Machine Learning-Based Classification of Disorders of Consciousness Using Multi-Dimensional EEG Features

Binghua Song*

School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China

Abstract. Disorders of Consciousness (DOC), including Vegetative State (VS) and Minimally Conscious State (MCS), present significant diagnostic challenges owing to the inherent subjectivity of behavioral assessments. Electroencephalography (EEG) provides a noninvasive and objective modality for evaluation. This study proposes a multi-dimensional EEG feature extraction framework encompassing temporal, spectral, and spatial domains. To classify VS and MCS, machine learning models including Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), and Convolutional Neural Network (CNN) were employed. Experimental results demonstrate that the SVM model achieved the highest performance, with a classification accuracy of 93% and an F1 score of 96%. Slow-wave and variability-related features were identified as the most significant contributors to model performance. These findings underscore the potential of multi-dimensional EEG features to enhance diagnostic accuracy, offering substantial clinical value for improving patient management and therapeutic outcomes.

1 Introduction

Disorders of Consciousness (DOC) represent a spectrum of severe neurological conditions following acquired brain injury, characterized by distinct disruptions in arousal and awareness. Clinically, the Vegetative State (VS) is marked by preserved wakefulness in the absence of discernible awareness, whereas the Minimally Conscious State (MCS) exhibits inconsistent but reproducible behavioral evidence of consciousness [1, 2]. Accurate discrimination between VS and MCS is imperative for prognosis, therapeutic decision-making, and resource allocation. However, conventional neurobehavioral assessments, such as the Glasgow Coma Scale (GCS), are limited by subjectivity, susceptibility to patient fatigue, and fluctuating responsiveness [3, 4].

To overcome these limitations, advanced neuroimaging techniques such as Computed Tomography (CT) and functional Magnetic Resonance Imaging (fMRI) have been utilized to assess structural and functional integrity. Despite their high spatial resolution, clinical utility is often restricted by low temporal resolution, high costs, and poor portability. Conversely, Electroencephalography (EEG) has emerged as a robust alternative due to its noninvasiveness, low cost, and high temporal resolution. Nevertheless, EEG signals in DOC patients frequently suffer from nonstationarity, low signal-to-noise ratios (SNR), and significant inter-individual variability, posing challenges for reliable automated analysis [5, 6].

Previous studies have explored diverse EEG biomarkers for DOC classification, spanning temporal, spectral, and spatial domains [7-9]. Although Machine Learning (ML) and Deep Learning (DL) models, such as Support Vector Machines (SVM) and Convolutional Neural Networks (CNN), have demonstrated promising performance, existing methods often lack a systematic approach for multi-domain feature integration, and may rely on fixed and non-scalable feature sets. For instance, Tolonen et al. [10] achieved an accuracy of 87% using alpha power features, whereas Vespa et al. [11] reported 86% with alpha variability features. Such a narrow focus may overlook the potential synergistic relationships among these domains. Furthermore, the field still lacks a standardized and adaptable data processing and analysis pipeline capable of accommodating the heterogeneity of clinical EEG data. To address this gap, the present study proposes a multi-dimensional EEG feature extraction framework that flexibly incorporates or excludes features of interest depending on data conditions.

This framework aims to establish a standardized paradigm for EEG data processing in DOC classification. Specifically, this study makes two primary contributions. First, we propose a feature fusion strategy to effectively integrate information derived from multiple feature domains. Second, we employ interpretability methods to analyze the contribution of individual features, thereby translating the predictions of black-box models into clinically interpretable insights and enhancing the reliability of automated diagnostic systems.

* Corresponding author: 220232485@seu.edu.cn

2 Methods

2.1 Data acquisition

The EEG data used in this study were obtained from a publicly available dataset: "Polysomnographic records of patients with disorders of consciousness" published on Mendeley Data in 2020 [12]. This dataset comprised 42 long-term polysomnographic (PSG) recordings, obtained from a total of 42 unique patients. Each record lasted for approximately 6 hours. The cohort consisted of patients with DOC, including VS and MCS. These patients were recruited from the Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitation, Russia.

Following a strict quality assessment, 6 recordings were excluded from the analysis due to excessive artifacts or poor signal quality (5 VS and 1 MCS). Consequently, a final cohort of 36 patients was included in this study, consisting of 25 patients in VS and 11 patients in MCS.

Data were recorded using the SOMNOScreen Plus PSG equipment at a sampling frequency of 256 Hz and stored in .edf format. Each record included two electrooculogram (EOG) channels, one electromyogram (EMG) channel, and a variable number of EEG channels depending on recording quality.

2.2 Data preprocessing

To ensure the consistency of channel configuration among different records, in this study, the 6 common channels F3, F4, C3, C4, O1 and O2 were uniformly selected for subsequent analysis, covering the frontal lobe, parietal lobe and occipital lobe regions.

The selected EEG recordings were quantitatively analyzed using Python 3.11 and the MNE-Python toolbox (version 1.6.0) [13]. Before feature extraction, a robust preprocessing pipeline was implemented to improve signal quality and reduce physiological and environmental artifacts.

First, a zero-phase notch filter was applied at 50 Hz and its harmonics to suppress power-line interference. Subsequently, the signals were processed using a 0.1-35 Hz band-pass finite impulse response (FIR) filter with zero-phase filtering to retain the physiologically relevant EEG frequency components.

Next, independent component analysis (ICA) was performed to decompose the EEG signals into independent components. Components associated with ocular artifacts, muscle activity, and cardiac interference were visually identified and removed by an experienced domain expert.

Finally, the continuous EEG recordings were segmented into 30-minute epochs for subsequent analysis. EEG recordings shorter than 30 minutes were excluded to ensure the relative stability and reliability of long-term physiological measurements.

2.3 Multi-domain feature extraction

The calculation of EEG data features was based on YASA (version 0.6.3) [14], and SciPy (version 1.11.4) [15] toolboxes. To comprehensively describe the EEG data characteristics of DOC patients, the features were extracted from three different domains:

2.3.1 Temporal domain

To quantify the amplitude and variability of the signal, including the mean, variance, skewness, kurtosis, peak-to-peak value, root mean square (RMS), and Shannon entropy [16], etc. These indicators measure the dynamic fluctuation characteristics of the time series from different aspects.

2.3.2 Spectral domain

Power spectral density (PSD) for each EEG channel was estimated using Welch's method. The estimation utilized a 10-second Hamming window with a 50% overlap. The frequency bands of interest were defined as follows: delta (0.1 - 4 Hz), theta (4 - 8 Hz), alpha (8 - 13 Hz), beta (13 - 20 Hz), and gamma (30 - 35 Hz). Absolute power for each frequency band was calculated by integrating the PSD over the corresponding frequency range. Total power was defined as the sum of power across all analyzed frequency bands, and relative power was subsequently determined as the ratio of the absolute power of a specific band to the total power.

Furthermore, two power ratios—the alpha-to-delta ratio (ADR) and the delta-theta to alpha-beta Ratio (DTABR)—were calculated to characterize the spectral balance.

To evaluate inter-electrode synchronization, spectral coherence was estimated. This analysis employed a 10-second Hamming window with a 5-second overlap to ensure estimation stability. A global coherence feature was then derived by averaging the coherence values across all possible channel pairs.

Finally, power variability was quantified to assess the temporal fluctuations of spectral power. This was measured by the ratio of the median absolute deviation (MAD) to the median power [10].

2.3.3 Spatial domain

To quantify the symmetry of EEG power distribution between the left and right hemispheres, the Brain Symmetry Index (BSI) was calculated for specific frequency bands [17]. The BSI is derived from the mean relative power difference between homologous electrode pairs across hemispheres and is defined as equation 1.

$$BSI(t) = \left| \frac{1}{N}\frac{1}{M}\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{R_{ij}(t)-L_{ij}(t)}{R_{ij}(t)+L_{ij}(t)} \right| \qquad (1)$$

Where N denotes the number of homologous electrode pairs, and M represents the number of discrete frequency samples within the analyzed band. L_{ij} and R_{ij} correspond to the spectral power of the electrode pair in the left and right hemispheres, respectively, at time t. A larger BSI value indicates stronger inter-hemispheric asymmetry in EEG power distribution, whereas values closer to zero reflect more symmetrical neural activity between the two hemispheres.

Furthermore, the Center of Gravity (COG) was employed to characterize the spatial topography of EEG power across the scalp [18]. The COG is defined as equation 2. where M denotes the total number of electrodes, d_{xy} represents the spatial coordinate of the each electrode and P_{ij} corresponds to the spectral power at each electrode.

$$COG(i) = \frac{\sum_{ij}^{M} P_{ij} \cdot d_{xy}}{\sum_{ij}^{M} P_{ij}} \quad (2)$$

This metric reflects the focal point of power distribution: negative values indicate a distribution bias toward the left temporal or occipital regions, while positive values signify a shift toward the right temporal or frontal regions. A COG value near 0 indicates a balanced, centrally-weighted power distribution.

2.3.4 Statistical analysis

Statistical analyses were performed to evaluate group differences in clinical outcomes and to conduct feature selection. Data processing was implemented using the SciPy toolbox in Python. The normality of continuous variables was assessed prior to group comparisons using the Shapiro-Wilk test. For variables that followed a normal distribution, independent-samples t-tests were used to compare differences between groups. For variables that did not meet the assumption of normality, the Mann-Whitney U test was applied to assess statistical significance. For all statistical analyses, a two-tailed p-value < 0.05 was considered to indicate statistical significance.

2.3.5 Classification models and evaluation

To differentiate between VS and MCS and evaluate prognostic performance, we implemented four classic pattern recognition algorithms: SVM, RF, LR and CNN. SVM was selected for its superior generalization capability in high-dimensional spaces by maximizing the classification margin. RF, an ensemble method based on decision trees, was utilized for its robustness against overfitting and its ability to handle complex nonlinear relationships. LR was included due to its high interpretability and efficiency in providing probabilistic outputs for linearly separable problems.

Based on the statistical analysis described above, a subset of EEG features with significant differences between groups was selected as input variables for the classification model. The resulting input features encompassed multiple domains, including temporal, spectral, and spatial characteristics of the EEG signals.

To assess the generalization performance of the proposed model, a 5-fold cross-validation strategy was employed.

Model performance was comprehensively evaluated using several standard classification metrics, including accuracy, area under the receiver operating characteristic curve (AUC-ROC), recall, and the F1 score. The inclusion of the F1 score ensures a robust evaluation despite the inherent class imbalance between VS and MCS groups.

3 Results

3.1 EEG features

For each EEG segment and for every channel, we computed a comprehensive feature set comprising 20 time domain features, 29 spectral domain features, and 18 spatial domain features. Figure 1 illustrates the distribution of two representative features: slow-wave proportion and gamma variability across six selected channels for the entire dataset. Statistical analysis of these distributions reveals that patients in VS exhibited a significantly higher proportion of slow-wave activity coupled with reduced variability in the high-frequency gamma band compared to those in MCS. These findings indicate a systematic alteration of neurophysiological oscillations across distinct frequency scales in DOC patients.

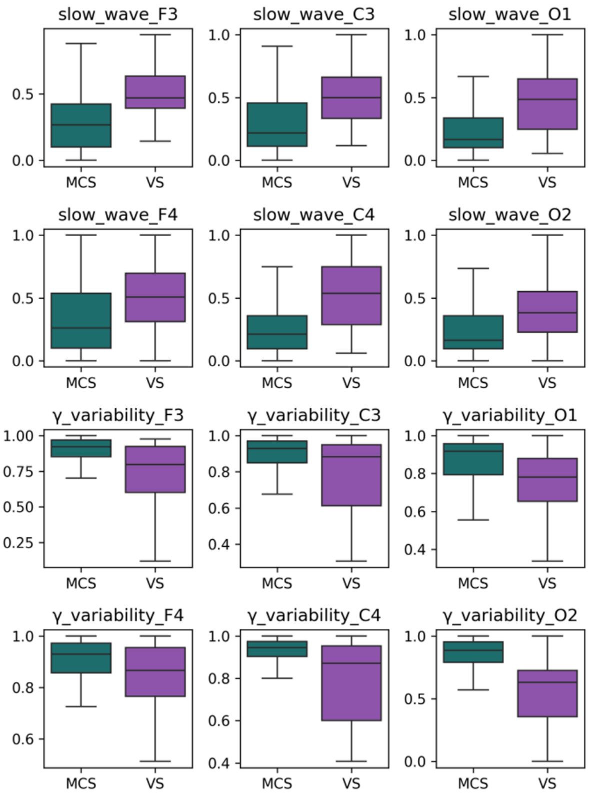

Figure 1. Boxplots of Slow-wave proportion and gamma variability across six EEG channels.

3.2 Comparative performance of classification models

Table 1 summarizes the results across multiple metrics. Among the tested algorithms, the SVM consistently outperformed the others, achieving the highest classification accuracy of 93% and an F1 score of 96%. In contrast, the LR model achieved an accuracy of 78% and an F1 score of 86%, while the RF model obtained 88%

and 93%, respectively. Although the CNN model achieved comparable accuracy of 92% and F1 score of 95%, we selected the SVM model for further analysis due to its superior balance between predictive power and interpretability.

Table 1. Evaluation results for each model

Model	Accuracy	AUC	Recall	F1
SVM	0.93	0.97	0.95	0.96
LR	0.78	0.77	0.79	0.86
RF	0.88	0.90	0.90	0.93
CNN	0.92	0.97	0.95	0.95

3.3. Feature significance and interpretability analysis

Figure 2 shows the permutation importance ranking of EEG features to the SVM classifier. The results indicated that slow-wave activity and variability features contributed most strongly to the model performance. In particular, slow-wave power in central and frontal regions showed the highest contribution, with the feature from C3 exhibiting the largest importance value.

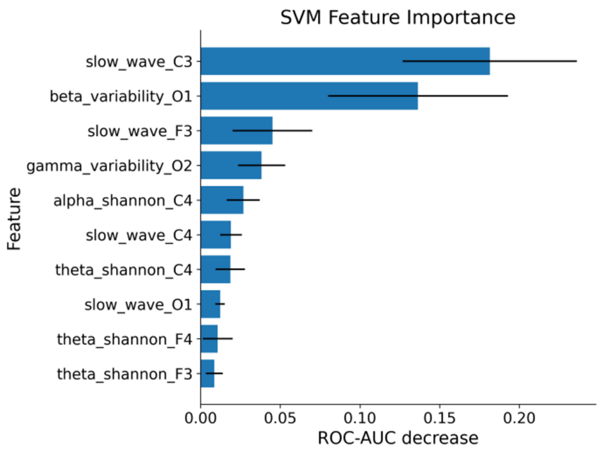

Figure 2. SVM Feature Importance

In addition, beta and gamma band variability in occipital channels also showed notable influence on the classifier output. Shannon entropy in the alpha and theta bands provided moderate contributions, mainly in central and frontal electrodes.

From a neurophysiological perspective, these results indicate that increased slow-wave activity and reduced high-frequency variability, particularly in fronto-central networks, play a key role in distinguishing consciousness states.

4 Discussion

The present results indicate that a multi-domain EEG feature set can effectively support DOC classification between VS and MCS. The SVM achieved the best performance among tested algorithms and provided greater interpretability for investigating feature contributions; by contrast, although the CNN achieved comparable accuracy, its internal representation lacks sufficient interpretability. Consequently, the SVM was preferred for subsequent analyses at the feature level.

The dominant contribution of slow wave features, particularly over fronto-central sites, aligns with established neurophysiological observations that EEG slowing and reduced cortical complexity are hallmarks of impaired consciousness, potentially reflecting thalamocortical dysconnectivity or suppression of cortical activity [19]. Complementary contributions from measures of spectral variability and entropy metrics suggest that both reductions in high-frequency variability and decreases in signal complexity play a role in discriminating consciousness states.

This study has several limitations. The dataset was drawn from public repositories and is relatively small, with class imbalance between diagnostic groups; such factors can introduce bias and limit the robustness and reproducibility of the models. Inter-subject variability in EEG further complicates external validation; therefore, additional validation on independent cohorts is required. Clinically, model performance could be enhanced by incorporating medical history or standardized behavioral scales. Moreover, real-time clinical deployment would require addressing latency, preprocessing and artifact-rejection constraints. Finally, future work should evaluate multi-modal integration and larger, more diverse datasets to improve generalizability and translational readiness.

5 Conclusion

This study shows that a multi-dimensional EEG framework combining temporal, spectral, and spatial features can effectively differentiate VS from MCS in patients with DOC Among the tested classifiers, the SVM achieved the best overall performance, indicating a strong balance between predictive accuracy and interpretability. Feature importance analysis revealed that slow-wave activity and variability-related measures were the most informative markers, underscoring their relevance to the neurophysiological characterization of impaired consciousness.

These findings suggest that the integration of quantitative EEG and interpretable machine learning may provide an objective and clinically valuable complement to conventional behavioral assessment. Despite limitations related to sample size, class imbalance, and the use of a single dataset, the proposed framework demonstrates promising potential for DOC evaluation. Future work should focus on validation in larger multicenter cohorts and on multimodal integration to improve generalizability and clinical translation.

References

1. J. Claassen, Y. Akbari, S. Alexander et al. Proceedings of the First Curing Coma Campaign NIH Symposium: Challenging the Future of Research for Coma and Disorders of Consciousness. Neurocrit Care **35** (Suppl 1), 4-23 (2021)
2. S. Laureys, G. G Celesia, F. Cohadon et al. Unresponsive wakefulness syndrome: a new name

for the vegetative state or apallic syndrome. BMC Med. **8**, 68 (2010)

3. D.Y. Cho, Y.C. Wang. Comparison of the APACHE III, APACHE II and Glasgow coma scale in acute head injury prognosis. Intensive Care Med **23**, 77-84 (1997)
4. J. Tian, Y. Zhou, H. Liu, Z. Qu, L. Zhang, L. Liu. Quantitative EEG parameters can improve the predictive value of the non-traumatic neurological ICU patient prognosis through the machine learning method. Front. Neurol. **13**, 897734 (2022)
5. M.C. Tjepkema-Cloostermans, F.B. van Meulen, G. Meinsma, M.J.A.M. van Putten. A Cerebral Recovery Index (CRI) for early prognosis in patients after cardiac arrest. Crit Care **17**, 1 (2013)
6. A. Alkhachroum, B. Appavu, S. Egawa et al. Electroencephalogram in the intensive care unit: a focused look at acute brain injury. Intensive Care Med **48**, 1443-62 (2022)
7. A. Alkhachroum, E. Fló, B. Manolovitz et al. Resting-State EEG Signature of Early Consciousness Recovery in Comatose Patients with Traumatic Brain Injury. Neurocrit Care **41**, 855-865 (2024)
8. M.E. Haveman, M.J.A.M. van Putten, H.W. Hom et al. Predicting outcome in patients with moderate to severe traumatic brain injury using electroencephalography. Crit Care **23**, 401 (2019)
9. J. Witsch, H. Frey, J. Schmidt et al. Electroencephalographic Periodic Discharges and Frequency-Dependent Brain Tissue Hypoxia in Acute Brain Injury. JAMA Neurol **74**, 301 (2017)
10. A. Tolonen, M.O.K. Särkelä, R.S.K. Takala, A. Katila et al. Quantitative EEG Parameters for Prediction of Outcome in Severe Traumatic Brain Injury: Development Study. Clin. EEG Neurosci. **49**, 248-57 (2018)
11. P M. Vespa , W J. Boscardin, D A. Hovda, et al. Early and persistent impaired percent alpha variability on continuous electroencephalography monitoring as predictive of poor outcome after traumatic brain injury. Journal of Neurosurgery **97**(1), 84-92 (2002)
12. J. Nekrasova, M. Kanarskii, D. Yankevich et al. Polysomnographic records of patients with disorders of consciousness. Mendeley Data V1 (2020), doi: 10.17632/6wx4n25h4v.1
13. A. Gramfort, M. Luessi, E. Larson, et al. MEG and EEG data analysis with MNE-Python. Front. Neurosci. **7**, 267 (2013)
14. R. Vallat, M.P. Walker. An open-source, high-performance tool for automated sleep staging. Elife **10** (2021)
15. P. Virtanen, R. Gommers, T.E. Oliphant et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods **17**, 261-272 (2020)
16. N.V. Thakor, J. Paul, S. Tong et al. Entropy of brain rhythms: normal versus injury EEG. IEEE Signal Processing Workshop on Statistical Signal Processing (2001)
17. M.J.A.M. van Putten. Extended BSI for continuous EEG monitoring in carotid endarterectomy. Clin. Neurophysiol. **117**, 2661-2666 (2006)
18. M.J.A.M. van Putten. The colorful brain: visualization of EEG background patterns. J. Clin.Neurophysiol. **25**, 63-68 (2008)
19. M.R. Nuwer, D.A. Hovda, L.M. Schrader, P.M. Vespa. Routine and quantitative EEG in mild traumatic brain injury. Clin. Neurophysiol. **116**, 2001-25 (2005)