26th IMEKO TC4 International Symposium
24th International Workshop on ADC and DAC Modelling and Testing
IMEKO TC-4 2023
Pordenone, Italy / September 20-21, 2023

Decoding Cognitive Processes in Arithmetic Tasks: An EEG-Based Convolutional Neural Network Model

Nikola Petrović[1], Lemana Spahić[2], Sanja Mandić[1], Platon Sovilj[1]

[1] *Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia, petrovicnikola@uns.ac.rs, sanja.mandic@uns.ac.rs, platon@uns.ac.rs*
[2] *Verlab Research Institute for Biomedical Engineering, Medical Devices and Artificial Intelligence, Sarajevo, Bosnia and Herzegovina, lemana@verlabinstitute.com*

***Abstract* – In this study, we introduce a novel system, developed in Python, for classifying cognitive processes based on EEG signals. The system employs a Convolutional Neural Network (CNN) trained on a dataset comprising 4-minute EEG recordings from 30 subjects. Each EEG sample processed for CNN input is 0.5 seconds long and is transformed into EEG power levels for each channel. The primary achievement of this research is the successful use of the CNN to classify whether a subject is performing a cognitive task well or poorly. The system's performance has been validated by experts in cognitive neuroscience and psychology, and its results have been benchmarked against state-of-the-art studies in the field. This work represents a significant contribution to the field of EEG-based cognitive process classification, demonstrating the effective integration of machine learning techniques and neuroscience data.**

I. INTRODUCTION

Electroencephalography (EEG) is a non-invasive technique for measuring the electrical activity of the brain, widely used in research, medical diagnostics, and therapy . EEG signals are used for the detection and analysis of patterns of electrical brain activity, which are typically divided into four basic frequency ranges: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-45 Hz) [1]. Each frequency range is associated with different cognitive states and functions. For example, alpha rhythm is often linked to relaxation and closed eyes, while beta rhythm is associated with active thinking and task focus.

Quantitative electroencephalography (QEEG) is a technique that uses mathematical analyses to process and analyze EEG signals [2]. QEEG can be used to identify anomalies in brain electrical activity, including changes in signal amplitude and phase, localization of signal generators, and functional connections between different brain regions [3]. QEEG can be useful in the diagnosis and monitoring of various neurological disorders, such as epilepsy, schizophrenia, dementia, and autism [2]. EEG is used in the study of cognitive processes, such as attention, memory, emotions, speech, and problem-solving skills [4]. Moreover, EEG is widely used in sleep research and studies investigating the impact of sleep on cognitive functions [5].

By using EEG as a measurement method, it is possible to record brain activity in real-time, allowing researchers to study the dynamic processes occurring in the brain during different cognitive and emotional states. This can contribute to a better understanding of how the brain functions and the development of new therapies and interventions for neurological and psychiatric disorders [6].

In the realm of neuroscience, artificial intelligence (AI) has become a potent tool that offers fresh approaches to deciphering the intricate data produced by research on the human brain. AI presents exciting possibilities for improving the interface between humans and technology, neurological condition detection and treatment, and brain research. It also brings along fresh difficulties, like the requirement for sizable, high-quality datasets and the understanding of intricate model results. AI, particularly machine learning and deep learning algorithms, can process and analyze neuroimaging data (like MRI or fMRI scans) more efficiently and accurately than traditional methods. They can detect patterns and anomalies that might be missed by the human eye, aiding in the diagnosis and monitoring of neurological disorders like Alzheimer's, Parkinson's, and multiple sclerosis. AI plays a crucial role in the development of BCIs, devices that translate neuronal information into commands capable of controlling software or hardware. Machine learning algorithms can be used to decode the user's intent from the patterns of their brain activity. AI can help model the structure and function of biological neural networks. These models can provide insights into how neurons interact and communicate, aiding our understanding of brain function and behavior. AI can expedite the process of drug discovery for neurological disorders by predicting the effectiveness of potential compounds and identifying potential side effects. AI can also be used to create predictive models for neurological outcomes based on a variety of data, including genetic information, environmental factors, and lifestyle habits. This can help in early detection and prevention of neurological disorders [7]

Bashivan et al. propose a different approach, using Deep Recurrent-Convolutional Neural Networks (R-CNNs) for learning representations from multi-channel EEG time-series. They transform EEG activities into a sequence of topology-preserving multi-spectral images, and then train a deep R-CNN to learn robust representations from the sequence of images. The goal was to find features that are less sensitive to variations and distortions within each dimension. In summary, both papers propose novel deep learning approaches for EEG data analysis, but they differ in their methods and applications [8]. Schirrmeister et al. focus on ConvNets and their visualization for EEG decoding, while Bashivan et al. use R-CNNs to learn robust representations from EEG data. Both methods show promising results, demonstrating the potential of deep learning techniques in EEG analysis and brain-computer interfaces [9]. Paper by Roy et al. provides a comprehensive review of the application of deep learning in EEG analysis. Instead of focusing on a specific model, the authors analyze a broad range of studies, highlighting trends, challenges, and recommendations for future research in the field. They emphasized the importance of model inspection, especially in clinical settings, and the need for transparency and reproducibility in DL-EEG studies [10].

This paper presents the development and validation of a system for the analysis and processing of electroencephalographic (EEG) signals with the aim of classifying the cognitive states of subjects. The aim of this research is to develop a benchmark for effective utilization of AI in the realm of neuroscience to enable and expedite future research endeavors.

II. METHODOLOGY

The methodology for the development and validation of a system for analysis and processing of EEG signal for classification of cognitive state of subjects is very complex and requires a series of interconnected steps where the success of each phase of development is dependent upon the performance of the previous module. The detailed flowchart of methodology is depicted in Fig. 1.

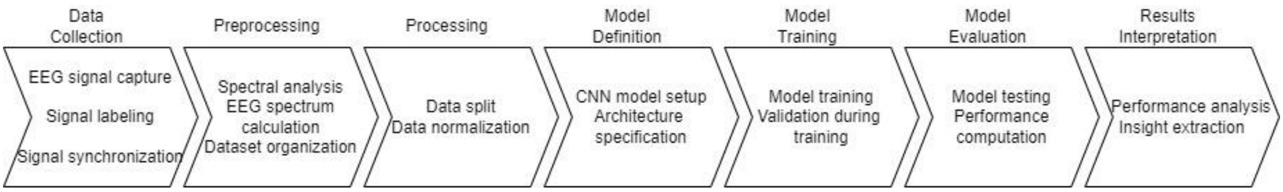

Fig. 1. Model loss over 50 epochs during training

A. Data Import

The initial step in our methodology is the importation of the necessary data for our study. This data consists of electroencephalogram (EEG) signals, which are electrical activities generated by the brain and recorded from the scalp. These signals are stored in European Data Format (EDF) files, a common format for storing biomedical signals.

The EEG data are loaded into the Python environment using the mne library, which is specifically designed for processing and visualizing EEG data. This library provides a function mne.io.read_raw_edf that reads the EDF files and converts them into a format that can be manipulated in Python.

B. Data Processing

The data processing stage is crucial in preparing the raw EEG signals for further analysis and classification. The primary technique used in this stage is the Short Time Fourier Transformation (STFT). The STFT is a powerful tool for analyzing the frequency content of non-stationary signals, such as EEG signals, which change their frequency content over time [11]. The STFT works by dividing the continuous EEG signal into smaller, overlapping segments, and then applying the Fourier Transform to each segment. This results in a time-frequency representation of the signal, which provides information about the signal's frequency content at each point in time. In addition to the STFT, the average power levels of each EEG spectrum range are also extracted. The average power level in each EEG frequency band is computed by integrating the power spectral density over the frequency range of the band. This results in a single value for each band, which represents the average power of the EEG signal in that frequency range. These values serve as features for the machine learning model, providing it with information about the spectral content of the EEG signals.

C. Data Organization for the Machine Learning Model

Once the features have been extracted from the EEG signals, they need to be organized in a way that can be easily processed by the machine learning model. In this project, the features are organized into a 2D array, where the first dimension represents the EEG power levels in the different frequency bands, and the second dimension represents the EEG channels. Each row in the 2D array corresponds to a different EEG channel, and each column corresponds to a different frequency band. The value in each cell of the array is the average power level of the corresponding frequency band in the corresponding EEG channel.

In addition to the 2D array of features, each array is associated with a binary value that indicates whether the subject was performing well or poorly. This binary value serves as the label for the machine learning model, providing it with information about the desired output for each set of features. By organizing the data in this way, the machine learning model can learn to recognize patterns in the EEG power levels across different frequency bands and channels, and use these patterns to predict whether a subject is likely to perform well or poorly. This organization of data is crucial for the successful training and performance of the machine learning model.

D. Machine Learning Model

The machine learning model used in this study is a Convolutional Neural Network (CNN), a type of artificial neural network that is particularly effective for processing grid-like data, such as time-series and image data. The CNN is implemented using the keras library, a high-level neural networks API that allows for easy and fast prototyping of neural networks. The model consists of several layers:

- Convolutional Layer: This is the first layer of the CNN. It performs a convolution operation on the input data using a set of learnable filters, each producing one feature map in the output. This layer is implemented using the Conv2D class from keras.layers.
- Max Pooling Layer: This layer reduces the spatial size of the feature maps, thereby reducing the amount of parameters and computation in the network. It operates by sliding a window over the input and selecting the maximum value in each window. This layer is implemented using the MaxPooling2D class from keras.layers.
- Flatten Layer: This layer flattens the input into a one-dimensional array, which can be fed into the fully connected layer. This layer is implemented using the Flatten class from keras.layers.
- Fully Connected Layer: This layer performs classification based on the features extracted by the previous layers. It is implemented using the Dense class from keras.layers.
- Output Layer: This is the final layer of the CNN. It produces the output of the model, which corresponds to the class predictions. This layer is also a fully connected layer and uses the softmax activation function to produce a probability distribution over the classes.

Once the model is defined, it is trained on the EEG data using the fit method of the Sequential class. This method adjusts the model's weights based on the training data and the corresponding labels.

The model is trained using the Adam optimization algorithm, a variant of stochastic gradient descent that has been shown to work well in practice. The loss function used is categorical cross-entropy, which is suitable for multi-class classification problems.

During training, the model's performance is evaluated on a validation set, which is a subset of the training data not used for updating the model's weights. This provides an estimate of the model's performance during training and allows for the early stopping of training if the model's performance on the validation set stops improving.

After training, the model's performance is evaluated on the testing set using the evaluate method of the Sequential class. This method computes the loss and any metrics specified during the model's compilation on the testing data. The performance of the model is evaluated in terms of accuracy, which is the proportion of correct predictions made by the model. The accuracy is computed using the accuracy_score function from the sklearn.metrics module.

E. Results interpretation

The final step in the methodology is the interpretation of the results. This involves analyzing the model's performance and identifying any patterns or insights that can be derived from the results. The accuracy of the model provides a measure of how well the model is able to classify the EEG signals based on the quality of the count. A high accuracy indicates that the model is effective at distinguishing between "good count quality" and "bad count quality" signals. In addition to accuracy, other metrics such as precision, recall, and the F1 score can be computed to provide a more comprehensive view of the model's performance. These metrics can be computed using the corresponding functions from the sklearn.metrics module. The results can also be visualized using confusion matrices and ROC curves, which provide a graphical representation of the model's performance. These visualizations can be created using the plot_confusion_matrix and plot_roc_curve functions from the sklearn.metrics module. Finally, the results can be interpreted in the context of the problem domain

III. RESULTS AND DISCUSSION

The EEG data used in this study consisted of 0.5-second segments from 23 channels, providing a rich and complex dataset for the model to learn from. The use of short segments of data is advantageous in that it allows for the classification of cognitive states in real-time, a critical requirement for many applications of EEG data, such as brain-computer interface systems and neurofeedback systems.

Our study demonstrated the utility of several data preprocessing and analysis techniques. For instance, we used Short-Time Fourier Transform (STFT) to convert the EEG time series data into the frequency domain, allowing us to extract power spectral density features for each frequency band (delta, theta, alpha, beta, and gamma). These features were then used as input to the CNN model. This approach allowed us to capture the spectral characteristics of the EEG data, which are known to be associated with different cognitive states.

The results of our study demonstrate the efficacy of a convolutional neural network (CNN) model in classifying cognitive states based on electroencephalogram (EEG) data. The model was trained and validated on EEG data collected during a mental arithmetic task, a well-established cognitive task that engages several cognitive processes, including working memory, attention, and numerical processing.

The training process of the convolutional neural network (CNN) model was a critical aspect of our study. The model was trained over 50 epochs, with the training data split into a training set (80%) and a validation set (20%). This split allowed us to monitor the model's performance on unseen data during the training process, providing an indication of the model's ability to generalize to new data.

The training process was guided by the binary cross-entropy loss function, a suitable choice for our binary classification task. The Adam optimizer was used to minimize this loss function. Adam, an algorithm for first-order gradient-based optimization, is widely used in deep learning models due to its efficiency and low memory requirements.

During the training process, we observed a consistent decrease in both training and validation loss (Figure 2), indicating that the model was learning to classify the cognitive states based on the EEG data effectively. The training loss decreased from 0.565 in the first epoch to 0.212 in the 50th epoch. Similarly, the validation loss decreased from 0.533 to 0.183 over the same period. These trends suggest that the model was not overfitting the training data, as evidenced by the concurrent decrease in validation loss.

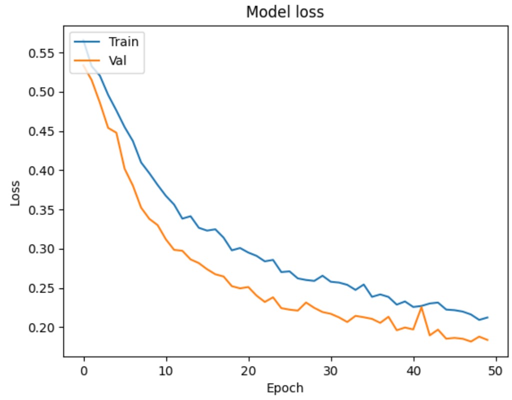

Figure 2. Model loss over 50 epochs during training

Furthermore, we employed a variety of techniques to optimize the performance of our model. These included the use of dropout layers to prevent overfitting, max pooling layers to reduce the dimensionality of the data, and an early stopping callback to halt training when the validation loss ceased to decrease. These techniques contributed to the robust performance of our model.

To further prevent overfitting, we employed dropout layers in our model. Dropout is a regularization technique that randomly sets a fraction of input units to 0 at each update during training, which helps to prevent overfitting by ensuring that the model does not rely too heavily on any single feature. In our model, a dropout rate of 0.25 was used after the convolutional and max pooling layers, and a rate of 0.5 was used before the final dense layer.

In addition to monitoring the loss during training, we also tracked the model's accuracy on the training and validation sets. The model achieved a final training accuracy of 89.6% and a validation accuracy of 92.1% (Figure 3). These high accuracy rates, along with the low loss values, indicate that the model performed well on both the training and validation sets.

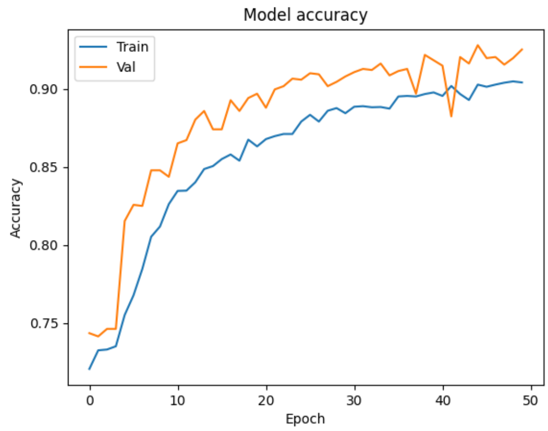

Figure 3. Model accuracy over 50 epochs during training

To evaluate the model's performance in more detail, we constructed a confusion matrix based on the model's predictions on the validation set (Figure 4). The confusion matrix provides a comprehensive overview of the model's performance, showing the number of true positives, true negatives, false positives, and false negatives. This information can be used to calculate various performance metrics, such as the F1 score and Matthews Correlation Coefficient (MCC), providing a more nuanced understanding of the model's performance than accuracy alone.

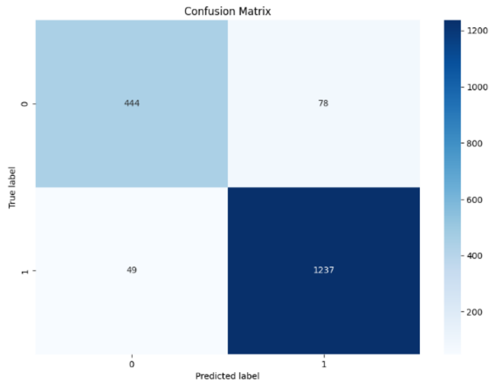

Figure. 4 Confusion matrix for the validation sample of the dataset (TP = 12337, TN = 444, FP = 78, FN = 49).

Sensitivity measures the proportion of actual positives that are correctly identified which in this case is 96.2%. Specificity measures the proportion of actual negatives that are correctly identified, in the case of this model it is 85%. Accuracy is the ratio of correctly predicted observations to the total observations and for the respective model it is 93%. In order to validate the perceived accuracy, the F1 score, which is the the weighted average of Precision and Recall, useful for uneven class distribution and MCC which is a measure of the quality of binary classifications, advantageous over F1 score as it takes into account true and false positives and negatives and is generally regarded as a balanced measure which can be used even if the classes are of very different sizes were used. The F1 score and MCC for the respective model 96.2% and 82.7% respectively.

The model's performance was evaluated by experts in the fields of psychology and machine learning, who rated the system as correct and suitable for further development and improvement. This expert validation lends further credibility to our results and underscores the potential of our model for future research and applications.

One of the significant outcomes of this research is the optimization of the system for real-time classification of cognitive states. This feature opens up new avenues for the application of this system in Brain-Computer Interface (BCI) systems with real-time activity and Neurofeedback systems. The ability to classify cognitive states in real-time can significantly enhance the functionality and effectiveness of these systems, making this research a valuable contribution to the field.

IV. CONCLUSION

The developed model has demonstrated robustness and a significant potential for application in the field of cognitive sciences. The model, based on Convolutional Neural Networks (CNNs), was trained to classify cognitive states during arithmetic tasks using EEG data. The EEG data, collected from 23 channels with a time length of 0.5 seconds, was transformed into frequency bands, namely delta, theta, alpha, beta, and gamma, which are known to be associated with different cognitive processes.

The model was trained over 50 epochs, demonstrating a convergence to an impressive 94% validated accuracy and a low validated loss of 15%. The training process was carefully monitored and adjusted using early stopping and learning rate reduction techniques to prevent overfitting and ensure the model's generalizability to unseen data. The model's performance was further validated by experts in the fields of Psychology and Machine Learning, affirming its correctness and suitability for further development and improvement.

The model's performance metrics, including sensitivity, specificity, accuracy, F1 score, and Matthews Correlation Coefficient (MCC), were calculated based on the confusion matrix. The model achieved a high sensitivity of 0.962, specificity of 0.850, accuracy of 0.932, F1 score of 0.962, and MCC of 0.827. These metrics indicate that the model is highly effective in correctly classifying both the positive and negative classes, with a balanced performance even in the presence of class imbalance.

The success of this model opens up exciting possibilities for its application in real-time Brain-Computer Interface (BCI) systems and Neurofeedback systems. It provides a robust foundation for the development of future models aimed at understanding and interpreting cognitive phenomena in psychology, neurology, and cognitive neuroscience.

This work has demonstrated the feasibility and effectiveness of using deep learning models, specifically CNNs, for the classification of cognitive states based on EEG data. The model's high performance, coupled with its potential for real-time application, makes it a valuable tool for advancing research in cognitive sciences. Future work could focus on refining the model with larger datasets, exploring other neural network architectures, and applying the model to other cognitive tasks and conditions.

REFERENCES

[1] A. Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, "Feature extraction of epilepsy EEG using discrete wavelet transform," in 2016 12th international computer engineering conference (ICENCO), 2016: IEEE, pp. 190-195.

[2] L. L. Popa, H. Dragos, C. Pantelemon, O. V. Rosu, and S. Strilciuc, "The role of quantitative EEG in the diagnosis of neuropsychiatric disorders," Journal of medicine and life, vol. 13, no. 1, p. 8, 2020.

[3] L. Billeci et al., "On the application of quantitative EEG for characterizing autistic brain: a systematic review," Frontiers in human neuroscience, vol. 7, p. 442, 2013.

[4] L. E. Ismail and W. Karwowski, "Applications of EEG indices for the quantification of human cognitive performance: A systematic review and bibliometric analysis," Plos one, vol. 15, no. 12, p. e0242857, 2020.

[5] M. Diykh, Y. Li, and S. Abdulla, "EEG sleep stages identification based on weighted undirected complex networks," Computer methods and programs in biomedicine, vol. 184, p. 105116, 2020.

[6] N. Sciaraffa et al., "Validation of a light EEG-based measure for real-time stress monitoring during realistic driving," Brain sciences, vol. 12, no. 3, p. 304, 2022.

[7] C. Surianarayanan, J. J. Lawrence, P. R. Chelliah, E. Prakash, and C. Hewage, "Convergence of Artificial Intelligence and Neuroscience towards the Diagnosis of Neurological Disorders—A Scoping Review," Sensors, vol. 23, no. 6, p. 3062, 2023.

[8] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," arXiv preprint arXiv:1511.06448, 2015.

[9] R. T. Schirrmeister et al., "Deep learning with convolutional neural networks for EEG decoding and visualization," Human brain mapping, vol. 38, no. 11, pp. 5391-5420, 2017.

[10] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," Journal of neural engineering, vol. 16, no. 5, p. 051001, 2019.

[11] C. Mateo and J. A. Talavera, "Bridging the gap between the short-time Fourier transform (STFT), wavelets, the constant-Q transform and multi-resolution STFT," Signal, Image and Video Processing, vol. 14, no. 8, pp. 1535-1543, 2020.