

UDC 004.93

METHOD FOR CLASSIFYING HUMAN PSYCHO-EMOTIONAL STATES BASED ON INTEGRAL ELECTROENCEPHALOGRAM METRICS

Holoborodko V.S.

postgraduate student

ORCID: 0009-0007-2385-9730

Lyfar V.O.

dr.e.s., as.prof.

ORCID: 0000-0002-7860-9663

Volodymyr Dahl East Ukrainian National University, St. John Paul 17, Kyiv, 01042

Abstract. The paper addresses the problem of classifying human psycho-emotional states (positive and negative) based on the intelligent analysis of electroencephalogram (EEG) data. Given that emotional states are characterized by non-stationarity and properties of fuzzy logic, a transition from dynamic amplitude metrics to stationary statistical data is proposed through the calculation of integral metrics over defined time intervals. A classification model employing comparative cluster analysis has been developed. Experimental results confirm the high efficiency of the methods, with identification accuracy reaching up to 95%.

Keywords: psycho-emotional states; cluster analysis; integral indicators; pattern recognition, encephalogram.

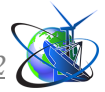
Introduction.

The current stage of development in artificial intelligence systems and digital environments presents new challenges in building effective human-computer interaction. Next-generation interfaces capable not only of executing given commands but also of adapting to the user's psycho-emotional state are becoming particularly relevant. This approach opens up opportunities for deeper, intuitive, and personalized interaction.

Objective monitoring of human psycho-emotional states can be realized by analyzing electroencephalography (EEG) data, which reflect the brain's electromagnetic activity in real time. In this paper, the human emotional sphere is chosen as the object of study.

At the same time, the identification of emotional states via EEG is accompanied by a number of fundamental problems that were taken into account when developing the proposed methodology [1–3].

First, there is the problem of emotion classification. The simplest division is into two basic categories:



- Positive (joy, sympathy);
- Negative (fear, rejection).

These emotions do not have clearly defined boundaries. This indicates that the model of human emotional functioning is characterized by fuzzy logic, which requires the application of appropriate analysis methods [4–6].

Second, a significant problem is the non-stationarity of psycho-emotional states. Emotions change over time and are characterized by inertia. During experiments, it was revealed that with the sequential alternation of stimulating images of opposite valence, a wave response with a frequency of about 1 Hz is recorded in EEG signals. This indicates the presence of transient processes lasting about one second, which must be taken into account to avoid classification errors [7–9].

The aim of this article is the development and verification of a model and methods for EEG data processing that ensure highly effective classification of human emotional states based on objective metrics of brain electromagnetic activity. To achieve this goal, a comprehensive methodology is proposed, based on transforming dynamic non-stationary signals into stationary integral metrics followed by comparative cluster analysis.

Main text.

In the experimental studies, a standardized electroencephalograph recording signals across 24 channels was used. Given the high sensitivity of the electroencephalograph, precise calibration of all leads was performed before each session. All experiments were conducted using an identical electrode placement scheme and standard calibration parameters.

To ensure the comparability of results between different sessions and subjects, a signal normalization procedure was applied. The goal of normalization is to obtain dimensionless quantities scaled to unity. This is achieved by dividing each time-specific value of the encephalogram by its maximum recorded value. As a result, the absolute value of the signal amplitude falls within the range [0; 1], while preserving the general relative characteristic of the signal shape and the dynamics of the extracted patterns. This approach allows the methodology to be applied to encephalograms with



different amplitude calibration settings [10-13].

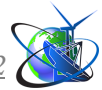
Figure 1 presents a table of statistical parameters illustrating a typical patient electroencephalogram in a resting state. The presented data characterize the main bioelectric brain rhythms: δ , θ , α , β . Specifically, the signal power and the corresponding rhythm frequency are recorded. Such data, structured into a matrix for each lead and rhythm type, form a unique signature of the brain state.

Statistics				
Power	Frequency	Dispersion		
...	δ	θ	α	β
Fp1	2838.420	2890.419	785.433	1467.986
Fpz	2496.836	2374.255	1652.447	1196.039
Fp2	7718.867	2804.228	1517.273	1942.498
F7	7444.301	7115.089	2183.560	2506.391
F3	1331.292	883.748	451.744	1008.829
Fz	4123.130	4925.735	3578.058	3801.853
F4	5608.317	2125.762	1358.728	1515.807
F8	4075.753	2920.763	1529.475	1341.613
T3	5029.992	2632.448	1088.879	1287.200
C3	3355.286	2346.212	1423.183	1426.372
Cz	4036.748	2616.777	362.841	575.528
C4	8981.411	6372.335	2296.585	1918.340
T4	1537.801	2727.796	1854.459	1543.709
T5	6832.660	3620.629	1940.840	2247.029
P3	4039.880	2155.337	1024.177	1767.155
Pz	5743.567	3799.399	2395.526	1813.478
P4	7092.104	3629.170	1975.142	2309.936
T6	2258.758	14622.958	7948.331	13773.455
O1	7010.016	4260.456	1774.078	2747.190
Oz	10083.892	3487.999	1738.273	1661.735
O2	8308.886	4463.490	2554.165	2493.598

Figure 1 - Matrix of statistical parameters of bioelectric rhythms (δ , θ , α , β) in a resting state.

The key objective of the experimental research is to identify stable patterns of electroencephalographic activity corresponding to emotional states of "acceptance" (positive valence) and "rejection" (negative valence) of events in which the subject participates. Correct determination of the actual emotional state is of fundamental importance, as erroneous interpretation during the training set formation stage leads to systematic errors and can significantly distort the results of further classification [14-17].

To ensure the statistical reliability of the results, it is recommended to train the



model on a sample containing at least 100 patterns for each class of emotional states.

Figure 2 presents graphs of the average signal power across each channel, illustrating four test experiments. It is evident that certain channels demonstrate consistently higher activity. Forming such an averaged vector allows for minimizing the influence of random noise bursts and isolating the stable core of the negative emotion pattern.

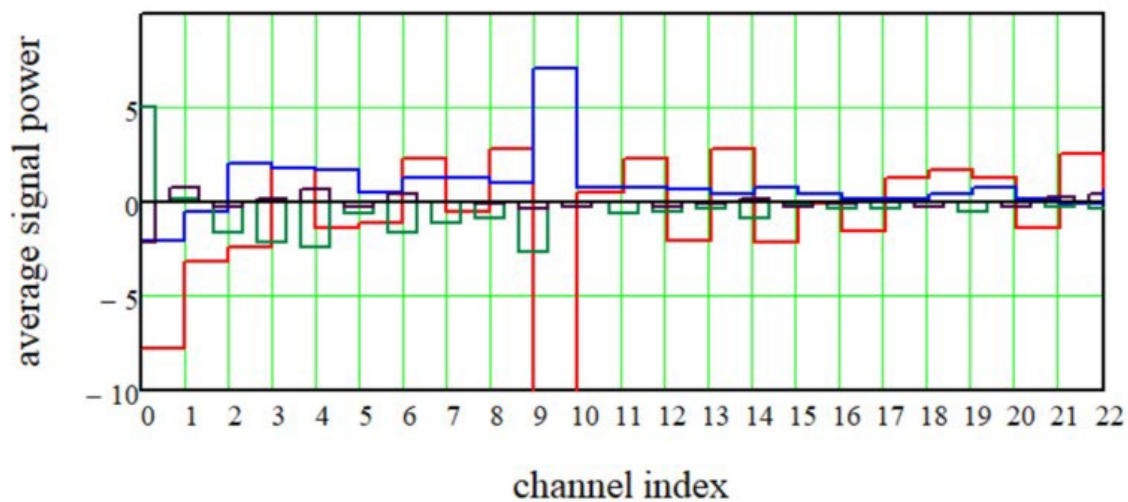


Figure 2 – Spatial distribution of average EEG signal power across channels for reference emotional states.

Averaging metrics from a sample of 50 recordings for the negative state allows for the creation of a typical reference vector for this state, minimizing the impact of random noise components and individual signal fluctuations.

A similar processing procedure is applied to signal power deviations across individual channels; this yields an additional feature set necessary for the complete and correct clustering of emotional state classes.

Figure 3 presents a graph of the average signal power deviations relative to the mean value of the corresponding channel, obtained through a comparative analysis of the first and last layers of the sample.

Analysis of the graph demonstrates that deviation magnitudes for the majority of channels differ minimally, indicating the stability of deviation parameters given a stable psycho-emotional state of the subject.

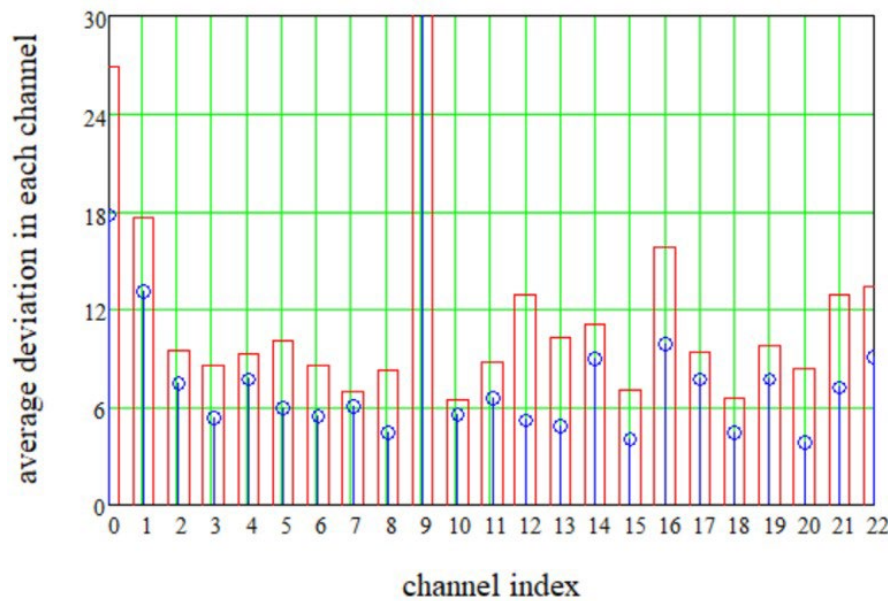
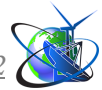


Figure 3 – Diagram of signal power deviations relative to the cluster centroid.

This experiment was conducted under controlled conditions of complete rest: in the absence of visual and auditory stimuli and any other external influences. The subject was instructed beforehand to refrain, as much as possible, from intense cognitive processes and emotionally charged memories capable of shifting the state towards acceptance or rejection of stimuli.

Under these conditions, representative data necessary for electroencephalograph calibration procedures were obtained. Minor discrepancies in deviation values between channels confirm the correct setup of the measurement system and the reproducibility of results.

The controlled process of brain activity is characterized by dynamic changes in the amplitude of electromagnetic oscillations within the spatiotemporal dimension.

Human psychophysiological metrics are in a state of continuous variability and adaptation. However, the classification of such states—whether regarding cognitive activity or emotional manifestations—is possible only through the use of stationary and invariant features.

A positive emotional state cannot be accurately described solely through instantaneous power values of the emotional manifestation, as it possesses temporal duration and pronounced inertia. Consequently, there is a need to represent the time-

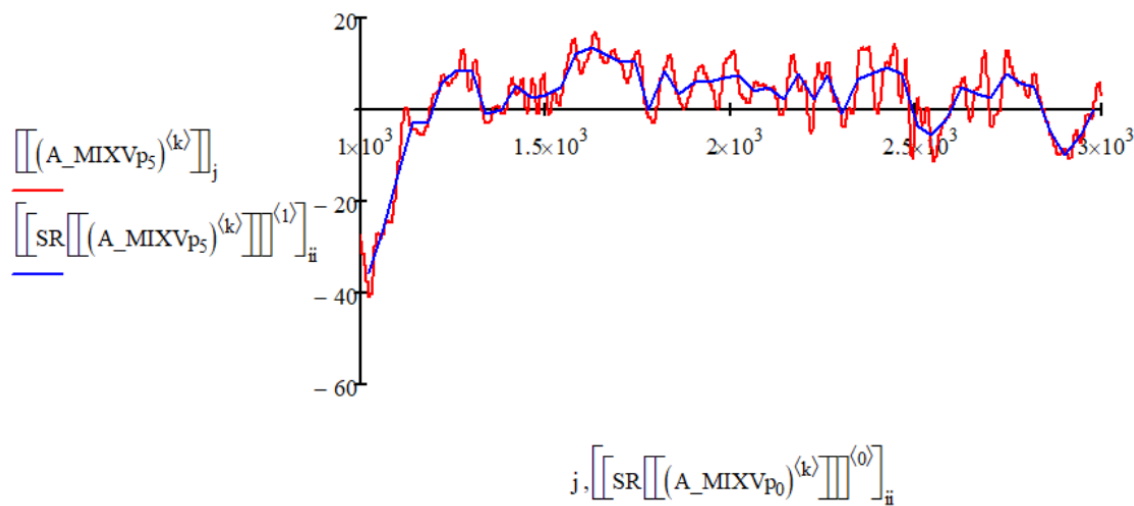
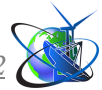


Figure 5 – Result of approximation and smoothing with the decimation factor.

As seen in Fig. 5, the decimation factor sets a time step larger than that of the original function. This allows for filtering out high-frequency noise metrics while preserving the general integral metric averaged over the time step.

Such processing cleans the data of unwanted artifacts without losing the useful signal responsible for slow changes in the emotional state.

It is important to note that isolating the useful signal from the superposition of harmonic oscillations using classical frequency analysis methods is ineffective in this context. Moreover, classifying the isolated harmonics by origin is practically impossible due to the complex stochastic nature of the signal.

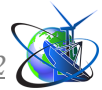
Therefore, this methodology proposes abandoning attempts at signal frequency decomposition in favor of comparative cluster analysis.

The method is based on comparing the investigated encephalograms with averaged metric patterns obtained through the clustering of uniform verified states.

Following the approximation stage, the integral value for each averaged signature is calculated according to the developed algorithm. This operation is performed iteratively for all recorded EEG leads. As a result, a unique vector of integral metrics is formed for each experimental session, serving as an identifier of the emotional state.

To train and verify the model, two data groups were formed:

- Training set (reference clusters): Cluster groups were created for polar states—3



clusters of the negative state and 3 clusters of the positive state. Each cluster is based on the aggregated data of 20 verified encephalograms.

- Test set (control groups): 10 negative and 10 positive encephalograms subject to classification were prepared; these did not participate in cluster formation.

The mathematical formalization of the process for creating arrays of averaged integral values uses the following indexing:

- $k = 2..22$ – indices of EEG leads;
- $m = 0..19$ – indices of signatures within the cluster (training set);
- $m_1 = 0..9$ – indices of analyzed signatures (test set).

To visually assess class separability, heatmaps were constructed (Fig. 6). Darker zones correspond to higher integral activity. The clear visual distinction of patterns confirms the hypothesis regarding the possibility of effective clustering of "positive" and "negative" states based on integral areas.

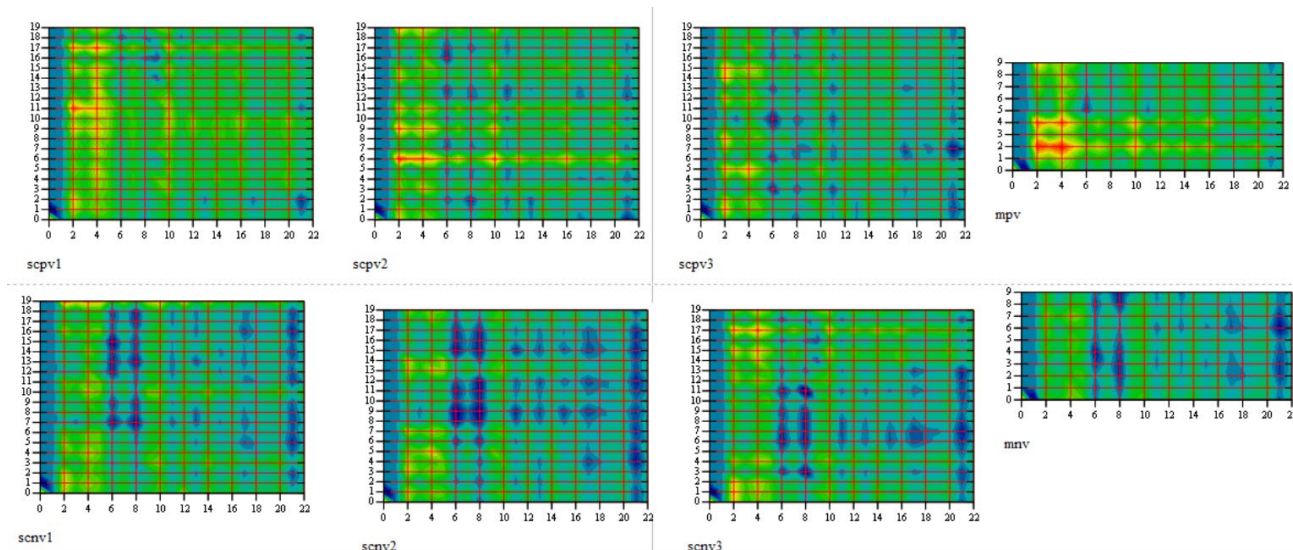
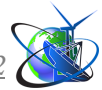


Figure 6 – Visualization of emotional state clusters using integral metric heatmaps.

To ensure high reliability and stability of emotional state recognition results, a classification method based on total integral distance has been developed.

This approach is grounded in the k-Nearest Neighbors (k-NN) principle; however, it is applied not to raw time series, but to pre-processed integral EEG metrics .



The core idea lies in determining the membership of the test EEG to a specific cluster by calculating the total deviation of integral signal values simultaneously across all leads. Unlike dynamic analysis, where graphs are compared in the time domain, this method involves the comparison of stationary characteristics—integral areas under the brain activity curve over a selected time interval.

Figure 7 presents an example of calculating the distance between a positive electroencephalogram from the control group and a reference positive electroencephalogram from the training cluster.

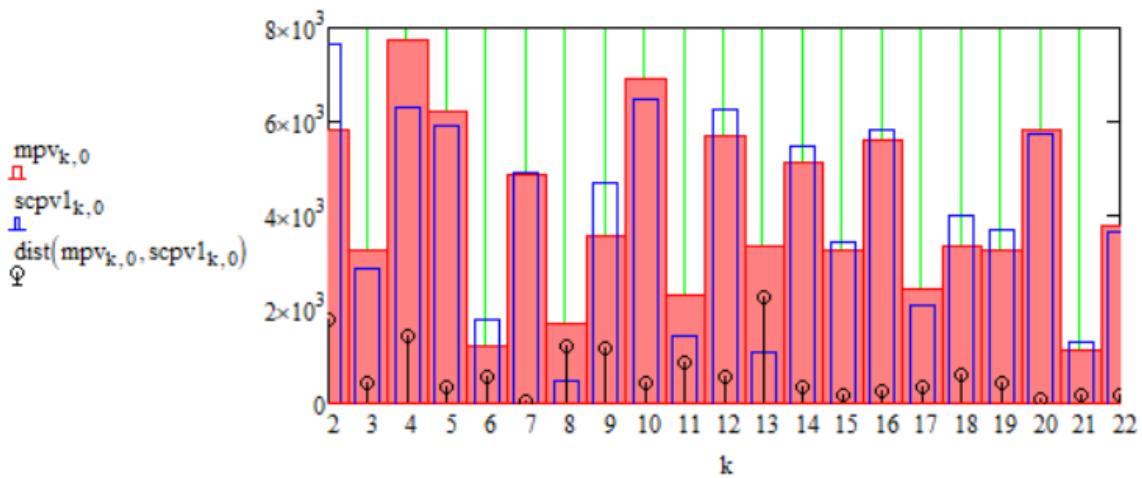


Figure 7 – Comparative analysis of integral curves of the test sample and the reference cluster.

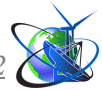
Legend:

- mpv (red color) – mixed positive electroencephalograms (control group);
- scp1 (blue color) – 1st positive EEG cluster (training set);
- distt (black dot) – distance magnitude (difference) between the mixed EEG and the 1st cluster.

The determination of the test signature's membership is provided by a classification function that identifies the cluster with the minimum total distance to the test EEG. Specifically, if the total difference in integral areas between the test sample and the sample from the positive cluster is smaller than that regarding the negative cluster, the system classifies the state as positive.

Summary and conclusions.

To verify the developed model, a test sample of 20 EEG recordings (10 positive



and 10 negative states) was formed; these were not used during the training phase.

Processing parameters:

- Integration start – 1000 ms, end – 3000 ms;
- Decimation factor $k_r = 50$.

Testing results demonstrated the following:

Positive states: All 10 test samples were correctly identified and assigned to the corresponding clusters of the positive spectrum. The accuracy in this subgroup was 100%.

Negative states: Classification results are presented in Table 1. Out of 10 test samples, 9 were correctly classified. One misclassification was detected for electroencephalogram No. 6 (index 5), which the system assigned to the 3rd positive cluster. This may be related to the individual characteristics of transient processes in the specific subject. The accuracy of the subgroup was 90%.

Table 1 – Classification results of negative electroencephalograms

	0	1
0	595514.9	-1
1	526206	-2
2	498354.4	-2
3	539725.3	-2
4	489761.6	-2
5	516428.1	3
6	604007.1	-2
7	531113.9	-2
8	508849.3	-2
9	532239.6	-2

The overall classification accuracy across the entire test sample was 95%.

The obtained results confirm the hypothesis that the use of integral metrics effectively mitigates the non-stationarity of psycho-emotional states.



The method successfully isolates stable, reproducible patterns of brain electromagnetic activity, ensuring reliable separation of positive and negative emotional responses even in the presence of noise components. Future research will focus on expanding the sample size and performing statistical validation of the model.

References:

1. Chen, Y.; Wang, W.; Yan, S.; Wang, Y.; Zheng, X.; Lv, C. Application of Electroencephalography Sensors and Artificial Intelligence in Automated Language Teaching. *Sensors* 2024, 24, 6969. <https://doi.org/10.3390/s24216969>
2. Leung C., Scarino A.: Reconceptualizing the nature of goals and outcomes in language/s education. *Modern Language Journal* 2016, 100, 81–95.
3. Zhang Y., Wa S., Liu Y., Zhou X., Sun P., Ma Q.: High-accuracy detection of maize leaf diseases CNN based on multi-pathway activation function module. *Remote Sensing* 2021, 13, 4218.
4. Sharma V.K., Holbah W.A.: Online Language Assessment the Exception, Not the Rule: For Inclusive Language Learning. *Arab World English Journal* 2022.
5. Li Q., Zhang Y., Ren J., Li Q., Zhang Y.: You Can Use But Cannot Recognize: Preserving Visual Privacy in Deep Neural Networks. *arXiv* 2024, arXiv:2404.04098.
6. Chappuis J., Stiggins R.J., Chappuis S., Arter J.: *Classroom Assessment for Student Learning: Doing It Right—Using It Well*. Pearson, New York, NY, USA, 2020.
7. Wei X., Saab N., Admiraal W.: Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review. *Computers & Education* 2021, 163, 104097
8. Ganepola D., Maduranga M.W.P., Tilwari V., Karunaratne I.: A Systematic Review of Electroencephalography-Based Emotion Recognition of Confusion Using Artificial Intelligence. *Signals* 2024, 5, 244–263. [<https://doi.org/10.3390/signals5020013>].
9. Bhattacharya S., Singh A., Hossain M.: Health system strengthening through massive open online courses (MOOCs) during the COVID-19 pandemic: An analysis from the available evidence. *Journal of Education and Health Promotion* 2020, 9, 195.



10. Ganepola D.: Assessment of Learner Emotions in Online Learning via Educational Process Mining. In: Proceedings of the 2022 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 8–11 October 2022, pp. 1–3.

11. El-Sabagh H.A.: Adaptive e-learning environment based on learning styles and its impact on development students' engagement. International Journal of Educational Technology in Higher Education 2021, 18, 53.

12. Wang H., Li Y., Hu X.S., Yang Y., Meng Z., Chang K.K.: Using EEG to Improve Massive Open Online Courses Feedback Interaction. CEUR Workshop Proceedings 2013, 1009, 59–66. [Available online: <https://experts.illinois.edu/en/publications/using-eeg-to-improve-massive-open-online-courses-feedback-interac> (accessed on 8 April 2024)].

13. Wang H.: Confused Student EEG Brainwave Data. Kaggle, 28 March 2018. [Available online: <https://www.kaggle.com/datasets/wanghaohan/confused-eeg/code> (accessed on 11 February 2023)].

14. Xu T., Wang J., Zhang G., Zhou Y.: Confused or not: Decoding brain activity and recognizing confusion in reasoning learning using EEG. Journal of Neural Engineering 2023, 20, 026018.

15. Chen X., Tao X., Wang F.L., et al.: Global research on artificial intelligence-enhanced human electroencephalogram analysis. Neural Computing and Applications 2022, 34, 11295–11333. [<https://doi.org/10.1007/s00521-020-05588-x>].

16. Chang W., Wang H., Yan G., Liu C.: An EEG based familiar and unfamiliar person identification and classification system using feature extraction and directed functional brain network. Expert Systems with Applications 2020, 158, 113448.

17. Becerra-Sánchez P., Reyes-Munoz A., Guerrero-Ibañez A.: Feature selection model based on EEG signals for assessing the cognitive workload in drivers. Sensors 2020, 20(20), 5881.

Scientific adviser: Doctor of Technical Sciences, prof. Lyfar V.O.

Article sent: 26.12.2025

© Holoborodko V.S.