Seventh International Conference on Recent Trends in Image Processing and Pattern Recognition (RTIP2R-2024)

Deep Learning Based EEG Based Mental Workload Detection with Discrete Wavelet Transform and Welch's Power Spectral Density

Shraddha Jain Sharma[a]*, Ratnalata Gupta[b]

[a]*Indian Institute of Technology, IIT, BHU, Varanasi (U.P), India,*
[b]*Shivajirao Kadam Institute of Technology & Management,Indore (M.P), India*

Abstract

EEG data provide a convenient and portable method of measuring mental workload. This work analyzes EEG data collected during rest intervals and arithmetic problems using deep learning and machine learning approaches. EEG signals were broken down into sub-bands using the Discrete Wavelet Transform (DWT), and the power levels of these bands were estimated using Welch's approach. Based on these power levels, feature vectors were identified using both conventional machine learning techniques including Linear Discriminant Analysis (LDA), k-Nearest Neighbors (k-NN), and Support Vector Machines (SVM) and deep learning algorithms like Long Short-Term Memory (LSTM). The findings show that LSTM performed better than conventional classifiers, with higher performance metrics and accuracy. This demonstrates how deep learning may be used to separate mental workload states from EEG signals with accuracy.

Keywords: Deep Learning; Welch's Power Spectral Density; Discrete Wavelet Transform; Mental Workload, EEG Signal.

1. Introduction

The amount of cognitive capacity needed to complete a psychological, cognitive, or motor task is known as the mental burden. It typically results from the interplay of an individual's abilities, behaviors, and perceptions with the task environment. Professionals including air traffic controllers, surgeons, military pilots, ship crew members, and drivers of motor vehicles frequently have their mental workloads assessed.

---

* Corresponding author.
E-mail address: jainshraddha1204@gmail.com

Recent technological developments have led to the application of mental workload evaluation by researchers to assess

user performance in brain-computer interface (BCI) systems [1-2]. Individuals' physiological measurements are frequently used in methods for evaluating mental workload. These techniques work by analyzing physiological data that was captured during both difficult and simple tasks. Heart rate, breathing rate, eye movements and blinking rate, heat activity, muscle activities, and brain activity are among the physiological measurement techniques utilized in mental workload evaluation [3]. The EEG technique is widely employed in assessments of mental workload because it can record bioelectrical changes in the brain that arise from different types of cognitive tasks [4]. Low-amplitude bioelectrical signals are acquired using the EEG technique from electrodes positioned on the scalp at predetermined intervals. These low-amplitude signals are captured in a computer environment after being boosted by amplifiers [5]. The amplified signals are subjected to a variety of statistical, spectral, or time-frequency analysis techniques to extract characteristics. For the purpose of classification, the collected characteristics are fed into any artificial neural network method. In this work, we utilize Welch's power spectral density and Discrete Wavelet Transform (DWT) to extract features in order to further the field of mental workload evaluation utilizing EEG signals. Our training process involved using these features to train many machine learning algorithms, such as Support Vector Machines (SVM), Linear Discriminant Analysis (LDA), k-Nearest Neighbors (KNN), Random Forest (RF), and Long Short-Term Memory (LSTM) networks. Our results show that the classification accuracy of the LSTM network is much better than that of other approaches.

The main contributions of this work are summarized as follows:
 1) Discrete Wavelet Transform (DWT) and Welch's technique were used to better extract EEG signals associated with mental workload.
 2) Long Short-Term Memory (LSTM) networks yielded the best classification accuracy when compared to other machine learning methods, such as SVM, LDA, KNN, and RF.
 3) Slightly enhanced categorization performance in brain-computer interface systems was demonstrated by using sophisticated EEG analysis to effectively quantify mental effort.

2. Related Work

Conventional techniques for classifying mental workload frequently depend on manually created features. Using SVM with a variety of statistical and frequency data, Wang et al. [6] were able to distinguish between workload levels with a high degree of accuracy. Using ensemble techniques and EEG subbands, Ladekar et al. [7] utilized KNN and achieved an accuracy of 83.65%, albeit with a small sample size. While Zammouri et al. [8] demonstrated the limitations of handmade feature-based approaches by achieving 79% accuracy using PSD in the theta band, Gómez et al. [9] integrated SVM, KNN, and RF with comprehensive feature extraction techniques. Deep learning techniques perform better because they automate feature extraction. By utilizing SDAE with PSD characteristics, Yang et al. [10] were able to reach 92% accuracy. TT-CBNTM was proposed by Qiao et al. [11] and achieved 96.3% accuracy on a limited dataset. Using a combination of 3D CNN and LSTM, Kwak et al. [12] achieved 93.9% binary classification accuracy. Time-domain data is underutilized in most studies, despite these developments, and is mostly focused on frequency domain information. In order to increase the precision of mental effort assessment, we propose in our study using both time and frequency domain data into EEG representation. We also present a regression framework for continuous mental workload estimation, which is a largely unexplored concern in the field.

3. Materials and Methods

The proposed method's block diagram is displayed in Figure 1. Raw EEG data were processed to extract characteristics using the Welch method and DWT. DWT was used to process the EEG data, which was obtained by obtaining four separate frequency bands (delta, theta, alpha, and beta) at a sampling frequency of 500 Hz from 21 channels. Next, the Welch method was used to determine the power values of these bands. Feature vectors were created using the obtained power values. Machine learning methods including the LSTM deep network, KNN, and decision trees were used to classify these feature vectors, which were obtained using DWT and the Welch method. The outputs from the classification results corresponded to the states of resting and mental workload.
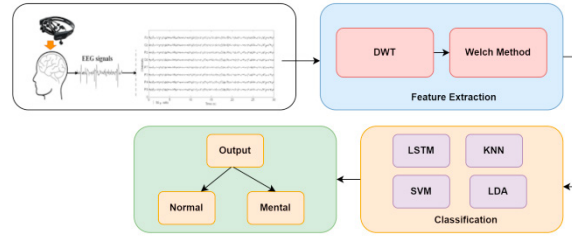
Fig. 1.Block Diagram of the Proposed Method

3.1 EEG Data Acquisition: EEG recordings produced by Zyma et al. while subjects completed mental math tasks were used in this investigation. An professional neurophysiologist recorded recordings from 36 individuals (27 females and 9 men) both before and during the performance of these activities. Written informed consent was obtained from each participant in accordance with the World Medical Association (WMA) Helsinki Declaration, and the EEG dataset used in this study was approved by the Bioethics Commission of the Education and Science Center "Institute of Biology and Medicine" of Taras Shevchenko National University of Kiev. In this investigation, EEG data were acquired using the Neurocom monopolar EEG 23-channel device. In accordance with the International 10-20 system, the electrodes were positioned symmetrically on the scalp [13].

3.2 Feature Extraction: Feature vectors were created using the power values for every frequency band that were obtained using the Welch technique. The power levels across the four EEG bands were represented by each feature vector. A 4-dimensional feature vector for each time window was the outcome of this. The amount of feature vectors for each condition (mental workload and rest) was multiplied by factors of 0.90, 0.95, and 1.1 to increase the dataset. This resulted in a total of 6048 feature vectors (3024 for mental workload and 3024 for rest). Because it strikes a compromise between computational efficiency and frequency resolution—a crucial factor for high temporal resolution EEG signal analysis—we selected the 'db8' wavelet for the Discrete Wavelet Transform (DWT) in our investigation. Our approach is well-suited to the 'db8' wavelet because it is especially good at capturing the complex properties of EEG data across many frequency bands.

A) *Discrete Wavelet Transform (DWT):* In a continuous wavelet transform, the parameters of shift and scale, represented by the letters b and a, respectively, fluctuate continually. The process of computing wavelet coefficients requires a lot of computing power for each potential value of the variable parameters a and b. In order to address this problem, the variables a and b in the continuous wavelet transform are selected as powers of two in the discrete wavelet transform (DWT). This transformation is called the discrete wavelet transform (DWT), and it is carried out with discrete values for shift and scale. When working with input signals that contain a lot of data points, DWT is a very effective and quick method. Equation 1 indicates the equation for any wavelet function with shift and scale values [14].

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-b}{a}\right) \tag{1}$$

Where $\varphi_{a,b}(t)$ represents the wavelet coefficient at scale a and shift b, while x(t) is the input signal and $\varphi(t)$ is the wavelet function. The scale parameter a adjusts the wavelet's compression or dilation, and b determines its position along the time axis. Substituting the scale parameter $a=2^r$ and the shift parameter $b=k.2^r$ into Equation 1 yields the wavelet function given in Equation 2 [14].

$$\varphi_{r,k}(t) = \frac{1}{\sqrt{|2^r|}} \varphi\left(\frac{t-k.2^r}{2^r}\right) \tag{2}$$

where $\varphi_{r,k}(t)$ is the discrete wavelet function, r is the discrete scale factor, and k is the discrete shift factor. If the function in Equation 2 is substituted into the formula for continuous wavelet transform, the resulting expression for discrete wavelet transform can be represented as shown in Equation 3 [15].

$$DWT_{r,k} = \frac{1}{\sqrt{|2^r|}} \int_{-\infty}^{+\infty} x(t)\varphi\left(\frac{t-k.2^r}{2^r}\right) \tag{3}$$

It calculates the discrete wavelet coefficient $DWT_{r,k}$ using the discrete scale 2^r and shift $k.2^r$, applying these parameters to the signal x(t) and the wavelet function $\varphi(t)$. These equations collectively illustrate the transformation from

continuous to discrete wavelet analysis, facilitating efficient computation for large data sets.

B) *Welch's Power Spectral Density:* The input signal is first split into segments of the required length in the spectral analysis method known as the Welch method. Any window function is applied to the signals corresponding to these segments, and the fast Fourier transform is then used to produce smoothed periodograms. Equation 4 represents the smoothed i-th periodogram expression [16].

$$\hat{p}_i(f) = \frac{1}{M}\frac{1}{H}\left[\sum_{n=0}^{M-1} w[n]x_i[n]e^{-j2\tau fn}\right]^2 \tag{4}$$

It defines the smoothed ith periodogram $\hat{p}_i(f)$, where $x_i[n]$ represents the signal segment, w[n] is the window function applied to the segment, and M is the segment length. The window function w[n], often a Hamming window, modifies the segment to reduce spectral leakage. The parameter w[n] in Equation 4 corresponds to the window function. Typically, a window called "Hamming" is chosen. The variable M represents the length of the segments. H is the normalized window function. The expression for the normalized window function is given in Equation 5 [16].

$$H = \frac{1}{M}\sum_{n=0}^{M-1} W[n]^2 \tag{5}$$

Where the normalized window function H, which is the window function W[n] adjusted for normalization to ensure accurate spectral density estimation. Welch's power spectral estimate is obtained by averaging the smoothed periodograms. Welch's spectral estimate is expressed as shown in Equation 6 [16].

$$\widehat{p_{welch}}(f) = \frac{1}{S}\sum_{i=1}^{s} \hat{p}_i(f) \tag{6}$$

It represents Welch's power spectral estimate $\widehat{p_{welch}}(f)$, which is obtained by averaging the smoothed periodograms from multiple segments to provide a more stable estimate of the power spectrum. The averaged result enhances the reliability of the PSD estimate by reducing the variance of the spectral estimation. Using Welch's approach, algorithm 1 calculates the Power Spectral Density (PSD) of EEG data and selects features for further analysis from the frequency domain.

Algorithm 1: Welch's Power Spectral Density (PSD) Estimation for Feature Extraction

1. **Input Signal :**
   Let v(t) represent the raw EEG signal sampled at a frequency f_s (e.g., 500 Hz).
2. **Define Parameters**:
   Segment Length N: Length of each segment (e.g., 512 samples).
   Overlap O: Number of overlapping samples between segments (e.g., 50% overlap).
   Window Function w(t): Choose a window function such as Hamming, Hanning, etc.
3. **Segmenting**:
   Divide the input signal v(t) into overlapping segments of length N. Each segment $v_d(t)$ starts at d. (N−O), where d ranges from 0 to M−1, and M is the number of segments.
   If the total number of samples in v(t) is T, then the number of segments M is computed as:

$$M = \left\lceil \frac{T - N + O}{N - O} \right\rceil + 1$$

4. **Apply Window Function**:
   Apply the window function w(t) to each segment $v_d(t)$:

$$v_{d,w}(t) = v_d(t).w_t$$

   The window function w(t) is typically a symmetric function of length N.
5. **Compute Periodogram for Each Segment**:
   For each windowed segment, compute the Fast Fourier Transform (FFT) and the squared magnitude to obtain the periodogram $P_d(f)$:

$$P_d(f) = |FFT(v_{d,w}(t))|^2$$

6. **Average Periodograms**:
   Average the periodograms from all segments to estimate the final power spectral density $P_{Welch}(f)$:

$$P_{Welch}(f) = \frac{1}{M}\sum_{d=0}^{M-1} P_d(f)$$

7. **Normalization Factor**:
   Compute the normalization factor U for the window function w(t):

$$U = \frac{1}{L}\sum_{l=-L/2}^{L/2-1} w^2(l)$$

   The normalization ensures that the PSD values are correctly scaled.
8. **Feature Extraction**:
   Extract features from the PSD estimate $P_{Welch}(f)$:
   Frequency Band Power: Compute the power in specific frequency bands (e.g., delta, theta, alpha, beta) by integrating the PSD over those bands:

$$Power_{band} = \int_{f_{start}}^{f_{end}} P_{welch}(f)df$$

Peak Power: Identify the peak power values in the PSD, which can indicate dominant frequency components.
Band Ratio: Compute ratios of power between different bands (e.g., alpha/beta ratio).

9. **Feature Vector Construction**:

Construct a feature vector for each segment or time window using the extracted features (e.g., power in different bands, peak power, band ratios). This results in a multi-dimensional feature vector representing the spectral characteristics of the EEG signal.

10. **Post-Processing** :

Normalize or standardize the feature vectors if needed, especially before applying machine learning models.

3.3 Classification

A) *Long Short-Term Memory (LSTM):* When Hochreiter and Schmidhuber uncovered Long Short-Term Memory (LSTM) in 1997, they called it a unique kind of recurrent neural network (RNN). The RNN model's training phase presents challenges, like the vanishing gradient problem, which LSTM tackles. An entity in the hidden layer of LSTM networks is referred to as a cell. Three gates, which regulate the information flow, make up an LSTM cell: the forget gate, the input gate, and the output gate [17]. An LSTM cell can be shown in Figure 2.
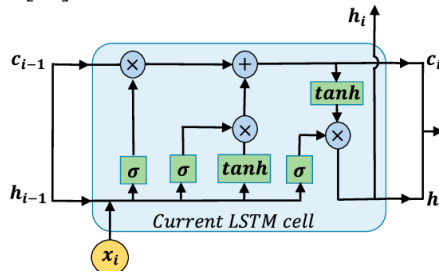

Fig 2. Structure of an LSTM Cell [11]

The information flow is controlled by the forget gate's sigmoid function, which outputs 0 or 1 depending on whether the present information is deleted or kept unchanged. The layer containing the tanh activation function is traversed by the updated cell state information at the output gate. The layer with the sigmoid activation function receives the input data. In the last stage, the cell's output is calculated by multiplying the data from these two layers [18-19]. On the basis of initial trials that showed optimal performance, we chose the precise hyperparameters for the LSTM network. A grid search strategy was used to establish the optimal trade-off between training time and accuracy, leading to the selection of 100 epochs, a mini-batch size of 128 and a learning rate of 0.005. It was discovered that this setup offers a stable training of the LSTM network by striking a compromise between computational efficiency and model complexity.

B) *K-Nearest Neighbors (KNN):* A supervised machine learning approach called k-Nearest Neighbors (KNN) classifies data according to a preset number of neighbors in the dataset. The algorithm calculates the separation between the data to be classified and the sample space's data clusters. The class with the closest distance receives the data. Here, it is important to remember that classes' properties need to be specified precisely before they may be classified [20-21].

C) *Support Vector Machine (SVM):* A supervised learning model used for regression analysis and classification issues is called Support Vector Machines (SVM). Class labels are usually assigned as -1 and +1 in binary classification using Support Vector Machines. Samples from distinct classes are separated by SVM using the decision function that is acquired from the training data. To divide data into two classes, a multitude of hyperplanes can be drawn. Finding the hyperplane that can most effectively divide the training data is the main goal, nevertheless. Put another way, the objective is to maximize the margin, or the distance, between the closest points in the various class data [22].

R. *Linear Discriminant Analysis (LDA):* A. Fisher created the classification algorithm known as linear discriminant analysis (LDA) in 1936 to address binary classification issues. It linearly divides features that correspond to several classifications. The LDA algorithm's main goal is to maximize the variance across classes. Consequently, the data distribution becomes more comprehensible when a decision boundary is drawn between classes [23].
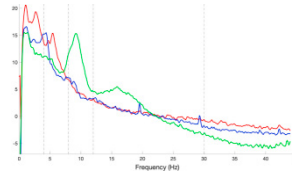
Fig 3. Power Spectral Density Graphs

4. Results and Discussion

4.1 Discrete Wavelet Transform and Power Spectral Density: The 'db8' wavelet function was used to perform a 6th level discrete wavelet transform on the raw EEG data, which were obtained at a sampling frequency of 500 Hz. By using this method, the raw EEG signals' delta (0–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), and beta (16–36 Hz) bands were extracted. Welch technique was used to determine the power spectral density values of these bands. The power levels that corresponded to the 0-36 Hz frequency range were concatenated to form feature vectors. The power spectral density graphs under resting and mentally taxed situations are shown in Figure 3.

4.2 Feature Vector Construction: The power spectral density graphs of a single channel displayed in Figure 3 indicate that mental effort is associated with a notable rise in the alpha band. There are 756 feature vectors for both mental workload and resting circumstances because the dataset used in this study has 36 subjects, each with 21 channels. However, the number of feature vectors was raised threefold by multiplying by factors of 0.90, 0.95, and 1.1 because deep learning techniques demand a considerable amount of data during the training phase. The enhanced dataset has 6048 feature vectors in total, of which 3024 are for mental workload and 3024 are for resting situations. Zero-meaning has been applied at this point. Applying data augmentation, which increases the size of the dataset for training deep learning models, involved multiplying the feature vectors. By using this method, the limited data sets would be addressed and a stronger training set would be provided. Our investigation did not, however, fully assess the impact of this augmentation on model performance or the possibility of overfitting. Subsequent studies comparing the outcomes of models trained with and without enhanced data should evaluate if data augmentation increases the generalizability of the model or introduces any biases. Analysis of this kind might shed light on how well data augmentation improves the robustness and performance of models.

4.3 Training and Testing of LSTM Network: Using 10-fold cross-validation, the dataset was split into subgroups for testing and training. With 100 epochs, a 128-mini-batch size, a 0.005 learning rate, and the Adam optimization technique, an LSTM network was trained. 5000 iterations were used in the training process, processing 128 feature vectors every epoch. The LSTM model outperformed other classifiers in the investigation, with a high classification accuracy of 95.40%. Table 1 gives an overview of the hyperparameters that were employed during the training procedure.

Table 1. Parameters Used in the LSTM Network

LSTM Parameters	Value
Number of Epochs	100
Mini-batch Size	128
Learning Rate	0.005
Optimization Algorithm	Adam

4.4 Performance Comparison: The Long Short-Term Memory (LSTM) network is more effective than other classification algorithms, as shown by the performance comparison of these algorithms in Table 2. Significantly surpassing SVM, LDA, and KNN classifiers, the LSTM attained the maximum accuracy of 98.40% and an outstanding F1 score of 97.40 %. This suggests that in addition to its superior performance in accurately diagnosing cases of mental stress, the LSTM network also demonstrated an effective balance between precision and recall. Also, the LSTM showed the highest specificity 97.93%, indicating its accuracy in detecting non-workload instances, and the highest sensitivity 96.77%, indicating its capacity to precisely identify positive occurrences of mental workload. In comparison with the other methods examined, these results highlight the overall superiority of the LSTM network

in detecting mental stress. Data augmentation was used to boost the amount of feature vectors and maybe enhance model performance, although its effect on the final outcomes was not fully investigated. Future research should compare model performance with and without augmented data in order to fully comprehend the advantages and possible disadvantages of data augmentation. This could aid in figuring out whether the additional data adds any biases or improves the model's capacity to generalize in other contexts.

Table 2. Mean Performance Metrics for All Channels

Classifier	Accuracy	F1 Score	Sensitivity	Specificity	Precision
LSTM	98.40 %	97.40 %	96.77 %	97.93 %	98.04 %
SVM	70.25 %	71.04 %	72.95 %	67.56 %	69.22 %
LDA	69.82 %	70.05 %	70.57 %	69.08 %	69.53 %
KNN	76.93 %	77.73 %	80.49 %	73.38 %	75.15 %

Our study compared LSTM with traditional classifiers like SVM, LDA, and k-NN, but expanding this comparison to incorporate other cutting-edge approaches like convolutional neural networks (CNNs) or attention mechanisms could yield a more thorough assessment of the state-of-the-art approaches in EEG-based mental workload detection. Future studies could benefit from incorporating these cutting-edge techniques to compare our suggested LSTM model to the most recent developments in the field and acquire a more comprehensive understanding of the efficacy of different strategies. The creation of more reliable models for the classification of mental workload may be guided by these comparisons, which may also highlight areas for improvement.

4.5 Comparison with Previous Studies: The LSTM based method we used in our study to classify mental workload from EEG signals produced an accuracy of 98.40%, which is similar to the 98.00% accuracy Roy et al. [27] obtained using ERP, CCA, LDA features. This outcome demonstrates how well LSTM performs when using spectral and temporal data to enable precise combined event-related potentials and canonical correlation analysis with LDA for classification, is marginally higher than our method. The LSTM model we have developed exhibits significant improvements when compared to other techniques like the Fast Fourier Transform (FFT) and support vector regression (SVR) used by Lim et al. [24] with an accuracy of 69.20%, and the LSTM approach combined with statistical features by Chakladar et al. [25] achieving 86.33%. With an accuracy of 93.80%, the SVM based strategy used by Aghajani et al. [26] likewise fared well. Overall, our study's LSTM approach performs admirably, matching or slightly underperforming the greatest outcomes documented in the literature. Next, a summary of findings from multiple studies on the identification of mental workload from EEG signals is shown in Table 3 .

Table 3. Comparison with Previous Studies

Author	Method	Accuracy
Lim et al. [24]	FFT, SVR	69.20 %
Chakladar et al. [25]	FFT, Statistical feature, LSTM	86.33 %
Aghajani et al. [26]	PSD, PLV, PAC, SVM	93.80 %
Aydın [3]	HFB, SVM	95.39%
Roy et al. [27]	ERP, CCA, LDA	98.00%
Our Study	**LSTM**	**98.40%**

5. Conclusion

Feature vector analysis of power spectral density values revealed considerable variances in power values, especially in the delta and alpha bands, which allowed EEG data to be used efficiently to discriminate between periods of mental exertion and relaxation. The Long Short-Term Memory (LSTM) network demonstrated superior performance in accurately identifying mental workload when compared to conventional classifiers like Linear Discriminant Analysis (LDA). The conclusions of the study, however, may not be as broadly applicable as they may be because they were derived from a dataset consisting of 36 participants with particular demographic traits. Incorporating larger and more diverse participant groups and investigating a range of mental workload scenarios are key areas of future study to enhance the robustness and usefulness of the model. Further analyses of other sophisticated feature extraction techniques, like sophisticated wavelet transforms or connectivity measurements, as well as comparisons of LSTM with cutting-edge methodologies like convolutional neural networks (CNNs) or attention mechanisms, may offer a

more thorough grasp of the model's functionality and increase its applicability to a wider range of users and tasks. To improve the robustness of the LSTM model, future studies could implement techniques such as dropout regularization, batch normalization, or early stopping. These methods could help mitigate overfitting and enhance the model's generalizability across different datasets and mental workload conditions.

References:

[1] Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. Applied ergonomics, 74, 221-232.

[2] Di Stasi, L. L., Antolí, A., & Cañas, J. J. (2013). Evaluating mental workload while interacting with computer-generated artificial environments. Entertainment Computing, 4(1), 63-69.

[3] Aydın, E. A. (2021). EEG sinyalleri kullanılarak zihinsel iş yükü seviyelerinin sınıflandırılması. Politeknik Dergisi, 24(2), 681-689.

[4] Di Flumeri, G., Borghini, G., Aricò, P., Sciaraffa, N., Lanzi, P., Pozzi, S., ... & Babiloni, F. (2018). EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings. Frontiers in human neuroscience, 12, 509.

[5] Siuly, S., Li, Y., Zhang, Y. (2016). Electroencephalogram (EEG) and Its Background. In EEG Signal Analysis and Classification (s. 3-21). Springer, Cham.

[6] Wang, S., Gwizdka, J., & Chaovalitwongse, W. A. (2015). Using wireless EEG signals to assess memory workload in the $ n $-back task. *IEEE Transactions on Human-Machine Systems*, *46*(3), 424-435.

[7] Ladekar, M. Y., Gupta, S. S., Joshi, Y. V., & Manthalkar, R. R. (2021). EEG based visual cognitive workload analysis using multirate IIR filters. Biomedical Signal Processing and Control, 68, 102819.

[8] Zammouri, A., Moussa, A. A., & Mebrouk, Y. (2018). Brain-computer interface for workload estimation: Assessment of mental efforts in learning processes. Expert Systems with Applications, 112, 138-147.

[9] Gómez, L. C., Hervás, R., Gonzalez, I., & Villarreal, V. (2021). Studying the generalisability of cognitive load measured with EEG. Biomedical Signal Processing and Control, 70, 103032.

[10] Yang, S., Yin, Z., Wang, Y., Zhang, W., Wang, Y., & Zhang, J. (2019). Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders. Computers in biology and medicine, 109, 159-170.

[11] Qiao, W., & Bi, X. (2020). Ternary-task convolutional bidirectional neural turing machine for assessment of EEG-based cognitive workload. Biomedical Signal Processing and Control, 57, 101745.

[12] Kwak, Y., Kong, K., Song, W. J., Min, B. K., & Kim, S. E. (2020). Multilevel feature fusion with 3d convolutional neural network for eeg-based workload estimation. IEEE access, 8, 16009-16021.

[13] Zyma, I., Tukaev, S., Seleznov, I., Kiyono, K., Popov, A., Chernykh, M., & Shpenkov, O. (2019). Electroencephalograms during mental arithmetic task performance. Data, 4(1), 14.

[14] Chen, D., Wan, S., Xiang, J., Bao, F. S. (2017). A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PloS one, 12 (3), e0173138.

[15] Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. Expert Systems with Applications, 36 (2), 2027-2036.

[16] Tosun, M., & Çetin, O. (2021). Ampirik Mod Ayrıştırması ve Welch Yöntemini Kullanarak Dört Sınıflı Motor Hayali EEG Sinyallerinin Derin Öğrenme ile Sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, (26), 284-288.

[17] Alhagry, S., Fahmy, A. A., & El-Khoribi, R. A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. International Journal of Advanced Computer Science and Applications, 8(10).

[18] Er, M.B., & İbrahim, I.Ş.I.K. (2021). LSTM tabanlı derin ağlar kullanılarak diyabet hastalığı tahmini. Türk Doğa ve Fen Dergisi, 10(1), 68-74.

[19] Yin, Y., Zheng, X., Hu, B., Zhang, Y., & Cui, X. (2021). EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. Applied Soft Computing, 100, 106954.

[20] Manjusha, M., & Harikumar, R., "Performance analysis of kNN classifier and k-means clustering for robust classification of epilepsy from EEG signals", International Conference on Wireless Communications, Signal Processing and Networking, IEEE, pp. 2412-2416, 2016.

[21] Türk, Ö., & Özerdem, M. S., "EEG İşaretlerinin k-NN ile Sınıflandırmasında Dalgacıklara İlişkin Performansların Karşılaştırılması", Tıp Teknolojileri Ulusal Kongresi, 25-27, 2014.

[22] Kavzoğlu, T., & Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. Harita Dergisi, 144(7), 73-82.

[23] Çağlıyan, B., & Utku, Köse., "Epilepsi EEG Verilerinin Makine Öğrenmesi Teknikleriyle Sınıflandırılması", Avrupa Bilim ve Teknoloji Dergisi, (23), 163-172, 2021.

[24] Lim W. L., Sourina O., Wang L. P., "STEW: Simultaneous Task EEG Workload Data Set", IEEE Transactıons on Neural Systems and Rehabılıtatıon Engıneerıng, 26 (11): 2106-2114, (2018).

[25] Chakladar D.D., Dey S., Roy P.P., Dogra D.P., "EEGbased mental workload estimation using deep BLSTMLSTMnetwork and evolutionary algorithm", Biomedical Signal Processing and Control, 60 101989, (2020).

[26] Aghajani, H., Garbey, M., & Omurtag, A. (2017). Measuring mental workload with EEG+ fNIRS. Frontiers in human neuroscience, 11, 359.

[27] Roy, R. N., Bonnet, S., Charbonnier, S., Jallon, P., & Campagne, A. (2015, August). A comparison of ERP spatial filtering methods for optimal mental workload estimation. In 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 7254-7257). IEEE.