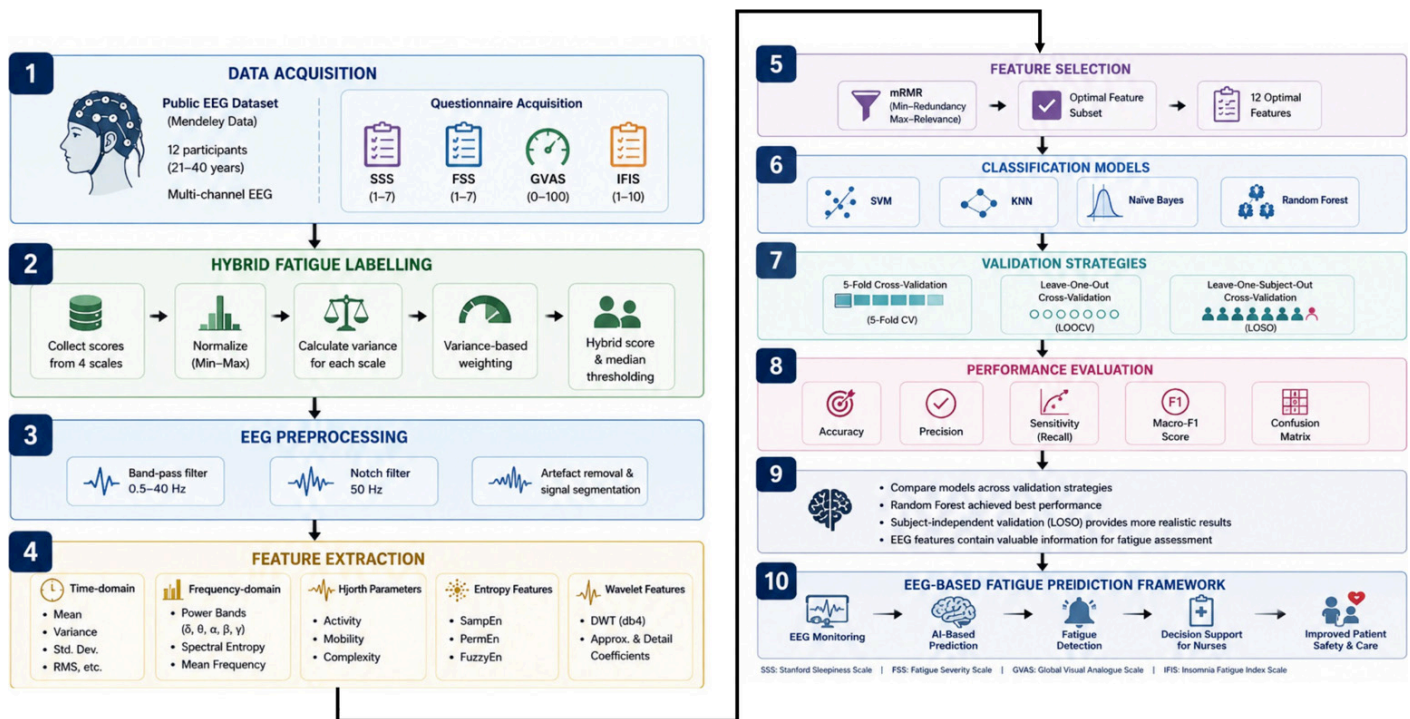


Figure 1. Flowchart of the EEG-based fatigue detection system.

AF8, and TP10 channels. The signals were then filtered using a 0.5–40 Hz Butterworth band-pass filter and a 50 Hz notch filter to remove noise.

The following feature sets were extracted for each participant and each recording:

- Frequency-based features: band powers in the delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz) frequency bands, along with theta/alpha and theta/beta power ratios.
- Hjorth parameters: activity, mobility, and complexity, computed separately for each frequency band.
- Entropy features: spectral entropy computed for each frequency band to capture frequency-specific signal complexity.
- Time-domain features: mean, standard deviation, skewness, kurtosis, and root mean square (RMS) of the signal.
- Wavelet features: normalized subband energies derived from a 4-level discrete wavelet transform using the Daubechies 4 (db4) wavelet, including approximation (A4) and detail coefficients (D1–D4).

All features were computed on a per-channel basis and subsequently averaged across channels to obtain a participant-level representation.

To reduce feature dimensionality and improve generalization performance, feature selection was conducted using the minimum redundancy maximum relevance (mRMR) method. The mRMR algorithm ranks features by maximizing their relevance to the target variable while minimizing redundancy among features.³⁰ Using this procedure, the optimal feature subset was determined to comprise 12 features.

To prevent information leakage, feature selection was conducted within each cross-validation fold using only the training data. The selected features were then applied to the corresponding test data.

Classification Models and Evaluation Metrics

The extracted features were classified using support vector machine (SVM), Naive Bayes (NB), k-nearest neighbors (KNN), and Random Forest (RF) classification models based on the generated label information. Three validation strategies were employed to evaluate model performance under different validation conditions: 5-fold cross-validation, LOOCV, and LOSO validation.

In 5-fold cross-validation, the dataset was randomly divided into five subsets. In each iteration, four subsets were used for training, and one subset was used for testing. LOOCV was implemented to maximize the use of the limited data, with each sample serving once as the test sample. Additionally, LOSO validation

Table 1. Fatigue scales and weighting coefficients

Scales	SSS (a)	FSS (b)	GVAS (c)	IFIS (d)
Weight	0.19	0.42	0.25	0.15

SSS: Stanford Sleepiness Scale; FSS: Fatigue Severity Scale; GVAS: Global Energy and Affect Scale; IFIS: International Fitness Scale.

was performed to evaluate subject-independent performance, with all samples from one participant used as the test set and samples from the remaining participants used for training.

This multilevel validation framework facilitates the evaluation of both sample-level performance and subject-level generalization.

Ethical Approval

This study utilized publicly available anonymized data. Therefore, ethics committee approval was not required. The results were classified in (MATLAB R2024b, The MathWorks, Inc., Natick, MA, USA) using machine learning techniques.

RESULTS

Results of Labeling

In this study, participants' overall fatigue was assessed using four different survey scales: the SSS, GVAS, FSS, and IFIS. Because each scale measures fatigue using different metrics, the scores were normalized and combined using a weighted-average method. The weights were determined based on the variance values of the scales, as shown in Table 1. The overall fatigue score was then calculated according to Equation 1.

Because no standardized clinical threshold for fatigue categorization was available in the literature, a data-driven

approach was adopted. The overall fatigue scores were divided into two classes using the median value as the threshold. Samples with scores less than or equal to the median were assigned to the low-fatigue class (Class 0), whereas those with scores above the median were assigned to the high-fatigue class (Class 1).

Statistical Structure of EEG Features

A feature matrix comprising 35 samples and 128 features was constructed. Features were extracted separately from each of the four EEG channels (TP9, AF7, AF8, and TP10), preserving channel-specific information rather than averaging across channels. The samples corresponded to 12 participants with 3 trials each, excluding one missing recording, while the feature set represented EEG characteristics extracted from multiple domains.

Frequency-domain features included relative band powers in the delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz) bands, along with theta/alpha and theta/beta power ratios. In addition, Hjorth parameters (activity, mobility, and complexity) and spectral entropy were computed separately for each frequency band to capture band-specific signal dynamics.

Time-domain features included the mean, standard deviation, skewness, kurtosis, and root mean square (RMS) values of the EEG signal.

For time–frequency analysis, the EEG signal was decomposed into four levels using the Daubechies 4 (db4) wavelet, resulting in one approximation subband (A4) and four detail subbands (D1–D4).

Table 2. Classification results

Validation	Model	Accuracy	Sensitivity	Precision	Macro-F1
5-Fold	SVM	0.77	0.78	0.76	0.73
5-Fold	KNN	0.83	0.83	0.86	0.82
5-Fold	NB	0.83	0.84	0.85	0.83
5-Fold	RF	1.00	1.00	1.00	1.00
LOOCV	SVM	0.77	0.39	0.39	0.39
LOOCV	KNN	0.89	0.44	0.44	0.44
LOOCV	NB	0.86	0.43	0.43	0.43
LOOCV	RF	0.97	0.49	0.49	0.49
LOSO	SVM	0.65	0.61	0.49	0.51
LOSO	KNN	0.85	0.78	0.76	0.75
LOSO	NB	0.83	0.76	0.76	0.74
LOSO	RF	0.97	0.90	0.92	0.91

SVM: Support vector machine; KNN: K-nearest neighbors; NB: Naive Bayes; RF: Random Forest; LOOCV: Leave-one-out cross-validation; LOSO: Leave-one-subject-out.

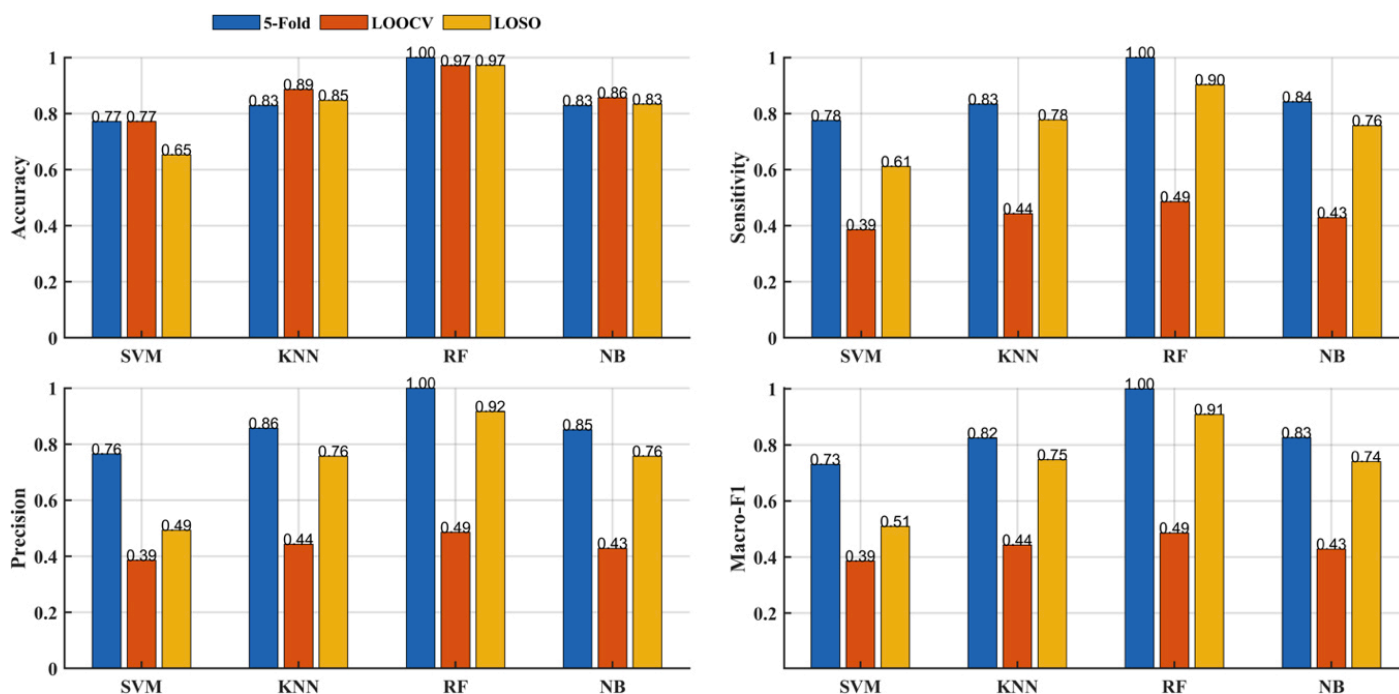


Figure 2. Performance metrics of classifiers across different validation strategies.

Results of Classification

The performance of four different classifiers (SVM, KNN, NB, and RF) was evaluated using multiple validation strategies, including 5-fold cross-validation, LOOCV, and LOSO validation.

For the SVM model, a radial basis function (RBF) kernel was employed; the KNN model used a neighbor count of $k=3$; the NB classifier was implemented assuming a Gaussian distribution; and the RF model was constructed using an ensemble of 100 decision trees.

The classification results are presented in Table 2, including accuracy, sensitivity, precision, and class-specific F1 scores. Additionally, the comparative performance of the classifiers across the different validation strategies is illustrated in Figure 2.

The classification performance of the four classifiers (SVM, KNN, NB, and RF) across the three validation strategies (5-fold CV, LOOCV, and LOSO) is presented in Table 2 and Figure 2. Under the 5-fold cross-validation scheme, the Random Forest (RF) classifier achieved perfect performance across all metrics (accuracy=1.00, macro-F1=1.00), while KNN and NB also demonstrated strong performance, with comparable macro-F1 scores of 0.82 and 0.83, respectively. In contrast, SVM exhibited relatively lower performance in this setting.

In the LOOCV scenario, although RF maintained high accuracy (0.97), sensitivity, precision, and the macro-F1 score decreased

substantially to approximately 0.49, indicating potential instability in generalization when evaluated using single-sample test folds. A similar trend was observed for SVM, KNN, and NB, for which accuracy remained relatively high, but class-specific performance metrics declined significantly.

LOSO validation, which provides a more realistic subject-independent evaluation, revealed a different performance pattern. RF again achieved the best overall performance (accuracy=0.97, macro-F1=0.91), followed by KNN and NB, which showed consistent and balanced results. In contrast, SVM demonstrated notably lower performance (macro-F1=0.51), indicating limited generalization across participants. Overall, RF emerged as the most robust classifier across all validation strategies, while KNN and NB provided competitive and stable alternatives, particularly under subject-independent evaluation conditions.

DISCUSSION

This study demonstrates the applicability of EEG-based machine learning models for fatigue detection. A notable discrepancy was observed among the validation strategies. While the Random Forest model achieved perfect performance under 5-fold cross-validation (accuracy=1.00), its performance decreased substantially under LOOCV and LOSO evaluations. This suggests that the 5-fold cross-validation results may have been influenced by data partitioning and may provide overly

Table 3. Comparison of related studies

Study	Year	Data	Classifier	Best accuracy (%)
Ren et al. ¹⁹	2021	EEG	RBF	92.71
Xu et al. ¹⁶	2023	EEG	CNN	94.5
Chen et al. ²³	2024	EEG	RF	87.0
Shen et al. ¹⁵	2025	Multimodal	Hybrid DL	96.1
Yang et al. ¹⁷	2024	EEG	LSTM	97.2
Geng et al. ²⁰	2025	EEG	SVM	94.32
Hafiz et al. ²²	2025	HRV	SVM	81.48
This study	2026	EEG	RF	100

ANN: Artificial neural network; CNN: Convolutional neural network; DL: Deep learning; HRV: Heart rate variability; LSTM: Long short-term memory network; RBF: Radial basis function; RF: Random forest; SVM: Support vector machine.

optimistic performance estimates, particularly given the limited sample size. In contrast, LOSO offers a more realistic subject-independent evaluation and revealed more moderate but reliable performance levels. Therefore, these results highlight the importance of employing subject-independent validation strategies when assessing EEG-based fatigue detection models. In addition to these findings, a notable decrease in performance was observed under LOOCV compared with 5-fold cross-validation. Although accuracy remained relatively high, sensitivity, precision, and macro-F1 scores declined substantially. This discrepancy may be attributed to the nature of LOOCV, in which each model is trained on a slightly reduced dataset and evaluated on a single sample, making performance more sensitive to data variability and noise.

In addition, the limited sample size and inherent variability of EEG signals may further amplify this effect, leading to unstable performance estimates. Therefore, LOOCV results should be interpreted with caution, as they may not fully reflect the generalization capability of the models. In contrast, LOSO validation provides a more realistic subject-independent evaluation and yielded more balanced performance results, suggesting that it may be a more appropriate validation strategy for EEG-based fatigue prediction tasks.

This study demonstrates a strong association between fatigue classes derived from self-report scales (SSS, FSS, GVAS, and IFIS) and EEG features, suggesting that EEG-derived features are associated with both the cognitive and physical aspects of fatigue.

The EEG-based classification results of this study provide a promising, physiologically based general methodological pilot framework with strong potential for future application in nursing. Given the need for real-time monitoring in high-risk settings (e.g., ICUs, emergency departments, operating rooms, and dialysis centers), these findings may inform the design of

future workload regulation systems. Consistent with the EEG-based fatigue framework of Liu et al.,³¹ the proposed modeling steps contribute methodologically to the development of algorithms for healthcare professionals working under high levels of stress. Moreover, in line with Alshammari et al.,³² the focus on fatigue-related factors highlights the role of such frameworks in improving worker well-being and reducing errors.

The applicability of this model in real-world clinical nursing settings presents several practical considerations. The comfort of wearable EEG devices, their suitability for prolonged use and user mobility, and their seamless integration into clinical workflows are critical issues. To this end, future systems should incorporate robust artifact removal techniques, real-time signal-processing algorithms, and user-friendly wearable designs.

The results obtained in this study reveal notable differences compared with the existing literature summarized in Table 3. Previous studies indicate that complex deep learning models (e.g., CNNs, LSTMs, and hybrid networks) generally achieve high accuracy in fatigue classification.

For instance, Xu et al.¹⁶ reported 94.5% accuracy using a CNN. Yang et al.¹⁷ achieved 97.2% accuracy using an LSTM.¹⁷ Xu et al.¹⁶ achieved high accuracy using CNN-based deep learning on multichannel EEG data, while Yang et al.¹⁷ demonstrated strong performance in fatigue classification using EEG data from athletes. In contrast, the present study differs by employing interpretable multidomain EEG feature extraction, introducing a composite fatigue-labeling strategy based on multiple self-report scales, and emphasizing subject-independent validation within a clinically oriented framework. Shen et al.¹⁵ reported 96.1% accuracy using a multimodal hybrid approach. However, more traditional machine learning models, such as those employed by Ren et al.¹⁹ and Geng et al.,²⁰ also demonstrated classification performance above 90%. Chen et al.²³ used a comprehensive feature-

engineering process involving multiscale entropy extraction from EEG signals and reported 87% accuracy with a Random Forest classifier. Nevertheless, this method is complex and computationally intensive because of the feature extraction process. Hafiz et al.²² focused solely on fatigue detection in nurses using HRV and achieved 81.48% accuracy with HRV features and an SVM classifier. Although HRV signals are relatively inexpensive to obtain, they only partially represent neurophysiological changes compared with EEG. Because HRV does not directly represent neural activity, it may not fully capture cortical-level processes such as cognitive load, reduced attention, or mental fatigue. In contrast to these studies, the present work emphasizes a balance between model simplicity and performance, offering a methodological framework that prioritizes interpretability and potential real-time applicability over maximum accuracy.

Among the evaluated classifiers, Random Forest generally demonstrated the highest performance across validation strategies, particularly under LOSO conditions, indicating its robustness to intersubject variability. KNN and NB also showed stable and competitive performance, whereas SVM exhibited comparatively lower generalization capability. These findings suggest that ensemble-based methods may be more suitable for EEG-based fatigue classification tasks.

Limitations

This study has several limitations. The small sample size ($n=12$), the use of an open-access dataset, and the fact that the EEG data were obtained from healthy individuals performing tasks in a laboratory setting rather than from nurses limit the generalizability of the findings and render them preliminary. The presence of two participants with a history of asthma should also be considered a potential limitation. The use of only four EEG channels may not capture all fatigue-related cortical activity, while self-report-based labeling using a percentile-based approach may introduce subjective bias because of the lack of clinically validated cutoff values. Accordingly, this study serves as a starting point for future nursing-focused research conducted in clinical settings with larger samples and multichannel EEG recordings.

CONCLUSION

This study demonstrates that fatigue can be objectively estimated by combining multidimensional EEG-derived features with machine learning models and presents a promising, physiologically based general methodological pilot framework with strong potential for future application in nursing. The pilot model achieved meaningful accuracy when integrated with self-report scales, indicating that EEG-derived features provide valuable information related to fatigue.

These findings highlight the potential of physiologically based fatigue-monitoring systems to enhance employee safety, patient safety, and continuity of care while providing a scientific basis for larger nursing-specific studies. Future research should include nurse samples, employ multichannel or high-density EEG, and incorporate performance and biophysiological measures to further improve model robustness and generalizability.

Ethics Committee Approval: Since this study utilized publicly available anonymized data, no ethics committee approval was required.

Informed Consent: Written informed consent was not required, as the study utilized publicly available anonymized data.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors declared that this study received no financial support.

Use of AI for Writing Assistance: Generative AI technology was used solely to generate Figure 1, which was subsequently edited.

Author Contributions: Concept – SC, ED; Design – SC, ED; Supervision – SC, ED; Resource – SC, ED; Materials – ED; Data Collection and/or Processing – SC; Analysis and/or Interpretation – SC, ED; Literature Review – SC, ED; Writing – SC, ED; Critical Review – SC, ED.

Acknowledgments: We sincerely thank the researchers who made the open-access EEG dataset publicly available, enabling the development of this methodological framework.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Chen J, Zhou Y, Rao H, Liu J. Mental Fatigue in Parkinson's Disease: Systematic Review and Evaluation of Self-Reported Fatigue Scales. *Parkinsons Dis* 2024;2024:9614163. [\[CrossRef\]](#)
2. Sari N, Tumanggor RD. Analysis of nurses' fatigue and stress factors in the COVID-19 isolation room in Indonesia: a cross-sectional research. *Caring: IJNS* 2023;5(1):1-8. [\[CrossRef\]](#)
3. Müller T, Apps MAJ. Motivational fatigue: A neurocognitive framework for the impact of effortful exertion on subsequent motivation. *Neuropsychologia* 2019;123:141-51. [\[CrossRef\]](#)
4. Penner IK, Paul F. Fatigue as a symptom or comorbidity of neurological diseases. *Nat Rev Neurol* 2017;13(11):662-75. [\[CrossRef\]](#)
5. Sarmento A, Fregonezi G, Lira M, Marques L, Pennati F, Resqueti V, et al. Changes in electromyographic activity, mechanical power, and relaxation rates following inspiratory ribcage muscle fatigue. *Sci Rep* 2021;11(1):12475. [\[CrossRef\]](#)

6. Skau S, Sundberg K, Kuhn HG. A Proposal for a Unifying Set of Definitions of Fatigue. *Front Psychol* 2021;12:739764. [\[CrossRef\]](#)
7. Kluger BM, Krupp LB, Enoka RM. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. *Neurology* 2013;80(4):409-16. [\[CrossRef\]](#)
8. Matthews J, Pisauro MA, Jurgelis M, Müller T, Vassena E, Chong TTJ, et al. Computational mechanisms underlying the dynamics of physical and cognitive fatigue. *Cognition* 2023;240:105603. [\[CrossRef\]](#)
9. Abbaszadeh R, Ahmadi F, Khoobi M, Kazemnejad A, Vaismoradi M. A Scoping Review of Fatigue Among Nurses in Critical Care Units. *Nurs Crit Care* 2025;30(4):e70088. [\[CrossRef\]](#)
10. Bagheri I, Salmani N, Mandegari Z, Pakcheshm B, Dadgari A. Evaluation of medication errors from the perspective of nurses in the intensive care units of Yazd City. *JSSU* 2021;29(3). [\[CrossRef\]](#)
11. Bernstrøm VH, Alves DE, Ellingsen D, Ingelsrud MH. Healthy working time arrangements for healthcare personnel and patients: a systematic literature review. *BMC Health Serv Res* 2019;19(1):193. [\[CrossRef\]](#)
12. Scott LD, Arslanian-Engoren C, Engoren MC. Association of sleep and fatigue with decision regret among critical care nurses. *Am J Crit Care* 2014;23(1):13-23. [\[CrossRef\]](#)
13. Farag A, Gallagher J, Carr L. Examining the Relationship Between Nurse Fatigue, Alertness, and Medication Errors. *West J Nurs Res* 2024;46(4):288-95. [\[CrossRef\]](#)
14. Liu Y, Xian J, Wang R, Ma K, Li F, Wang FL, et al. Factoring and correlation in sleep, fatigue and mental workload of clinical first-line nurses in the post-pandemic era of COVID-19: a multi-center cross-sectional study. *Front Psychiatry* 2022;13:963419. [\[CrossRef\]](#)
15. Shen D, Tao X, Končar V, Wang J. A hybrid deep learning framework combined with intelligent garment system for real-time fatigue monitoring in cycling. *AATCC Journal of Research* 2025;12(4). [\[CrossRef\]](#)
16. Xu L, Li J, Feng D. Miner Fatigue Detection from Electroencephalogram-Based Relative Power Spectral Topography Using Convolutional Neural Network. *Sensors (Basel)* 2023;23(22):9055. [\[CrossRef\]](#)
17. Yang Z. Fatigue driving detection technology based on electroencephalogram signal. In: Arthur H, Huang F, editors. 2nd International Conference on Artificial Intelligence, Database and Machine Learning (AIDML 2024); 2024 July 6-7; Rotterdam, Netherlands; 2024.pp. 1022-29. [\[CrossRef\]](#)
18. Gao Z, Li Y, Yang Y, Ma C. A recurrence network-based convolutional neural network for fatigue driving detection from electroencephalogram. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 2019;29(11):113126. [\[CrossRef\]](#)
19. Ren Z, Li R, Chen B, Zhang H, Ma Y, Wang C, et al. EEG-Based Driving Fatigue Detection Using a Two-Level Learning Hierarchy Radial Basis Function. *Front Neurobot* 2021;15:618408. [\[CrossRef\]](#)
20. Geng X, Hu W, Liang Q, Chen C, Zhang X, Yu P, et al. An improved feature extraction algorithm for electroencephalogram-based driving fatigue recognition. *Sci Rep* 2025;15:33226. [\[CrossRef\]](#)
21. Wang L, Johnson D, Lin Y. Using electroencephalogram to detect driving fatigue based on common spatial pattern and support vector machine. *Turk J Elec Eng & Comp Sci* 2021;29(3):1429-44. [\[CrossRef\]](#)
22. Hafiz WS, Puspasari MA, Fitriani DY, Hanowski RJ, Syaifullah DH, Arista SA. Developing a fatigue detection model for hospital nurses using HRV measures and machine learning. *Safety* 2025;11(2):48. [\[CrossRef\]](#)
23. Chen Z, Zhang Z, Zhang Z, Lin S, Li X, Liu H. Research on the analysis method of mental fatigue in medical personnel based on EEG signals. *Research Square*. 2024 Jan 11. doi: 10.21203/rs.3.rs-3848113/v. [Epub ahead-of-print]. [\[CrossRef\]](#)
24. Kalanadhabhatta M, Min C, Montanari A, Kawsar F. FatigueSet: a multimodal dataset for modeling mental fatigue and fatigability. In: Lewy H, Barkan R, editors. *Pervasive Computing Technologies for Healthcare. Proceedings of International Conference on Pervasive Computing Technologies for Healthcare*; 2021 Dec 6-8; Tel Aviv, Israel. Springer Nature; 2022.pp.204-17. [\[CrossRef\]](#)
25. MacLean AW, Fekken GC, Saskin P, Knowles JB. Psychometric evaluation of the Stanford Sleepiness Scale. *J Sleep Res* 1992;1(1):35-9. [\[CrossRef\]](#)
26. Monk TH. A Visual Analogue Scale technique to measure global vigor and affect. *Psychiatry Res* 1989;27(1):89-99. [\[CrossRef\]](#)
27. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989;46(10):1121-3. [\[CrossRef\]](#)
28. Armutlu K, Korkmaz NC, Keser I, Sumbuloglu V, Akbiyik DI, Guney Z, et al. The validity and reliability of the Fatigue Severity Scale in Turkish multiple sclerosis patients. *Int J Rehabil Res* 2007;30(1):81-5. [\[CrossRef\]](#)
29. Ortega FB, Ruiz JR, España-Romero V, Vicente-Rodriguez G, Martínez-Gómez D, Manios Y, et al.; HELENA study group.

- The International Fitness Scale (IFIS): usefulness of self-reported fitness in youth. *Int J Epidemiol* 2011;40(3):701-11. [\[CrossRef\]](#)
30. Dayan V, Chileshe N, Hassanli R, Parvaneh A. Sustainable bridge management using refined feature selection for machine learning-aided bridge condition prediction: Incorporation of Pareto distribution in MRMR method. *Applied Soft Computing* 2025;184(Part B):113878. [\[CrossRef\]](#)
31. Liu J, Zhou Y, He Q, Gao Z. ASFT-transformer: a fast and accurate framework for electroencephalogram-based pilot fatigue recognition. *Sensors* 2025;25(19):6256. [\[CrossRef\]](#)
32. Alshammari B, AlEnazy A, Alshammari F, Madkhali N, Al-Masaeed M. Investigation of the Level and Factors Influencing Emergency Department Nurses Fatigue: A Case Study of the Saudi Arabian Context. *Healthcare (Basel)* 2022;10(7):1294. [\[CrossRef\]](#)