ELECO 2025 International Conference on Electrical and Electronics Engineering, Bursa, Turkiye, November 27-29, 2025

Towards Adaptive ERP-Based BCIs: EEG and fNIRS Guided Flashing Strategies to Account for Cognitive Fatigue

Onur Erdem Korkmaz[1,2*], Tolga Turay[1], Hasan Kilickaya[1], Hasan Ayaz[3], Riccardo Poli[4]

[1]Department of Electrical and Electronics Engineering, Atatürk University, Erzurum, Turkey,
tolgaturay@atauni.edu.tr, hasankilickaya@ogr.atauni.edu.tr
[2*]Sports Sciences Application and Research Center, Atatürk University, Erzurum, Turkey
onurerdem.korkmaz@atauni.edu.tr
[3]School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, PA, United States of America
ayaz@drexel.edu
[4]School of Computer Science and Electronic Engineering, University of Essex, Colchester, United Kingdom
rpoli@essex.ac.uk

Abstract

This pilot study presents a hybrid EEG–fNIRS framework to jointly assess speller performance and cognitive fatigue. Speller paradigms often face a trade-off: increasing stimulus flash repetitions improves ERP signal-to-noise ratio (SNR) but prolongs task duration, while fewer flashes reduce efficiency under fatigue. To address this limitation, EEG and fNIRS signals were recorded simultaneously during speller and cognitive task, arithmetic problem-solving blocks. Behavioral findings showed a decline in performance of the cognitive task, with fewer correct answers and longer response times across blocks, suggesting increased fatigue. ERP analyses revealed overall strong target responses but with attenuated amplitudes over time, especially in parietal and occipital channels. Classification confirmed these patterns: target vs. non-target discrimination reached 98.6% accuracy with three EEG channels, cognitive block classification achieved 92.4% using EEG, and 87.4% using fNIRS. These results demonstrate the feasibility of hybrid EEG–fNIRS systems for monitoring user states and provide a foundation for adaptive speller BCIs that dynamically adjust stimulus repetitions.

1. Introduction

Brain–computer interface (BCI) systems are commonly classified into three categories [1]: motor imagery-based, steady-state visual evoked potential (SSVEP)-based, and event-related potential (ERP)-based BCIs. ERP-based systems are widely used since they require little user training, impose low cognitive load, and provide competitive performance [2–3]. EEG is the main modality in ERP-based BCIs for its millisecond temporal precision in capturing neural activity [4], whereas fNIRS measures oxygenated and deoxygenated hemoglobin changes reflecting cerebral hemodynamic response [5]. Although slower, fNIRS is less affected by electromagnetic or muscle artifacts and provides localized cortical information. Its complementary features make it a valuable combination with EEG for mental state assessment [6] including cognitive fatigue [7].

Recent studies have highlighted the use of EEG and fNIRS for detecting cognitive fatigue. Shoaib et al. found that blue light increased HbO and beta-band activity in sleep-deprived participants, while red and green light showed no effect [8]. Dehasis et al. developed a passive EEG–fNIRS BCI for pilots, achieving 87.2% accuracy in simulation and 87.6% in real flights [9]. Similarly, Haman et al. reported that high workload reduced performance, increased frontal theta, and decreased HbO/HbR levels [10]. In ERP-based spellers, performance decline over time has been consistently observed [11,12]. This deterioration is attributed not to task familiarity but to cognitive fatigue, which reduces attention and P300 amplitudes, thus lowering classification accuracy and information transfer rate [11,12]. Increasing stimulus repetitions can mitigate these effects, with accuracy rising from about 57% (two repetitions) to over 80% (ten repetitions) as ERP averaging and signal-to-noise ratio improve [11].

This study addresses an important gap in the literature [13,14] by focusing on the integration of EEG and fNIRS signals to better capture users' cognitive states during speller BCI operation. While previous research has shown that performance decline can be mitigated by increasing stimulus repetitions, conventional systems rely on fixed flash numbers. This static approach may cause unnecessary time loss when users are alert and insufficient repetitions under fatigue. To overcome this, we propose a preliminary framework that evaluates performance across different fatigue levels using EEG and fNIRS. Although real-time adaptation has not yet been implemented, the study establishes a methodological foundation for adaptive speller BCIs that dynamically adjust flash numbers based on users' cognitive states.

2. Material and Method

This section provides a detailed description of the experimental design and implementation. The paradigm, data acquisition setup and subsequent EEG and fNIRS data processing steps are explained in the following subsections. In addition, the overall experimental framework is outlined to clarify how these components were integrated within a coherent methodological structure.

2.1. Paradigm

The experimental protocol was designed to evaluate speller performance while inducing cognitive fatigue. Participant completed four speller blocks and three cognitive fatigue blocks, each lasting about 10 minutes, for a total duration of roughly 1.5 hours including short breaks.

In the speller sessions, stimuli were presented with randomized flash speeds and repetition counts. Flash speeds were defined as slow (0.15 s flash, 0.075 s ISI), medium (0.10 s, 0.05 s ISI), and fast (0.05 s, 0.025 s ISI). Repetitions were low (12), medium (15), or high (17). The number of characters was adjusted so each session lasted about 10 minutes. Across all sessions, participants were presented with a total of 41 target region–character pairs. After each target sequence, participants verbally reported the perceived flash count. Each speller trial followed the regional P300 speller paradigm [15]. In the first phase, four regions each containing seven characters were flashed in random order. Once the target region was detected, the characters within that region were flashed in random order in the second phase. Participants reported the flash count in both phases. This two-stage protocol reduces perceptual errors and improves accuracy compared to the conventional row–column speller.

In the cognitive fatigue sessions, participants solved mental math problems without external aids, responding as quickly and accurately as possible. New questions appeared immediately after each response, and performance measures included number of attempts, correct responses, and reaction times. Breaks were provided between sessions: 3 minutes after speller blocks and 4 minutes after cognitive fatigue blocks. After each fatigue session, participants also completed the NASA-TLX workload questionnaire.

Table 1. Speller configuration and participant reports

block1					block2				
region	character ind	character	flash label	repetition	region	character ind	character	flash label	repetition
1	6	F	fast	17	4	7	Sp	slow	15
4	5	-	fast	12	3	8	X	slow	15
3	3	S	slow	15	4	5	-	fast	17
2	4	L	medium	15	2	1	I	medium	17
3	2	R	fast	12	1	2	B	medium	12
2	7	O	medium	15	4	6	Ok	fast	15
1	8	H	medium	17	1	4	D	medium	15
4	1	Y	fast	12	3	2	R	medium	12
3	4	T	medium	12	2	3	K	slow	17
1	3	C	medium	17	3	7	W	fast	15
2	6	N	slow	17					
block3					block4				
region	character ind	character	flash label	repetition	region	character ind	character	flash label	repetition
1	8	H	slow	15	3	5	U	medium	17
4	5	-	slow	17	1	8	H	fast	12
2	1	I	slow	12	3	6	V	fast	15
1	3	C	slow	12	4	4	,	fast	17
4	4	,	slow	17	2	7	O	medium	17
2	7	O	fast	15	3	3	S	fast	17
3	1	Q	slow	17	2	1	I	medium	15
2	6	N	fast	12	4	5	-	medium	15
1	8	H	slow	15	1	2	B	slow	17
4	2	Z	slow	12	3	1	Q	medium	17

The summary list of predefined flash speed and repetition parameters for each block, target region, and target character are listed in table 1 above. These parameters were fixed prior to the experiment and did not change during the sessions.

2.2. Data Acquisition

EEG and fNIRS data were recorded simultaneously. Stimuli were presented using MATLAB/Psychtoolbox on a 1920×1080 LED monitor positioned ~1 m in front of the participant in a quiet, dimly lit room. EEG was collected with a 32-channel actiCHamp amplifier (Brain Products GmbH, Germany) at 250 Hz, with electrodes placed according to the 10–20 system, reference at Fz, ground on the forehead, and impedances kept below 5 kΩ. fNIRS was acquired from the prefrontal cortex using an NIRx NIRSport2 system with 8 sources and 8 detectors forming multiple channels.

EEG and fNIRS were temporally synchronized through hardware event markers delivered via the parallel port using the io64 interface. Distinct TTL codes indicated block onsets/offsets and region/character flashes, enabling precise stimulus-locked alignment across modalities.

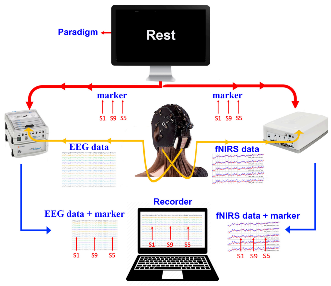


Fig. 1. Experimental setup for EEG–fNIRS recording

Data were collected from a single right-handed, neurologically healthy participant. To improve data quality, the participant abstained from caffeine for 8 h and from smoking, food, and beverages for 3 h prior to the session.

As shown in Figure 1, the setup included a stimulus presentation computer, EEG and fNIRS acquisition devices, and a recording computer. Stimulus events (block, region, and character flashes) were transmitted in parallel to both EEG and fNIRS systems and integrated into the ongoing data streams. The combined dataset ensured synchronized neural, hemodynamic, and behavioral responses for subsequent ERP and hemodynamic analyses.

2.3. EEG and fNIRS Data Processing

EEG signals from speller sessions were preprocessed in EEGLAB, band-pass filtered (0.1–10 Hz), and epoched (–200 to +1000 ms). Baseline correction was applied, and target/non-target flashes were labeled using event markers. ERPs were averaged per block and combined across four blocks. For classification, features were constructed from mean amplitudes of parietal-occipital channels (Pz, Oz, O1) and fed into an artificial neural network (ANN) with one hidden layer (55 neurons). Data were normalized, split into training (70%), validation (15%), and testing (15%), and classification was repeated 50 times to obtain average test accuracies.

For cognitive fatigue sessions, EEG was filtered (1–45 Hz), and channels Fz, Cz, and Pz were selected due to their sensitivity to fatigue. Data were segmented into 3-second windows, and power features were computed for delta, theta, alpha, and beta bands. Features were log-transformed, standardized, and labeled according to block (1, 2, or 3). A linear SVM with five-fold cross-validation was used for three-class classification.

fNIRS signals were preprocessed using nirsLAB, Homer2, and MATLAB scripts. Light intensity was converted to optical density, corrected for motion artifacts, band-pass filtered (0.0001–0.1 Hz), and transformed into HbO/HbR concentrations. PCA regression was applied to reduce systemic noise, and only HbO was retained. Signals were segmented into three blocks and further divided into 25-second epochs with 10-second overlap. Absolute power features were extracted across channels and used in a linear SVM with five-fold cross-validation. Accuracy was averaged across runs to ensure robust classification.

2.4. Experimental Framework

In this study, results from the speller and cognitive fatigue sessions were reported in a structured way, starting with behavioral performance, followed by electrophysiological and hemodynamic analyses, and finally classification outcomes. This organization provided a comprehensive evaluation of participant performance, ERP dynamics, and the discriminative capacity of EEG and fNIRS signals. The results are presented as follows: (1) Behavioral performance in cognitive blocks: number of questions answered, correct responses, and average response times. (2) Block-wise average ERP waveforms: grand averages of target and non-target conditions were compared for a representative channel. (3) Multi-channel ERP analyses: average target ERP responses were shown across channels on a block-by-block basis. (4) Grand average ERP across blocks: all speller blocks were combined to compare target and non-target responses across electrodes. (5) EEG classification of target vs. non-target responses: results demonstrated the discriminability of ERP signals. (6) EEG classification of cognitive blocks: EEG features were used to distinguish the three fatigue blocks. (7) fNIRS classification of cognitive blocks: HbO features were analyzed to classify the three fatigue blocks. This stepwise structure ensured clear and interpretable presentation of behavioral, electrophysiological, and classification results. manner.

3. Results

The results of this study are presented in a stepwise manner, beginning with behavioral findings from the cognitive fatigue sessions, followed by electrophysiological and hemodynamic analyses, and concluding with classification outcomes.

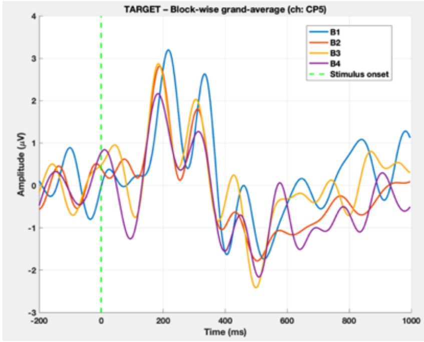
In the first cognitive block, the participant answered 29 questions, 27 of which were correct. In the second block, 26 questions were answered, with 23 correct responses. In the third block, 23 questions were answered, and 20 of them were correct. Both the total number of answered questions and the number of correct responses decreased across consecutive blocks, which may indicate the presence of cognitive fatigue over time.

Table 2. Average response times across cognitive blocks.

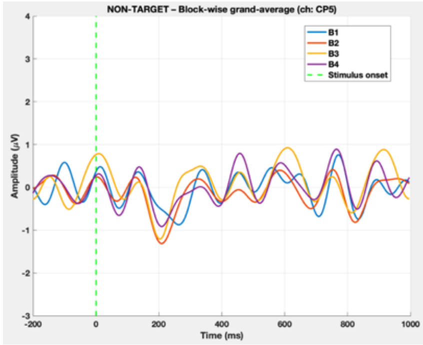
	block1	block2	block3
Average Response Time	20,55s	20,77s	24,07s

Table 2 presents the average response times for answering questions in each cognitive block. As observed, response times gradually increased from block 1 to block 3. This progressive increase in average response duration may indicate rising cognitive load and the possible onset of cognitive fatigue in the participant.

Figure 2 illustrates the superimposed average ERP signals for the CP5 channel across all speller blocks, where (a) shows target responses and (b) shows non-target responses. While ERP latencies appear consistent across blocks, clear differences in amplitude are observed. The highest amplitude is seen in block 1, whereas the lowest amplitude appears in block 4; blocks 2 and 3 fall in between. Previous studies have shown that larger ERP amplitudes are generally associated with increased attention, whereas reduced amplitudes are linked to decreased attention [16,17]. Accordingly, the decline in ERP amplitude from block 1 to block 4 may reflect reduced attentional resources, suggesting the onset of cognitive fatigue. Figure 2(b) presents the superimposed non-target averages for the same channel. As expected, no systematic differences are observed across blocks. This finding supports the notion that non-target responses mainly reflect background neural activity and do not exhibit attentional modulation.



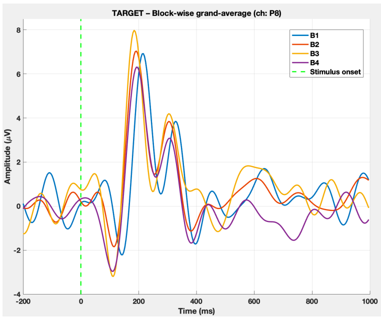
a.



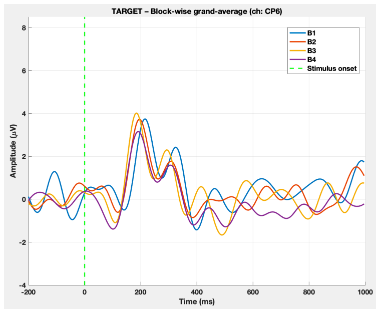
b.

Fig. 2. Block-wise comparison of target (a) and non-target (b) average ERP waveforms for the CP5 channel

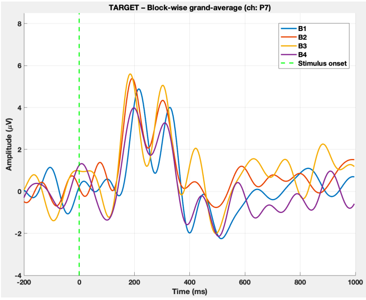
Figure 3 extends the analysis presented in Figure 2 by showing block-wise average ERP responses to target stimuli for four additional channels: P8 (Figure 3a), CP6 (Figure 3b), P7 (Figure 3c), and PO9 (Figure 3d). As illustrated in the graphs, ERP amplitudes are lowest in block 4 across all channels. This pattern suggests a decline in attentional engagement over time, indicating the presence of cognitive fatigue in the participant.



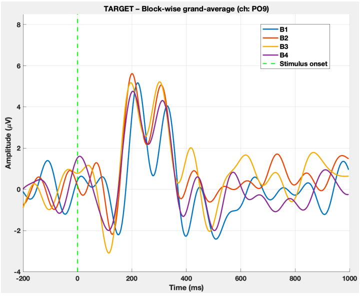
a.



b.



c.

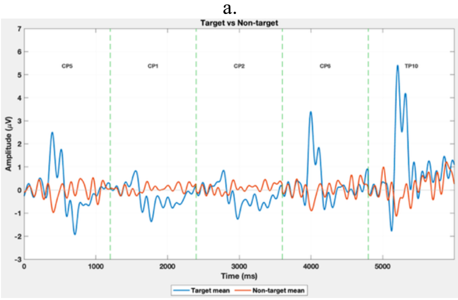


d.

Fig. 3. Block-wise ERP averages across multiple channels

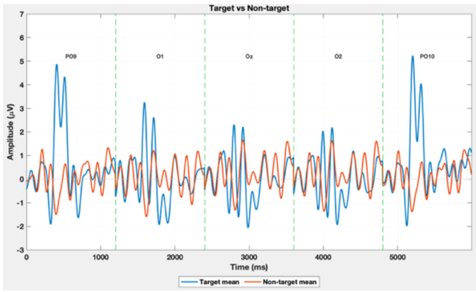
Figure 4 presents the grand average ERP responses obtained by averaging across all four speller blocks (block 1, block 2, block 3, and block 4). In these plots, blue lines represent target responses, while orange lines indicate non-target responses. This visualization allows for a direct comparison of average responses across channels, highlighting where the distinction between target and non-target conditions is most evident. The results show that differences between target and non-target waveforms are particularly pronounced in channels T8, TP9, TP10, CP5, CP6, P7, P8, PO9, and PO10. These findings suggest that such channels provide stronger discriminative power and are likely to be more effective for classification. Indeed, the subsequent classification analyses were informed by these observations, indicating that selecting these channels may improve target versus non-target discrimination.

a.



b.

c.



d.

Fig. 4. Grand average ERPs across channels.

Figure 5 presents the confusion matrix of the average test accuracy obtained from the binary classification of target and non-target data across the entire experiment. An ANN was employed, and the results shown represent the averages over 50 runs. The classification was performed using ERP responses from only three channels (TP9, TP10, and P8), concatenated to form the feature vector. This highlights that even with a limited number of channels, a high classification performance can be achieved. The confusion matrix indicates that target and non-target conditions were classified with very high accuracy, with errors remaining at minimal levels. An average test accuracy of %98.6 ± 1.2 was reported, computed as the mean over 50 runs.

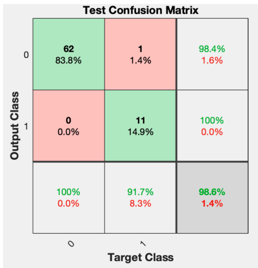


Fig. 5. Confusion matrix of binary classification.

Figure 6 presents the confusion matrix for the three-class classification task aimed at distinguishing between block 1, block 2, and block 3 of the cognitive fatigue sessions based on EEG features. The classification was performed using only a single channel and a single feature, demonstrating the discriminative power of EEG signals even with minimal input. As seen in the confusion matrix, the model successfully differentiated the three blocks with high accuracy across all classes. The overall average test accuracy reached 92.44%, which is considered a strong result for a three-class classification problem. Misclassification rates remained relatively low, indicating that the extracted EEG features captured meaningful distinctions between cognitive states across the different blocks.

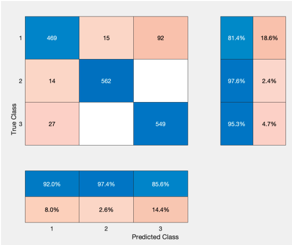


Fig. 6. EEG-Based Cognitive Block Classification

Figure 7 presents the confusion matrix for the classification of cognitive blocks using fNIRS signals, similar to the results previously obtained with EEG. For each block (block 1, block 2, and block 3), band power features were extracted and subjected to three-class classification. The results demonstrate that fNIRS signals can also be effectively used to distinguish cognitive blocks, supporting the findings derived from EEG data. The average test accuracy was 87.38%.

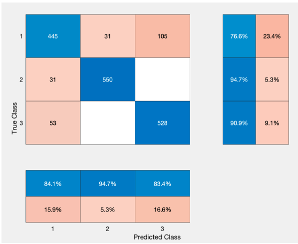


Fig. 7. fNIRS-Based Cognitive Block Classification

6. Conclusions

This pilot study introduced a hybrid EEG–fNIRS framework to evaluate speller performance and cognitive fatigue. Speller systems face a trade-off: more flashes improve ERP signal-to-noise ratio (SNR) and accuracy but prolong trials, whereas fewer flashes are efficient when users are alert but insufficient under fatigue. This study aimed to address these limitations by exploring real-time estimation of cognitive state and adaptive adjustment of flash numbers. Behavioral results showed a decline across cognitive blocks: 27 correct out of 29 in block 1, 23 out of 26 in block 2, and 20 out of 23 in block 3, with increasing response times, indicating fatigue. ERP analyses confirmed reduced amplitudes across blocks, strongest in block 1 and weakest in block 4, especially at parietal and occipital sites (e.g., TP9, TP10, P8, CP5, CP6, P7, PO9, PO10). Classification results further supported these findings. Target vs. non-target discrimination reached 98.6% accuracy with three channels (TP9, TP10, P8). Cognitive block classification achieved 92.4% with EEG and 87.4% with fNIRS, showing EEG's temporal precision and fNIRS' complementary hemodynamic sensitivity.

References

[1] Li, R., Yang, D., Fang, F., Hong, K. S., Reiss, A. L., & Zhang,Y. (2022). Concurrent fNIRS and EEG for brain function investigation: a systematic, methodology-focused review. Sensors, 22(15), 5865.

[2] Lee, M. H., Kwon, O. Y., Kim, Y. J., Kim, H. K., Lee, Y. E., Williamson, J., ... & Lee, S. W. (2019). EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy. GigaScience, 8(5), giz002.

[3] Rezeika, A., Benda, M., Stawicki, P., Gembler, F., Saboor, A., & Volosyak, I. (2018). Brain–computer interface spellers: A review. Brain sciences, 8(4), 57.

[4] Wu, L., Liu, A., Ward, R. K., Wang, Z. J., & Chen, X. (2023). Signal Processing for Brain–Computer Interfaces: A review and current perspectives. IEEE Signal Processing Magazine, 40(5), 80-91.

[5] Ayaz, H., Baker, W. B., Blaney, G., Boas, D. A., Bortfeld, H., Brady, K., Brake, J., Brigadoi, S., Buckley, E. M., Carp, S. A., Cooper, R. J., Cowdrick, K. R., Culver, J. P., Dan, I., Dehghani, H., Devor, A., Durduran, T., Eggebrecht, A. T., Emberson, L. L.,…Zhou, W. (2022). Optical imaging and spectroscopy for the study of the human brain: status report. Neurophotonics, 9(S2), S24001. https://doi.org/10.1117/1.NPh.9.S2.S24001.

[6] Liu, Y., Ayaz, H., & Shewokis, P. A. (2017). Multisubject "Learning" for Mental Workload Classification Using Concurrent EEG, fNIRS, and Physiological Measures. Frontiers in Human Neuroscience, 11(389). https://doi.org/10.3389/fnhum.2017.00389.

[7] Varandas, R., Lima, R., Bermúdez i Badia, S., Silva, H., & Gamboa, H. (2022). Automatic cognitive fatigue detection using wearable fNIRS and machine learning. Sensors, 22(11), 4010.

[8] Shoaib, Z., Akbar, A., Kim, E. S., Kamran, M. A., Kim, J. H., & Jeong, M. Y. (2023). Utilizing EEG and fNIRS for the detection of sleep-deprivation-induced fatigue and its inhibition using colored light stimulation. Scientific Reports-Nature, 13(1), 6465.

[9] Dehais, F., Dupres, A., Di Flumeri, G., Verdiere, K., Borghini, G., Babiloni, F., & Roy, R. (2018, October). Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. In 2018 IEEE international conference on systems, man, and cybernetics (SMC) (pp. 544-549). IEEE.

[10] Hamann, A., & Carstengerdes, N. (2022). Investigating mental workload-induced changes in cortical oxygenation and frontal theta activity during simulated flights. Scientific Reports-Nature, 12(1), 6449.

[11] Marassi, A., Budai, R., & Chittaro, L. (2018). A P300 auditory brain-computer interface based on mental repetition. Biomedical Physics & Engineering Express, 4(3), 035040.

[12] Schembri, P., Pelc, M., & Ma, J. (2020). Impact of Mental Fatigue during Repetitive Exercises of a Visual P300 Speller. In BIOSIGNALS (pp. 156-163).

[13] Finnis, R., Mehmood, A., Holle, H., & Iqbal, J. (2025). Exploring Imagined Movement for Brain–Computer Interface Control: An fNIRS and EEG Review. Brain Sciences, 15(9), 1013.

[14] Shi, X., Wang, H., Li, B., Qin, Y., Peng, C., & Lu, Y. (2025). Fusion analysis of EEG-fNIRS multimodal brain signals: a multitask classification algorithm incorporating spatial-temporal convolution and dual attention mechanisms. IEEE Transactions on Instrumentation and Measurement.

[15] R. Fazel-Rezai and K. Abhari, "A region-based P300 speller for brain-computer interface," Can. J. Elect. Comput. Eng., vol. 34, no. 3, pp. 81–85, 2009.

[16] Koivisto, M., Kainulainen, P., & Revonsuo, A. (2009). The relationship between awareness and attention: evidence from ERP responses. Neuropsychologia, 47(13), 2891-2899.

[17] Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: an ERP study. Cognitive brain research, 25(1), 107-116.ma