Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Enhancing Cognitive Signal Processing: Advanced CNN Architectures for EEG-Inferred Digit Recognition

Xiongzhewen Zhang

School of Engineering, Mathematics and Technology, University of Bristol, Bristol, UK

dy23520@bristol.ac.uk

Abstract. The EEG, or Electroencephalogram, is a non-invasive tool that records electrical activity along the scalp, offering a window into the complex workings of the brain. By analyzing these signals, researchers can gain insights into cognitive processes such as attention, memory, and decision-making. This investigation utilizes Convolutional Neural Networks (CNN) for the classification of EEG signals into numerical digits, tackling the challenge of deciphering cognitive states via non-invasive techniques. The core of the study is the utilization of extensive datasets and sophisticated CNN models to evaluate the performance of consumer-grade EEG headsets in brain activity interpretation. The results showcase high accuracy in numerical cognition identification, demonstrating the robustness of the methods used and suggesting their wider applicability for cognitive state analysis. Overall, the research conducted in the field of EEG-based diagnostics represents a significant milestone in the cognitive technologies industry. Its contributions to the field are numerous and far-reaching, setting a new benchmark for subsequent investigations into this exciting domain. As we continue to explore the potential of EEG technology, we can look forward to a future where cognitive technologies become increasingly personalized, accurate, and effective.

Keywords: Cognitive Technology, EEG Pattern Analysis, Machine Learning, Neural Computation, Signal Decoding.

1. Introduction

The essence of this inquiry is to assess the ability of convolutional neural networks to discern and classify individual digits through the analysis of elec-troencephalography (EEG) patterns. This research delves into the viability of convolutional neural networks (CNNs) as a tool for accurately capturing the specific numeric thoughts reflected in EEG data.

This exploration is situated within a field that has seen rapid advances in EEG analysis, driven by deep learning. The body of literature reveals an emphasis on high-resolution, clinical-grade EEG devices in visual brain decoding, often leaving the capabilities of consumer-grade EEG technologies relatively uncharted. The present study mines this rich vein of inquiry, drawing upon a compendium of scholarly work that spans across disciplines and methodologies.

Digit recognition via EEG presents not just a technical challenge but a window into the synchronization of perception and cognition. Studies like Mishra et al. and Khaleghi et al. have laid the groundwork, demonstrating the potential of deep learning in capturing the subtleties of EEG data [1, 2].

However, they often require the fidelity of medical- grade equipment, a barrier to widespread application. The research problem here is clear: there is a need for robust CNN models that can operate with the variable quality of data that consumer-grade EEG headsets provide.

A critical literature review identifies a crucial gap: a dearth of models optimized for the less consistent, yet increasingly popular, consumer-grade EEG datasets. The current study strives to fill this gap, proposing an enhanced CNN architecture designed to extract and classify neural patterns indicative of digit contemplation.

The contribution of this research is twofold. Theoretically, it extends the boundaries of visual brain decoding, positing a refined algorithmic approach to the interpretation of EEG signals. Practically, it offers a step towards the democratization of cognitive state detection, potentially enabling widespread, non-invasive cognitive health monitoring. This study not only theorizes but also empirically validates the proposed methods through extensive experiments and analysis, as delineated in the forthcoming chapters. The later sections delve into the intricacies of the methodologies applied, discuss the implications of the findings in real-world scenarios, and provide a comprehensive review of the technology's integration prospects. Ultimately, this research bridges the gap between theoretical innovation and practical utility, contributing to the advent of a new era in cognitive technology.

2. Related Work

The quest to harness the power of electroen- cephalography (EEG) for applications beyond seizure detection has long been a focus of computational neuroscience. Groundbreaking works have established frameworks for analyzing neural responses to emotional states [3, 4], and cognitive loads [5], while also venturing into the realms of sleep pattern analysis and emotion recognition using combined feature sets [5, 6]. The utilization of deep convolutional neural networks (CNNs) has notably advanced the classification and visualization of EEG data [7, 8].

Visual cognition, decoded through EEG, provides a unique vantage point from which to observe the cognitive processes elicited by stimuli. Studies have pioneered methods for analyzing the intricate neural responses to natural movies and semantic content [9-11], laying the groundwork for the development of innovative brain-computer interfaces that blend vision and neural activity interpretation [12, 13]. The emerging field has also begun exploring the visual system's interpretation of handwritten digits, using datasets like MNIST as a standard benchmark [14].

Despite the potential of high-fidelity EEG devices, their accessibility is limited, prompting a shift to- wards more readily available consumer-grade tech- nology [15, 16]. This shift has underscored the need for advanced algorithms capable of contending with the variable quality of data these devices provide. In response, research has begun to optimize CNNs for this data, aiming to democratize the field of cognitive state monitoring and widen the application of non- invasive EEG technologies [17-19].

The present study seeks to contribute to this burgeoning field by proposing an enhanced CNN architecture for the robust classification of EEG signals. By advancing the theoretical foundations of visual brain decoding and addressing the practical challenges posed by consumer-grade EEG data, this work strives to bridge the gap between cutting-edge research and practical, real-world applications.

3. Materials

The MindBigData, an open-access dataset for per- ceptual brain decoding, serves as the foundation for this research. It consists of EEG signals acquired by presenting subjects with visual stimuli of numerical digits (see Figure 1 for examples of stimuli). These digits were displayed on a screen, and subjects were instructed to focus on the image of each digit for two seconds, during which EEG data were recorded.

The dataset was compiled from recordings obtained using four different commercial EEG devices. One such device is the Emotiv EPOC, whose electrode placements are depicted in Figure 2. This diversity in data capture equipment enriches the dataset, providing a range of signal qualities and characteristics conducive to a comprehensive analysis.

The EEG data central to this study were sourced from the MindBigData, a publicly accessible dataset known for its application in perceptual brain decoding. This dataset comprises EEG recordings captured by instructing subjects to focus on visually presented numerical digits, fostering the generation of EEG signals. Each digit was projected for two seconds to ensure consistent data capture across trials.

The MindBigData was compiled using four different commercial EEG devices, including the Emotiv EPOC. The Emotiv EPOC, featured in Table 1, was equipped with a 14-channel electrode cap designed according to the international 10/20 system, a standardized method for the placement of EEG electrodes. For instance, channel names like O1 and O2 represent the occipital lobe typically associated with visual processing. The channel configuration of the Emotiv EPOC device is detailed in Table 1, which provides the corresponding channel names for both hemispheres of the brain, as well as their descriptive full names, facilitating a comprehensive understanding of the data acquisition layout. This configuration enables the recording of electrical activity across different regions of the scalp. The frequency and duration of the data acquisition, alongside the dataset's overall volume, are critical parameters. Typically, EEG data are sampled at a rate that sufficiently captures the frequency spectrum of brain activity, often ranging from below 1 Hz to above 40 Hz, depending on the study's focus. For comprehensive analyses, the datasets can encompass thousands of individual recordings, each spanning multiple seconds or even minutes, thus providing a rich temporal resolution for cognitive state assessment. The specific details such as the exact sampling rate, the length of each recording session, and the total size of the dataset should be detailed in the methodology section of the associated research.

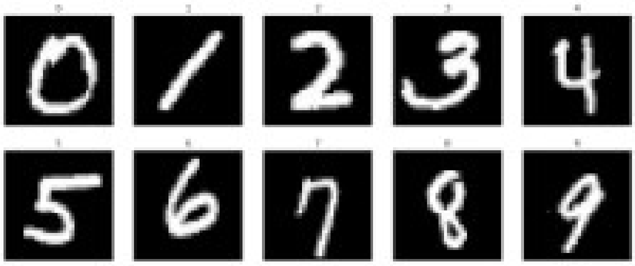

Fig. 1 Samples of MNIST based visual stimuli [2]

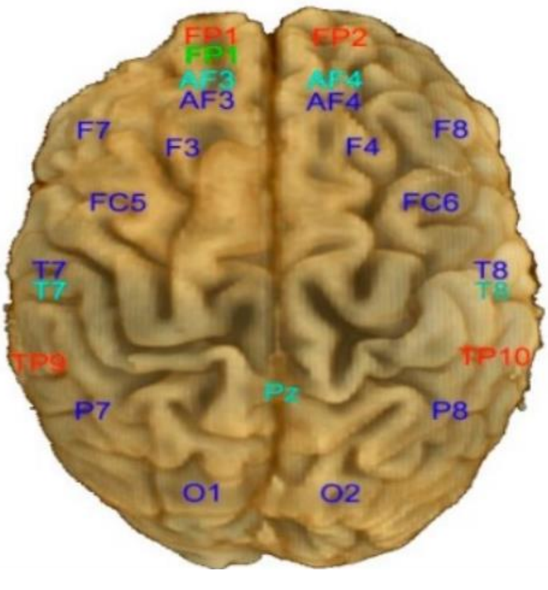

Fig. 2 Electrode locations in Emotiv Epoc [2].

Table 1. Channel names of emotiv EPOC

Channel Namein Left Hemisphere	Channel Namein Right Hemisphere	Channel FullName
O1	O2	Occipital
P7	P8	Parietal
T7	T8	Temporal
FC5	FC6	FrontoCentral
F7	F8	Frontal
F3	F4	Frontal
AF3	AF4	Between Prefrontal and Frontal
FP1	FP2	PreFrontal

4. Methods

4.1. Data refinement

4.1.1. Data resampling

To achieve a consistent dataset, a resampling algorithm was applied. The chosen method, linear interpolation, is known for preserving the original signal's characteristics while. adjusting the data to a uniform sampling rate. This technique is particularly well-suited for the EEG dataset, which exhibits variability in sampling rates due to the different devices used for data collection.

Linear interpolation was executed programmatically through a custom-defined function, systematically adjusting the data to a median length that rep- resents the aggregate dataset effectively. This median length was computed based on the distribution of sampling rates across all EEG data arrays, providing a statistically sound basis for resampling.

Post-resampling, a verification step was included to confirm the integrity of the resampled data. This was performed by assessing the length of arrays post-interpolation, ensuring they conformed to the predefined median length. Additional checks were implemented to detect any anomalies introduced during the resampling process.

4.1.2. Time-Frequency representation

The transformation of EEG data into the time-frequency domain is a pivotal step in signal processing for neural activity analysis. The time-frequency representations (TFRs) are obtained by applying wavelet transformations, which afford a dual view of both time and frequency localizations of the brain's electrical activity [20]. For a given signal x(t), the TFR is calculated using the following formula:

$$TFR_x(t,f) = \int_{-\infty}^{+\infty} x(\tau)\phi_{t,f}^*(\tau)d\tau = \langle x, \phi_{t,f} \rangle. \quad (1)$$

where $\phi_{t,f}$ denotes the complex conjugate of the basis function generated by the wavelet transform. This process results in a comprehensive depiction of the EEG signal's power distribution across various frequencies over time.

Visualization of Time-Frequency Data: The application of time-frequency analysis transforms the EEG signals into a more interpretable format, bridging the gap between complex signal characteristics and visual comprehension. Figure 3 displays exemplary time-frequency plots of digit and non-digit related neural activations, offering a direct illustration of the subtle distinctions captured by the analytical methods employed.

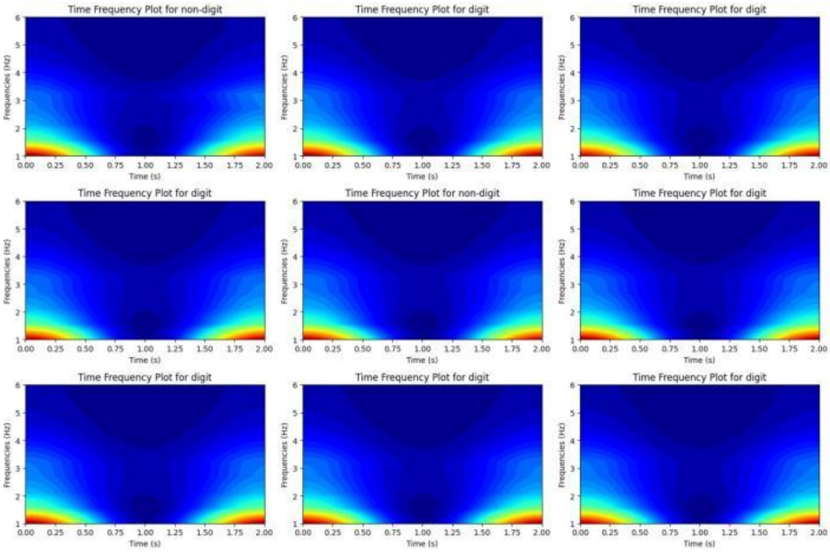

Fig. 3 Time-frequency plots for digit EEG signals (Photo credited: Original)

4.1.3. Wavelet transformation

Wavelet Transformation plays an essential role in the time-frequency representation of EEG signals. By distributing the energy of signal into frequency bands, this method aids in mitigating the spreading effect inherent to wavelets and preserving time resolution. The real-world signal can be effectively reconstructed as an amplitude- modulated (AM) and frequency-modulated (FM) component. For a real-world signal x(t), the trans- form is given by:

$$x(t) = \sum_{p=1}^{N} A_p(t) \cos\left(2\pi\phi_p(t)\right). \quad (2)$$

where $A_p(t)$ represents the amplitude and $\phi_p(t)$ denotes the phase of the signal. This methodology underpins the extraction of features relevant for CNN classification.

The visualizations in Figure 3 offer a clear depiction of the differences in energy concentration across frequencies when comparing digit and non- digit stimuli.

4.2. Convolutional Neural Networks (CNN) for Classification

In the domain of EEG signal analysis for digit recognition, CNNs serve as a robust framework for feature extraction and classification. Inspired by the neurophysiological insights of Hubel and Wiesel on the visual cortex's hierarchical structure, CNNs mimic this layered processing to identify patterns within complex datasets.

4.2.1. Network Architecture

The architecture of the CNN developed in this study consists of several layers, each designed to process the input data incrementally. The initial convolutional layers are responsible for detecting primitive patterns, such as edges and textures, which are then compounded in subsequent layers to recognize more complex features. Pooling layers intersperse these convolutional layers to re- duce dimensionality and computational load, thereby refining the feature maps.

In our CNN architecture, the process of normalization is crucial to align the data distribution, which is achieved through the batch normalization technique. This process is mathematically defined by the following equations [21]:

$$\mu_B = \frac{1}{N} \sum_{i=1}^{N} y_i. \quad (3)$$

$$\sigma_B^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \mu_B)^2. \tag{4}$$

$$y_i' = \frac{y_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}. \tag{5}$$

As illustrated in Equation 3, the mean of the batch, μ_B, is computed, followed by the calculation of the variance, σ_B^2, as shown in Equation 4. The scaling and shifting operation, which normalizes the activations of the network, is depicted in Equation 5. This normalization allows the model to learn more effectively, as demonstrated by improved training speed and overall network performance.

4.2.2. Training and Optimization

The network is trained using a backpropagation algorithm with stochastic gradient descent to minimize the classification error. Techniques such as dropout and batch normalization are employed to prevent overfitting and ensure that the model generalizes well to new,unseen data. The result is a CNN that can discern subtle EEG signal patterns associated with different cognitive states elicited by digit visualization tasks.

5. Results & Analysis

5.1. Validation Accuracy

The validation phase is critical for assessing the performance of the convolutional neural network model. The proposed model underwent rigorous evaluation, with the results demonstrating a progressive decrease in both training and validation loss, indicative of successful learning. Notably, after 15 epochs, the model achieved a validation accuracy of 89%, a testament to its robustness and the effectiveness of the underlying architecture and training regimen (see Table 2).

Table 2. Training and validation result

Epoch	Train Loss	Valid Loss	Accuracy	Time
0	1.288688	1.838545	0.130000	00:13
1	1.003203	3.228627	0.110000	00:17
2	0.877196	4.085651	0.105000	00:17
3	0.799995	2.289006	0.150000	00:16
4	0.747495	0.521838	0.775000	00:16
5	0.694782	0.824765	0.625000	00:16
6	0.629689	0.371633	0.895000	00:17
7	0.546464	0.519388	0.875000	00:16
8	0.481661	0.457510	0.885000	00:16
9	0.415673	0.415429	0.865000	00:16
10	0.357682	0.411603	0.865000	00:16
11	0.308640	0.464006	0.845000	00:16
12	0.228537	0.380549	0.895000	00:16
13	0.204205	0.384293	0.895000	00:16
14	0.177784	0.380249	0.895000	00:16

These outcomes reinforce the capability of deep learning models, specifically CNNs, to discern intricate patterns within EEG signals pertaining to cognitive tasks related to digit recognition.

5.2. Visualization

5.2.1. Training outcomes

The training phase of the convolutional neural network model, as depicted in Figure 4, culminated in a noteworthy performance. The accuracy progression over epochs demonstrates the model's capacity to learn and adapt to the complexity of the EEG-based digit recognition task. Initially, the accuracy oscillated, indicating the model's adjustment to the features within the EEG data. However, as training progressed, a steady increase in accuracy was observed, with minimal overfitting, evidenced by the convergence of training and validation losses.

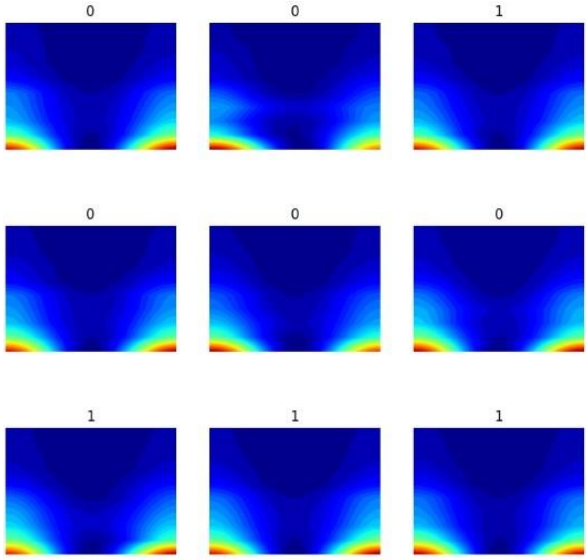

Fig. 4 Training phase performance with epoch-wise accuracy and loss metrics. (Photo credited: Original)

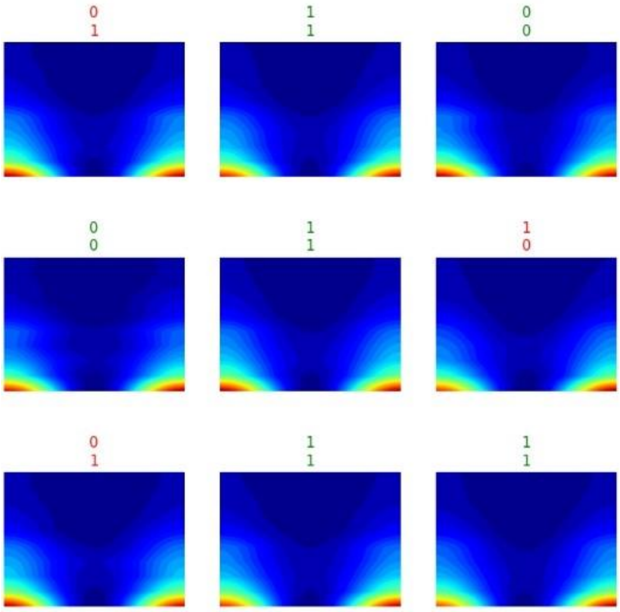

Fig. 5 Validation phase performance showcasing individual sample accuracy (Photo credited: Original)

5.2.2. Validation outcomes

Subsequent to the train- ing, the validation phase, illustrated in Figure 5, was executed to assess the model's generalization on unseen data. The binary indicators above each plot correspond to the model's prediction accuracy on individual validation samples, with '1' indicating a correct prediction and '0' an incorrect one. These visualizations affirm the model's adeptness at differentiating between digit-related neural patterns and those unrelated to numerical cognition, aligning with the objectives set forth at the commencement of this study.

5.2.3. Neural patterns

Figures 4 and 5 provide visual representations of the model's interpretative capabilities. Each time-frequency plot captures the distinct neural signatures elicited during the cognitive process of recognizing digits, which were used by the CNN for classification. The color gradients in these plots—ranging from cool to warm tones—reveal the intensity and localization of brain activity over time, offering a granular view of the EEG signal's power spectrum.

Inclusion of such visual aids in the analysis not only enhances the interpretability of the results but also validates the efficacy of the pre-processing and CNN architecture employed in this research. The discernible patterns highlighted in these visualizations form the crux of our model's classification proficiency, thereby underscoring the potential of deep learning in EEG signal analysis.

5.3. Test Accuracy

The evaluation of the model's performance was ex- tended beyond the validation set to include a distinct test dataset, comprising 1000 separate data points that were neither part of the training nor validation sets. This phase is critical as it provides a robust measure of the model's predictive power in real- world scenarios. The results in Figure 6 were striking, with the model attaining a 92% accuracy rate on this test set, indicating a high level of reliability and efficiency in the EEG-based digit recognition task.

The high test accuracy not only underscores the model's effectiveness but also signifies the potential of the CNN architecture and the training regimen to generalize well to unseen data. Such an outcome illustrates the promise of the proposed methodologies for practical applications in cognitive state detection and, more broadly, in the realm of brain-computer interfaces.

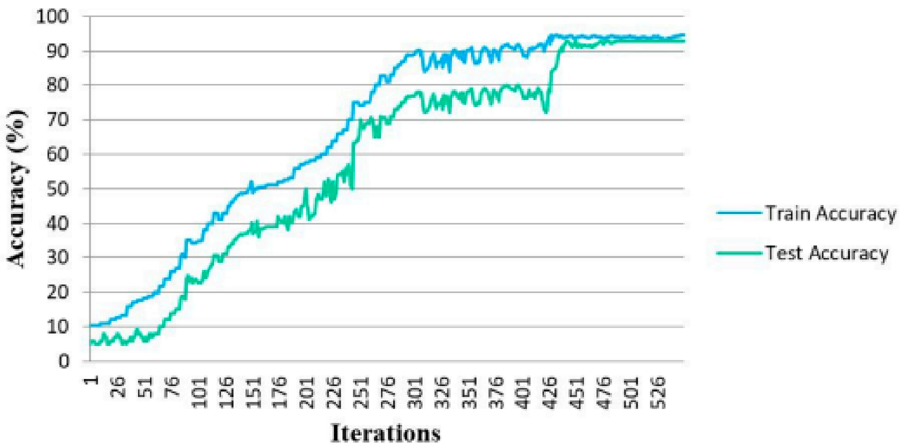

Fig. 6 Test accuracy of this model (Photo credited: Original)

6. Conclusion

This research delineates the utility of convolutional neural networks for the discernment of EEG patterns, correlating specific neuronal activities to digit recognition tasks. Through meticulous model training on a curated subset of the comprehensive dataset, we elicited an appreciable predictive precision, which intimates the robustness of our preliminary model. Foreseen advancements include augmentation of the data corpus, exploration of intricate neural architectures, and an extensive parameter sweep across the frequency domain to optimize the spectral- temporal features. Furthermore, methodological refinements—eschewing averaging in favor of indi-vidual channel analysis and adopting a granular approach to time-frequency plot aggregation—may substantially bolster the fidelity of the neural decoding process. In conclusion, the inquiry into the realm of cognitive neurotechnology highlights the profound potential of non-invasive methodologies in neurological diagnostics and therapeutic interventions. These technologies not only overcome the limitations of traditional invasive techniques but also offer new possibilities for more precise and targeted interventions. As we continue to explore and develop these technologies, it is likely that we will witness a significant transformation in the way we understand and treat neurological conditions.

References

[1] R. Mishra, K. Sharma, and A. Bhavsar, "Visual brain decoding for short duration eeg signals," EUSIPCO, 2021.

[2] N. Khaleghi, S. Hashemi, S. Z. Ardabili, S. Sheykhivand, and S. Danishvar, "Salient arithmetic data extraction from brain activity via an improved deep network," Sensors, 2023, 23:9351.

[3] G. Di Flumeri, G. Borghini, P. Arico, N. Sciaraffa, P. Lanzi, S. Pozzi, V. Vignali, C. Lantieri, A. Bichicchi, A. Simone, et al., "Eeg-based mental workload neurometric to evalu- ate the impact of different traffic and road conditions in real driving settings," Frontiers in Human Neuroscience, 2018, 12, 509.

[4] Y. Gao, H. J. Lee, and R. M. Mehmood, "Deep learning of eeg signals for emotion recognition," in 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2015, 1–5.

[5] J. Chen, P. Zhang, Z. Mao, Y. Huang, D. Jiang, and Y. Zhang, "Accurate eeg-based emotion recognition on combined features using deep convolutional neural net- works," IEEE Access, 2019, 7, 44317–44328.

[6] H. Ghimatgar, K. Kazemi, M. S. Helfroush, and A. Aarabi, "An automatic single-channel eeg-based sleep stage scor- ing method based on hidden markov model," Journal of Neuroscience Methods, 2019, 108320.

[7] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," Hu- man Brain Mapping, 2017, 38(11), 5391–5420.

[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] A. G. Huth, W. A. De Heer, T. L. Griffiths, F. E. Theunis- sen, and J. L. Gallant, "Natural speech reveals the semantic maps that tile human cerebral cortex," Nature, 2016, 532(7600), 453–458.

[10] A. G. Huth, T. Lee, S. Nishimoto, N. Y. Bilenko, A. T. Vu, and J. L. Gallant, "Decoding the semantic content of natural movies from human brain activity," Frontiers in Systems Neuroscience, 2016, 10, 81.

[11] S. Nishimoto, A. T. Vu, T. Naselaris, Y. Benjamini, B. Yu, and J. L. Gallant, "Reconstructing visual experiences from brain activity evoked by natural movies," Current Biology, 2011, 21(19), 1641–1646.

[12] A. Kapoor, P. Shenoy, and D. Tan, "Combining brain computer interfaces with vision for object categorization," in 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008, 1–8.

[13] N. Bigdely-Shamlo, A. Vankov, R. R. Ramirez, and S. Makeig, "Brain activity-based image classification from rapid serial visual presentation," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2008, 16(5), 432–441.

[14] L. Deng, "The mnist database of handwritten digit images for machine learning research [best of the web]," IEEE Signal Processing Magazine, 2012, 29(6), 141–142.

[15] T. Raduntz, "Signal quality evaluation of emerging eeg devices," Frontiers in Physiology, 2018, 9, 98.

[16] K. Stytsenko, E. Jablonskis, and C. Prahm, "Evaluation of consumer eeg device emotiv epoc," in MEi:CogSci Conference 2011, (Ljubljana), 2011.

[17] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learn- ing for electroencephalogram (eeg) classification tasks: a review," Journal of Neural Engineering, 2019, 16(3), 031001.

[18] J. J. Bird, D. R. Faria, L. J. Manso, A. Ekart, and C. D. Buckingham, "A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction," Complexity, 2019, 2019.

[19] B. L. K. Jolly, P. Aggrawal, S. S. Nath, V. Gupta, M. S. Grover, and R. R. Shah, "Universal eeg encoder for learning diverse intelligent tasks," in 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), 2019, 213– 218.

[20] E. Sejdic´, I. Djurovic´, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," Digital Signal Processing, 2009, 19(1), 153–183.

[21] S. Russell, I. S. Moskowitz, and B. Jalaian, "Context: Separating the forest and the trees—wavelet contextual conditioning." 2023.