Analyzing Shallow and Deep CNNs with Gold Standard and Random EEG Segment Selections for Coma Prognosis

Gustavo B. S. Oliveira[1], **Paulo D. S. Souza**[1], **Tiago S. Silva**[1],
João B. D. Filho[2], **Murillo G. Carneiro**[1]

[1]Faculty of Computing
Federal University of Uberlândia, Uberlândia, MG, Brazil

[2]Faculty of Electrical Engineering
Federal University of Uberlândia, Uberlândia, MG, Brazil

{gustavo.oliveira4, pdodonto, tiago.souza, mgcarneiro, jbdestro}@ufu.br

Abstract. *This study focuses on deep learning models for predicting coma outcomes using electroencephalogram (EEG) data, exploring convolutional neural networks (CNNs), particularly Shallow and Deep ConvNets, based on Filter Bank Common Spatial Patterns. A dataset of 121 EEG samples (42 favorable, 79 unfavorable) was analyzed. EEG segments were selected using two strategies and frequencies. Models were trained with 10-fold cross-validation and FTSurrogate for class balance. Shallow ConvNet showed stable performance across frequencies, while Deep ConvNet excelled at 200Hz. Simple segment selection and sampling frequency methods improved CNN performance. The findings offer insights for future research and potential clinical applications.*

1. Introduction

This paper presents the development of deep learning models for the prognosis of patients in coma using electroencephalogram (EEG) data. In recent years, advances in machine learning techniques, especially deep neural networks, have provided new possibilities for analyzing biomedical signals, enabling the extraction of complex patterns that are difficult to identify with traditional methods [LeCun et al. 2015, Mathew et al. 2021]. The application of these techniques to EEG data, which measures brain electrical activity, shows considerable promise for obtaining more accurate and reliable prognoses for patients in coma [Hossain et al. 2023, Saeidi et al. 2021, Xu et al. 2023].

The electroencephalogram (EEG) is a valuable tool for monitoring comatose patients, as it captures brain electrical activity in a non-invasive and relatively accessible manner [İnce et al. 2021, Iwama et al. 2023]. However, interpreting EEG traces poses significant challenges due to the high variability and noise present in the recorded signals [Aellen et al. 2023, Baldo Júnior et al. 2023]. Thus, the increasing length of recordings and number of electrodes significantly amplify the complexity and volume of data to be analyzed. This study conducts an exploratory analysis of the potential of convolutional neural networks (CNNs), specifically Deep ConvNet and Shallow ConvNet [Schirrmeister et al. 2017] for processing and analyzing EEG signals in the context of coma prognosis. Our main contributions are:

- Design of a random selection method of EEG segments as a fast and simple alternative to gold standard (experts) selection;

- Evaluation of Shallow and Deep ConvNets based on Filter Bank Common Spatial Patterns (FBCSP) for different sampling frequencies and segment selection strategies using real EEG data related to the coma prognosis task.

The subsequent sections of the paper are organized as follows. Section 2 presents related work, discussing relevant research in the context of this study. Section 3 describes the materials and methods, including the EEG-Coma Dataset, implementation details, preprocessing, CNN models, hyperparameter optimization, and performance evaluation. Section 4 presents the results for both Expert-Selected Segments and Randomly Selected Segments, along with a discussion of the experimental findings. Finally, Section 5 provides the conclusions, summarizing the findings, contributions, limitations, and proposals for future work.

2. Related Works

[Schirrmeister et al. 2017] analyzed various ConvNet architectures for decoding imagined or executed tasks from raw EEG data. They found that deep ConvNets outperformed conventional methods and could learn spectral power modulations across different frequencies, highlighting the growing interest and potential of deep ConvNets for advanced EEG analysis and brain mapping without predefined features.

[Ramos et al. 2022] investigated the classification of comatose patient prognoses using time and frequency domain quantifiers in EEG signals. Applying classical machine learning algorithms like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Logistic Regression (LR), their study highlighted the importance of comprehensive signal analysis in predicting patient outcomes by improving performance through the incorporation of both time-domain and frequency-domain features.

The study by [Aellen et al. 2023] investigated the prognosis of comatose patients after cardiac arrest using auditory stimulation and deep learning. They hypothesized that Convolutional Neural Networks could extract interpretable patterns from EEG responses to auditory stimuli during the first day of coma, predicting awakening and survival chances at three months. The use of CNNs resulted in a positive predictive value of 0.83 and an area under the curve of 0.69, demonstrating the potential of deep learning to improve coma prognostication.

[Bissaro et al. 2023] addressed EEG signal decoding for coma prognosis using Echo-State Networks (ESNs) and CNNs. Introducing a spatial dimension to the EEG data by modeling electrode placement and relationships, the transformed data fed into CNN architectures surpassed state-of-the-art approaches for predicting two and three possible outcomes, demonstrating the robustness of their method.

[Carneiro et al. 2023] explored high-level classification techniques for coma prognosis, focusing on assortativity and shortest path metrics. Compared with nine other approaches, including CNNs, their method showed potential to enhance predictive performance, emphasizing the value of advanced classification techniques in improving prognostic models.

Unlike previous related works, this paper assesses the potential of Shallow and Deep ConvNets based on FBCSP, considering both Gold-standard and Random EEG segment selection.

3. Material and Methods

This section presents the materials and methods used in this study. The process is summarized in Figure 1, beginning with data collection by specialists. The second step involves selecting the dataset, considering gold and random selection, resampling, and data augmentation as pre-processing techniques. In the third phase, the training-evaluation pipeline is configured, along with the definition of the hyperparameter optimization setup. The fourth step consists of splitting the dataset into training and validation sets and instantiating the CNN models under study. As part of this step, data augmentation is applied exclusively to the training set in order to improve model generalization. Next, the models are evaluated using the validation set, and partial results are collected and fed back into the pipeline. In the final step, partial results from all pipelines are aggregated and ranked, leading to the final results.

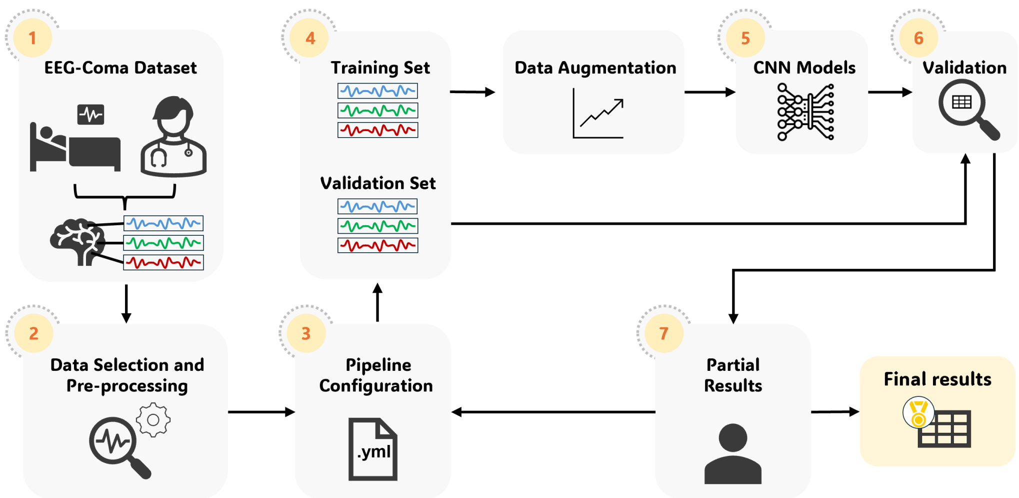

Figure 1. Overview of the study training-evaluation workflow.

3.1. EEG-Coma Dataset

The EEG-coma dataset was clinically collected at the adult intensive care unit of the Clinical Hospital of the Federal University of Uberlândia (HCU-UFU). The collection was approved by the Research Ethics Committee of the Federal University of Uberlândia. The aim of this study was to develop models to assist in the prognosis of patients in coma by categorizing EEG signals into favorable and unfavorable outcomes. A favorable outcome is defined as the patient being discharged from the intensive care unit in satisfactory health, while an unfavorable outcome includes either death from various clinical causes or a diagnosis of brain death.

A total of 121 EEG samples were used, consisting of 42 favorable and 79 unfavorable outcomes, with sampling frequencies ranging from 100Hz to 600Hz. Within each record, EEG signal experts segmented ten significant 2-second segments deemed crucial for prognosis, referred to as **Gold-standard EEG segment selection**. Additionally, the study explored an automated strategy, randomly selecting non-overlapping segments from the original EEG data, termed **Random EEG segment selection**. This dual approach aimed to evaluate whether deep learning models are capable of extracting discriminative patterns that are comparable to, or potentially more informative than, those identified by experts, even when models were fit on data that is presumably not significant for the outcome [Costa et al. 2022, de Paiva et al. 2018].

3.2. Implementation Details

This study utilized Python as the programming language. Key libraries included NumPy for numerical operations, Pandas for data manipulation, SciPy for scientific computing, and Scikit-learn for machine learning algorithms. For deep learning models, Braindecode was employed, providing robust tools for building and training neural networks.

3.3. Pre-processing

In this study, two approaches for pre-processing EEG data were investigated. The first approach utilized 9 out of 10 segments of EEG traces, previously separated by specialists. The second approach involved randomly selecting 9 segments from the original exam trace, ensuring no data overlap. This process divided the raw EEG into 9 equal segments, followed by the selection of a 2-second window with a randomly chosen starting point. If overlap occurred, a new point was selected until the non-overlapping condition was met. In both cases, the samples underwent resampling to standardize the dataset sampling frequency at 100Hz and 200Hz.

The implementation, facilitated by the MNE library [Gramfort et al. 2013], applied a low-pass filter before point selection or interpolation for downsampling and upsampling, respectively. The data distribution revealed approximately 2 unfavorable outcomes for every favorable outcome, indicating a class imbalance issue. To address this, the FTSurrogate [Caza-Szoka and Massicotte 2022] data augmentation technique was used, similar to [Al-Hussaini and Mitchell 2023]. This technique was applied to each electrode's EEG traces, using standard parameters for phase noise magnitude and channel independence, resulting in a balanced training set with equal numbers of favorable and unfavorable outcome samples. Notice that FTSurrogate is applied exclusively over the training set, as the validation set remains separated.

3.4. CNN Models

This study utilized different approaches of CNN models to classify EEG signals and predict coma outcomes. The architecture was chosen for its ability to effectively process temporal and spatial data inherent in EEG signals. The models implemented included Deep ConvNet and Shallow ConvNet architectures, which were selected for their varying depths and complexities to capture a broad range of features from the EEG data.

Deep ConvNet, proposed by [Schirrmeister et al. 2017] is a comprehensive architecture designed to efficiently extract features from EEG signals through deep layers. The first block of the Deep ConvNet incorporates the FBCSP (Filter Bank Common Spatial Patterns) technique. This stage employs a convolutional layer to perform convolution over the temporal channels of the EEG signal. The filters generated are then passed to another convolutional layer that extracts spatial filters capable of discriminating the signal characteristics. The resulting features are processed by a Max Pooling layer. Subsequent blocks operate similarly to traditional convolutional networks, with an increasing number

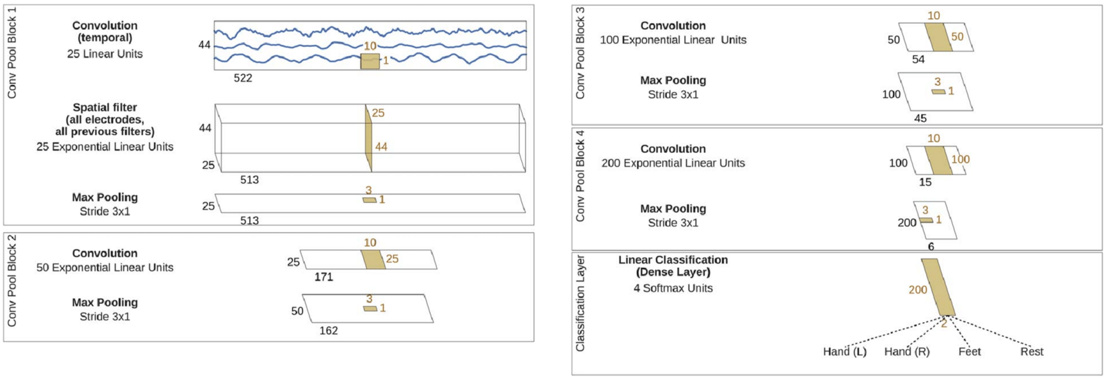

Figure 2. Deep ConvNet Architecture [Schirrmeister et al. 2017].

of filters at each stage to capture low-level local features initially and evolving to high-level global features as the network depth increases. The final layer is densely connected, aimed at performing the final classification of the EEG record.

Shallow ConvNet, proposed by [Schirrmeister et al. 2017] is a simpler alternative to the Deep ConvNet that does not include the three convolutional blocks that precede the densely connected layer. Instead, Shallow ConvNet adopts the Average Pooling to summarize the block of temporal and spatial filter extraction. This architecture is designed to capture more localized patterns with fewer layers, making it computationally less intensive compared to its deeper counterpart.

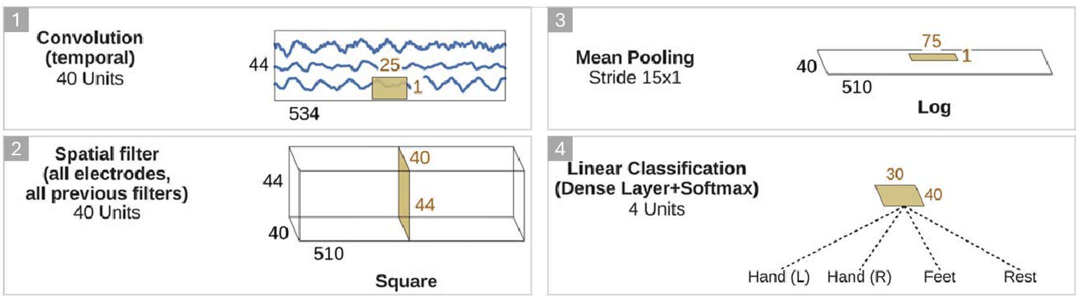

Figure 3. Shallow ConvNet Architecture [Schirrmeister et al. 2017].

Both architectures were implemented using the Braindecode library [Schirrmeister et al. 2017], which is optimized for EEG deep learning applications. Braindecode's integration with PyTorch allowed development and training of these models. The library also contributes in the application of standard preprocessing steps and data augmentation techniques, ensuring that the models could generalize well to new data.

3.5. Hyperparameter Optimization

Unlike model weights, hyperparameters require fine-tuning before training, a computationally intensive process due to numerous possible combinations [Yang and Shami 2020]. The parameters studied in this work were stored as key-value pairs in YAML text files, with configurations determined through manual experimentation.

These files were processed by a Python pipeline running in an environment integrated with Optuna [Yu and Zhu 2020], a versatile framework designed for integration and optimization across various Python-based machine learning libraries [Akiba et al. 2019]. A MySQL database was utilized for external data storage and management of the results of optimization experiments.

The optimization parameters for various architectures were systematically defined and evaluated. Table 1 lists the hyperparameters considered for optimizing the Deep ConvNet architecture.

Table 1. Parameters considered for optimizing the Deep ConvNet architecture.

Architecture	Parameter	Values
	n_filters_spat	8, 12, 16, 20, 24, 28, 32
	n_filters_time	8, 12, 16, 20, 24, 28, 32
	n_filters_2	20, 24, 28, 32, 36
	n_filters_3	42, 46, 50, 54, 58
Deep ConvNet	n_filters_4	62, 66, 70, 74, 78
	drop_prob	0.1, 0.2, 0.5
	batch_size	16, 24, 32, 40, 48, 56, 64
	learning_rate	low: 0.0001; 0.0001; 0.0001
		high: 0.099; 0.099; 0.099
		step: 0.0001; 0.001; 0.01

For the Shallow ConvNet, Table 2 details the hyperparameter setup, focusing on parameters like *pool_time_stride*, *filter_time_length*, and *pool_time_length*. These settings allowed a broader search due to the architecture's lower computational cost compared to Deep ConvNet.

Table 2. Parameters considered for optimizing the Shallow ConvNet architecture.

Architecture	Parameter	Values
	n_filters_spat	8, 12, 16, 20, 24, 28, 32
	n_filters_time	8, 12, 16, 20, 24, 28, 32
	filter_time_length	15, 20, 25, 30, 35, 40, 45
		50, 55, 60
	pool_time_stride	10, 15, 20
Shallow ConvNet	pool_time_length	40, 50, 60, 70, 80, 90, 100
	drop_prob	0.1, 0.2, 0.5
	batch_size	16, 24, 32, 40, 48, 56, 64
	learning_rate	low: 0.0001; 0.0001; 0.0001
		high: 0.099; 0.099; 0.099
		step: 0.0001; 0.001; 0.01

3.6. Performance Evaluation

For a more reliable evaluation of the models' performance, the stratified K-Fold cross-validation technique was used. This approach was chosen to preserve the class imbalance during model evaluation, a common condition in healthcare-related problems. After partitioning into folds, the training set underwent data augmentation using the FTSurrogate technique.

The final patient prognosis was determined by the mode of the classifications across all its corresponding EEG segments. Model performance was ranked primarily based on sensitivity and F1-score, alongside accuracy and specificity. This emphasis aimed to identify models that effectively detect patients with a higher probability of coma recovery while ensuring balanced classification of both outcomes. Finally, the mean and standard deviation of all metrics were computed, stored, and sent to the optimization tool for extracting new values for further experimentation.

4. Results and Discussion

The results obtained in this study are presented in this section and consider the different convolutional neural network architectures applied to the dataset. Each architecture is specified with the sampling frequency (Samp.Freq) considered during the dataset preprocessing (100Hz or 200Hz). The results are presented as percentages, including the mean and standard deviation, in relation to various performance metrics, including accuracy, specificity, sensitivity, and F1-Score. The findings are detailed for both expert-selected (gold standard) segments and randomly selected segments, highlighting the performance differences across these preprocessing techniques.

4.1. Results with Expert-Selected Segments

Table 3 presents the classification results of EEG trace segments selected by experts, considering sampling frequencies of 100Hz and 200Hz. The *Shallow ConvNet* architecture demonstrated consistent performance across both sampling frequencies. However, the *Deep ConvNet* architecture showed lower performance at 100Hz but improved significantly at 200Hz.

Table 3. Results of CNN models using gold standard segments (selected by experts), with mean and standard deviation.

Architecture	Samp.Freq	Accuracy	Specificity	Sensitivity	F1-Score
Shallow ConvNet	100Hz	71.15 ±08.93	70.00 ±16.01	**74.50 ±18.90**	**69.08 ±09.64**
Shallow ConvNet	200Hz	71.79 ±14.15	72.50 ±21.51	71.00 ±13.93	68.96 ±13.04
Deep ConvNet	100Hz	64.49 ±09.01	66.25 ±14.84	62.50 ±13.28	62.45 ±08.74
Deep ConvNet	200Hz	**71.92 ±12.94**	**74.82 ±20.05**	67.50 ±16.62	67.56 ±13.24

The *Shallow ConvNet* architecture demonstrated solid performance, achieving an F1-Score of 69.08% and a sensitivity of 74.5% at 100Hz, and an F1-Score of 68.96% and a sensitivity of 71% at 200Hz. This indicates its robustness and stability across different sampling frequencies. On the other hand, the *Deep ConvNet* architecture exhibited lower performance at 100Hz, with an F1-Score of 62.45% and sensitivity of 62.50%. However, it improved significantly at 200Hz, achieving an accuracy of 71.92%, specificity of 74.82%, and an F1-Score of 67.56%, demonstrating its potential when higher temporal resolution data is available.

The *Shallow ConvNet* consistently outperformed *Deep ConvNet* at 100Hz, achieving the highest mean F1-Score (69.08%) and sensitivity (74.50%). This suggests that *Shallow ConvNet* may be more effective for lower sampling frequencies, maintaining both high accuracy and stability in performance metrics.

At 200Hz, *Deep ConvNet* showed a considerable improvement, achieving the highest accuracy (71.92%) and specificity (74.82%). This indicates that the increased dimensionality provided by the higher sampling frequency was better leveraged by the deeper architecture of *Deep ConvNet*. Nevertheless, the *Shallow ConvNet* still maintained competitive performance, with all key metrics (accuracy, specificity, sensitivity, and F1-Score) above 70%.

These findings highlight the significant impact of sampling frequency on the performance of these architectures. The *Shallow ConvNet* demonstrated strong performance and stability across both frequencies, while *Deep ConvNet* benefited from higher temporal resolution, suggesting its suitability for applications where higher sampling rates are feasible. Overall, the results underscore the importance of considering both model architecture and data resolution in EEG signal classification for coma prognosis.

4.2. Results with Randomly Selected Segments

Table 4 presents the results for randomly selected segments, which may represent a simple and fast alternative for segment selection in some EEG applications. At a sampling frequency of 200Hz, the *Shallow ConvNet* demonstrated the best overall performance, achieving the highest mean accuracy (72.76%), mean sensitivity (76.50%), and F1-Score (69.03%). This indicates a strong ability to correctly identify true positive patients. Additionally, *Shallow ConvNet* maintained good specificity at 71.07%.

Table 4. Results of models using randomly selected segments, with mean and standard deviation.

Architecture	EEG-Coma	Accuracy	Specificity	Sensitivity	F1-Score
Shallow ConvNet	100Hz	71.15 ±16.13	**71.25 ±25.65**	71.50 ±14.15	67.24 ±15.25
Shallow ConvNet	200Hz	**72.76 ±09.03**	71.07 ±15.69	**76.50 ±24.19**	**69.03 ±14.34**
Deep ConvNet	100Hz	67.76 ±10.84	70.00 ±26.93	66.00 ±31.13	55.68 ±17.46
Deep ConvNet	200Hz	66.92 ±11.81	66.25 ±20.19	69.00 ±19.08	63.83 ±09.84

The *Deep ConvNet* also presented satisfactory indicators when data were sampled at 200Hz, but its performance was not as strong as *Shallow ConvNet*. It achieved a mean accuracy of 66.92%, specificity of 66.25%, and an F1-Score of 63.83%, demonstrating that while it performed adequately, it was outperformed by *Shallow ConvNet* in this configuration.

At a sampling frequency of 100Hz, *Deep ConvNet* exhibited the lowest mean F1-Score (55.68%) and mean sensitivity (66.00%), indicating difficulty in correctly identifying true positive and true negative samples. This poorer performance could be attributed to the high standard deviation of 31.13% for sensitivity and 26.93% for specificity, highlighting instability in the patterns learned within each K-Fold iteration.

In contrast, the *Shallow ConvNet* maintained solid performance at 100Hz with a mean accuracy of 71.15%, specificity of 71.25%, sensitivity of 71.50%, and an F1-Score of 67.24%. These results underscore the robustness of *Shallow ConvNet* across different sampling frequencies and highlight its effectiveness in learning from randomly selected EEG segments.

Overall, these findings suggest that the *Shallow ConvNet* achieved better overall performance when exposed to randomly selected EEG traces compared to *Deep ConvNet*. The simpler structure and fewer parameters of *Shallow ConvNet* likely contributed to its superior pattern learning capabilities across various preprocessing techniques and sampling frequencies, as evidenced by its consistent performance metrics.

4.3. Discussion of Experimental Results

Based on the analysis of the results in Tables 3 and 4, for expert-selected segments, the *Shallow ConvNet* achieved the most consistent and superior results in terms of mean accuracy, specificity, sensitivity, and F1-Score when analyzing all architectures and sampling frequencies. This architecture demonstrated the highest stability among the expert-selected segments, showing good consistency in mean accuracy and F1-Score across each K-Fold iteration. Overall, *Shallow ConvNet* performed best for both segment selection approaches, particularly excelling when the data was sampled at 200Hz and randomly selected. This superior performance can be attributed to the higher amount of captured information, greater temporal resolution, and the architecture's adaptability to the problem domain.

The *Deep ConvNet* architecture showed the highest mean accuracy (71.92%) and specificity (74.82%) among all expert-selected segment experiments. However, it exhibited the lowest mean F1-Score (55.68%) when segments were randomly selected and sampled at 100Hz. For expert-selected segments, higher temporal resolution (200Hz) improved the results of *Deep ConvNet*, highlighting that its complex and deeper structure benefits from finer-grained data.

When considering all results, *Shallow ConvNet* at 200Hz with random selection showed the best overall performance. This emphasizes the potential of these architectures to uncover hidden patterns in the data, which might be overlooked by experts. On the other hand, *Deep ConvNet* struggled to consistently learn patterns that discriminated between classes across all tested sampling frequencies and random selections. The lowest mean F1-Score of 55.68% and the highest standard deviation for sensitivity (31.13%) were observed in this architecture when exposed to randomly selected data at 100Hz. This suggests that in deeper architectures gold standard selection may contribute significantly to identifying more relevant, less noisy, and reliable data, leading to more accurate machine learning model results.

The application of CNN architectures combined with densely connected layers for learning patterns in EEG signals was also explored by [Bissaro 2021], adding a new dimension to model the electrode positioning on the patient's scalp, with relationships modeled based on averages between neighboring electrodes. For this work, temporal and sequential modeling was performed using the FBCSP technique, which enabled pattern learning in different significant bandwidths for better class discrimination. Although [Bissaro 2021] simulations were not reproduced, the *Shallow ConvNet* model at 200Hz demonstrated superior overall results compared to the standard CNN architecture with attention.

In general, models showed a standard deviation close to or above 10% across all evaluation metrics, with the highest standard deviation observed in the sensitivity and specificity of the *Deep ConvNet-100Hz* architecture with random segments. Several factors may contribute to this issue, including the inherent complexity of EEG data, as noted by [Bissaro et al. 2023] in their work. [Delorme 2023] and [Coelli et al. 2024] explored the application of automated pipelines for preparing EEG trace signals. This type of automated approach, applied to all records in the dataset, could be explored in future work to potentially improve model performance and reliability.

5. Conclusion

This study investigated the use of convolutional neural networks to classify EEG signals and predict coma outcomes from gold standard and random EEG segment selections. The *Shallow ConvNet* consistently outperformed the *Deep ConvNet* across various sampling frequencies and segment selection methods. Particularly, the *Shallow ConvNet* excelled at a 200Hz sampling frequency with random segment selection, showcasing its robustness and adaptability to different data preprocessing techniques.

The results demonstrated the impact of sampling frequency and segment selection strategy. The *Shallow ConvNet* maintained high accuracy, specificity, sensitivity, and F1-Score across both 100Hz and 200Hz frequencies. In contrast, the *Deep ConvNet* performed better at higher temporal resolutions (200Hz) but struggled with lower sampling rates and random segment selections, resulting in lower F1-Scores and higher variability in performance metrics.

These findings emphasize the need to consider both model architecture and data resolution in EEG signal classification for coma prognosis. The *Shallow ConvNet*'s simpler structure and fewer parameters contributed to its superior pattern learning capabilities, making it a valuable tool for clinical applications. Moreover, this architecture performed very well with our fast and simple random selection of EEG segments, which achieved competitive results when compared to the use of gold standard segments selected by experts, which are time-consuming and expensive.

To address class imbalance, we employed the FTSurrogate technique. However, the use of randomly selected EEG segments, although presumably not significant for prognosis, may help deep learning architectures identify and learn relevant patterns to coma outcome prediction, as this strategy shifts part of the interpretative burden from human experts to the model itself, allowing it to effectively filter noise and uncover patterns that may be overlooked by experts. This approach can be explored in future work using various length and stride configurations, as it offers the potential to extract a greater number of segments from the positive outcome.

Despite the promising results, the study had several limitations, including the inherent variability in EEG data, which poses challenges to model generalization. The random selection strategy aimed to replicate the segment length and quantity used in the gold standard approach but was not extensively explored across different window lengths and stride configurations. Additionally, the size of the EEG-Coma dataset constrains the capacity of the explored deep learning architectures and hinders experimentation with an environment closer to real-world clinical applications. Future research should explore larger and more diverse datasets, advanced hybrid models combining CNNs with techniques like LSTM or attention mechanisms, and automated pre-processing pipelines to enhance model reliability and generalization.

Acknowledgements

Authors thank the financial support given by Brazilian National Council for Scientific and Technological Development - CNPq (grant number 420212/2023-0). We also thank NVIDIA Corporation by the donation of a Titan V GPU used in this research.

References

Aellen, F. M., Alnes, S. L., Loosli, F., Rossetti, A. O., Zubler, F., De Lucia, M., and Tzovara, A. (2023). Auditory stimulation and deep learning predict awakening from coma after cardiac arrest. *Brain*, 146(2):778–788.

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631.

Al-Hussaini, I. and Mitchell, C. S. (2023). Seizft: interpretable machine learning for seizure detection using wearables. *Bioengineering*, 10(8):918.

Baldo Júnior, S., Carneiro, M. G., Destro-Filho, J.-B., Zhao, L., and Tinos, R. (2023). Classification of coma etiology using convolutional neural networks and long-short term memory networks. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Bissaro, L. Z. (2021). Aprendizado de padrões eeg para prognóstico precoce de pacientes em coma usando redes echo state e redes neurais convolucionais. Master's thesis, Universidade Federal de Uberlândia, Uberlândia.

Bissaro, L. Z., Junior, O. O. N., Destro Filho, J. B., Jin, Y., and Carneiro, M. G. (2023). Towards the prognosis of patients in coma using echo state networks for eeg analysis. *Procedia Computer Science*, 222:509–518.

Carneiro, M. G., Ramos, C. D., Destro-Filho, J.-B., Zhu, Y.-t., Ji, D., and Zhao, L. (2023). High-level classification for eeg analysis. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Caza-Szoka, M. and Massicotte, D. (2022). Windowing compensation in fourier based surrogate analysis and application to eeg signal classification. *IEEE Transactions on Instrumentation and Measurement*, 71:1–11.

Coelli, S., Calcagno, A., Cassani, C. M., Temporiti, F., Reali, P., Gatti, R., Galli, M., and Bianchi, A. M. (2024). Selecting methods for a modular eeg pre-processing pipeline: An objective comparison. *Biomedical Signal Processing and Control*, 90:105830.

Costa, P. G. d. et al. (2022). Base de dados em eletroencefalografia (eeg).

de Paiva, M. R., Gobbo Jr, M., Rufim, E. S., Campos, M., and Destro Filho, J. B. (2018). Avaliação visual dos padrões eletroencefalográficos de pacientes clinicamente em coma. *Brazilian Journal of Health Review*, 1(2):447–455.

Delorme, A. (2023). Eeg is better left alone. *Scientific reports*, 13(1):2372.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., et al. (2013). Meg and eeg data analysis with mne-python. *Frontiers in neuroscience*, 7:70133.

Hossain, K. M., Islam, M. A., Hossain, S., Nijholt, A., and Ahad, M. A. R. (2023). Status of deep learning for eeg-based brain–computer interface applications. *Frontiers in computational neuroscience*, 16:1006763.

İnce, R., Adanır, S. S., and Sevmez, F. (2021). The inventor of electroencephalography (eeg): Hans berger (1873–1941). *Child's Nervous System*, 37:2723–2724.

Iwama, S., Morishige, M., Kodama, M., Takahashi, Y., Hirose, R., and Ushiba, J. (2023). High-density scalp electroencephalogram dataset during sensorimotor rhythm-based brain-computer interfacing. *Scientific Data*, 10(1):385.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.

Mathew, A., Amudha, P., and Sivakumari, S. (2021). Deep learning techniques: An overview. *Advanced Machine Learning Technologies and Applications*, page 599.

Ramos, C. D. et al. (2022). Aprendizado de máquina como ferramenta para o prognóstico de pacientes em coma usando sinais eletroencefalográficos no espectro de 1 a 100 hz.

Saeidi, M., Karwowski, W., Farahani, F. V., Fiok, K., Taiar, R., Hancock, P. A., and Al-Juaid, A. (2021). Neural decoding of eeg signals with machine learning: A systematic review. *Brain Sciences*, 11(11):1525.

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., and Ball, T. (2017). Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, 38(11):5391–5420.

Xu, M., Ouyang, Y., and Yuan, Z. (2023). Deep learning aided neuroimaging and brain regulation. *Sensors*, 23(11):4993.

Yang, L. and Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415:295–316.

Yu, T. and Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. *arXiv preprint arXiv:2003.05689*.