Enhancing Mental Workload Prediction through LightGBM during Multitasking

Syed Saad Azhar Ali [1] , Maged S. AL-Quraishi [2] ,Sami El Ferik [3] and Aamir Saeed Malik [4]

Abstract— **Multitasking is an essential aspect of daily life; however, it significantly increases mental workload (MWL), which can affect cognitive performance, decision making, and overall effectiveness. Thus, accurately assessing MWL is significant in various fields, including human-computer interaction, aviation, and healthcare, where cognitive overload can lead to unsuitable decisions. The brain computer interface (BCI) based on electroencephalography (EEG) presents a viable, non-invasive option for real-time monitoring of MWL, allowing an adaptive system to improve performance and user experience. However, because EEG patterns vary widely among individuals, it is still challenging to develop a generalized MWL prediction model. Therefore, Light Gradient Boosting Machine (LightGBM) with manually extracted features is proposed. Our analysis was based on the "STEW" dataset, which includes two task conditions: "No task" and a multitasking activity using the SIMKAP framework. The proposed model achieved an average accuracy of 84.0% (±14.4%) and an average F1-score of 83.1% (±18.2%), showcasing its strong predictive performance while maintaining computational efficiency compared to deep learning methods. These results highlight LightGBM's potential as a fast, subject-independent MWL classification tool, therefore enabling the design of scalable and flexible cognitive monitoring systems for practical use.**

I. INTRODUCTION

The mental health of individuals influences their perspectives, emotions, and interaction with their environment, thereby impacting their overall quality of life [1], [2]. Additionally, multitasking is a key factor that influences cognitive efficiency and MWL. Therefore, it is important to assess MWL to monitor mental health and prevent issues that could lead to mental disorders. BCI technologies have had an impactful development in recent years, in particular those employing EEG signals, which provide innovative ways to assess and manage MWL[3]. Although, there are various brain imaging techniques utilized for detecting brain activities such as functional magnetic resonance imaging (fMRI), functional near infrared spectroscopy (fNIRS), Magnetoencephalography (MEG), EEG is the most affordable technique to use for brain activity measurements [4], [5], [6]. EEG measures electrical activity in the brain, providing real-time insights into cognitive states [7]. Additionally, by analyzing specific EEG patterns, such as frontal theta, alpha, or other EEG band activity, researchers can accurately gauge MWL levels. This information is invaluable in various settings, from optimizing workplace environments to enhancing learning experiences, as it allows for the adjustment of task demands to align with an individual's cognitive capacity, thereby promoting better mental health and performance [5], [8], [9].

Previous studies have developed experimental frameworks that integrate controlled tasks with real-time events to investigate how mental workload and mental stress impact brain function. Popular methods include visual and auditory reaction tests, problem-solving tasks such as arithmetic challenges, and specialized assessments like the NASA-TLX questionnaire and the multitasking MATB system [10],[11]. These approaches simulate everyday cognitive demands like managing competing priorities or processing sensory information to measure shifts in concentration, mental strain, and fatigue levels. Additionally, by observing how participants adapt to these challenges, scientists can track fluctuations in alertness and identify patterns linked to cognitive exhaustion [12]. Furthermore, measuring MWL through EEG offers numerous benefits, including assessment in real-time, without subject bias, portable methods, high temporal resolution, and being noninvasive [13]. To predict the MWL levels, several studies have applied machine learning algorithms such as linear discriminant analysis (SVM), k-nearest neighbors (kNN), and other well-known approaches [14], [15]. On the other hand, other researchers employed deep learning techniques to improve the classification performance of MWL status [16], [17], [18].

Two distinct approaches were employed for the training and validation, including subject-dependent and subject-independent. The first approach, which is a commonly implemented method in the EEG-based classification, uses data from each participant individually for training and testing. Particularly, a separate model is trained and assessed, and then the overall performance is reported as the average across all participants. In comparison, the subject-independent method involves training the classification model on the data from a subset of participants and then validating it on the data from different, unseen participants. Thus, this approach evaluates the generalizability of the classification model

*This work was supported by King Fahd University of Petroleum & Minerals

[1] Syed Saad Azhar Ali is with the Interdisciplinary Research Center for Smart Mobility and Logistics, Interdisciplinary Research Center for Aviation and Space Exploration, and the Aerospace Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia syed.ali@kfupm.edu.sa

[2] Maged S. AL-Quraishi is with Interdisciplinary Research Center for Smart Mobility and Logistics King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia maged.quraishi@kfupm.edu.sa

[3] Sami El Ferik is with the Interdisciplinary Research Center for Smart Mobility and Logistics and The Control and Instrumentation Engineering Department, King Fahd University of Petroleum & Minerals Dhahran, Saudi Arabia selferik@kfupm.edu.sa

[4] Aamir Saeed Malik is with the Department of Computer Systems, Faculty of Information Technology Brno University of Technology, Bozetechova 2, 612 00 Brno, Czech malik@fit.vutbr.cz

among individuals. Therefore, when using EEG signals to classify mental workload (MWL), individual differences such as unique brain structures, individual experiences, and varying cognitive strategies can significantly influence the EEG patterns recorded [19] .This variability poses challenges in developing universal models for MWL classification.

Nemours studies conducted MWL classification using a subject-dependent approach and achieved high classification accuracy. For instance, G. Zhu et al [20], the authors aimed to improve the assessment of cognitive load during multitasking by utilizing graph methods applied to single-channel electroencephalography (EEG) data. The study involved 48 subjects who performed two tasks: one without any task and another with a simultaneous capacity task. Their main finding was that the right hemisphere channels, particularly F8, P8, and T8, showed significant differences in activity levels between high and low cognitive loads, with an accuracy of 89.6% in identifying mental workload using the graph features. Additionally, D. Das et al [21] implemented a comparison of different machine learning and deep learning algorithms. They utilized the "STEW" dataset, which includes tasks like "No task" and "SIMKAP-based multitasking activity," to evaluate different workload levels. The classification performance ranged from 83.1% to 91.15%. Moreover, T. Taori reported the highest classification accuracy, 97.1% to 100% for a two-class problem of the MWL status using the STEW data set [9].

On the other hand, for the subject-independent it is more challenging, where the classification performance declined drastically compared to the subject-dependent scheme. For instance, V. Pandey et al [22] conducted a comparative study on various machine learning algorithms to estimate mental workload using EEG data. The study implemented various classification models, including KNN, Random Forest, MLP, CNN+LSTM, and LSTM networks, reporting their performance metrics such as sensitivity, specificity, and precision. The main finding of the research is that evolutionary machine learning and deep learning models are effective in classifying workload and other mental states from EEG data, with an accuracy ranging from 57.19% to 61.08%. Moreover, C. Fan et al [23], reported the enhancement in the MWL status classification. They presented EEG-TNet, an end-to-end Brain-Computer Interface (BCI) framework designed for estimating mental workload (MWL). The results reported in subject-dependent experiments, the EEG-TNet achieved estimation accuracies of 99.82% for dual-task scenarios and 99.21% for triple-task scenarios. Nevertheless, the accuracy was lower in subject-independent studies, achieving 82.78% and 66.83% for various tasks, suggesting that it is challenging to generalize the model across diverse participants. Similarly, M. Safari et al. [18], presented a method for estimating (MWL) using EEG signals, focusing on a hybrid model that combined Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The findings indicated that the proposed hybrid model achieved an accuracy of 83.12% in classifying different workload levels.

In this study, we used Light Gradient Boosting Machine (lightGBM) with manually extracted features, including power band ratio, statistical features, and nonlinear features. Furthermore, we adopted a Leave-one-subject-out approach for the model training and evaluation, which ensures robust generalization. Additionally, lightGBM was chosen for its efficiency and strong predictive performance.

II. METHODS

Fig. 1 illustrates the general flow diagram of the proposed method. First, raw EEG data are loaded and preprocessed, windowed, and features extracted, and fed to the classifier.

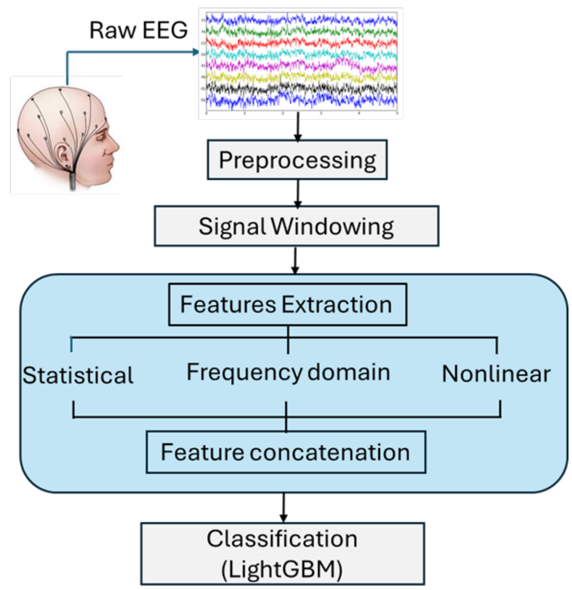

Fig. 1. General flow diagram for the proposed method

A. Data Description and Preprocessing

Simultaneous Task EEG Workload (STEW) is a publicly available data set published by W. Lim et al. STEW utilizes EEG signals to measure MWL [24]. It includes EEG recordings from 48 male participants, all from a university's graduate population, engaged in a multitasking workload experiment. The experiment consisted of two conditions: a "No Task" condition, in which participants were at rest for 2.5 minutes (representing low mental workload), and a multitasking scenario using the SIMKAP test, where EEG data were collected for another 2.5 minutes (representing high mental workload). The EEG signals were recorded using the Emotiv EPOC EEG headset, a portable 14-channel system with 128 Hz sampling frequency and 16-bit resolution. The electrodes were placed according to the 10–20 international electrode placement system, covering key frontal, temporal, parietal, and occipital regions. The dataset underwent a rigorous preprocessing pipeline, including high-pass filtering at 1 Hz, artifact removal using independent component analysis, and rereferencing to an average reference. All of these processes were performed using the EEGLAB toolbox [25].

B. Feature Extraction

To find important patterns that represent MWL levels, this study analyzed EEG signals using a variety of feature extraction approaches. Specifically, we evaluated the overall distribution and variability of the signals using a range of statistical features, including mean, variance, standard deviation, skewness, and kurtosis. Furthermore, we examined the band power ratio in the frequency domain for the various EEG bands (gamma-beta, beta-alpha, theta-delta, and alpha-theta). This approach helps reveal the rhythmic dynamic and cognitive states of the brain. Furthermore, to capture the nonlinear characteristics of brain activity, Hjorth parameters were also computed, providing insights into signal complexity, mobility, and activity. Thus, all the features mentioned were extracted from a window length of 4 seconds with 50% overlap of the processed EEG signals. Thus, by combining these diverse feature sets, the study aimed to enhance the robustness of EEG-based mental workload classification, ensuring a more comprehensive understanding of how the brain responds to varying cognitive demands.

C. Classification

The classification was performed using the LightGBM model, an efficient boosting method. Additionally, LightGBM stands out due to its ability to manage large-scale data efficiently, thanks to its histogram-based learning approach, which speeds up training while maintaining accuracy [26]. Furthermore, previous studies reported the superiority of LightGBM over other machine learning algorithms [27], [28]. In this study, specific hyperparameters shown in Table I were fine-tuned to balance performance and generalizability: Number of leaves (30) was adjusted to prevent overfitting. At the same time, a smaller learning rate (0.05) with 200 boosting rounds ensured steady and refined learning. The balanced class weighting helped counteract any imbalances in the data, making classification fairer across workload conditions. To reduce subject variability and enhance generalization, the model was trained and validated using a leave-one-subject-out (LOSO) approach, ensuring that each subject's data was tested independently, without being part of the training set. This approach makes the model more robust and subject-independent, a crucial factor for real-world applications where EEG-based workload classification needs to generalize across different individuals.

TABLE I
PARAMETER SETTING OF THE LIGHTGBM

Parameter	Value
Number of leaves	30
Learning rate	0.05
Maximum depth	-1
Number estimators	200
Class weight	balanced

III. RESULTS AND DISCUSSION

A. Band power ratio analysis

It is worth demonstrating the brain activity behavior during the MWL before delving to the classification performance findings. Fig. 2shows the boxplot of the band power ratio comparison in different brain areas. The results revealed an interesting pattern describing how the EEG band power ratios change across MWL conditions. Particularly, the alpha/theta ratio stands out as the most distinctive feature, illustrating a significant increase in the high MWL condition across all three channels (AF3, F7, and F8). This suggests that as cognitive demand increases, theta activity decreases relative to alpha, especially in the frontal lobe, which plays a key role in cognitive control. Additionally, the beta/alpha ratio also rose with the increase of MWL, and this result is aligned with the idea that beta activity is related to focus and task engagement. However, the difference between conditions is not as strong as in the alpha/theta ratio, particularly in F7 and F8. Interestingly, the theta/delta followed a different pattern, with slightly higher values in the low MWL condition. This could indicate that in less demanding tasks, theta fluctuations are more prominent compared to delta activity, whereas under high MWL, theta power decreases relative to delta. Thus, these outcomes revealed that alpha/theta and beta/alpha ratios could be valuable features for differentiating various cognitive workload levels, particularly in the frontal and prefrontal brain areas. Moreover, these findings confirm that EEG spectral features reflect cognitive processing characteristics and might be a key to developing robust workload monitoring systems.

B. Classification Results

The confusion matrix and ROC are represented in Fig. 3 and Fig. 4, highlighting the performance of the LightGBM and manual feature extraction approach. Particularly, the model correctly classified 81.08% of low MWL and 86.60% of high MWL instances, with relatively low misclassification rates (18.92% and 13.40%, respectively). Furthermore, the ROC supports this finding with an impressive AUC of 0.91, indicating discrimination between the two MWL conditions.

As mentioned in the methodology section, LOSO was employed for model training and validation. The overall classification performance of the proposed approach (various feature extraction methods and LightGBM) slightly outperformed the previous studies [18], [22], [29]. A mean accuracy of 84.0% and an F1-score of 83.1% were achieved. The LOSO method, which iteratively trains on all but one subject and tests on the excluded individual, offers a rigorous assessment of cross-subject performance. This approach minimizes data leakage and ensures the model is not overfitting to specific individuals, which is critical for real-world deployment, where new users' data will differ. However, the high standard deviations (±14.4% for accuracy, ±18.2% for F1) highlight inherent challenges, including performance variability across subjects, from biological differences, noise in EEG signals, or task engagement inconsistencies.

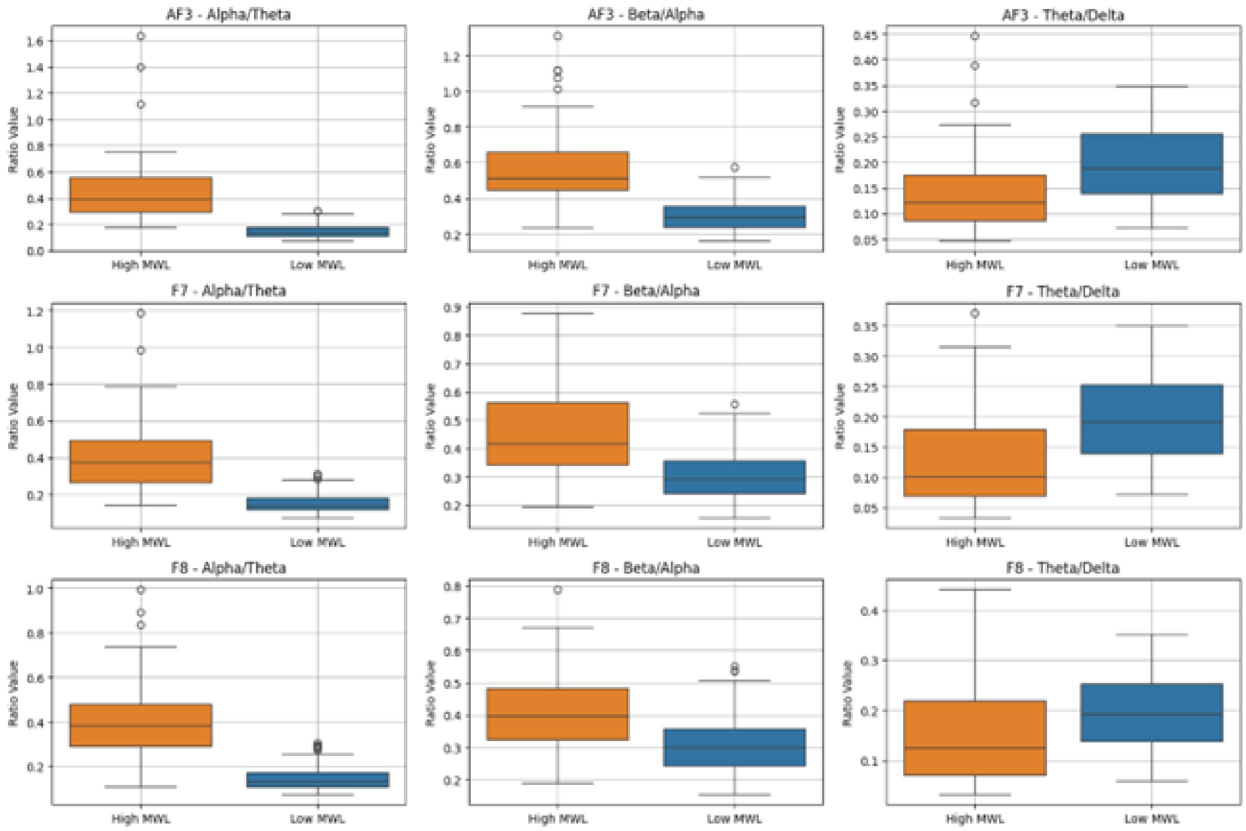

Fig. 2. Power band ratio comparison across different brain areas.

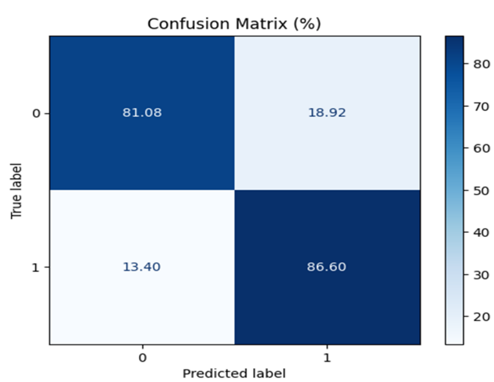

Fig. 3. Confusion matrix which demonstrates the classification performance of the MWL levels

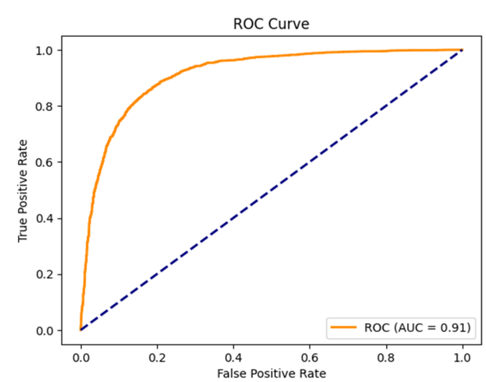

Fig. 4. ROC curve which demonstrates the classification performance of the MWL levels

In contrast, grouped leave-subjects-out schemes, where multiple subjects are excluded per fold, reduce computational costs and may better simulate scenarios involving cohorts of new users. However, this approach risks masking individual outliers or underrepresented patterns, potentially inflating perceived stability. The use of LOSO, while computationally demanding, provides a more granular understanding of model robustness, revealing where personalized calibration might enhance performance.

IV. CONCLUSIONS

A framework for classifying MWL levels was developed by integrating LightGBM with salient feature extraction techniques. The proposed approach achieved an average accuracy of 84.0% (±14.4%) and an F1-score of 83.1% (±18.2%) under a rigorous leave-one-subject-out validation scheme, demonstrating strong generalizability across individuals. These results slightly outperformed prior benchmarks in mental workload classification and underscore the efficacy of combining domain-specific feature engineering with gradient-boosted models for EEG-based tasks. Including band power ratios, which capture neurophysiological dynamics between frequency bands, proved particularly valuable in distinguishing high and low cognitive demands. This work contributes to the growing field of neuroadaptive technologies by offering a practical, interpretable solution for real-time mental workload monitoring.

ACKNOWLEDGMENT

The authors thank King Fahd University of Petroleum & Minerals for the support received under the University Funded Grant INML 2508. The authors would also like to thank Mr. Karim Asif Sattar for his valuable assistance.

REFERENCES

[1] M. Mingardi, P. Pluchino, D. Bacchin, C. Rossato, and L. Gamberini, "Assessment of implicit and explicit measures of mental workload in working situations: implications for industry 4.0," *Applied Sciences*, vol. 10, no. 18, p. 6416, 2020.

[2] A. Othmani, B. Brahem, Y. Haddou *et al.*, "Machine-learning-based approaches for post-traumatic stress disorder diagnosis using video and eeg sensors: A review," *IEEE Sensors Journal*, vol. 23, no. 20, pp. 24 135–24 151, 2023.

[3] P. Aricò, G. Borghini, G. Di Flumeri, A. Colosimo, S. Bonelli, A. Golfetti, S. Pozzi, J.-P. Imbert, G. Granger, R. Benhacene *et al.*, "Adaptive automation triggered by eeg-based mental workload index: a passive brain-computer interface application in realistic air traffic control environment," *Frontiers in human neuroscience*, vol. 10, p. 539, 2016.

[4] M. I. Al-Hiyali, N. Yahya, I. Faye, M. S. Al-Quraishi, and A. Al-Ezzi, "Principal subspace of dynamic functional connectivity for diagnosis of autism spectrum disorder," *Applied Sciences*, vol. 12, no. 18, p. 9339, 2022.

[5] A. Abdalhadi, N. Koundal, M. Z. Yusoff, M. S. Al-Quraishi, F. Merienne, and N. M. Saad, "Study of the acute stress effects on decision making using electroencephalography and functional near-infrared spectroscopy: a systematic review," *IEEE Access*, 2024.

[6] S. S. A. Ali, "Brain mri sequence and view plane identification using deep learning," *Frontiers in Neuroinformatics*, vol. 18, p. 1373502, 2024.

[7] M. S. Al-Quraishi, W. H. Tan, I. Elamvazuthi, C. P. Ooi, N. M. Saad, M. I. Al-Hiyali, H. Karim, and S. S. A. Ali, "Cortical signals analysis to recognize intralimb mobility using modified rnn and various eeg quantities," *Heliyon*, vol. 10, no. 9, 2024.

[8] N. Koundal, A. Abdalhadi, M. S. Al-Quraishi, I. Elamvazuthi, M. S. Moosavi, C. Guillet, F. Merienne, and N. M. Saad, "Effect of interruptions and cognitive demand on mental workload: A critical review," *IEEE Access*, 2024.

[9] A. Sengupta, A. Dasgupta, A. Chaudhuri, A. George, A. Routray, and R. Guha, "A multimodal system for assessing alertness levels due to cognitive loading," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 7, pp. 1037–1046, 2017.

[10] Y. Santiago-Espada, R. R. Myer, K. A. Latorella, and J. R. Comstock Jr, "The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide," Tech. Rep., 2011.

[11] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[12] A. Sengupta, A. Tiwari, A. Chaudhuri, and A. Routray, "Analysis of loss of alertness due to cognitive fatigue using motif synchronization of eeg records," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 1652–1655.

[13] Y. Zhou, S. Huang, Z. Xu, P. Wang, X. Wu, and D. Zhang, "Cognitive workload recognition using eeg signals and machine learning: A review," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 3, pp. 799–818, 2021.

[14] F. Dehais, A. Duprès, S. Blum, N. Drougard, S. Scannella, R. N. Roy, and F. Lotte, "Monitoring pilot's mental workload using erps and spectral power with a six-dry-electrode eeg system in real flight conditions," *Sensors*, vol. 19, no. 6, p. 1324, 2019.

[15] N. Sciaraffa, P. Aricò, G. Borghini, G. D. Flumeri, A. D. Florio, and F. Babiloni, "On the use of machine learning for eeg-based workload assessment: Algorithms comparison in a realistic task," in *Human Mental Workload: Models and Applications: Third International Symposium, H-WORKLOAD 2019, Rome, Italy, November 14–15, 2019, Proceedings 3*. Springer, 2019, pp. 170–185.

[16] Y. Zhang and Y. Shen, "Parallel mechanism of spectral feature-enhanced maps in eeg-based cognitive workload classification," *Sensors*, vol. 19, no. 4, p. 808, 2019.

[17] W. Qiao and X. Bi, "Ternary-task convolutional bidirectional neural turing machine for assessment of eeg-based cognitive workload," *Biomedical Signal Processing and Control*, vol. 57, p. 101745, 2020.

[18] M. Safari, R. Shalbaf, S. Bagherzadeh, and A. Shalbaf, "Classification of mental workload with eeg analysis by using effective connectivity and a hybrid model of cnn and lstm," *Computer Methods in Biomechanics and Biomedical Engineering*, pp. 1–15, 2024.

[19] T. J. Taori, S. S. Gupta, S. S. Gajre, and R. R. Manthalkar, "Cognitive workload classification: Towards generalization through innovative pipeline interface using hmm," *Biomedical Signal Processing and Control*, vol. 78, p. 104010, 2022.

[20] G. Zhu, F. Zong, H. Zhang, B. Wei, and F. Liu, "Cognitive load during multitasking can be accurately assessed based on single channel electroencephalography using graph methods," *IEEE Access*, vol. 9, pp. 33 102–33 109, 2021.

[21] D. D. Chakladar, S. Dey, P. P. Roy, and D. P. Dogra, "Eeg-based mental workload estimation using deep blstm-lstm network and evolutionary algorithm," *Biomedical Signal Processing and Control*, vol. 60, p. 101989, 2020.

[22] V. Pandey, D. K. Choudhary, V. Verma, G. Sharma, R. Singh, and S. Chandra, "Mental workload estimation using eeg," in *2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE, 2020, pp. 83–86.

[23] C. Fan, J. Hu, S. Huang, Y. Peng, and S. Kwong, "Eeg-tnet: an end-to-end brain computer interface framework for mental workload estimation," *Frontiers in neuroscience*, vol. 16, p. 869522, 2022.

[24] W. L. Lim, O. Sourina, and L. P. Wang, "Stew: Simultaneous task eeg workload data set," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2106–2114, 2018.

[25] 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*, vol. 134, no. 1, pp. 9–21, 2004.

[26] P. Jain, J. Yedukondalu, H. Chhabra, U. Chauhan, and L. D. Sharma, "Eeg-based detection of cognitive load using vmd and lightgbm classifier," *International Journal of Machine Learning and Cybernetics*, vol. 15, no. 9, pp. 4193–4210, 2024.

[27] H. Pan, Z. Li, C. Tian, L. Wang, Y. Fu, X. Qin, and F. Liu, "The lightgbm-based classification algorithm for chinese characters speech imagery bci system," *Cognitive Neurodynamics*, vol. 17, no. 2, pp. 373–384, 2023.

[28] H. Zeng, C. Yang, H. Zhang, Z. Wu, J. Zhang, G. Dai, F. Babiloni, and W. Kong, "A lightgbm-based eeg analysis method for driver mental states classification," *Computational intelligence and neuroscience*, vol. 2019, no. 1, p. 3761203, 2019.

[29] U. M. Al-Saggaf, S. F. Naqvi, M. Moinuddin, S. A. Alfakeh, and S. S. A. Ali, "Performance evaluation of eeg based mental stress assessment approaches using machine learning techniques," *Sensors*, vol. 23, no. 20, p. 8134, 2023.