SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

A DEEP LEARNING ARCHITECTURE FOR THE DENOISING OF ELECTROENCEPHALOGRAM SIGNALS AFFECTED BY ELECTROMYOGRAPHY AND ELECTROOCULOGRAPHY NOISE

Laurea Magistrale in Biomedical Engineering - Ingegneria biomedica

Author: Annarita Barone

Advisor: Prof. Anna Maria Maddalena Bianchi

Co-advisor: Elisabetta Farella, Francesco Paissan

Academic year: 2023-2024

1. Introduction

The Electroencephalogram (EEG) is a crucial tool in both clinical diagnostics and neuroscience research, providing insights into brain function and activity. EEG signals, representing the electrical activity of neurons in the brain, are complex and can be challenging to interpret due to various artifacts. These artifacts, arising from both biological sources, such as muscle activity, (electromyography - EMG) and eye movements (electrooculography - EOG) and non-biological sources, such as electromagnetic interference, can obscure the true neural signals and compromise the accuracy of EEG data analysis. To address these challenges, researchers employ signal processing techniques like filtering and artifact removal algorithms. However, advanced techniques, including machine learning approaches, are emerging because of their ability to adapt to the complexity of EEG signals and effectively remove artifacts. Deep learning, especially convolutional neural networks, has shown promise in denoising EEG signals. These models can automatically learn relevant features from noisy EEG data without explicit feature engineering, leading to improved denoising performance. Despite these advancements, challenges persist, such as the scarcity of labeled datasets for training deep learning models. Haoming Zhang et al. aimed to address this limitation in their work EEGdenoiseNet by providing structured datasets specifically designed for EEG denoising tasks [1]. Overall, the integration of deep learning techniques into EEG denoising pipelines holds promise for enhancing the quality and reliability of neurophysiological data, paving the way for more accurate interpretations and broader applications in brain research and clinical diagnostics.

2. Methods and Materials

2.1. Artifacts Modeling

Accurate interpretation of EEG signal artifacts depends on effective modeling and classification of the artifacts, a task that involves analyzing their frequency content through spectrum analysis. This approach exploits the unique spectral characteristics of EEG, EOG, and EMG signals to develop a classification algorithm. The spectra of EEG and EOG signals exhibit a decreasing trend and can be effectively modeled using a $1/f^{\alpha}$ distribution [2] with an additional offset term for flexibility. The equation used for fitting the spectra in a logarithmic scale is:

$$PSD(f) = offset - \log\left(f^{exp}\right) \tag{1}$$

To model EMG spectra, which display both decreasing and slightly increasing trends, a more complex function is required. The equation used for fitting EMG spectra is:

$$PSD(f) = y_1 + C^{f-D} \times y_2 \tag{2}$$

where y_1 and y_2 are defined as:

$$y_1 = 10^{\text{offset}} \times \frac{1}{f^{\text{exp}}}$$
$$y_2 = af^2 + bf + c$$

Figures 1 and 2 show some examples of spectra fitting.

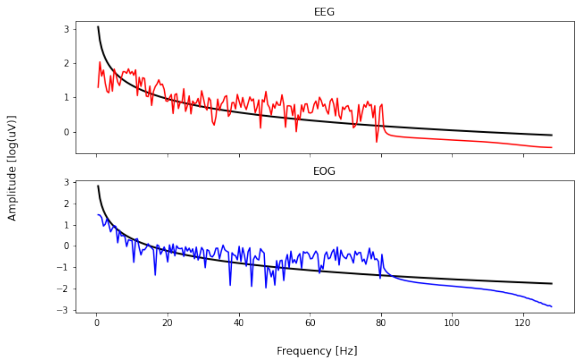

Figure 1: Examples of fitting EEG and EOG spectra using equation (1).

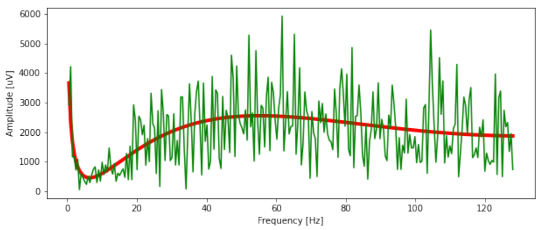

Figure 2: Example of fitting EMG spectrum using equation (2).

2.2. Artifacts Classification

A multilayer perceptron neural network was used to classify artifact types (EOG or EMG). The network architecture included nine input neurons representing parameters derived from spectrum fitting with both Function 1 and Function 2, a hidden layer with ReLU activation, and an output layer with Softmax activation for multiclassification. Data were split into training and test sets using the hold-out method. Training parameters included a batch size of 64, a learning rate of 0.2, and a Mean Squared Error (MSE) loss function, resulting in 99.8% accuracy.

2.3. Dataset Creation

A robust dataset is crucial for deep learning applications. The EEGdenoiseNet dataset used for this study consists of 4514 EEG segments, 3400 EOG segments, and 5598 EMG segments. Each segment is 2 seconds long, a duration selected for its efficacy in capturing signal characteristics without noise interference. Recorded at 512 Hz, the signals underwent filtering and notching at the power line frequency before being resampled to 256 Hz. EEG noise contamination is then performed according to the following equation:

$$y = x + \lambda \cdot n \tag{3}$$

where λ is obtained according to the ratio between Root Mean Square (RMS) of the clean signal and RMS of the noisy one:

$$SNR = 20log\frac{RMS(x)}{RMS(\lambda \cdot n)} \tag{4}$$

The fisiological range of signal-to-noise ratio (SNR) is from -7dB to 2 dB for EOG artifacts [3], and from -7 dB to 5 dB for EMG artifacts [4]. Equation 4 was mathematically corrected, ensuring accurate signal contamination, as the one proposed in the paper incorrectly used a multiplication factor of 10 instead of 20.
To further diversify the dataset and improve model performance, this study explored different methods for dataset creation like pairing each EEG signal with a single noise signal at different SNRs, as well as contaminating the same EEG signal with multiple noise signals, aiming to enhance model robustness and generalization. Additionally, online data augmentation techniques was also explored.

2.4. Development Environment

The project used a development environment featuring an NVIDIA GTX1650 4GB GPU on

Ubuntu 22.04, supported by CUDA version 11.2 for parallel processing. Python 3.7, alongside essential libraries like NumPy and Pandas for data manipulation and PyTorch 2.2.1 for deep learning, was employed. GitHub served as the version control system, facilitating collaboration and iterative improvements.

3. Experimental Design

3.1. Evaluation Metrics

The evaluation metrics employed in the EEGDenoisenet paper include three objective measures to the denoised data: the Pearson correlation coefficient (CC), the Relative Root Mean Squared Error (RRMSE) in the temporal domain, and the RRMSE in the spectral domain. CC measures the linear relationship between predicted and ground truth values, while RRMSE provides normalized measures of error in both temporal and spectral domains. The equations for CC, RRMSE in the temporal domain, and RRMSE in the spectral domain are as follows:

$$CC = \frac{Cov(f(y), x)}{\sqrt{Var(f(y))Var(x)}}$$

$$RRMSE_{\text{temporal}} = \frac{RMS(f(y) - x)}{RMS(x)}$$

$$RRMSE_{\text{spectral}} = \frac{RMS(PSD(f(y)) - PSD(x))}{RMS(PSD(x))}$$

where $f(x)$ represents the predicted signal (EEG cleaned by the network), and x represents the ground truth signal (clean EEG).
Ideally, CC should approach 1 and RRMSE both in temporal and spectral domain, should be minimized for optimal performance.

3.2. Neural Network Architectures

The neural network architecture consists of three primary models designed for denoising EEG signals: SimpleCNN, SimpleAE, and an autoencoder featuring dilated convolutions.

3.2.1 SimpleCNN: Baseline Network

The SimpleCNN model was used as a baseline architecture, replicating the structure proposed in the EEGDenoiseNet paper and the respective results. It comprises four 1D-convolution layers with 1x3 kernels and 64 feature maps each. Batch normalization and ReLU activation functions follow each convolutional layer. The final layer is a dense layer for signal reconstruction. The model employs MSE as the loss function and is trained for 10 epochs. The paper's results were faithfully reproduced using their same signal contamination formula. This ensured consistency and allowed for a reliable comparison of model performance.

3.2.2 Simple AE

The SimpleAE model is a scalable autoencoder architecture designed for signal denoising. It consists of an arbitrary number of convolutional layers, a max-pooling layer, batch normalization, and ReLU activation functions. The encoder compresses input data into a low-dimensional latent space, while the decoder reconstructs the original signal.
Unlike SimpleCNN, SimpleAE ends with a convolutional layer instead of a dense layer, allowing for flexible input and output dimensions.

3.2.3 Autoencoder with Dilated Convolution

The autoencoder with dilated convolution, proposed in the ECGDenoiseNet paper [5], incorporates dilated convolutions and residual blocks. Dilated convolutions expand the receptive field without increasing parameters, making them suitable for capturing long-range dependencies in signals. Residual blocks facilitate information flow between encoder and decoder components, enhancing reconstruction accuracy. The architecture includes skip connections to combine encoder and decoder outputs effectively, facilitating also the training process and mitigating the vanishing gradient problem associated with very deep networks.

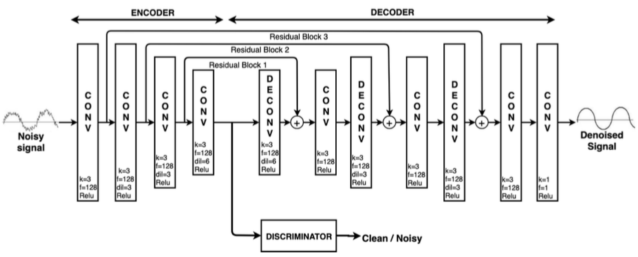

Figure 3: Autoencoder network architecture proposed by ECGDenoiseNet.

3.3. Training Strategy

The training parameters were systematically explored to optimize model performance. Feature scaling, using min-max normalization, ensured smoother convergence during training and maintained correct signal amplitudes during evaluation. Weight decay, set at 0.001, helped prevent overfitting by penalizing large weights in the model. Mean Absolute Error emerged as the preferred loss function due to its robustness to outliers and superior performance compared to MSE. A batch size of 64 was found to yield the best results, balancing computational efficiency and training effectiveness.

The original dataset included EEG and EOG signals sampled at different frequency compared to the EMG signals. To ensure compatibility and enhance temporal resolution, both the EEG and EOG signals were resampled to match the sampling frequency of the EMG signals, set at 512 Hz. Data augmentation techniques, including random and fixed contamination, were employed to expand the dataset lenght and enhance model robustness. The dataset was partitioned into training and testing subsets, with 75% of the data allocated for training and the remaining 25% for testing. Within the training set, an 80% portion was used for actual model training, while the remaining 20% served as validation data.

3.4. Testing Phase

During the testing phase, the model's performance was evaluated by contaminating the test set with different physiological SNR levels associated with each artifact type. This comprehensive assessment allowed for a thorough evaluation of the model's denoising efficacy across various levels of contamination. After contaminating the test set with each SNR value, the model was employed to denoise the signals, and metrics such as loss, correlation coefficient (cc), rrmse of the power spectral density (psd), and rrmse of the temporal (temp) signals were calculated. These metrics were then averaged across all SNR levels to derive the final results. This approach ensured a robust evaluation of the model's performance under different levels of contamination, providing valuable insights into its overall effectiveness in denoising EEG signals.

3.5. Experiments

In the experiments conducted to remove EMG and EOG artifacts from EEG signals, several methodologies and architectures were explored. Initial results were obtained from experiments conducted with a randomized dataset creation approach. This method involves the dynamic generation of the dataset during training by randomly assigning noise types (EOG or EMG) and SNR to each EEG segment. However, these experiments showed unexpected trends since increasing the dataset size did not lead to significant improvements in denoising metrics. This trend persisted across multiple experiments, indicating that the issue lay within the dataset creation methodology rather than the architecture itself. Further investigations revealed that changing the dataset during training introduced instability, leading to deteriorating performance. It was hypothesized that this instability hindered the model's ability to learn effectively, causing it to struggle with adapting to varying data distributions. As a solution, the fixed dataset creation method was adopted, where the dataset remained constant throughout the training process. Experiments conducted using the fixed dataset method showcased improved performance, especially in the autoencoder with dilated convolution architecture, particularly when combined with dropout regularization. Dropout, with a probability of 0.2 applied to all layers of the autoencoder, proved to be effective in reducing overfitting and improving the model's generalization ability. Interestingly, with the fixed dataset and dropout regularization, the model's performance improved with an increase in dataset size, aligning with the initial expectation. The main issue encountered during the experiments was the instability introduced by changing the dataset during training, even though the change in the dataset was adopted after a certain number of epochs, to provide the model with the necessary time to learn from the data, as can be seen in Figure 4. This instability hindered the model's learning process, leading to suboptimal performance. By adopting a fixed dataset creation method and incorporating dropout regularization, the model's performance improved significantly, addressing the instability issue and enabling better generalization to unseen data.

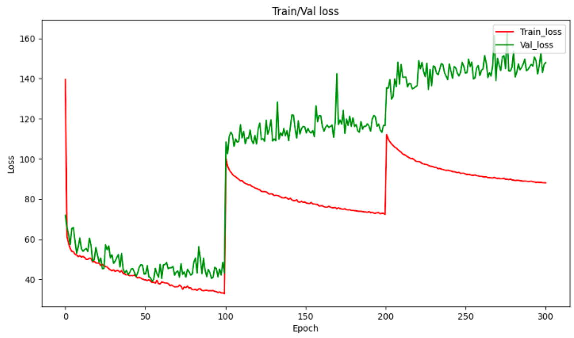

Figure 4: Training and validation losses of one experiment conducted with the autoencoder with dilated convolution architecture. As can be seen, the performance starts to decrease when the dataset is updated.

4. Results

The autoencoder with dilated convolution consistently outperformed the Simple AE architecture, demonstrating superior denoising capabilities across varying dataset lengths and noise levels. While the Simple AE architecture showed promise in mitigating EOG noise, it was eventually surpassed by the autoencoder with dilated convolution due to its computational efficiency and robustness in handling both types of artifacts. The final results underscored the effectiveness of the autoencoder with dilated convolution in addressing EMG artifacts, with stable and satisfactory performance across different dataset lengths. Conversely, denoising EOG artifacts exhibited improved performance with larger training datasets, indicating the importance of dataset size in training the network effectively. Table 1 illustrates the improvements in metrics observed throughout the entirety of the thesis work, encompassing various experiments conducted, along with their respective percentage improvements.

Table 1: Performance metrics measured at the beginning and end of experiments

Metrics	Initial Results	Final Results	Improvements (%)
CC (EMG)	0.814	0.920	13.02
RRMSE_temp (EMG)	0.578	0.338	41.52
RRMSE_psd (EMG)	0.428	0.252	41.12
CC (EOG)	0.520	0.809	55.58
RRMSE_temp (EOG)	1.145	0.662	42.18
RRMSE_psd (EOG)	0.980	0.488	50.20

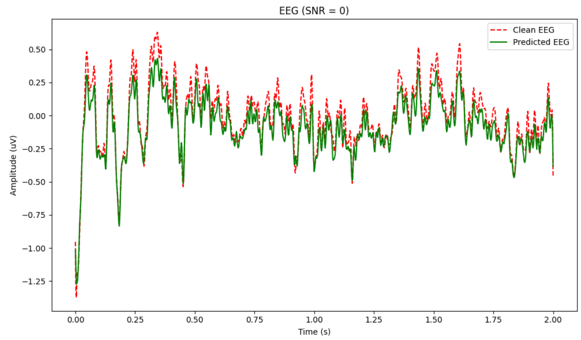

Figure 5: Example of reconstruction of ground-truth EEG signal affected by EMG noise.

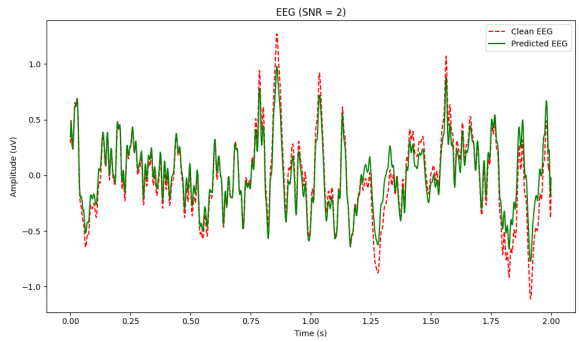

Figure 6: Example of reconstruction of ground-truth EEG signal affected by EMG noise.

4.1. Comparison with Previous Work

Direct comparison with state-of-the-art results has been prevented by two discrepancies. Firstly, there was a disparity in the physiological range considered, as it was typically limited to 2 dB for both types of artifacts in the majority of papers. While this dissimilarity could be mitigated by averaging the metrics obtained down to 2 dB, a more substantial issue emerged from the use of an incorrect formula for signal construction.

4.2. Artifacts Fitting

The study focuses on evaluating the efficacy of a neural network in removing but also identifying EMG and EOG noise from EEG signals. This evaluation involves classifying predicted noise using a neural network architecture by fitting each segment accordingly. To ensure successful fitting, a rescaling process was performed given the reduction in amplitude inherent in the construction of the noisy signal as additive noise. The denoising and classification performance of the models were evaluated using biased and unbiased estimation approaches, respectively, on the test data.

4.2.1 Biased Estimation

The denoised model's performance was assessed by classifying the noise removed by the autoencoder using a classification model trained on the same dataset. This approach measures the effectiveness of the denoised model in removing noise. Biased estimation shows consistently high accuracy for both EMG and EOG noise, with optimal performance at lower SNR levels, as can be seen in Table 2.

Table 2: Accuracy classification of estimated noise according to SNR (biased estimation).

SNR	EMG Accuracy (%)	EOG Accuracy (%)
-7	99.02	100
-5	97.07	100
-2	95.66	100
0	91.49	99.91
2	88.56	99.91
5	84.13	-

4.2.2 Unbiased Estimation

In Table 3, the results derived from unbiased estimation, in which the classification model was trained on a dataset different from the test set (i.e., the data used for assessing the denoised performance was excluded from the training set of the classification model), are presented. These results provide an authentic assessment of classification performance, since the model is tested on unseen data. In particular, the classification accuracy for EMG noise shows a slight decrease compared to the biased estimation, but remains satisfactory. Interestingly, the results for EOG noise exceed expectations, demonstrating even better performance than previously observed.

Table 3: Accuracy classification of estimated noise according to SNR (unbiased estimation).

SNR	EMG Accuracy (%)	EOG Accuracy (%)
-7	95.30	100
-5	91.76	100
-2	85.64	100
0	80.14	100
2	76.42	100
5	71.72	-

5. Conclusion

This study explored advanced deep learning techniques for EEG signal denoising, focusing on the removal of EMG and EOG artifacts. The autoencoder with dilated convolution was found to perform best, showing solid denoising capabilities. By stabilizing the dataset during training and employing dropout regularization, model performance improved significantly. Evaluation through unbiased estimation highlighted the effectiveness of the classification model in real-world scenarios. In conclusion, this thesis emphasizes the effectiveness of deep learning in automating EEG signal denoising, providing interpretable results without the need for manual intervention. Future research directions include refining models on diverse datasets, deploying them for real-time applications, and exploring synthetic data generation to enhance model robustness and advance EEG signal processing techniques. The success of these approaches not only improves EEG-based diagnostics and monitoring, but also promises to deepen understanding of brain function and activity.

References

[1] et al. Haoming Zhang, Mingqi Zhao. Eegdenoisenet: A benchmark dataset for end-to-end deep learning solutions of eeg denoising. *arxiv*, 2021.

[2] Matar Haller et al. Thomas Donoghue. Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, 23:1655–1665, 2020.

[3] Kuo Li et. al Gang Wang, Chaolin Teng. The removal of eog artifacts from eeg signals using independent component analysis and multivariate empirical mode decomposition. *IEEE Journal of Biomedical and Health Informatics*, 20(5):1301–1308, 2016.

[4] Fengqiong Yu Kai Wang Xun Chen, Hu Peng. Independent vector analysis applied to remove muscle artifacts in eeg data. *IEEE Transactions on Instrumentation and Measurement*, pages 1–10, 2017.

[5] Attila Klimmek Leslie Casas. Adversarial signal denoising with encoder-decoder networks. *arxiv*, 2020.