Feasibility of an EEG-based dynamic suboptimal cognitive monitoring for field neuroergonomics

Eren (M.) Shahrokhi[1], S. Nizam Ahmed[2], Gaang Lee[1] and SangHyun Lee[3]

[1]Hole School of construction Engineering and Management, Dept. of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada
[2] Clinical Neurophysiology Laboratory, Division of Neurology, Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada
[3] Tishman Construction Management Program, Dept. of Civil and Environmental Engineering, University of Michigan, 2350 Hayward St., Ann Arbor, MI 48109, USA
shahrok1@ualberta.ca, snahmed@ualberta.ca, gaang@ualberta.ca, shdpm@umich.edu

Abstract

Suboptimal cognitive states among construction workers significantly impact safety and productivity, with mental workload playing a key role in triggering these states. Determining if the mental workload fluctuation is leading to an error is challenging as the relationship between mental workload and suboptimal cognitive states is complex and non-linear, with traditional theories failing to map their fluctuations effectively. Recently, a two-dimensional space has been introduced to theoretically map mental workload fluctuations and suboptimal cognitive states using task engagement and arousal. However, there is currently no framework in place to continuously apply this theoretical knowledge in practical settings. To address this gap, this study investigates the feasibility of EEG-based frameworks for classifying four different cognitive states, namely comfort zone, mind wandering, effort withdrawal, and inattentional blindness, based on mental workload fluctuations. EEG signals were collected from 10 participants using a headset with dry electrodes, processed to extract relevant features, and classified using Support Vector Machine (SVM) and Artificial Neural Network (ANN) models. The ANN achieved superior performance in k-fold and leave one period out validation methods, though accuracy declined in leave one subject out validation. These findings underscore the potential of EEG-based differentiation of cognitive suboptimalities to enhance safety and productivity in construction by providing crucial information about when construction workers are most likely to make cognitive errors, which is essential for timely and appropriate interventions. Also, the low subject independent accuracy emphasizes the need to address individual differences in EEG signals for broader applicability.

Keywords

EEG, Wearables, Biosensors, Neuroergonomics, Construction safety

1 Introduction

Managing workers' mental workload is a critical aspect of construction management [1]. The dynamic nature of construction sites often subjects workers to varying levels of mental workload, ranging from excessively low to overly high. These suboptimal cognitive states increase the likelihood of unsafe or erroneous work behavior, making it essential to monitor and manage mental workload for both safety and productivity.

Traditionally, mental workload monitoring has been conducted through self-assessment methods, such as questionnaires and interviews [2,3]. However, these methods are highly subject to individuals' recall bias and sporadic due to their invasive nature, making them unsuitable for continuous and dynamic monitoring, which is critical for understanding the dynamics of mental workload oscillation during the ongoing task. To address these limitations, neuroergonomics, which is a field of research to study human brain function in the workplace [4], has been integrated with various wearable-based techniques capable of continuous monitoring and assessing the mental workload. These studies have applied different wearable biosensors, such as electrocardiogram (ECG) [5], electroencephalogram (EEG) [6], photoplethysmography (PPG) [7], and Electrodermal activity (EDA) [8], to measure individuals' mental workload levels during their ongoing tasks.

The previous efforts have enabled the monitoring of changing mental workload levels. For example, [9] applied EEG to quantify the construction workers mental workload and the findings suggest that Gamma band activity of specific EEG channels strongly correlates with tasks with high mental demands. However, challenges remain in distinguishing problematically low or high levels—those that involve suboptimal cognitive states, such as mind wandering and blindness, and increased susceptibility to cognitive errors—from acceptable/appropriate levels during ongoing field tasks. The current literature does not establish thresholds for problematic mental workload, leaving no clear reference for identifying activities that may induce suboptimal cognitive states in workers. This limitation hinders the operationalization of biosensor-based mental workload monitoring to improve workers' cognitive states by identifying and addressing problematic moments and environmental factors likely to cause suboptimal cognitive states.

A recently introduced framework [10] has shown potential not only in addressing this limitation, but also in enabling the differentiation of different suboptimal states for interventions, by characterizing problematically low and high mental workload in relation to neurophysiological states. The framework, proposed by Dehais [10] and his team, reconceptualizes mental workload as a transactional interaction between an individual and a task's cognitive demands, influenced by intra-individual factors (e.g., emotional states, fatigue) and inter-individual variability (e.g., skill levels, personality). It categorizes neurophysiological states along with their associated suboptimal cognitive states and their symptoms that emerge when mental workload becomes problematically low or high. Specifically, an individual's neurophysiological state, depending on their mental workload, can be mapped onto a two-dimensional space defined by task engagement (the cognitive and emotional effort invested in achieving task goals) and arousal (the physiological readiness to respond to goals and external task demands, influenced by inherent traits and situational factors). This two-dimensional space is divided into four regions: three suboptimal cognitive states—Mind Wandering, Effort Withdrawal, and Mind Blindness—and one optimal cognitive state (Figure 1).

Mind Wandering occurs when mental workload is excessively low, leading to boredom and a lack of focus on the task. Effort Withdrawal and Inattentional Blindness, on the other hand, arise from excessively high mental workload levels, representing two distinct suboptimal states characterized by reduced task engagement and diminished awareness of unexpected stimuli, respectively. Differentiating these three mental workload-related suboptimal cognitive states not only enables the detection of problematic mental workload, but also provides insights for "state-specific" interventions, thereby significantly contributing to the field application of the wearable biosensor-based neuroergonomic approach. This framework takes a more detailed look at mental workload by considering both the intensity and quality of cognitive engagement. It offers a clearer way to understand how individual cognitive states influence performance on complex tasks.

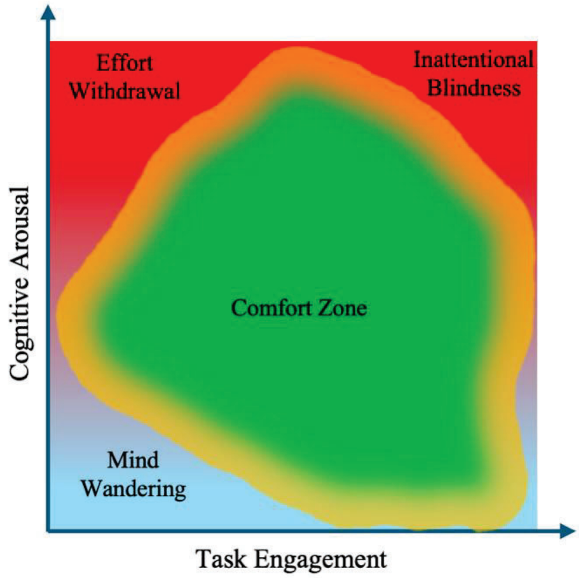

Figure 1. Dehais's framework characterizing mental workload-related suboptimal states [10]

To effectively implement this sophisticated framework in practical settings, an equally advanced method of analyzing the collected neurophysiological data is essential. In the field of neuroergonomics, traditional approaches to analyzing the signals collected during cognitive monitoring typically employ either post-hoc analysis or machine learning methods. While post-hoc analysis has been valuable for in-depth, after-the-fact evaluations, it does not facilitate the continuous and dynamic adjustments needed for ongoing cognitive state monitoring. Therefore, the integration of machine learning analysis and the new perspective on mental workload shows significant potential for effective and dynamic monitoring of field works. Despite this potential, the feasibility of this framework in accurately identifying distinct cognitive states remains untested. This gap signifies a crucial missed opportunity for enhancing operational efficiency, safety, and worker well-being through state-specific cognitive interventions. To address this knowledge gap, this study aims to investigate the feasibility of developing an EEG-based machine learning model that can differentiate between four cognitive states derived from the novel perspective on operationalized mental workload.

2 Methods

To achieve the research objective, a machine learning model was developed for classifying different suboptimal cognitive states (i.e., Mind Wandering, Effort

Withdrawal, and Mind Blindness), along with the optimal. To this end, an in-lab data collection was structured incorporating the characteristics of the suboptimal cognitive states stated in Dehais's framework and conducted. Following this, the data preprocessing techniques were employed to ensure the EEG signals were clean and reliable. The focus then shifted to the feature extraction process, which is crucial for preparing the data for effective machine learning analysis and classification model development.

2.1 Data collection

2.1.1 Participants

Data were collected from 10 participants aged 20 to 39 years, comprising 5 females and 5 males. All participants reported being in good physical and mental health, with no history of neurological disorders. To minimize confounding factors, participants were instructed to abstain from alcohol for 24 hours before the session. Each participant completed only the data collection session, with no additional activities scheduled before or after. The study protocols were approved by the University of Alberta's Research Ethics Board (Reference: Pro00135533), and participants provided informed consent via an electronic form before starting.

2.1.2 Tools

The study employed the DSI-24 dry-electrode EEG headset from Wearable Sensing for data acquisition. This device was chosen for its practical advantages, including reduced setup time, stable signal quality without the need for conductive gel, and comfort during extended use. In addition, the patented common-mode follower technology in this headset offers near immunity to electrical and motion artifact signals during data collection [11]. The headset features 19 channels based on the 10-20 system, a sampling rate of 300 Hz, and an active shielding system to filter environmental noise. Wireless data collection ensured flexibility and minimized interference during the session.

The data collection session was designed using the PsyToolkit platform [12,13], while the iMotion platform was utilized to integrate biosensors and stimuli for improved synchronization. This setup enabled the efficient execution of complex data collection designs.

2.1.3 Procedures

The data collection session involved three cognitive tests, each designed to target specific suboptimal cognitive states characterized by Dehais's framework [10]. First, an adapted Task-Switching test was used to induce inattentional blindness [14]. Participants alternated between rules for responding to stimuli, such as letters and numbers in a grid, under time constraints. The frequent rule changes introduced cognitive demands, simulating real-world multitasking challenges. Participants had only 4 seconds to respond to each stimulus; failing to do so triggered an error message. This setup required participants to concentrate intensely at the beginning of each block to apply the newly presented rule effectively and perform well. The intense focus and task engagement accompanied with high cognitive arousal due to time constraints, can lead to Inattentional blindness and deafness as shown in Figure 1. Each participant completed six Task-Switching sessions, each lasting 2.5 minutes, with randomized rules for each session. Breaks were allowed between sessions if needed.

The second test was the n-back task [15], designed to induce effort withdrawal. Participants were required to identify matches in a sequence of stimuli presented a specific number of steps (n) earlier, with task difficulty increasing incrementally from 1-back to 7-back. When the number of steps is higher than 5 the task looks impossible to individuals [16] and discourages them from fully engaging in the task since achieving the goal seems very unlikely. This procedure can induce Effort Withdrawal behavior among participants. Four n-back tasks were administered, each lasting 2.5 minutes, with no breaks between levels to maintain cognitive load and push participants toward a mental breaking point.

The third test was the Sustained Attention to Response Task (SART) [17], conducted to induce mind-wandering. Participants responded to any digit from 1 to 9 while withholding responses to an exception, the digit "3". Participants had 0.9 seconds to respond to the digits, and since the target digit was infrequent, they frequently had to respond to other stimuli. Over time, this repetitive response to non-target digits could become an automatic process requiring minimal attention, potentially allowing their freed-up focus to drift to other tasks. This cognitive shift increases the likelihood of Mind Wandering. Five SART sessions were presented toward the end of the session, leveraging reduced cognitive arousal after the more demanding preceding tasks. Each session lasted 4 minutes.

Upon arrival, participants were introduced to the biosensors and provided informed consent before setup. The tests were conducted in a quiet, isolated room to minimize distractions. The entire data collection session, including setup and breaks, lasted between 1.5 and 2 hours. This duration was carefully optimized to minimize fatigue while ensuring sufficient data quality and variability for the machine learning model.

2.2 Data preprocessing

EEG data were continuously recorded throughout the session without interruptions. To prepare the data for analysis, a 60 Hz notch filter was first applied (to account for the powerline frequency in Canada), followed by bandpass filtering between 0.5 Hz and 50 Hz to reduce extrinsic artifacts. Despite these preprocessing steps, the data remained prone to noise from intrinsic artifacts such as eye blinks and muscle movements. To mitigate these, Independent Component Analysis (ICA) was employed to isolate and remove artifact-related components, enhancing the overall signal-to-noise ratio.

After preprocessing, the data were prepared for labelling. Since no self-reported measurements were used, the labelling strategy was based on the design and cognitive mechanisms of the tasks, supported by previous research. Each section's labelling was validated separately to ensure the tasks effectively induced the intended cognitive suboptimalities.

For Task Switching, it was hypothesized that participants initially struggle with new rules under time constraints but adapt as the session progresses, feeling more at ease. Therefore, first 30 seconds of each new trial was labelled as inattentional blindness. For effort withdrawal, the Task Engagement Index (TEI) was calculated during the 1-back, 3-back, 5-back, and 7-back tasks. The TEI increased from 1-back to 5-back but significantly dropped for the 7-back task, reflecting a state of effort withdrawal as participants found the task too demanding to complete effectively. Hence, the 7-back task data were used as effort withdrawal data.

For SART, errors were captured using a setup where an error message was displayed alongside a white box with a photodetector attached to the screen. This system pinpointed the exact moments when participants made errors. Following the methodology of previous studies [18], the 5 seconds preceding each error were labeled as mind-wandering states. After the error message appeared, participants were prompted to refocus and attempt better performance, and these moments were labeled as focus states.

2.3 Feature Extraction

After labeling the data, features were extracted from the EEG signals to prepare inputs for the machine learning model and analyze signal patterns linked to various suboptimal cognitive states. An effective window size for EEG data segmentation in cognitive state classification is typically determined through trial and error and generally ranges from 1 to 12 seconds [19]. The optimal window size in this study was found at using a 2-second window with a 25% overlap to ensure sufficient temporal resolution.

Feature extraction is a critical step in developing robust EEG-based machine learning models for classifying cognitive states. Selecting features that effectively capture the underlying neural dynamics is essential for accurate and meaningful classification. From the literature, features derived from both time and frequency domains, as well as non-linear analyses, have demonstrated their effectiveness in distinguishing between mental states and cognitive processes. Specifically, power spectral features, such as Alpha, Beta, Theta, and Delta Power, offer insights into the brain's frequency-specific activities related to relaxation, focus, drowsiness, and disengagement, respectively. Complementing these, non-linear and time-domain measures, such as Higuchi Fractal Dimension and Hjorth parameters (Activity, Mobility, and Complexity), provide valuable information on neural complexity, brain activity levels, cognitive flexibility, and adaptability. Together, these features encapsulate the multifaceted nature of brain activity, making them a suitable choice for the classification of various cognitive states [20].

A total of eight features were computed for each EEG channel: alpha, beta, theta, and delta power bands; Higuchi fractal dimension; and Hjorth parameters (activity, mobility, and complexity). With 19 channels in the EEG system, this resulted in 152 features (8 features × 19 channels) being extracted for each window. A summary of the selected features, along with the rationale for their selection and their significance in EEG signals, is presented in Table 1.

Table 1. Summary of feature selection

Feature	Description
Alpha Power	Represents relaxed but wakeful states and reduced sensory input, typically decreasing with increased cognitive load
Beta Power	Linked to focus, problem-solving, and sustained attention, often increasing with cognitive effort
Theta Power	Associated with drowsiness and working memory, helpful in identifying low vigilance or mental fatigue
Delta Power	Indicates low-frequency activity often tied to relaxation or cognitive disengagement in awake individuals
Higuchi Fractal Dimension (HFD)	Captures neural complexity, with higher values reflecting increased cognitive processing
Hjorth Activity	Reflects overall brain activity levels, useful for assessing mental workload or stress
Hjorth Mobility	Represents frequency variability, linked to cognitive flexibility
Hjorth	Indicates adaptability and mental engagement, highlighting the temporal

Complexity	intricacy of brain activity

These features provide a comprehensive representation of brain dynamics, enabling the model to differentiate between distinct cognitive suboptimalities effectively.

2.4 Multiclass classification of suboptimal cognitive states

After completing the feature extraction and labelling stages, a feature matrix and a target matrix were constructed to train machine learning models for classifying various suboptimal cognitive states: Mind Wandering, Effort Withdrawal, Inattentional Blindness, and the Comfort Zone.

In this study, Support Vector Machines (SVM) and Artificial Neural Networks (ANN) were utilized to classify EEG-based cognitive states due to their complementary capabilities in processing complex, high-dimensional data. SVM was employed for its effectiveness in identifying nonlinear decision boundaries and its suitability for smaller datasets, leveraging kernel functions to map EEG data into higher-dimensional spaces for enhanced separability. Fully connected ANN was selected for its capacity to model intricate patterns and automatically learn features from EEG signals, accommodating the data's non-stationary and multivariate characteristics. Both models were trained and evaluated on preprocessed EEG data to compare their classification performance across different cognitive states.

In machine learning research, validation methods play a crucial role in assessing the generalizability of models to unseen data [21,22]. Among the commonly used approaches is K-Fold Cross-Validation (KFCV), which splits the dataset into k equal-sized subsets or "folds." The model is trained on k-1 folds and tested on the remaining fold, cycling through until each subset has been used for testing. While this method is efficient and widely applicable, it assumes independence between data points. This assumption can lead to overestimation of model performance, particularly when applied to time-series or biosignal data, where neighboring data points are inherently correlated. The result of this validation method is used as a benchmark to compare the developed model with existing studies since this validation method is frequently used in the current literature [19,23].

The second method, Leave-One-Period-Out Cross Validation (LOPOCV), is designed to evaluate the generalizability of machine learning models across different environmental or contextual factors. In this approach, the dataset is divided so that data from one specific context—such as a particular environmental condition, task type, or experimental setup—is completely excluded during training and reserved for testing. This process is repeated iteratively, with each context taking a turn as the test set, ensuring that every context is evaluated independently.

Finally, Leave-One-Subject-and-Context-Out Cross-Validation (LOSCOCV) [24] was used to ensures both subject and context independence. By excluding data from one subject and one context during training and using them for testing, LOSCOCV provides a more stringent evaluation of generalizability. This method is particularly suited for biosignal research, where both individual and contextual variability significantly influence the data. Incorporating these three validation methods into the methodology ensures a comprehensive evaluation of model performance and generalizability under varying conditions.

To optimize performance, hyperparameter tuning was conducted for both the SVM and ANN models. The holdout method was employed for this process, with 8 subjects used for training and 2 set aside as the validation dataset. Models were trained and tested on this subset to identify the best hyperparameters, which were subsequently applied to the full dataset of 10 subjects. The selected hyperparameters ensured the models achieved robust performance across all validation methods.

3 Results

The classification models were trained using the SVM and ANN algorithms, with hyperparameters optimized to achieve the best performance. The implemented ANN has a fully connected architecture with 3 hidden layers and 2 drop out layers. The hyperparameters were determined using a grid search method. Final architecture and hyperparameter setting of SVM and ANN are presented presented in Table 2.

Table 2. Hyper parameter tunning results

Model	Hyperparameter and Architecture
SVM	Kernel: RBF C: 10 / Gamma: 0.1
ANN	Input layer: 152 neurons 1st Hidden layer: 100 neurons 1st Dropout layer: 15 percent 2nd Hidden layer: 100 neurons 2nd Dropout layer: 15 percent 3rd Hidden layer: 75 neurons Output layer: 4 neurons

Model performance was evaluated using three validation methods—KFCV, LOPOCV, and LOSCOCV—and the results are presented in Table 3. The ANN outperformed the SVM across all validation methods, achieving the highest accuracy in both KFCV and LOPOCV validations. These results align with previous studies that differentiate between levels of specific cognitive suboptimalities, providing benchmarks for comparison. Importantly, this study highlights the challenge of achieving high accuracy in LOSCOCV models, a critical area for future research and practical applications.

Table 3. Accuracy of different models for each validation method

Model	Validation method					
	KFCV		LOPOCV		LOSCOCV	
	Acc	F1	Acc	F1	Acc	F1
SVM	0.62	0.62	0.58	0.59	0.24	0.10
ANN	0.72	0.71	0.68	0.65	0.27	0.24

* ACC: classification accuracy

4 Discussion

Traditional approaches to mental workload have faced challenges in distinguishing various suboptimal cognitive states, as the relationship between mental workload and cognitive errors induced by these states remained unclear. Recently, new perspectives have emerged, focusing on measuring task engagement and arousal to map suboptimal cognitive states related to mental workload, opening new opportunities to enhance field applicability. This study explores the feasibility of this approach by employing a neuroergonomic framework to map suboptimal cognitive states resulting from excessive mental workload. A machine learning model was developed to classify four distinct cognitive states using features extracted from EEG signals.

The study classifies cognitive comfort zone, mind wandering, inattentional blindness, and effort withdrawal with acceptable accuracy. Given that the task involves a 4-class classification, accuracies above 0.7 can be considered promising. Both classification accuracy and F1-scores were reported to mitigate the effects of data imbalance and validate the performance of the models. Among the models tested, the ANN trained with extracted features outperformed the SVM, achieving the highest accuracy. This result aligns with expectations, given the ANN's ability to handle the high number of features and capture complex relationships between input data and cognitive state labels. While the SVM demonstrated reasonable performance, its simpler architecture limits its ability to manage the intricacies of the dataset compared to the ANN.

Accuracy decreased as validation methods introduced greater independence across time periods or subjects, highlighting the challenges of generalization. However, the ANN demonstrated robust performance under LOPOCV, showcasing its potential for generalization within the same participant group. This validation method ensures that testing occurs on unseen time periods and data collection sessions, meaning the model can differentiate classes with acceptable precision even when the participant remains the same but the data collection setup changes. However, the accuracy for LOSCOCV was notably low across all models, underscoring the difficulty of achieving generalizability across different individuals. This limitation highlights the need for additional labeled data from new subjects to train models effectively, a process that is both time-consuming and inefficient.

This study demonstrates the feasibility of using machine learning combined with EEG-extracted features to classify cognitive states, including both the comfort zone and cognitive suboptimalities. This framework holds significant potential for continuous, dynamic monitoring of construction field workers with minimal invasiveness and without disrupting their tasks. Its neuroergonomic foundation enhances reliability by leveraging data directly from the central nervous system.

The benefits of such a framework are multifaceted. It enables timely interventions to mitigate risks associated with cognitive suboptimalities, thereby reducing cognitive errors and enhancing worker safety and performance. Furthermore, it offers indirect applications, such as integration with geographic information systems (GIS) or eye-tracking data for hotspot analyses. This integration can help identify areas or environmental factors where cognitive suboptimalities frequently occur, facilitating targeted interventions to improve workplace conditions.

This study sets a benchmark for multiclass classification of cognitive suboptimalities but has some limitations that need further exploration. First, the tested framework shows low generalizability across different subjects, which means there is a need to individualize the model for each user or increase the dataset diversity by data augmentation. Second, the proposed technique requires extensive labelled data collection eliciting cognitive suboptimalities, highlighting the need for a combination of improved protocols and techniques that reduce the reliance on learning from scratch to quickly build more reliable datasets. This prolonged data collection process means the developed model may require many hours of data acquisition and hyperparameter tuning whenever a new user is introduced, significantly increasing implementation time and effort. Third, as the models were developed and tested with data collected in a controlled environment, the additional finetuning and/or validation might be required in more realistic work scenarios to examine the field applicability.

Given the lack of a clear solution in the current literature for developing a generalizable model across individuals and the absence of a comprehensive dataset with sufficient representation from diverse demographic groups, future research should explore strategies to simplifying the data collection and hyperparameter tuning process for new users. Reducing the time required for these tasks would enhance the framework's cost-effectiveness and scalability, opening new possibilities for industrial applications.

Addressing this technical challenge is crucial because the practical use of current EEG-based models is hindered by their limited subject independence. Future studies need to focus on enhancing subject independence to facilitate wider use and real-world implementation of these models.

5 Conclusion

This study demonstrates the feasibility of using mobile EEG-based machine learning models to detect and classify suboptimal cognitive states, highly relevant to construction workers' cognitive errors in the field. By leveraging neuroergonomic principles and advanced validation methods, the proposed framework successfully differentiates between four distinct cognitive states: mind wandering, inattentional blindness, effort withdrawal, and the comfort zone. Among the models tested, the ANN consistently outperformed the SVM in classification tasks, demonstrating superior capability in handling high-dimensional EEG data and capturing complex relationships. The findings underscore the potential of EEG-based frameworks for real-time and continuous monitoring of cognitive states in demanding environments like construction, with minimal invasiveness and without disrupting workers' tasks.

Despite its success, the study also identifies significant challenges, particularly in achieving generalizability across individuals. Accuracy decreased when validation methods introduced greater independence across subjects or contexts, highlighting the need for extensive retraining with new user-specific data. This limitation points to a critical area for future research: enhancing model generalizability to enable broader applicability without compromising accuracy. Additionally, the time-intensive process of collecting labeled data and tuning model hyperparameters further underscores the necessity of more efficient protocols and comprehensive datasets representing diverse demographics.

The implications of this work are both immediate and far-reaching. In the short term, the framework offers a robust tool for identifying cognitive suboptimalities in real-time, enabling timely interventions to mitigate risks, reduce cognitive errors, and improve worker safety and performance. In the long term, its integration with other technologies, such as GIS or eye-tracking systems, could provide actionable insights into environmental factors that contribute to cognitive errors. This integration could facilitate targeted interventions to optimize workplace conditions and enhance overall productivity.

Future research should focus on addressing these limitations by improving data collection, developing generalizable models that work across individuals and contexts, and validating the framework in larger and more diverse populations. Expanding its practical applications to other industries and refining its usability for real-world conditions could further enhance its impact. This study establishes a critical foundation for EEG-based neuroergonomic frameworks, advancing our ability to monitor and manage mental workload dynamically and effectively in the field settings.

Acknowledgement

This work was supported by the University of Alberta Research Exploration Funds, the Construction Innovation Centre, and the New Frontiers in Research Fund (NFRF), grant number NFRFE-2023-00174. The authors also wish to thank their industry partners for their support in data collection, as well as the anonymous participants who contributed their time and insights.

References

[1] J. Chen, X. Song, Z. Lin, Revealing the "Invisible Gorilla" in construction: Estimating construction safety through mental workload assessment, Autom Constr 63 (2016) 173–183. https://doi.org/10.1016/J.AUTCON.2015.12.018.

[2] W. Umer, H. Li, Y. Yantao, M.F. Antwi-Afari, S. Anwer, X. Luo, Physical exertion modeling for construction tasks using combined cardiorespiratory and thermoregulatory measures, Autom Constr 112 (2020) 103079. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103079.

[3] H.J. Michielsen, J. De Vries, G.L. Van Heck, Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale, J Psychosom Res 54 (2003) 345–352. https://doi.org/10.1016/S0022-3999(02)00392-6.

[4] R. Parasuraman, M. Rizzo, Neuroergonomics: The brain at work, Oxford University Press, 2006.

[5] H. Qu, X. Gao, L. Pang, Classification of mental workload based on multiple features of ECG signals, Inform Med Unlocked 24 (2021) 100575. https://doi.org/https://doi.org/10.1016/j.imu.2021.100575.

[6] W.K.Y. So, S.W.H. Wong, J.N. Mak, R.H.M. Chan, An evaluation of mental workload with

frontal EEG, PLoS One 12 (2017) e0174949. https://doi.org/https://doi.org/10.1371/journal.pone.0174949.
[7] W.-K. Beh, Y.-H. Wu, A.-Y. Wu, Robust PPG-based mental workload assessment system using wearable devices, IEEE J Biomed Health Inform 27 (2021) 2323–2333. https://doi.org/10.1109/JBHI.2021.3138639.
[8] C. Lee, M.J. Shin, D. Eniyandunmo, A. Anwar, E. Kim, K. Kim, J.K. Yoo, C. Lee, Predicting Driver's mental workload using physiological signals: A functional data analysis approach, Appl Ergon 118 (2024). https://doi.org/10.1016/j.apergo.2024.104274.
[9] J. Chen, J.E. Taylor, S. Comu, Assessing Task Mental Workload in Construction Projects: A Novel Electroencephalography Approach, J Constr Eng Manag 143 (2017). https://doi.org/10.1061/(asce)co.1943-7862.0001345.
[10] F. Dehais, A. Lafont, R. Roy, S. Fairclough, A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance, Front Neurosci 14 (2020). https://doi.org/10.3389/fnins.2020.00268.
[11] Y. Mahdid, U. Lee, S. Blain-Moraes, Assessing the quality of wearable EEG systems using functional connectivity, IEEE Access 8 (2020) 193214–193225. https://doi.org/10.1109/ACCESS.2020.3033472.
[12] G. Stoet, PsyToolkit: A software package for programming psychological experiments using Linux, Behav Res Methods 42 (2010) 1096–1104. https://doi.org/10.3758/BRM.42.4.1096.
[13] G. Stoet, PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments, Teaching of Psychology 44 (2017) 24–31. https://doi.org/10.1177/0098628316677643.
[14] S. Monsell, Task switching, Trends Cogn Sci 7 (2003) 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7.
[15] M.J. Kane, A.R.A. Conway, T.K. Miura, G.J.H. Colflesh, Working memory, attention control, and the N-back task: a question of construct validity., J Exp Psychol Learn Mem Cogn 33 (2007) 615. https://doi.org/https://psycnet.apa.org/doi/10.1037/0278-7393.33.3.615.
[16] S.H. Fairclough, L.J. Moores, K.C. Ewing, J. Roberts, Measuring task engagement as an input to physiological computing, in: 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009: pp. 1–9. https://doi.org/10.1109/ACII.2009.5349483.
[17] I.H. Robertson, T. Manly, J. Andrade, B.T. Baddeley, J. Yiend, Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects, Neuropsychologia 35 (1997) 747–758. https://doi.org/https://doi.org/10.1016/S0028-3932(97)00015-8.
[18] C. Braboszcz, A. Delorme, Lost in thoughts: Neural markers of low alertness during mind wandering, Neuroimage 54 (2011) 3040–3047. https://doi.org/10.1016/J.NEUROIMAGE.2010.10.008.
[19] H. Jebelli, S. Hwang, S.H. Lee, EEG-based workers' stress recognition at construction sites, Autom Constr 93 (2018) 315–324. https://doi.org/10.1016/J.AUTCON.2018.05.027.
[20] J. Ruiz de Miras, A.J. Ibáñez-Molina, M.F. Soriano, S. Iglesias-Parro, Schizophrenia classification using machine learning on resting state EEG signal, Biomed Signal Process Control 79 (2023). https://doi.org/10.1016/j.bspc.2022.104233.
[21] M. Shahrokhishahraki, M. Malekpour, S. Mirvalad, G. Faraone, Machine learning predictions for optimal cement content in sustainable concrete constructions, Journal of Building Engineering 82 (2024). https://doi.org/10.1016/j.jobe.2023.108160.
[22] V. Asghari, M. Hossein Kazemi, M. Shahrokhishahraki, P. Tang, A. Alvanchi, S.-C. Hsu, Process-oriented guidelines for systematic improvement of supervised learning research in construction engineering, Advanced Engineering Informatics 58 (2023) 102215. https://doi.org/10.1016/j.aei.2023.102215.
[23] I. Mehmood, H. Li, Y. Qarout, W. Umer, S. Anwer, H. Wu, M. Hussain, M. Fordjour Antwi-Afari, Deep learning-based construction equipment operators' mental fatigue classification using wearable EEG sensor data, Advanced Engineering Informatics 56 (2023). https://doi.org/10.1016/j.aei.2023.101978.
[24] G. Lee, S. Lee, Importance of testing with independent subjects and contexts for machine-learning models to monitor construction workers' psychophysiological responses, J Constr Eng Manag 148 (2022) 04022082. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002341.