Accessible EEG Classification with Attention-Based Neural Networks

Mercy Doan
Queen's University
mercy.doan@queensu.ca

Leopold Ehrlich
Queen's University
leopold.ehrlich@queensu.ca

Daryan Fadavi
Queen's University
21dkf1@queensu.ca

Arya Farivar
Queen's University
arya.farivar@queensu.ca

Colin McLaughlin
Queen's University
mclaughlin.colin@queensu.ca

***Abstract*—To mitigate problems with noisy electroencephalogram (EEG) data and financially inaccessible medical-grade EEG devices, we present 2 NLP-inspired attention-based neural networks to improve classification accuracy, tested on 3 unique datasets. View our code here.**

I. Introduction

An electroencephalogram (EEG) is a device commonly used for medical purposes. By placing electrodes on a subject's head in specific areas, we can record their brain activity separated into channels from the different electrical signals. Medical uses of EEGs include diagnosis of epilepsy [1], diagnosis of parasomnias [2], and determination of cerebral death [3]. In the field of artificial intelligence, EEG data is often used in classification tasks, such as emotion recognition [4]. While there exist many different EEG devices, from consumer-friendly devices with 4 to 8 channels, to medical and research grade devices with 64 channels and more, classifying these signals into meaningful insights is a task that does not have a 'best' solution yet. Our paper explores ways to improve classification accuracy by testing 2 different models on 3 different datasets.

A. Motivation

Electroencephalography (EEG) serves as a pivotal tool in neuroscience, allowing for the non-invasive monitoring of brain activity for both clinical and research applications. Traditional high-density EEG systems, equipped with numerous channels, provide detailed resolution but are often accompanied by significant financial and logistical constraints. The cost of these professional-grade EEG systems can range from approximately $1,000 to over $25,000, depending on the number of electrodes and additional features [5]. This cost poses a barrier for many researchers and clinicians operating under limited budgets.

In contrast, low-cost, portable EEG devices with fewer channels have emerged as accessible alternatives. Such systems offer a balance between affordability and functionality, making EEG technology more accessible to a broader range of users. Furthermore, studies have shown that 8-channel EEG setups can reliably detect expected neural patterns. For instance, an exploration of different EEG configurations revealed that the 8-channel setup was reliable in detecting expected trends, with 100% reliability in certain measures [6]. This finding shows the potential of 8-channel systems to provide meaningful data.

The primary motivation for this project is to develop an accessible and effective model for classifying motor imagery using 8-channel EEG data. This has significant implications for assistive technologies, particularly for individuals with disabilities such as locked-in syndrome, who rely on brain-computer interfaces (BCIs) for communication and interaction with their environment. By developing reliable classification of intentions through an affordable and accessible EEG setup, this project seeks to empower disabled individuals, enhancing their autonomy and quality of life.

B. Related Works

Zhang et al. [7] proposed two deep learning models—Cascade and Parallel Convolutional Recurrent Neural Networks (CRNNs)—to enhance EEG-based intention recognition. The cascade model applies a 2D-CNN for spatial feature extraction, followed by an LSTM for temporal dynamics, while the parallel model processes spatial and temporal features simultaneously before fusion. Their approach mitigates the need for extensive preprocessing by learning directly from raw EEG data, achieving an accuracy of 98.3% in cross-subject validation and 93% in a real-world BCI system. Despite its robustness, the study highlights challenges related to EEG noise and inter-subject variability.

EEG data is generally contaminated with voltage sources other than neuronal action potentials due to heavy amplification and low signal to noise ratios. The various sources of noise are well studied. Muscle and eye movements both cause electrical dipoles that can be transmitted to the sensors [8]–[10]. Power line interference is a primary source of 60 or 50 hz noise [11], [12]. Thermal artifacts and the slow accumulation of sweat can contribute to low frequency noise [12]. Furthermore, small shifts in the electrode position, unstable contact, and the half-cell effect are all sources of noise that can arise from sensors [12].

Data preprocessing is crucial to remove noise from the signal. Xu et al. [10] proposed a preprocessing framework that removes artifacts while preserving desired frequency ranges. Their method combines adaptive filtering and statistical analysis to retain relevant signal components for downstream processing. This approach showed signs of enhanced signal clarity, though it discusses the challenge of distinguishing between low-amplitude brain signals and artifacts.

Similarly, Sweeney et al. [13] provided a comprehensive review of artifact removal techniques in EEG signal processing. The study compared methods such as Independent Component Analysis (ICA), wavelet decomposition, and regression-based techniques, while discussing the trade-offs between computational complexity and artifact removal efficacy.

In the area of classification, recent work by Lee et al. [14] utilized an autoencoder for feature extraction combined with a ResNet architecture featuring a double augmented attention mechanism for ADHD classification from EEG data. This approach enhanced the model's ability to focus on informative signal segments, achieving high classification accuracy. However, the study also noted the challenge of generalizing across diverse subject data and the need for robust augmentation techniques to mitigate overfitting.

Overall, these studies emphasize the importance of both effective data processing and advanced model architectures in improving EEG-based classification performance. Both of these aspects will be crucial in addressing remaining challenges such as noise variability, artifact distinction, and subject generalization.

C. Problem Definition

Our research aimed to tackle common EEG problems, such as noisy data and cross-subject accuracy, by drawing inspiration from NLP. We segmented our project into the following research goals:

1) Will incorporating NLP techniques like attention into biomedical data classification improve a model's accuracy by focusing on important temporal features?
2) Can we mitigate problems in low resolution EEG data by 'filling in gaps' with masked autoencoding and using existing high resolution datasets?
3) Can we replicate professional research results using a beginner EEG device?

Our team was in the possession of an OpenBCI 8-channel EEG. OpenBCI is an initiative that promotes accessible EEG technology, selling affordable EEG sensors, headsets, and circuits. Additional challenges we explored throughout our design process included:

1) Recording our own dataset with an OpenBCI EEG to test the contrasts of a low resolution dataset vs. a high resolution dataset
2) Investigating whether different brain activities (mental/emotional vs. physical/motor) could have similar classification accuracy with the same model

Overall, we wanted to improve classification accuracy for EEG data using NLP-inspired techniques, which could then be applied to a wide range of functions, such as controlling wheelchairs with one's mind or allowing consumers to benefit from low-resolution EEG devices to the same degree as if they could afford a high-resolution device.

II. Methodology

Our work consisted of analyzing 3 datasets, each with unique properties and one of which we recorded ourselves. We then developed 2 classification models and tested them with all 3 datasets. The metadata for each is presented in Table I.

TABLE I: Metadata for Emotion Recognition, Motor Imagery, and OpenBCI Motor Imagery Datasets

	ER	**MI**	**MI OpenBCI**
# electrodes	14	64	8
Subjects	28	109	6
Trials per subject	4	14	5x30 or 3x30
Trials total	112	1500+	720
# classes	4	3 or 4	3
EEG	Emotiv Epoc+	BCI2000	OpenBCI
Trial duration	5 mins	2 mins	5 secs

A. Data Collection

Our first dataset was an Emotion Recognition (ER) dataset [4], where 28 subjects would play video games. They used the Emotiv Epoc+ device with 14 channels. Their brain activity was recorded for 5 minutes during the gameplay of boring, calm, horror, and funny video games, and the dataset was created to classify emotional states.

The second dataset was a Motor Imagery (MI) dataset [15]. Motor imagery classification with EEGs typically consists of subjects making some physical movement or imagining making a physical movement. In this case, 109 participants were asked to either open and close their right, left, or both fists, or imagine doing so. The dataset includes data from 64 EEG channels for over 1500 recordings, each either one or two minutes in length.

The third dataset was a replication of the MI dataset with our EEG, an OpenBCI Cyton board with 8 channels. The placement of the electrodes are shown in Figure 1. These were chosen as left and right hand MI brain activity is generated from the C3 and C4 areas [9].

For each subject, a video was generated with randomized prompts. These were either:

1) Text only
2) Audio only
3) Both text and audio

Examples of the visual prompts are shown in Figure 2.

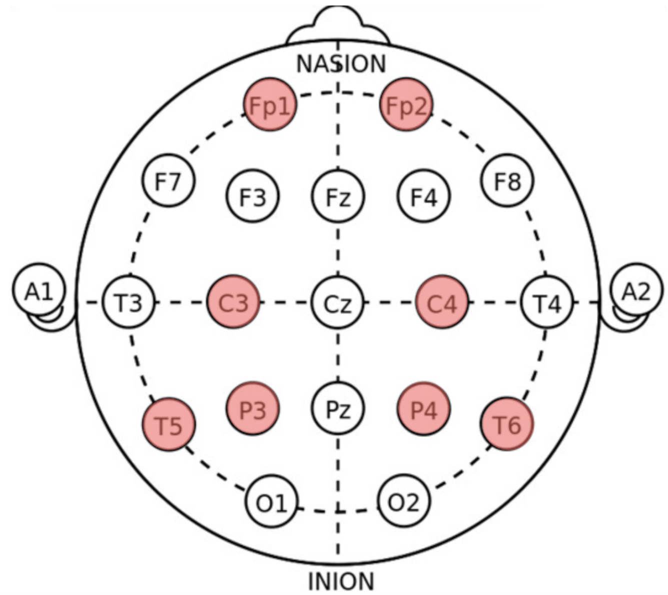

Fig. 1: Electrode placement on OpenBCI EEG for data collection

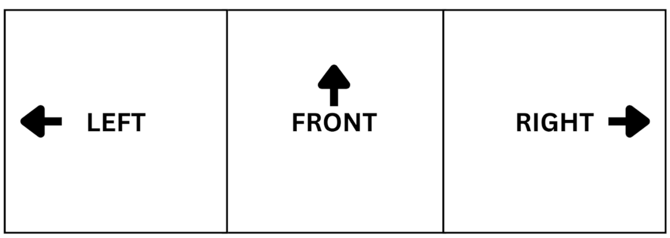

Fig. 2: Visual prompts for subjects in data collection

The audio prompts had a guitar strum sound in either the right, left, or both ears. We recorded the brain activity of 6 subjects - 2 male and 4 female - between 18-22 years old. There were 5 runs, each with 30 five-second prompts. In order, they were:

1) Physically opening/closing fists with audio and text prompt
2) Imagining opening/closing fists with text only prompt
3) Imagining direction with audio and text prompt
4) Imagining direction with audio only prompt
5) Imagining direction with text only prompt

If the subject did not have earbuds, only trials 1, 2, and 5 were played. Before each run, we started recording so that we could ensure the EEG device was accurately responding to blinks. When each trial started, we attempted to do a large movement, usually a loud clap, to create a spike in the data and see when a trial started.

We chose to replicate the MI dataset since it was simple to set up, and we predicted that our 8-channel EEG would be more responsive to a physical task than a mental task. We were also curious to see if the resulting 8-channel readings would resemble the 64-channel readings, and if so, whether we could use the higher resolution data to predict the lower resolution classification.

B. Data Processing

The OpenBCI MI dataset that we obtained was first processed to crop out the non-experimental recorded numbers, using accelerometer data to indicate the start of the tests.

The data that we collected required processing before use. A band pass filter filter on the 0.1 hz to 30 hz interval was applied. ICA was employed to remove artifacts. To train the ICA, the data was copied, then processed to allow better component extraction. A high-pass filter was applied at 2hz as proposed in [16], and the outlier epoch rejection algorithm presented in [17] was employed to allow stronger ICA results. Algorithms from the MNE library were employed to identify artefactual ICA components, isolating muscle artifacts and eye blink artifacts [18]. Since an inexpensive electroencephalogram was used, no EOG channels were available. As such the Fp1 and Fp2 channels were used as analogues. From there, the original data could be processed with this ICA, leaving out the identified artefactual components.

C. Model Creation

The first model uses a convolutional neural network (CNN) with a masked autoencoder (MAE) to process the time-series EEG data and classify target labels. The MAE model architecture is inspired by the work of Pulver et al. [19] and shown in Figure 3.

The data was collected from multiple subjects and pre-processed before training. Missing values were forward-filled to maintain continuity, and each feature was normalized to have a mean of zero and a standard deviation of one. To capture temporal dependencies, the data was segmented into overlapping windows of 100 time steps with a step size of 50, ensuring that each window served as an independent training sample while preserving the sequential nature of the EEG data.

The overall model architecture consists of two main components: a masked autoencoder for feature extraction and a CNN for classification. The autoencoder applies random masking to 25% of the input data before passing it through a convolutional encoder with convolutional layers, max pooling layers, and a dense layer to encode any meaningful feature representations. A decoder reconstructs the original input using transposed convolutional layers and a final convolutional layer with sigmoid activation. The autoencoder is trained using mean squared error (MSE) loss and the Adam optimizer. After pretraining, the encoder is used in the CNN classifier, which consists of a fully connected layer with ReLU activation, a dropout layer (0.3 probability) to reduce overfitting, and a final softmax layer that outputs class probabilities. The classifier is trained with categorical cross-entropy loss and the Adam optimizer.

A leave-one-subject-out (LOSO) cross-validation strategy is used to evaluate the generalization performance of the model.

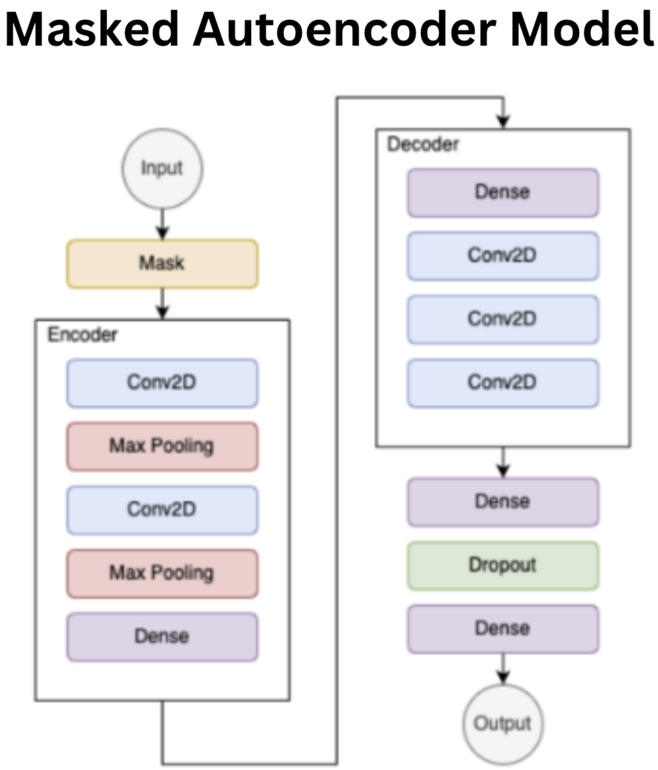

Fig. 3: Masked Autoencoder Model Architecture

In each iteration, one subject is left out for testing, while the model is trained on the remaining subjects. Training is carried out for 10 epochs with a batch size of 32, and validation is carried out on the left-out subject.

The second models were created based on adding attention layers to deep neural networks. The architecture is shown in Figure 4. The models were evaluated using LOSO cross validation and holdout validation.

For the ER dataset, we implemented a sliding window, with a size of 100 time steps and a step size of 50, to segment the continuous recordings. After preprocessing and filtering, windows were labeled based on their corresponding game.

For the classification model, a CNN-LSTM architecture with a custom attention layer was developed. Its key components include:

1) Convolutional layers to extract local temporal features
2) Batch normalization and max pooling to stabilize and downsample the activations
3) LSTM layers to capture sequential dependencies
4) A custom attention mechanism to focus on the most informative time steps
5) Dense layers culminating in a softmax output for four-class emotion classification

We chose a slightly different approach for the MI dataset, opting for an EEGNet-inspired architecture, which was then augmented with a transformer-based attention mechanism. The model includes:

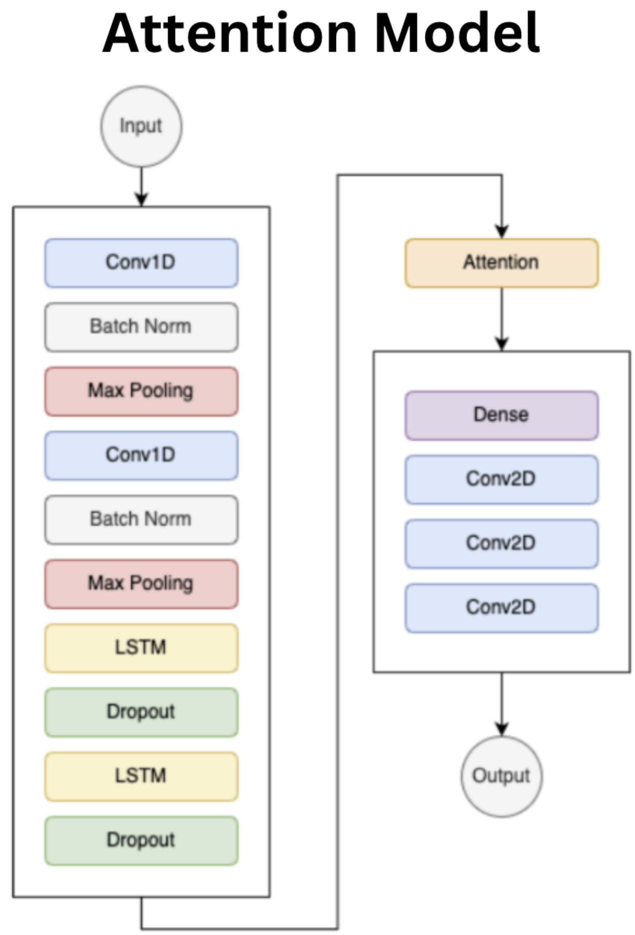

Fig. 4: Attention Model Architecture

1) A temporal convolutional layer to capture time-dependent features
2) A depthwise convolution block for spatial filtering
3) A separable convolutional layer to combine temporal features efficiently
4) An adaptive average pooling layer
5) A transformer encoder layer to emphasize the most important features via attention
6) A final fully connected layer for binary classification.

III. Results

Both models were evaluated using three datasets: Emotion Recognition, Motor Imagery, and OpenBCI Motor Imagery, using a leave-one-subject-out (LOSO) cross-validation approach. Performance was assessed using accuracy, precision, recall, F1-score, and loss to evaluate the model's classification capabilities. The CNN model with a masked autoencoder (MAE) performed very well on the Emotion Recognition dataset, achieving high classification accuracy and balanced precision and recall scores, indicating its effectiveness in classifying emotional states. In contrast, performance on the OpenBCI Motor Imagery dataset was lower, likely due to increased noise and variability in the EEG signals, as this dataset was collected independently rather than from an external source; the model's ability to generalize was impacted by inconsistencies in signal quality, making classification more challenging. These results highlight the strengths of the CNN with MAE approach while also identifying challenges associated with working with noisier, independently collected EEG data. Table 2 shows the exact metric scores of the CNN model, with tables 4 and 5 as accuracy matrices.

TABLE II: CNN + MAE Model Performance Metrics

Dataset	Accuracy	Precision	Recall	F1-score	Loss
Emotion Recognition	0.972	0.978	0.972	0.968	0.243
Motor Imagery	0.647	0.419	0.647	0.508	0.649
OpenBCI Motor Imagery	0.334	0.113	0.336	0.169	1.099

TABLE III: MAE model accuracy matrix for the Emotion Recognition dataset.

Accuracy Matrix for Emotion Recognition Dataset				
	Calm	Boring	Funny	Horror
Calm	0.954	0.034	0.011	0.001
Boring	0.001	0.995	0.004	0.000
Funny	0.004	0.009	0.982	0.004
Horror	0.003	0.009	0.033	0.955

TABLE IV: MAE model accuracy matrix for the Motor Imagery dataset.

Accuracy Matrix for Motor Imagery Dataset		
	Left Hand (T1)	Right Hand (T2)
Left Hand (T1)	1.00	0.00
Right Hand (T2)	1.00	0.00

TABLE V: MAE model accuracy matrix for the OpenBCI Motor Imagery dataset.

Accuracy Matrix for Motor Imagery Dataset			
	Right	Left	Front
Right	0.202	0.126	0.672
Left	0.210	0.125	0.665
Front	0.216	0.125	0.660

IV. Conclusion

We achieved high accuracy without the LOSO protocol, indicating that our models generally work in a typical machine learning pipeline. In contrast, we found that our models were not generalizable, as evidenced by the decrease in classification accuracy while performing LOSO experiments. In addition, our models performed better on the emotion recognition dataset. Although this may be due to high quality datasets, it could also be a sign that our models are more suited to mental tasks like emotion classification. Finally, although we aimed to use these models to improve classification accuracy for low resolution datasets, our own data was not well suited to our models as indicated by the low accuracies. However, this may simply be due to problems in the initial recording of the data itself.

A. Discussion

Key limitations of this work include narrow demographics for data collection, electrode placement inconsistencies. Trial participants were recruited from the Queen's University undergraduate student body, and thus over represent associated demographics compared to the general populace. Our headset was a rigid 3D printed "one size fits all" model, which did not uniformly fit each trail participant. Therefore, channels do not perfectly correspond to their intended locations.

B. Future Work

While this study showed the potential of EEG to classify motor imagery, there are still several areas for future exploration. Originally, one of the project's main goals was to demonstrate the viability of cheaper 8-channel EEG devices for classification tasks. While our device showed some ability to record viable data, there were also numerous limitations of the hardware that could be addressed in the future. During the data recording process, the observed signals were sometimes unexpected and did not match the behaviour of the subject, or were simply extremely noisy. Some common calibration methods were attempted to address these issues, but to little effect. If more time were allotted to the project, various other techniques could be used to make the meaningful data more visible. Some of these include individual channel calibration, ensuring proper grounding, and mitigating electrical interference.

Furthermore, different deep learning architectures could be explored to improve feature extraction in the model. For instance, a Graph Neural Network (GNN) could be used to better capture the spatial and temporal patterns in the EEG data, thus improving classification accuracy. Alongside different architectures, data augmentation techniques could be used to reduce the impact of having a low-resolution EEG. Generative Adversarial Networks (GANs) would allow for the creation of synthetic EEG data which could lead to a more robust model performance.

Overall, although our research goal of achieving accurate classification with a beginner EEG did not perform better than random sampling, we still created models which were successful on mid and high resolution data as per our other research goals. With more resources, we would re-evaluate our data collection process to gain clearer data, and improve our model's accuracies on low resolution datasets.

REFERENCES

[1] S. J. M. Smith, "Eeg in the diagnosis, classification, and management of patients with epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii2–ii7, 2005. [Online]. Available: https://jnnp.bmj.com/content/76/suppl_2/ii2

[2] M. S. Aldrich and B. Jahnke, "Diagnostic value of video-eeg polysomnography," *Neurology*, vol. 41, no. 7, pp. 1060–1060, 1991. [Online]. Available: https://www.neurology.org/doi/abs/10.1212/WNL.41.7.1060

[3] J. Praline, J. Grujic, P. Corcia, B. Lucas, C. Hommet, A. Autret, and B. de Toffol, "Emergent eeg in clinical practice," *Clinical Neurophysiology*, vol. 118, no. 10, pp. 2149–2155, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1388245707003653

[4] W. Waheed, "Emotion recognition using eeg and computer games," August 2024, retrieved March 18, 2025. [Online]. Available: https://www.kaggle.com/datasets/wajahat1064/emotion-recognition-using-eeg-and-computer-games

[5] iMotions, "The real cost of eeg systems," 2023, accessed: 2024-06-10. [Online]. Available: https://imotions.com/blog/eeg-system-cost/

[6] A. Smith and B. Jones, "Exploration of eeg channel configurations for reliable neural pattern detection," *Journal of Neuroscience Methods*, vol. 387, p. 109948, 2023.

[7] S. Zhang, L. Yao, and X. Sheng, "Cascade and parallel convolutional recurrent neural networks on eeg-based intention recognition for brain computer interface," *arXiv preprint arXiv:1708.06578*, 2018.

[8] P. Jadhav, D. Shanamugan, A. Chourasia, A. Ghole, A. Acharyya, and G. Naik, "Automated detection and correction of eye blink and muscular artefacts in eeg signal for analysis of autism spectrum disorder," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 1881–1884.

[9] M. Rashid, N. Sulaiman, P. P. Abdul Majeed, R. M. Musa, A. F. Ab Nasir, B. S. Bari, and S. Khatun, "Current status, challenges, and possible solutions of eeg-based brain-computer interface: A comprehensive review," 2020.

[10] M. Xu and X. Zhang, "Artifact removal from eeg with a preprocessing for preserving desired bands," in *Proceedings of the International Conference on Computer Application and System Modeling*. Atlantis Press, 2012, pp. 100–103.

[11] J. Mateo, E. Sánchez-Morla, and J. Santos, "A new method for removal of powerline interference in ecg and eeg recordings," *Computers Electrical Engineering*, vol. 45, pp. 235–248, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790614003152

[12] E. Habibzadeh Tonekabony Shad, M. Molinas, and T. Ytterdal, "Impedance and noise of passive and active dry eeg electrodes: A review," *IEEE Sensors Journal*, vol. 20, no. 24, pp. 14 565–14 577, 2020.

[13] K. Sweeney, T. Ward, and S. McLoone, "Removal of artifacts from eeg signals: A review," *Sensors*, vol. 12, no. 2, pp. 1481–1507, 2012.

[14] H. Lee and S. Kim, "Eeg-based adhd classification using autoencoder feature extraction and resnet with double augmented attention mechanism," *Sensors*, vol. 24, no. 4, p. 1932, 2024.

[15] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Eeg motor movement/imagery dataset," 2009. [Online]. Available: https://physionet.org/content/eegmmidb/

[16] M. Jas, D. A. Engemann, Y. Bekhti, F. Raimondo, and A. Gramfort, "Autoreject: Automated artifact rejection for meg and eeg data," *NeuroImage*, vol. 159, pp. 417–429, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917305013

[17] I. Winkler, S. Debener, K.-R. Müller, and M. Tangermann, "On the influence of high-pass filtering on ica-based artifact reduction in eeg-erp," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 4101–4105.

[18] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.

[19] D. Pulver, P. Angkan, P. Hungler, and A. Etemad, "Eeg-based cognitive load classification using feature masked autoencoding and emotion transfer learning," in *Proceedings of the 25th International Conference on Multimodal Interaction*, ser. ICMI '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 190–197. [Online]. Available: https://doi.org/10.1145/3577190.3614113