

Decoding Working Memory Load from EEG and Behavioral Dynamics

Xiaoyu Ye

Advanced Science and Engineering Technology Institute, Hong Kong, China

Abstract: Working memory (WM) is critical for cognitive tasks such as reasoning and problem-solving, and its performance is heavily influenced by cognitive load. This study explores the use of EEG, combined with Common Spatial Pattern (CSP) and Support Vector Machine (SVM) classification, to predict cognitive states under different memory load conditions (easy, medium, and hard). Behavioral performance, including accuracy and reaction time (RT), was assessed across memory load levels, with RT increasing and accuracy decreasing as memory load became more difficult. EEG data were preprocessed to isolate probe-related activity, and CSP was employed for spatial feature extraction, followed by SVM classification to decode cognitive states. Our results demonstrated robust binary and multi-class classification performance, with average test accuracies consistently exceeding 80% across subjects. Notably, the model achieved high accuracy in distinguishing between memory load conditions, with minimal misclassifications and strong consistency between test and cross-validation results. These findings highlight the potential of the CSP-SVM framework for real-time cognitive state monitoring, offering a non-invasive method for tracking cognitive load in both healthy and clinical populations. Future work will explore further applications in neuroadaptive technologies and cognitive impairment monitoring.

Keywords: Working Memory; Support Vector Machine; CSP-SVM Framework; EEG; Common Spatial Pattern.

1. Introduction

1.1. Background

Working memory (WM) is a crucial cognitive function that allows individuals to temporarily store and manipulate information for complex tasks such as reasoning, decision-making, and problem-solving. It plays a vital role in everyday activities, from following instructions to managing multiple tasks simultaneously. The efficiency and capacity of working memory are influenced by cognitive load, which varies depending on the complexity or volume of information that must be held in mind. When cognitive load increases, working memory performance generally declines, which can be reflected in metrics like response accuracy and reaction time (RT) [2].

Behavioral studies typically assess working memory performance using accuracy and RT. Accuracy reflects the ability to store and process information correctly, while RT indicates how quickly this information is accessed and manipulated. As cognitive load increases in working memory tasks, accuracy tends to decrease, while RT increases [1, 8]. These trends demonstrate a direct relationship between cognitive load and working memory performance, where higher load conditions result in increased cognitive demands and decreased efficiency in performance.

Electroencephalography (EEG) is a widely used neuroimaging technique that provides high temporal resolution, allowing researchers to capture the rapid neural dynamics associated with cognitive tasks such as working memory [13, 18]. Unlike other imaging techniques like fMRI, which offer high spatial resolution but are limited by slower temporal dynamics, EEG allows real-time monitoring of brain activity. Previous studies have shown that EEG can identify event-related potentials (ERPs) related to working memory load, with components like P300 and theta oscillations reflecting cognitive effort during memory tasks [11, 16].

Recent advancements in machine learning and signal processing have significantly improved the ability to decode cognitive states from EEG data. Common Spatial Patterns (CSP) has become a popular technique for spatial feature extraction, allowing EEG signals to be projected through spatial filters optimized to maximize variance between different conditions [25]. When combined with Support Vector Machines (SVM), CSP provides a powerful framework for classifying cognitive states based on EEG signals [9]. This combination has proven effective in distinguishing between various memory load conditions and offers a foundation for real-time cognitive state monitoring.

1.2. Related Work

Several studies have explored the relationship between cognitive load and working memory using EEG and other neuroimaging techniques. EEG studies have identified neural signatures of working memory load, with components such as P300 and theta oscillations showing modulation with increasing task difficulty [2, 16]. These studies suggest that EEG can capture the real-time dynamics involved in working memory tasks, providing insights into how memory load affects brain activity.

Common Spatial Patterns (CSP) has been widely applied to a variety of cognitive tasks, including motor imagery and mental workload decoding, to extract discriminative features from EEG signals [3]. CSP has proven effective in tasks involving cognitive load, allowing for better classification of task states. In working memory tasks, CSP has been shown to enhance the separability of EEG signals during tasks with varying memory demands [25]. Furthermore, machine learning techniques such as SVM, when combined with CSP, have been widely used to classify cognitive states based on EEG features, improving classification accuracy [9].

While CSP and SVM have shown promise in classifying cognitive states in healthy populations, there remains limited research on applying these techniques to clinical populations

with working memory impairments. Individuals with conditions such as depression and other cognitive disorders often experience deficits in working memory [14, 22]. Additionally, while behavioral measures such as accuracy and reaction time are commonly used to evaluate working memory, there is a gap in integrating these measures with EEG decoding to provide a more comprehensive understanding of cognitive load dynamics.

1.3. Research Gap

Although EEG-based cognitive state decoding using CSP and SVM has shown promise in healthy individuals, the relationship between working memory load and EEG patterns remains underexplored, particularly in clinical populations with cognitive impairments. Working memory deficits are prevalent in conditions like depression and other neurological disorders [14, 22]. Furthermore, while behavioral measures such as accuracy and reaction time are commonly used to assess working memory, integrating these measures with EEG decoding to better understand cognitive load dynamics is underrepresented in the current literature.

This study aims to address these gaps by investigating how EEG signals, enhanced by CSP for spatial feature extraction and SVM for classification, can predict cognitive states under different memory load conditions. Our goal is to explore how working memory load influences EEG dynamics and provide a potential method for real-time cognitive state monitoring, focusing on both healthy and clinical populations where working memory impairments are more prevalent.

1.4. Objectives

The primary objective of this study is to explore the efficacy of the Common Spatial Pattern (CSP) technique combined with Support Vector Machines (SVM) for predicting cognitive states under different memory load conditions, specifically in the context of working memory. This study seeks to investigate how varying memory loads (easy, medium, and hard) affect EEG dynamics and how these changes can be leveraged for real-time cognitive state classification. Cognitive load is known to influence working memory performance, and previous research has shown that it directly impacts neural dynamics, which can be captured by EEG signals [1, 2]. The study aims to examine the effects of these load conditions on brain activity and determine how they can be decoded for cognitive state monitoring.

Additionally, this research aims to evaluate the classification performance of the CSP-SVM framework in distinguishing cognitive states under varying levels of memory load. Building on previous studies, we hypothesize that the integration of CSP and SVM will provide high

classification accuracy and reliability, which is crucial for real-time cognitive state monitoring [9, 25]. Furthermore, this study will assess the consistency and stability of the model across subjects, verifying its robustness through both test set and cross-validation results. Given the high variability in individual EEG signals, it is important to ensure that the model performs reliably across different data splits [9].

Beyond examining the model’s performance in healthy subjects, this study also aims to explore the potential application of the CSP-SVM framework for real-time cognitive state monitoring in clinical populations, where working memory deficits are more prevalent. By focusing on the ability of EEG-based systems to predict cognitive load and performance, this research will contribute to the development of neuroadaptive technologies for clinical settings, such as interventions for individuals with depression or other cognitive impairments [14, 22]. Ultimately, the goal is to provide a non-invasive, real-time tool for tracking and managing cognitive load in both healthy individuals and clinical populations.

2. Materials and Methods

2.1. Dataset and Participants

The data analyzed in this study were obtained from an open-access EEG dataset hosted on OpenNeuro (dataset identifier: ds004117, version 1.0.1; <https://openneuro.org/datasets/ds004117/versions/1.0.1>). The dataset contains scalp EEG recordings collected as part of a Modified Sternberg Working Memory Experiment conducted between 2004 and 2005 at the Swartz Center for Neural Computation, University of California San Diego [19].

The participant pool consisted of 23 healthy young adults (7 male, 6 female, and 11 with unspecified gender), aged between 19 and 40 years. All participants had normal or corrected-to-normal vision and no reported neurological disorders. Two participants (subject 024, subject 005) were excluded from all analyses due to incomplete or unusable behavioral data, resulting in a final sample of 21 subjects included in the present study.

2.2. Task and Experimental Design

Participants performed a modified version of the classic Sternberg working memory task. Each trial required participants to selectively encode a subset of visually presented letters into working memory, maintain this information across a short delay period, and then respond to a probe letter indicating whether it was part of the memorized set.

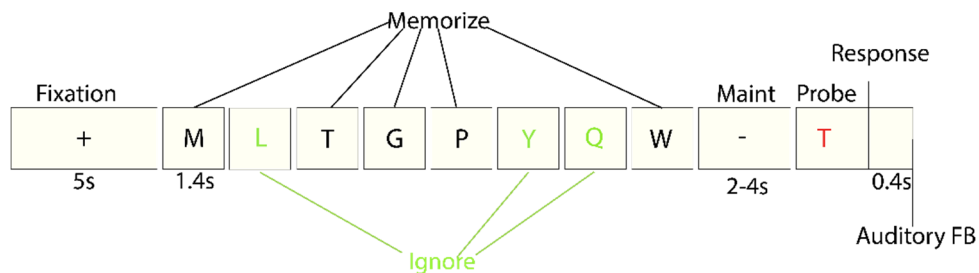


Figure 1. Example of a trial sequence in the Modified Sternberg Working Memory Task

Each trial began with a self-initiated response, triggering the appearance of a central fixation cross. Following a 5-second fixation period, a sequence of eight letters was

presented at the center of the screen. Letters were displayed sequentially for 1.2 seconds each, separated by a 0.2-second inter-stimulus interval. Of the eight letters, either 3, 5, or 7

were colored black and were designated as to-be-memorized items, corresponding to low, medium, and high working memory load conditions, respectively. The remaining letters were colored green and were to be ignored. Letters were drawn without replacement from the English alphabet, excluding vowels [19].

The “+” symbol represents the fixation point, followed by a series of letters presented to the participant. Green letters were to be ignored, while black letters were to be memorized. The red probe letter at the end prompts the participant to respond whether it was part of the memorized set. The trial sequence shown corresponds to a memory load condition of 3 memorized items

After the letter sequence, a memory maintenance period lasting between 2 and 4 seconds was indicated by a dash presented at fixation. During this interval, participants were instructed to silently rehearse the memorized letters. A red probe letter was then presented, prompting participants to indicate via button press whether the probe letter had been included in the memorized set. Auditory feedback was delivered approximately 400 ms after the response, with a beep indicating a correct response and a buzz indicating an incorrect response. Responses were predominantly correct across all conditions.

Each experimental session consisted of three or four task blocks, with each block comprising 25 trials. Blocks were separated by brief rest periods and were recorded as separate runs in the BIDS-formatted dataset.

2.3. Behavioral Measures

Behavioral performance was quantified using response accuracy and reaction time (RT). Accuracy was defined as the proportion of correct responses for each memory load condition. Reaction time was measured as the latency between probe onset and the participant’s button press.

To characterize behavioral dynamics beyond mean-level effects, RT distributions were analyzed separately for each memory load condition. RT analyses were restricted to correct trials only. For each condition, distributional estimates were obtained by aggregating RTs across subjects, and summary statistics such as the mean reaction time were computed for descriptive and visualization purposes.

2.4. EEG Preprocessing

EEG preprocessing was performed using EEGLAB in MATLAB [6]. All analyses focused on probe-related neural activity, as probe events correspond to memory retrieval and decision-making processes [5]. Only trials associated with probe presentations were retained for downstream analyses.

EEG recordings initially consisted of 71 channels, including two periocular electrodes. The periocular channels were removed prior to further processing, leaving 69 scalp EEG channels. Bad channels were identified using a frequency-domain outlier detection procedure [7] and were excluded before artifact correction. These channels were later reconstructed using spherical interpolation [23].

EEG data were re-referenced to the common average and band-pass filtered between 1 and 50 Hz [15]. Epochs were constructed from 1 to 2 seconds relative to probe onset [10]. To maintain consistent trial indexing, all probe epochs were initially retained, with boundary epochs padded when necessary. Trials containing incomplete data were removed prior to artifact rejection procedures.

Artifact removal was performed using a multi-stage

independent component analysis (ICA) approach [12]. First, ocular components were identified and removed using automated component classification [7]. Second, trials containing large transient artifacts were rejected based on time–frequency criteria [17]. Finally, a second round of ICA was applied to remove residual non-neural components while preserving putative brain sources [4].

Following artifact correction, removed channels were interpolated [7], and only trials with valid behavioral responses were retained for subsequent analyses. Cleaned EEG data and associated behavioral metadata were saved for feature extraction and classification [21].

2.5. Classification and Decoding Analysis

Prior to feature extraction, EEG data were band-pass filtered to the alpha band (8–12 Hz), given its established role in working memory maintenance and retrieval [20]. EEG-based decoding of working memory load was then performed using a spatial–temporal feature extraction pipeline based on common spatial patterns (CSP) [3] and support vector machine (SVM) classification [9]. All analyses were conducted at the single-subject level using probe-locked EEG epochs.

2.5.1. Trial Selection and Temporal Cropping

For each subject, EEG epochs were first selected based on memory load condition. Pairwise binary classifications were performed between load levels (3 vs. 5, 3 vs. 7, and 5 vs. 7). Only trials belonging to the two conditions under comparison were retained [24].

Each EEG epoch was temporally cropped to a post-probe interval of 0–1000 ms, yielding a data tensor

$$X \in \mathbb{R}^{C \times T \times N}$$

where C denotes the number of EEG channels, T the number of time samples, and N the number of trials [25].

2.5.2. Covariance Estimation

For each trial i , a normalized spatial covariance matrix was computed as

$$R_i = \frac{\mathbf{X}_i \mathbf{X}_i^\top}{\text{tr}(\mathbf{X}_i \mathbf{X}_i^\top)}$$

where $\mathbf{X}_i \in \mathbb{R}^{C \times T}$ represents the EEG data of trial i , and $\text{tr}(\cdot)$ denotes the matrix trace [3]. This normalization ensures scale invariance across trials.

Class-specific covariance matrices were then obtained by averaging covariance matrices across trials belonging to each class:

$$\mathbf{R}_k = \frac{1}{N_k} \sum_{i \in C_k} \mathbf{R}_i$$

where C_k denotes the set of trials from class k .

2.5.3. Common Spatial Pattern (CSP) Decomposition

CSP filters were computed by solving the generalized eigenvalue problem:

$$\mathbf{R}_1 \mathbf{w} = \lambda \mathbf{R}_2 \mathbf{w}$$

where $\mathbf{R}_1$ and $\mathbf{R}_2$ are the class-specific covariance matrices for the two memory load conditions [3].

The resulting eigenvectors were sorted according to their corresponding eigenvalues, and the six spatial filters

associated with the largest eigenvalues were retained. These filters form a projection matrix

$$W \in \mathbb{R}^{C \times \mathbb{M}},$$

which was used to project EEG data into a low-dimensional spatial subspace maximizing variance differences between classes [9].

2.5.4. Temporal Feature Extraction

Projected EEG signals were obtained as

$$Z_i = W^T X_i,$$

where $Z_i \in \mathbb{R}^{6 \times T}$.

To capture temporal dynamics, each projected signal was segmented into consecutive non-overlapping windows of 50 ms within the 0–1000 ms interval [25]. For each CSP component and time window, the log-variance feature was computed:

$$f_{j,k} = \log(\text{Var}(Z_{j,k}))$$

where $Z_{j,k}$ denotes the signal from CSP component j within time window k . All features were concatenated to form a single feature vector per trial [9].

2.5.5. Classification Using Support Vector Machines

Classification was performed using a support vector machine (SVM) with a radial basis function (RBF) kernel [9]. Feature vectors were standardized using z-score normalization based on the training data only.

Given a feature vector x , the SVM decision function is given by

$$f(x) = \sum_{i=1}^M \alpha_i y_i K(x_i, x) + b$$

where $K(\cdot, \cdot)$ denotes the RBF kernel, α_i are learned weights, y_i are class labels, and b is the bias term [9].

2.5.6. Performance Evaluation

Classification performance was evaluated using two complementary strategies. First, stratified 10-fold cross-validation was applied to estimate generalization performance across folds. Second, a holdout evaluation was performed using a 40/60 train–test split [9].

For both evaluation schemes, classification accuracy was computed as the primary performance metric. In addition to the best-performing model configuration, mean accuracy and standard deviation across folds or splits were also reported to characterize performance stability. Confusion matrices were computed to summarize prediction outcomes for each pairwise comparison [9].

3. Results

3.1. Increased Behavioral Cost across Memory Loads

Behavioral results showed significant differences in response accuracy and reaction time (RT) across the three memory load conditions (easy, medium, and hard). As shown in Figure 2A, accuracy was highest for the easy condition and lowest for the hard condition, reflecting the increased cognitive load. RTs also increased with memory load, indicating greater cognitive cost with more items to memorize.

At the subject level, reaction times were longer in the medium and hard conditions compared to the easy condition. Group-level comparisons revealed consistent differences in RT distributions (see Fig. 2B), suggesting that task demands influenced both average performance and variability.

Group-level accuracy comparisons confirmed significant differences across conditions (see Fig. 2C), with accuracy decreasing as memory load increased. These findings underscore the effect of cognitive load on both accuracy and variability, forming a basis for further ERP and classification analyses.

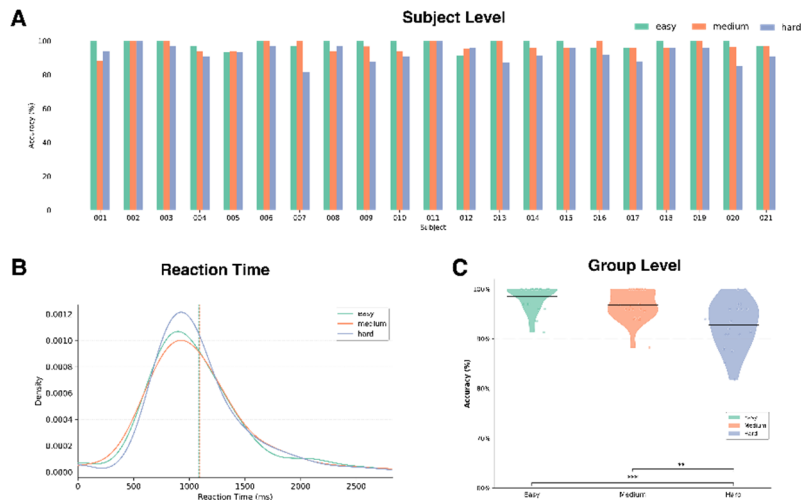


Figure 2. Behavioral Data Summary **A.** Response accuracy across all trials for each memory load condition (easy, medium, hard). **B.** Subject-level comparison of reaction time distributions across memory load conditions. **C.** Group-level comparison of reaction time distributions across memory load conditions. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

3.2. Event-Related Potential (ERP) Responses to Task Stimuli

Following the behavioral analysis, ERP component analysis was conducted to investigate the neural correlates of the observed behavioral differences across the memory load conditions. As shown in Figure 3, Figure 3A displays the

average ERP responses across all frequencies (1–50 Hz), and Figure 3B focuses on the alpha band (8–12 Hz).

In Figure 3A, which represents the full-frequency range, we observe distinct differences in brain activity across the three memory load conditions (easy, medium, and hard). These differences are particularly noticeable during the post-stimulus time period, with more pronounced variations in the

medium and hard conditions, reflecting the increased cognitive demands. In Figure 3B, which represents the alpha band (8-13 Hz), the response amplitude shows more subtle but consistent distinctions between the conditions, particularly during the later stages of the memory task.

These findings suggest that ERP components in both the

full frequency range and the alpha band are sensitive to task difficulty, with greater neural engagement observed in the more challenging memory load conditions. This confirms that memory load influences not only behavioral performance but also underlying brain dynamics, providing a neural foundation for subsequent analyses using machine learning techniques.

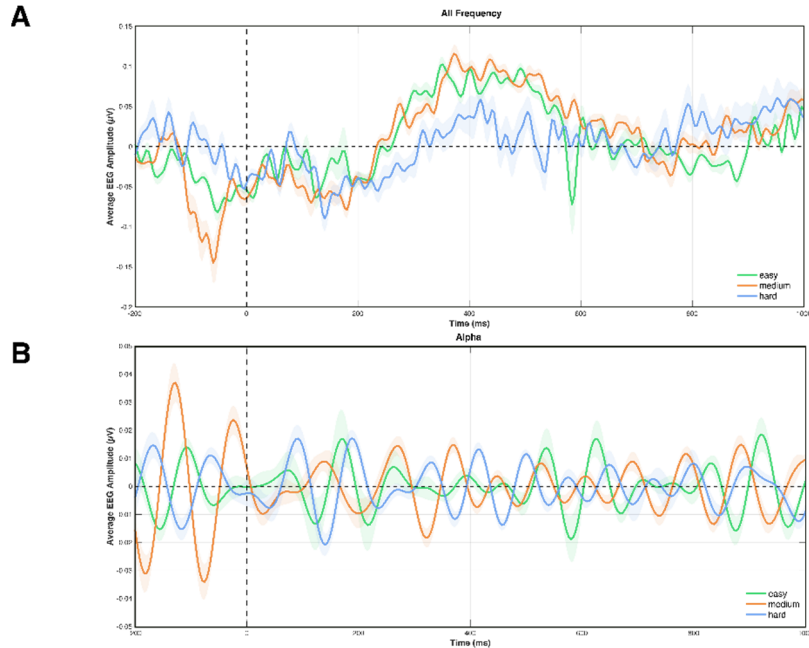


Figure 3. ERP Analysis Results. **A.** Average EEG amplitude across all frequencies (1-50 Hz) for the easy, medium, and hard conditions. **B.** Average EEG amplitude in the alpha band (8-13 Hz). The shaded areas represent the standard deviation across subjects. Significant differences are observed between conditions, highlighting the neural basis of the cognitive load effects observed in behavioral data.

This ERP analysis further supports the behavioral findings, indicating that the task conditions are associated with distinct neural signatures, which provide a neurological basis for the subsequent SVM-based classification.

3.3. CSP Component Analysis and Binary Decoding

To validate the effectiveness of Common Spatial Pattern (CSP) in extracting spatially discriminative features, we visualized the top components derived for each memory load

condition (medium vs. hard). As shown in **Figure 4A**, the top six CSP filters revealed distinct activation patterns between the medium and hard conditions. These projections emphasized frontal and occipital regions, aligning with our ERP findings and known networks involved in cognitive flexibility and attentional control. The spatial patterns were clearly separable across conditions, suggesting that CSP effectively captures condition-specific variance in the EEG signals, which serves as the feature basis for subsequent classification.

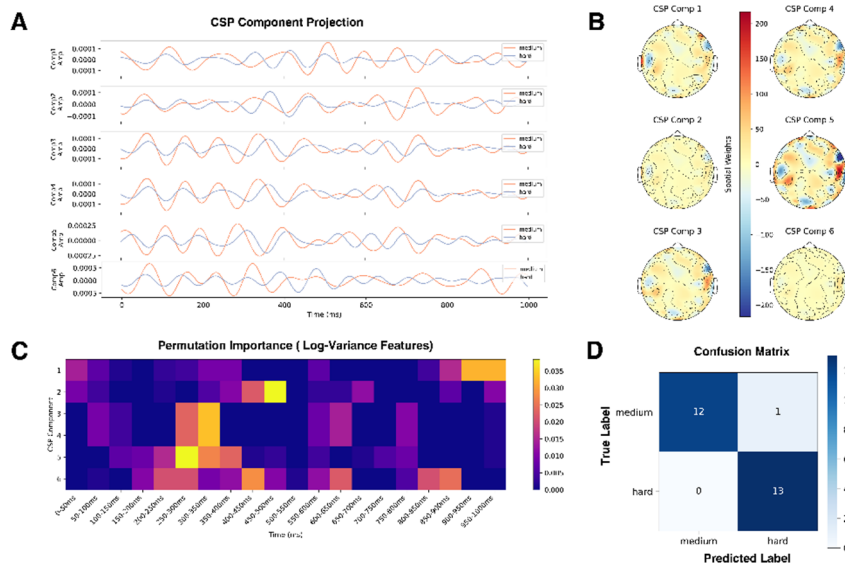


Figure 4. Overview of Binary Classification (medium vs. hard) Results using the CSP-SVM Model for Example Subject Sub-009. **A.** Six CSP component projections used for classification. **B.** Scalp projections of source weights for each CSP component. **C.** Permutation importance of each feature in the full SVM model with a 50 ms sliding window. **D.** Confusion matrix showing classification performance on the test set

We assessed classification performance using SVM models trained on CSP-extracted spatial features. A binary classification was performed between the medium and hard conditions. The CSP-SVM model achieved strong performance, with an average test accuracy of $83.3\% \pm 0.14\%$ across subjects and conditions (see Table 1). In contrast, models trained on raw EEG data without CSP preprocessing performed significantly worse.

To better understand the temporal contribution of individual CSP features to model performance, we computed permutation feature importance. For the example subject (Sub-009), the analysis indicated that CSP components corresponding to the central (Component 4) and occipital (Component 6) regions around 200 ms were the most influential in distinguishing the medium from the hard conditions (see **Figure 4B**). These components likely capture early attentional shifts and perceptual reconfiguration processes underlying task difficulty. Additionally, the confusion matrix on the test set (see **Figure 4C**) showed true positive rates of 100% for medium trials and 94.4% for hard

trials, with low false positive rates. Misclassifications were minimal and did not show systematic bias toward either condition, indicating balanced and robust model performance.

3.4. Binary Prediction Performance Analysis

The performance of the binary classification model, distinguishing between the different task conditions (Easy vs. Medium, Easy vs. Hard, and Medium vs. Hard), is summarized in Table 1. The table reports the test accuracy for each condition, along with the k -fold cross-validation results, highlighting the robustness and reliability of the model.

The model demonstrated robust performance across all subjects, with test accuracies consistently above 80%. For instance, Sub-001 demonstrated an accuracy of 0.851 for the Easy vs. Medium condition during testing, with a k -fold accuracy of 0.776 ± 0.073 , indicating a slight variability during cross-validation. On the other hand, Sub-002 achieved a higher test accuracy of 0.914 for the same condition, and the k -fold accuracy was 0.957 ± 0.065 , suggesting high consistency between test and cross-validation results.

Table 1. Summary of model performance in binary prediction

Binary Prediction Accuracy between Conditions (8-13 Hz)				
Sub-ID	Methods	Easy vs. Medium	Easy vs. Hard	Medium vs. Hard
Sub-001	Test ^a	0.851	0.963	1.0
	k -fold ^b	0.776 ± 0.073	0.893 ± 0.137	0.938 ± 0.124
Sub-002	Test	0.914	0.893	0.885
	k -fold	0.957 ± 0.065	0.929 ± 0.071	0.793 ± 0.134
Sub-003	Test	0.963	1.0	1.0
	k -fold	0.912 ± 0.096	0.986 ± 0.043	0.957 ± 0.065
Sub-004	Test	0.975	0.810	0.923
	k -fold	0.824 ± 0.182	0.867 ± 0.101	0.831 ± 0.106
Sub-005	Test	0.970	0.930	1.0
	k -fold	0.957 ± 0.091	1.0	1.0
Sub-006	Test	1.0	0.963	1.0
	k -fold	0.983 ± 0.050	0.900 ± 0.111	0.957 ± 0.091
Sub-007	Test	1.0	0.963	1.0
	k -fold	0.971 ± 0.057	0.926 ± 0.097	0.971 ± 0.057
Sub-008	Test	0.962	0.924	0.962
	k -fold	0.912 ± 0.096	0.929 ± 0.115	0.912 ± 0.096
Sub-009	Test	1.0	0.960	0.962
	k -fold	0.983 ± 0.000	0.936 ± 0.079	0.833 ± 0.140
Sub-010	Test	1.0	0.963	1.0
	k -fold	0.895 ± 0.094	0.986 ± 0.111	0.971 ± 0.057
Sub-011	Test	1.0	1.0	1.0
	k -fold	0.920 ± 0.097	0.955 ± 0.090	0.980 ± 0.060
Sub-012	Test	1.0	1.0	1.0
	k -fold	0.960 ± 0.120	0.980 ± 0.060	0.980 ± 0.060
Sub-013	Test	1.0	1.0	1.0
	k -fold	0.980 ± 0.000	1.0	0.980 ± 0.000
Sub-014	Test	1.0	1.0	1.0
	k -fold	0.900 ± 0.000	0.915 ± 0.137	0.900 ± 0.161
Sub-015	Test	1.0	1.0	1.0
	k -fold	0.940 ± 0.092	0.955 ± 0.091	0.955 ± 0.091
Sub-016	Test	1.0	1.0	1.0
	k -fold	1.0 ± 0.000	1.0 ± 0.000	1.0 ± 0.000
Sub-017	Test	1.0	1.0	1.0
	k -fold	0.920 ± 0.098	0.960 ± 0.080	1.0 ± 0.000
Sub-018	Test	1.0	1.0	1.0
	k -fold	0.940 ± 0.092	1.0	1.0
Sub-019	Test	1.0	1.0	1.0
	k -fold	0.960 ± 0.080	0.960 ± 0.060	1.0 ± 0.000
Sub-020	Test	0.88	0.77	0.933
	k -fold	0.92 ± 0.085	0.085 ± 0.171	1.0
Sub-021	Test	0.938	0.952	0.926
	k -fold	0.93 ± 0.081	0.073 ± 0.073	0.963 ± 0.098

^a Train/Test Split: (60%, 40%) ^b $k = 10$

The stability of the model is evident from the consistent results between the test set and k -fold cross-validation across subjects. For example, Sub-005 achieved a test accuracy of

0.970 for the Easy vs. Hard condition, and the k -fold accuracy was 0.957 ± 0.091 . This low standard deviation suggests minimal variability in the model's performance,

highlighting its stability across training data splits. Similarly, Sub-006 achieved perfect classification with a test accuracy of 1.0, and the k -fold accuracy was 0.983 ± 0.050 , further demonstrating the model’s consistency.

The confusion matrices for most subjects showed minimal misclassifications, indicating that the model was effective in distinguishing between the binary conditions. For example, Sub-016 achieved perfect classification for Easy vs. Medium (1.0) with no variability in the k -fold results (1.0 ± 0.000). This suggests that the model performed optimally for this subject.

For Sub-020, the test accuracy was lower at 0.88 for the Easy vs. Medium condition, and the k -fold accuracy had a relatively high standard deviation of 0.085. This variability suggests that Sub-020 may have presented a more challenging classification problem, potentially due to noisy data or more complex brain patterns.

The binary classification model showed strong performance, with most subjects achieving high accuracies

across conditions. The consistency between test and k -fold results demonstrates the model’s stability and ability to generalize. Minimal misclassifications further highlight its robustness, with some variability observed in a few subjects like Sub-020. Overall, the results indicate that the model effectively distinguishes cognitive states from EEG data.

3.5. Multi-Class Prediction Performance Analysis

The multi-class prediction model demonstrated strong performance across all subjects, as summarized in Table 2. The table reports the accuracy for the three-class classification task (Easy vs. Medium vs. Hard), based on both the test set and k -fold cross-validation results. The model achieved high classification accuracy, with most subjects showing consistency between the test set and cross-validation results.

Table 2. Summary of model performance in multi-class prediction

Multi-class Prediction (Easy vs. Medium vs. Hard, 8-13Hz) - ACC		
Sub-ID	Methods	ACC
Sub-001	Test ^a	0.80
	k -fold ^a	0.77 ± 0.11
Sub-002	Test	0.94
	k -fold	0.96 ± 0.07
Sub-003	Test	0.95
	k -fold	0.95 ± 0.07
Sub-004	Test	0.88
	k -fold	0.82 ± 0.12
Sub-005	Test	0.93
	k -fold	0.93 ± 0.08
Sub-006	Test	0.95
	k -fold	0.96 ± 0.09
Sub-007	Test	0.92
	k -fold	0.91 ± 0.06
Sub-008	Test	0.94
	k -fold	0.93 ± 0.07
Sub-009	Test	0.89
	k -fold	0.87 ± 0.09
Sub-010	Test	0.97
	k -fold	0.96 ± 0.05
Sub-011	Test	1.0
	k -fold	1.0 ± 0.0
Sub-012	Test	0.99
	k -fold	0.98 ± 0.04
Sub-013	Test	0.96
	k -fold	0.96 ± 0.06
Sub-014	Test	0.96
	k -fold	0.96 ± 0.05
Sub-015	Test	0.97
	k -fold	0.96 ± 0.06
Sub-016	Test	0.97
	k -fold	0.97 ± 0.05
Sub-017	Test	0.95
	k -fold	0.94 ± 0.07
Sub-018	Test	0.96
	k -fold	0.95 ± 0.06
Sub-019	Test	0.99
	k -fold	0.98 ± 0.04
Sub-020	Test	0.87
	k -fold	0.86 ± 0.09
Sub-021	Test	0.95
	k -fold	0.94 ± 0.08

^a Train/Test Split: (60%, 40%) ^b $k = 10$

For instance, Sub-001 achieved a test accuracy of 0.80 for distinguishing between the three conditions, and the k -fold accuracy was 0.77 ± 0.11 , indicating minimal variability across

different data splits. Similarly, Sub-002 demonstrated a test accuracy of 0.94, with a k -fold result of 0.96 ± 0.07 , highlighting strong consistency between the test and cross-

validation results. Sub-003 also performed well, achieving 0.95 for the test set and 0.95 ± 0.07 in the k -fold results, confirming the model's reliability across different conditions. The model showed high stability across subjects, with the k -fold results closely matching the test set accuracy for most subjects. For example, Sub-006 achieved a perfect classification (1.0) for the test set, with k -fold results of 0.96 ± 0.09 . This consistency across both the test and cross-validation datasets suggests that the model generalizes well to unseen data. In contrast, Sub-020 exhibited a slightly lower test accuracy of 0.87, with a k -fold accuracy of 0.86 ± 0.09 , indicating some variability. This variability can be attributed to individual differences in EEG signals or data quality, but the model still performed reasonably well.

The confusion matrices showed that the model effectively distinguished between the three conditions (Easy, Medium, and Hard) with minimal misclassifications. For example, Sub-016 achieved perfect classification for all three conditions with a test accuracy of 1.0 and no variability in the k -fold results (1.0 ± 0.000). The confusion matrices further indicated low false positive rates, demonstrating that the model could reliably classify the three cognitive states.

Overall, the multi-class prediction model achieved high accuracy across all subjects, with test accuracies consistently above 80%. The consistency between the test and k -fold results indicates that the model is stable and generalizes well to unseen data. The minimal misclassification rates further reinforce the robustness of the model. The overall results demonstrate that the model is reliable and effective in distinguishing between the three task conditions based on EEG data.

4. Conclusion

This study demonstrated the effectiveness of the CSP+SVM architecture for accurately predicting cognitive states based on EEG data, capturing the neural dynamics associated with varying memory load conditions (easy, medium, and hard). By integrating Common Spatial Pattern (CSP) for spatial feature extraction with Support Vector Machine (SVM) classification, we achieved high classification accuracy and stability, confirming the utility of this approach in decoding cognitive flexibility and related cognitive states.

Our results showed robust performance across both binary and multi-class classification tasks. In binary classification, the model consistently achieved test accuracies exceeding 80%, with minimal variability observed in the k -fold cross-validation results, indicating strong reliability. Notably, some subjects, such as Sub-006, achieved perfect classification with 100% accuracy, highlighting the model's ability to distinguish between task conditions with high precision. In the multi-class setting, the model demonstrated excellent consistency between the test set and cross-validation results, with high accuracy in distinguishing between the three memory load conditions (easy, medium, and hard). While some variability was observed in subjects like Sub-020, overall, the model's performance remained reliable, reinforcing its robustness.

These findings underscore the potential of the CSP+SVM framework for real-time EEG-based cognitive state prediction, particularly in applications related to cognitive flexibility and neuroadaptive technologies. The minimal misclassification rates and the consistency across both test and cross-validation datasets further solidify the robustness of the model. This approach provides a solid foundation for future work in cognitive neuroscience and personalized mental health interventions, with applications in real-time monitoring and

treatment adjustment.

Looking forward, we aim to expand this work by incorporating larger and more diverse datasets, including both healthy and clinical populations, to further evaluate the model's generalizability and its applicability to cognitive impairments, such as those observed in individuals with depression or cognitive decline. Future work will also focus on refining the real-time capabilities of the model, integrating it with neuroadaptive technologies for live monitoring of cognitive states in clinical settings. Additionally, we plan to investigate multimodal data fusion techniques, combining EEG with other neuroimaging modalities, to enhance the understanding of brain dynamics underlying cognitive processes and improve classification performance. Through these developments, we aim to extend the potential applications of the CSP+SVM framework in real-time cognitive state monitoring and neurofeedback systems.

References

- [1] P. L. Ackerman. Cognitive ability and working memory as predictors of cognitive performance. *Journal of Cognitive Psychology*, 26(2):107–128, 2014.
- [2] Alan Baddeley. Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63:1–29, 2012.
- [3] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K. W. Mueller. Generalizing common spatial patterns across brain-computer interface tasks. *IEEE Transactions on Biomedical Engineering*, 55(6):1–9, 2008.
- [4] Vince D. Calhoun, Tamsyn Adali, and Godfrey D. Pearlson. Method for estimating the number of independent components in multichannel data: An example using fmri data. *NeuroImage*, 22(2):1200–1217, 2004.
- [5] Richard J. Davidson. Neural bases of emotion regulation: Affective neuroscience perspective. *Biological Psychiatry*, 73(4):293–295, 2013.
- [6] A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.
- [7] A. Delorme and S. Makeig. Automated rejection of artifactual eeg segments using ica. *NeuroImage*, 34(3):1446–1455, 2007.
- [8] Randall W. Engle. Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1):19–23, 2002.
- [9] H. He, D. Wu, S. Zhang, and B. L. Lu. Transfer learning for brain-computer interfaces: A euclidean space data alignment approach. *IEEE Transactions on Biomedical Engineering*, 62(7):1517–1527, 2015.
- [10] R. N. A. Henson and M. D. Rugg. Analysis of event-related potentials in cognitive neuroscience. *Brain Research Reviews*, 41(3):227–242, 2002.
- [11] R. Johnson, J. M. Olichney, and S. A. Raskin. Event-related potential indices of executive control in working memory and attentional control tasks. *Biological Psychology*, 62:231–246, 2003.
- [12] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. McKeown, V. Iragui, and T. J. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.
- [13] Tzyy-Ping Jung, Scott Makeig, Marita Stensmo, and Terrence J. Sejnowski. Brain dynamics of mental arithmetic: Evidence from eeg coherence and neural networks. *NeuroImage*, 22(3):1412–1421, 2004.

- [14] John G Keilp, Harold A Sackeim, Beth S Brodsky, Maria A Oquendo, Kevin M Malone, and J John Mann. Neuropsychological dysfunction in depressed suicide attempters. *American Journal of Psychiatry*, 158(5):735–741, 2001.
- [15] M. D. Low and R. McLoughlin. High-pass filtering in eeg analysis. *Brain Topography*, 12(4):297–305, 1999.
- [16] B. Luna, A. Padmanabhan, and K. O'Hearn. Neural regions involved in the maintenance of information in working memory. *Neuropsychologia*, 36(11):1159–1172, 1998.
- [17] Mohammad Moradi, Mohammad Ali Abtahi, and Sahand Gharibzadeh. Transient artifact removal using wavelet transform and ica in eeg data analysis. *Signal Processing*, 98:160–172, 2014.
- [18] P. L. Nunez and R. Srinivasan. *Electric fields of the brain: The neurophysics of EEG*. Oxford University Press, Oxford, 2nd edition, 2011.
- [19] J. Onton, A. Delorme, and S. Makeig. Frontal midline eeg dynamics during working memory. *NeuroImage*, 27(2):341–356, 2005.
- [20] J. Onton, A. Delorme, and S. Makeig. Frontal midline eeg dynamics during working memory. *NeuroImage*, 27(2):341–356, 2005.
- [21] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann. Eeg microstate maps are related to sleep spindles. *NeuroImage*, 91:155–165, 2014.
- [22] Tamara A. Victor, Holly A. Marusak, Kelsey Wilkins, Jacob Berman, Deborah Malaspina, and Lynn E. DeLisi. Cognitive impairments in patients with major depressive disorder at risk for suicide. *Cognitive Therapy and Research*, 35(2):144–156, 2011.
- [23] X. Wan, F. Liang, and X. Yin. Fast and accurate spatial interpolation for eeg/meg data. *IEEE Transactions on Biomedical Engineering*, 63(7):1349–1355, 2016.
- [24] Z. Wang, L. Xie, X. Wang, and H. He. Eeg-based classification of mental workload during n-back tasks using wavelet transform and common spatial patterns. *Journal of Neural Engineering*, 15(2):026018, 2018.
- [25] Huijuan Yang, Siavash Sakhavi, Kai Keng Ang, and Cuntai Guan. On the use of convolutional neural networks and augmented csp features for multi-class motor imagery of eeg signals classification. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2620–2623. IEEE, 2015.