

Received October 20, 2021, accepted November 2, 2021, date of publication November 8, 2021, date of current version November 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125728

Automatic Electroencephalogram Artifact Removal Using Deep Convolutional Neural Networks

FÁBIO LOPES^{1,2}, ADRIANA LEAL¹, JÚLIO MEDEIROS¹, MAURO F. PINTO¹,
ANTÓNIO DOURADO¹, (Member, IEEE), MATTHIAS DÜMPELMANN²,
AND CÉSAR TEIXEIRA¹

¹Center for Informatics and Systems of the University of Coimbra, Department of Informatics Engineering, University of Coimbra, 3030-290 Coimbra, Portugal

²Epilepsy Center, Medical Center, Department of Neurosurgery, University of Freiburg, 79106 Freiburg, Germany

Corresponding author: Fábio Lopes (fadcl@dei.uc.pt)

This work was supported by the National Funds through the FCT—Foundation for Science and Technology, I.P., under Project UID/CEC/00326/2020 and Project PTDC/EEI-EEE/5788/2020. The work of Fábio Lopes was supported in part by the Portuguese Funding Institution Foundation for Science and Technology (FCT), in part by the Human Capital Operational Program (POCH), and in part by the European Union (EU) for his Ph.D. research work under Grant 2020.04537.BD.

ABSTRACT Scalp electroencephalogram (EEG) is a non-invasive measure of brain activity. It is widely used in several applications including cognitive tasks, sleep stage detection, and seizure prediction. When recorded over several hours, this signal is usually corrupted by noisy disturbances such as experimental errors, environmental interferences, and physiological artifacts. These may generate confounding factors and, therefore, lead to false results. Models able to minimise EEG artifacts are then necessary for improving further analysis and application. In this work, we developed an EEG artifact removal model based on deep convolutional neural networks. The proposed approach was applied on long-term EEGs, acquired from epileptic patients, available in the EPILEPSIAE database. The main goal of our work is to develop a model able to automatically and quickly remove artifacts from EEGs. To develop it, we used EEG segments, manually preprocessed by experts and named target EEG segments. Our approach was evaluated comparing denoised segments with the target segments. Furthermore, we compared our approach with other artifact removal models. Results show that the developed model was able to attenuate the influence of artifacts, present in long-term EEG signals, in a similar way to that performed by experts. Additionally, results evidence that our approach performs better than other artifact removal models, combining a minor reconstruction error with a fast processing. Being a fully automatic and fast model that does not require reference artifact templates, turns it suitable, for example, for continuous preprocessing of long-term electroencephalogram for sleep staging or seizure prediction.

INDEX TERMS Artifact removal, automatic reconstruction, deep convolutional neural networks, electroencephalogram, preprocessing.

I. INTRODUCTION

Electroencephalogram (EEG) is a nonlinear and nonstationary signal that measures the electrical activity of the brain [1], [2]. It is widely used in tasks involving the study of the brain dynamics such as cognitive tasks, development of epileptic seizure prediction models, and sleep stage detection.

The associate editor coordinating the review of this manuscript and approving it for publication was Ludovico Minati¹.

Brain potentials propagate over the entire scalp. Therefore, several electrodes are required to capture them with high spatial resolution [3]. Beyond brain information, these electrodes often capture noise, such as environment interference, experimental errors, and physiological artifacts [4].

Environmental interference is generated by external disturbances, e.g., main power leads and electromagnetic waves [5]. Experimental errors are usually related with poor electrode adhesion, incorrect scalp cleansing, and subject motion resulting from daily life routine. These errors, that

frequently distort the EEG signal, are quite difficult to remove, even with artifact removal approaches [4], [6]. Physiological artifacts are alterations generated from other physiological processes, such as eye movements, muscle activity (chewing, swallowing, talking, and scalp contraction), and cardiac activity. Therefore, these artifacts cannot be fully avoided even in controlled environments. Physiological artifacts generally present a spectrum overlapping the frequencies of interest of the EEG signals [5]–[10].

In general, EEG artifacts can be reduced or even avoided when signal acquisition is performed under controlled conditions. However, in tasks such as epileptic seizure prediction, EEG signals have to be continuously acquired over several days [11], [12] being practically impossible to avoid artifacts. Although a possible solution would be to detect and to remove noisy segments, this removal would result in a high loss of information. Thus, researchers have developed artifact removal techniques to eliminate, or at least attenuate, noisy data from the EEG signals while preserving neural information [4], [6].

Simple digital filtering is a highly used technique for removing undesired frequency spectrum bands from the EEG signals, e.g., power-line component. However, these filters may not be used to separate EEG from artifacts with overlapped frequency spectra, as is the case of experimental errors and physiological artifacts. For this reason, other techniques have been considered to improve EEG filtering [4], [5].

Linear regression algorithms were the most used methods for artifact removal until the 1990s due to their lower computational complexity. However, these methods present two major drawbacks: the linear behaviour does not fully adapt to the nonlinearity of the physiological processes, and a template signal is required [5], [6].

Filtering methods conduct artifact elimination by adapting filter weights to minimise the mean squared error (MSE) between the target and denoised signals. Adaptive, Wiener, and Bayes filters are examples of filtering methods which consider different optimisation techniques for achieving the minimum MSE. The major drawback of these algorithms is the requirement of *a priori* user input [4], [6].

Source decomposition methods, such as wavelets and empirical mode decomposition (EMD) approaches, aim at separating the neural information from the artifacts by decomposing each signal channel into different waveforms. However, similarly to simple digital filtering, wavelets cannot remove artifacts that overlap frequencies of interest without removing important data. Also, although the EMD is able to adapt itself to nonlinear and nonstationary signals, such as EEG, it is computationally complex and thus difficult to be used in real-time. Furthermore, both approaches require thresholds tuning to select the components of interest [4]–[6].

Linear blind source separation (BSS) methods are the most used for artifact attenuation [5], [6], [13]. These methods focus on the separation of the signals into their independent sources, by assuming that the measured signals result from the sum of the linear mixture of sources. Generally, EEG

signals are considered to be generated by independent dipolar sources that linearly mix together. Thus, linear BSS algorithms tend to perform well when separating brain signals from artifacts [13], [14]. These methods do not require any external information about the type of artifacts, making them an important solution whenever artifact template signals are not available. However, these methods require visual inspection to distinguish between brain and noisy sources. Some authors have overcome this drawback by developing classifiers to label the independent sources [15]–[21]. Although these classifiers may solve the visual inspection task, linear BSS approaches still require expensive computational time, which makes these ones difficult to be used in real-time scenarios.

Recently, authors have reported new EEG artifact removal methods, based on deep learning architectures, that aim at solving the drawbacks of the aforementioned methods [22]–[26]. Ghosh *et al.* [22] and Yang *et al.* [23] developed autoencoders (AEs), based on fully connected layers, to automatically remove ocular artifacts from EEG signals. Later on, Leite *et al.* [24], Zhang *et al.* [25] and Sun *et al.* [26] proposed models, based on deep convolutional neural networks (DCNNs), which are able to extract spatio-temporal features, and, therefore, are more robust than the traditional fully connected neural networks. Leite *et al.* developed a deep convolutional autoencoder (DCAE) for removing eye blink and jaw clenching artifacts that were previously added to clean EEG signals. Zhang *et al.* [25] proposed a DCNN, that gradually increases its width, to remove muscle artifacts from EEG signals. They reported that this architecture prevents the occurrence of overfitting. Their model was trained using a publicly available benchmark dataset [27]. Sun *et al.* [26] presented a DCNN, based on residual connections, for removing ocular, muscle and cardiac artifacts from noisy EEG signals. These signals were generated from summing clean epileptic EEG segments, from the CHB-MIT Scalp EEG Database, with electromyogram (EMG), electrocardiogram (ECG) and electrooculogram (EOG) signals from Physionet. To develop these approaches, authors require large datasets and considerable computational time. However, comparing to the current state-of-the-art BSS methods, these approaches present some main advantages: minor loss of relevant information; faster signal estimation; no need for several channels to get cleaner signals; and fully automatic output.

In summary, researchers are currently exploring the potential of deep neural networks to eliminate artifacts from EEG signals. They report that these approaches can learn the complex patterns and the high-dimensional characteristics of the EEG signals, being able to separate these ones from noisy disturbances. However, despite the high performances obtained using deep learning methods, studies were evaluated using either simulated data or data acquired under controlled environments. Therefore, these studies do not completely simulate artifact removal from realistic long-term EEG signal acquisitions.

To overcome this drawback, this article proposes an original EEG artifact removal model, based on DCNN, developed using long-term data acquired from epileptic patients, in pre-surgical monitoring. Specifically, firstly two experts removed artifacts present in these signals by visually inspecting the independent sources of the signals. Then, using these data, we developed a model that tries to automatically remove artifacts present in long-term EEG signals. The model was evaluated using root mean squared error, relative root mean squared error, Pearson correlation coefficient, and signal-to-noise ratio difference. Finally, we compared it with the 1D-ResCNN model from [26] and with an automatic ICA model based on extended Infomax ICA and MARA classifier [16].

The main goal of our approach is to develop a model able to automatically and quickly remove artifacts from long-term EEG signals without human intervention, making it suitable to be applied in real-time long-term scenarios such as epileptic seizure prediction.

The remainder of this document contains the following sections: Section II describes the dataset, the preprocessing methods performed to remove the EEG artifacts, and the development and evaluation of the proposed EEG artifact removal model; Section III presents the results and subsequent analysis; Section IV discusses the obtained results and presents the advantages and the limitations of the developed approach; Section V concludes the paper by presenting the completed objectives and directions for future work.

II. MATERIALS AND METHODS

This section presents the methods considered to prepare the dataset used in this study as well as the procedures followed to develop and evaluate our approach.

A. DATASET

EPILEPSIAE database [28], [29] contains long-term epileptic EEG signals from 275 patients, along with seizure meta-data acquired during presurgical monitoring. From these 275 datasets, 222 contain scalp EEG, 49 contain intracranial EEG, and 4 contain both types of EEG recordings. These recordings were obtained with different sampling rates, which vary from 250 Hz to 2500 Hz, over several days.

Data were acquired at Universitätsklinikum Freiburg (Germany), Centro Hospitalar e Universitário de Coimbra (Portugal) and Hôpital de la Pitié-Salpêtrière, Paris (France). The use of these data for research purposes has been authorised by the Ethical Committee of the three hospitals involved on the EPILEPSIAE database development (Ethik-Kommission der Albert-Ludwigs-Universität, Freiburg; Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé, Pitié-Salpêtrière University Hospital; and Comité de Ética do Centro Hospitalar e Universitário de Coimbra). All methods were performed following the relevant guidelines and regulations. Informed written consent was obtained from the patients.

Our dataset comprises data, acquired at Universitätsklinikum Freiburg, from 25 patients (13 males and 12 females, aged 39.6 ± 16.8 years) who suffer from temporal lobe epilepsy which is the most prevalent, difficult to treat epilepsy. These data were acquired using a sampling rate of 256 Hz and the electrodes were organised according to the 10-20 international system [3] with the following channels: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz and Pz.

This study is framed in the context of epileptic seizure prediction. To develop epileptic seizure prediction models, we considered data ranging from 4.5 hours before the beginning of the leading seizure [30] until its onset. This selection was made based on the assumption that EEG signals, within the mentioned period, contain information from both normal and pre-seizure brain states [31]–[33].

Before proceeding to the development of seizure prediction models, we believe it is crucial to remove artifacts that may be present over the long-term EEG signals. However, performing visual inspection of the ICs of the EEG signals is a tough and time-consuming task, therefore demanding for an automatic procedure to remove noise from this type of data. Based on this, we used the aforementioned data to develop EEG denoising models, which can be later used to preprocess data before applying further specific methods in any EEG-based application including epileptic seizure prediction models.

B. DATA PREPARATION

We filtered the 4.5-hour EEG signals using a 0.5-100 Hz bandpass 4th-order Butterworth filter and a 50 Hz 2nd-order notch filter, with the purpose of removing DC component, high frequency noise and powerline interference, respectively. Then, we removed noise generated by experimental errors, such as flatlines, saturated segments, and abnormal peaks. Afterwards, we divided the 4.5-hour EEG signals in 10-minute segments. Later, we identified channels with experimental errors that were not removed earlier and fixed them using spherical interpolation method [34] available in EEGLAB toolbox [35]. More details are available in Supplementary Material.

Table 1 presents the duration of both raw data and data after the described preprocessing steps (preprocessed EEG data). 35.32 hours of the initial EEG data (5.45%) were removed. Regarding interpolation steps, 18.25 hours of the preprocessed EEG data (2.98%) contain, at least, one interpolated sample.

After removing experimental errors, we re-referenced the EEG segments to average reference and processed them using extended-infomax ICA [36] available in EEGLAB. Finally, the resulting independent components (ICs) were visually inspected by experts (see Supplementary Material). The entire dataset was divided in training and test sets. The first one contains 3399 segments (486.03 hours), from 20 patients, whereas the second one includes 910 segments (126.65 hours), from the remaining 5 patients. Afterwards, two experts visually inspected the ICs of the EEG segments

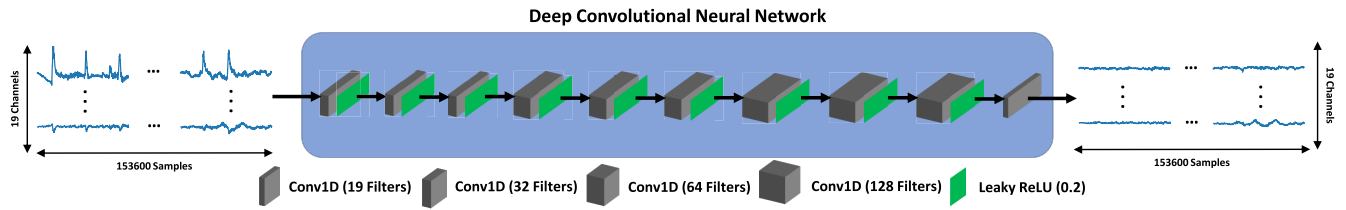


FIGURE 1. DCNN proposed to automatically remove artifacts from EEG segments. Input and output data contain 153600×19 samples. This size corresponds to the number of samples that a 10-minute segment with 19 channels, acquired using a sampling rate of 256 Hz, contains. Convolutional layers are presented as grey parallelipeds. The larger the number of filters in the layer, the larger the width of the paralleliped. All convolutional filters were of size 3. Leaky ReLU activation layers are presented green rectangles. All activation layers uses a α value of 0.2.

TABLE 1. Duration of all raw EEG signals, all preprocessed EEG signals, removed EEG signals, and preprocessed EEG signals with at least one interpolated sample. These values describe the data corresponding to all 25 patients.

Data	Duration (Hours)
Raw EEG Data	648.00
Preprocessed EEG Data	612.68
Removed EEG Data	35.32
EEG Data with at least one interpolated sample	18.25

of both training and test sets with the purpose of eliminating noisy ICs. However, two different procedures were performed for both sets. EEG segments that were already analysed by one expert, were not analysed by the other, i.e. each expert analysed different segments from the training set. Test set was, firstly, analysed by both experts, independently. Then, discordant samples were inspected by the two experts together with the purpose of producing a set, validated by both, to evaluate our approach. After the visual inspection, the segments from training and test sets were reconstructed using the non-noisy ICs. Finally, we had a training set and a test set with two different versions for the same EEG segment: the segment before visual inspection of the ICs (noisy segment), and the segment after the visual inspection of the ICs (target segment).

C. EEG ARTIFACT REMOVAL DEEP CONVOLUTIONAL NEURAL NETWORK

The proposed EEG artifact removal method, based on deep convolutional neural networks (DCNNs), was designed to automatically remove noise from EEG segments. Although the ICA reconstruction is linear, the decisions performed by the experts to classify the ICs are nonlinear. Therefore, a nonlinear model is required to automatically remove noisy artifacts from the EEG segments.

DCNNs contain convolutional layers and layers with several possible activation functions. Convolutional layers [37] include several filters, used for extracting features from the input data, optimised during learning process. Layers with activation functions are used for controlling the information which is transferred to the following layer. Rectified linear unit (ReLU) function is commonly used given its nonlinear

behaviour and fast computation [37]. However, this nonlinear function can produce *dead neurons*, which means that some neurons of the network will output a zero value for different inputs. Leaky ReLU function was introduced in order to overcome this disadvantage [38]. It solves the problem by outputting a smaller portion of the negative inputs instead of nullifying them.

As seen in Figure 1, we developed an architecture based on three convolutional blocks, i.e., three sets of three convolutional layers followed by leaky ReLU activation function. The convolutional layers, used in each block, become wider as DCNN depth increases. ICA may be viewed as a single convolutional layer with a linear filter that covers all channels at a time. Therefore, we consider that more than one nonlinear convolutional layer is required to allow the model to better learn such task.

Since the various scalp EEG channels are not independent from each other, and as ICA processing covers all channels at the same time, we decided to produce a model able to remove artifacts from all the channels, simultaneously.

Researchers report that deep learning models improve with the increasing of depth and width [39], [40]. Thus, we developed an architecture that combine both factors taking into account the available computational resources (4 GPU NVIDIA Quadro P5000 with 16 GB GDDR5 RAM). The number of filters per layer starts at 32 and doubles from one block to the next. The last convolutional layer is used for converting the data back to the initial dimensions. Small filters are useful for exploring fine details of the data and have less computational cost than large filters [41]. Filters with size 1 were not considered because these ones are not able to analyse the values around the unit under analysis. Filters with an even size were also not used because these ones cannot maintain the symmetry around the unit under analysis resulting in data distortions across the layers. Finally, we performed grid-search experiments using filters with size 3 and filters with size 5 and verified that the results were similar. Therefore, all convolutional layers comprise filters with size 3 making the training of the model faster and less prone to overfit. As we did not want to reduce the sample size across the layers, we used a stride of 1 for every convolutional layer.

All activation layers use leaky ReLU function. All the used leaky ReLU functions consider an α of 0.2 as suggested by Xu et al. [42].

D. TRAINING AND VALIDATION

The training set was further filtered by the number of eliminated ICs. Therefore, the EEG segments, with more than half of their ICs classified as noise, were discarded. This step was performed in order to remove segments with few brain independent sources, which would not provide enough information for reliable EEG segment reconstruction. After this filtering step, 2900 EEG segments remained. It is worth noting that this filtering step was not performed in the test set.

We split the training set in training and validation subsets by performing a random 70/30 holdout partition. Validation aims to prevent overfitting in training. Therefore, training subset contains 2030 samples whereas the validation subset contains 870 samples. Each sample consists of one noisy segment and one target segment. After that, the samples lasting less than 10 minutes were zero padded. Thereafter, both subsets were standardised using the average and standard deviation calculated using all noisy segments belonging to the training subset.

For training the DCNN, we used Adam optimisation function [43], with an initial learning rate of 3.0e-4. Regarding the loss function, the usually used root mean squared error (RMSE) gives more significance to larger reconstruction errors, thus leading the algorithm to focus in artifacts with larger amplitude, independently from the range of values of the target signal (see Equation 1). For reducing this bias, we replaced the RMSE by the relative root mean squared error (RRMSE). RRMSE [6] normalises the RMSE by dividing it by the root mean square (RMS) of the target EEG segment (see Equations 2 and 3).

$$RMSE(x, y) = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \tag{1}$$

$$RMS(x) = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}} \tag{2}$$

$$RRMSE(x, y) = \frac{RMSE(x, y)}{RMS(y)} \tag{3}$$

where:

- x_i i-th Input Value
- y_i i-th Target Value
- N Number of Samples

The model was trained for 500 epochs. Simultaneously, the model was evaluated, every new epoch, using the validation subset, with the purpose of saving the one that obtained the lowest validation loss.

The aforementioned procedures were performed ten times. This was intended to decrease the randomness of the training process. At the end of each run, the best model was saved with the intention of being tested with the completely independent test set.

Table 2 summarises the hyperparameters used for training the models.

TABLE 2. Hyperparameters used to train the DCNN models.

Hyperparameter	Value
Dataset Partition	Holdout Validation 70/30
Optimisation Function	Adam
Learning Rate	3.0e-4
Loss Function	RRMSE
Epochs	500
Number of Runs	10

E. EVALUATION METRICS

We evaluated the model using standard statistical metrics. As standard statistical metrics we used RMSE, for measuring reconstruction error (see Equation 1), RRMSE, for measuring normalised reconstruction error (see Equation 3), Pearson correlation coefficient (PCC), for measuring the linear correlation between the denoised and the target segments (see Equation 4), and signal-to-noise ratio (SNR) difference [18], [44], [45] for measuring the noise attenuation.

We calculated RMSE, RRMSE, and PCC for both noisy and denoised segments. In other words, we compared the noisy segments and the denoised segments with the target segments. In this way, we could inspect whether the DCNN model approximates the noisy data to the target data.

SNR difference is the difference between input and output SNRs (see Equations 5 and 6). Input SNR was computed under the assumption that noise is equal to the difference between the noisy and target segments. Output SNR was performed under the assumption that noise is equal to the difference between the denoised and target segments.

$$PCC(x, y) = \frac{Covariance(x, y)}{\sigma_x \sigma_y} \tag{4}$$

$$SNR(x, y) = 10 \times \log_{10} \frac{\sum_{i=1}^N y_i^2}{\sum_{i=1}^N (x_i - y_i)^2} \tag{5}$$

$$SNR_{diff} = SNR(\hat{y}, y) - SNR(x, y) \tag{6}$$

where:

- x Noisy segment
- y Target segment
- \hat{y} Denoised segment
- N Number of samples

We computed RMSE, RRMSE, PCC, and SNR difference for each EEG channel, independently. In this way, we can analyse the alterations that the model performed in each channel.

It is worth noting that the SNR difference cannot be performed when there is no difference between noisy and target segments. Thus, for implementing this evaluation metric, we removed test segments containing only brain ICs.

F. COMPARISON WITH DIFFERENT ARTIFACT REMOVAL MODELS

We compared our DCNN model with 1D-ResCNN model from [26] and with an automatic ICA model based on

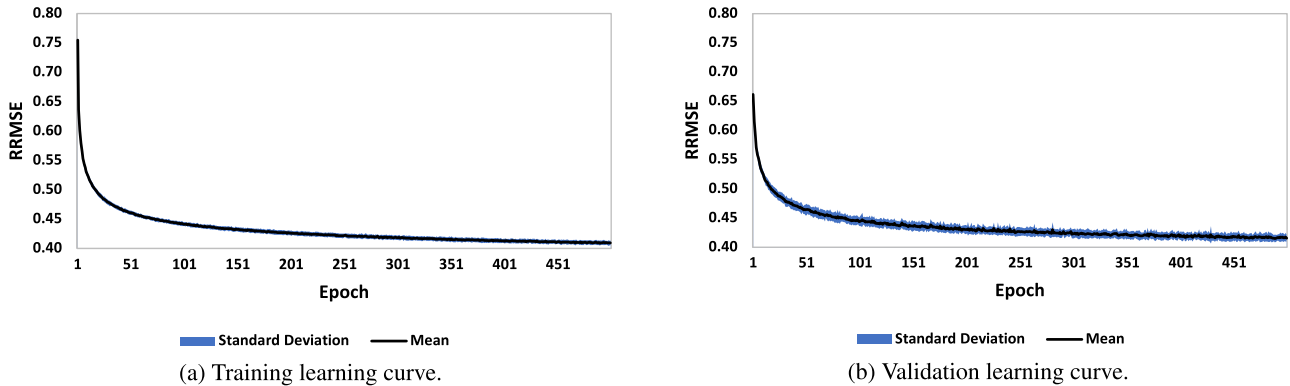


FIGURE 2. Mean and standard deviation of training and validation learning curves obtained from averaging all ten developed models.

extended Infomax ICA and MARA classifier [16]. As the 1D-ResCNN is not publicly available, we developed it following the procedures presented by the authors. The MARA model is publicly available in EEGLAB toolbox. These models were chosen because they are also able to automatically remove several different artifacts from the EEG signals.

All models were tested in a computer with an AMD Ryzen 5 2600 CPU 3.4 GHz, 64 GB of RAM, NVIDIA RTX 2060 Super, and Linux Ubuntu 20.04 LTS. The extended Infomax ICA-MARA was tested in Matlab 2019b whereas the DCNN and 1D-ResCNN models were tested using Tensorflow 2.0 and Keras 2.3 from Python 3.8 in Anaconda Spyder 4.¹ The inference phase of the DCNN models was performed using CPU rather than GPU with the purpose of comparing it with the extended Infomax-MARA model, which has to be performed in CPU. Additionally, testing the models on the CPU allows to approximate the simulation to a real environment where GPUs are usually less available.

III. RESULTS

This section describes the results obtained for the developed deep convolutional neural network (DCNN).

A. TRAINING AND VALIDATION LEARNING CURVES

Figure 2 show the mean and standard deviation of the training and validation learning curves for all the developed models. Figures 2a and 2b show that the validation learning curve follows the training learning curve. This suggests that the developed models did not overfit the training data. Furthermore, it is seen that the models started to stabilise around the 300th epoch which means that the number of epochs was not a limiting factor to the learning procedure. Moreover, the low standard deviation indicates that all the ten models perform similarly. Therefore, we randomly selected one of them for further analysis.

B. EXAMPLES OF EEG SEGMENTS RECONSTRUCTED BY OUR APPROACH

In order to demonstrate how our approach performed for the various types of artifacts found in segments, we present some

examples of the noisy segments along with the target and obtained denoised segments. We show one example for each type of EEG artifact present in our data. More examples are available in Supplementary Material for further exploration.

Figure 3a shows three types of artifacts: eye blinks, eye movements, and muscle activity. Figures 3b, 3c, and 3d show these artifacts in detail whereas Figures 3e, 3f, and 3g present the power spectrum density (PSD) of the time series in the selected time windows.

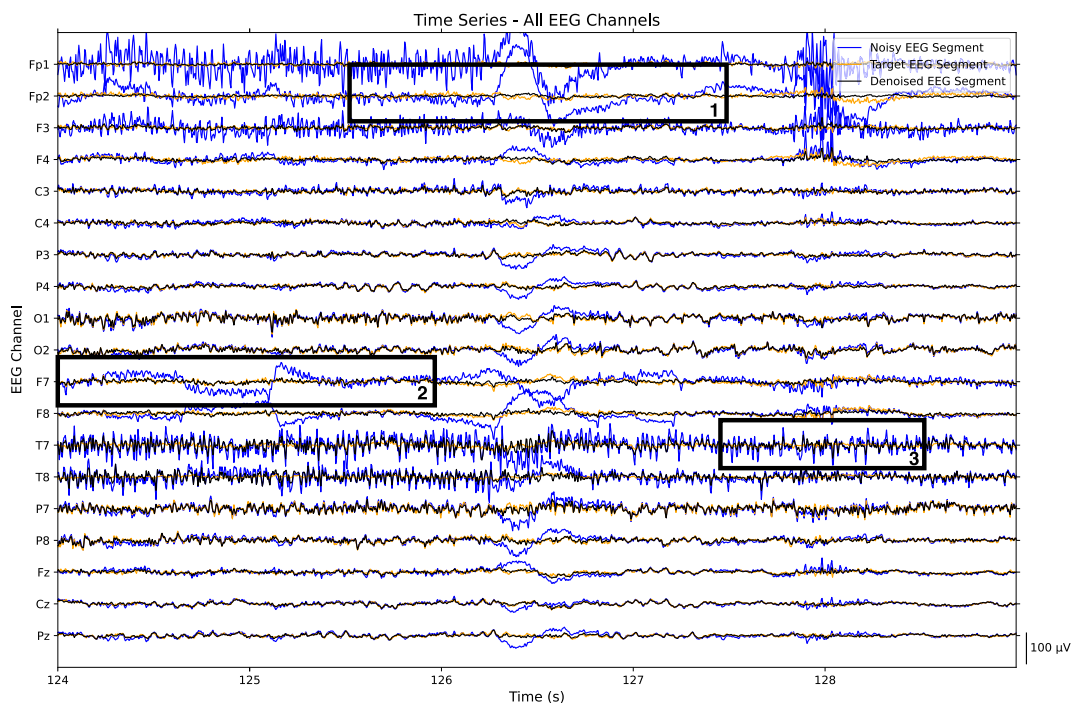
As can be observed in Figures 3b and 3c, the model was able to remove the ocular artifacts and returned a denoised segment similar to the target segment. However, Figure 3e evidences a loss of information in high frequencies. Eye blinks are typically artifacts with large amplitude and low frequency. Therefore, as the used loss function most strongly penalises the larger differences between the denoised and the target segments, the training of the model tries to find out how to reduce these artifacts before learning how to correct the small details of the data. As the EEG amplitude is, in most cases, inversely proportional to its frequency, in the case of an incomplete training there may be a loss of high frequency information. Figure 3d shows that the model attenuated the presence of the muscle activity, but did not remove it completely. This behaviour is confirmed by Figure 3g, i.e., there was only an attenuation of the PSD of the noisy channel. This may happen as a result of the difficulty of eliminating this artifact even by visual inspection of the independent components (ICs).

Figure 4 shows cardiac peaks in channel O1, which were not removed by the DCNN model. As these artifacts appeared rarely in the training set, the model may have had difficulty in considering them as noise.

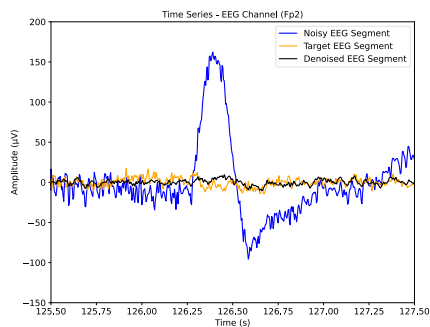
Figure 5 shows pulse artifacts in channel C4. These artifacts resulted from having the EEG electrode on a pulsating vessel on the scalp. It can be observed that the model was not able to remove this interference from the noisy segment. These artifacts also did not occur frequently in the training set. Therefore, similarly to cardiac artifacts, the model may not had learned to consider them as noise.

Figure 6 shows electrode movement in all channels. These artifacts usually appear when there is a disturbance in the

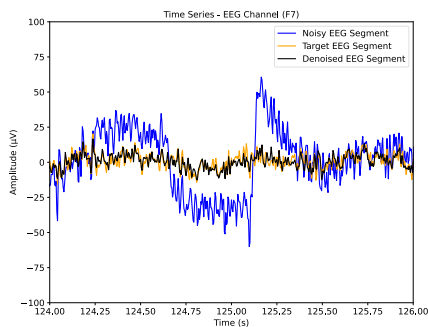
¹<https://docs.spyder-ide.org/4/>



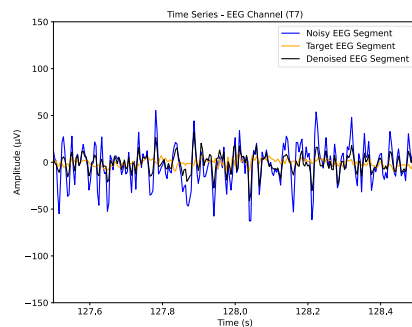
(a) Five seconds of all channels of an example EEG segment containing eye blinks (1), eye movements (2) and muscle activity (3).



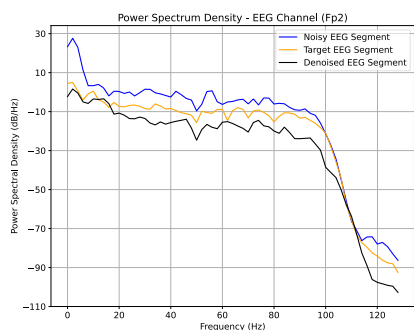
(b) EEG time series of the selected portion of the Fp2 channel of Figure 3a.



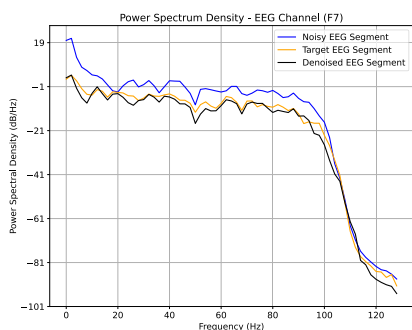
(c) EEG time series of the selected portion of the F7 channel of Figure 3a.



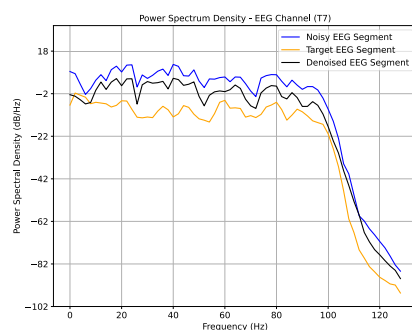
(d) EEG time series of the selected portion of the T7 channel of Figure 3a.



(e) PSD of the selected portion of the Fp2 channel of Figure 3a.



(f) PSD of the selected portion of the F7 channel of Figure 3a.



(g) PSD of the selected portion of the T7 channel of Figure 3a.

FIGURE 3. Example EEG segment from the test set. The noisy segment, target segment and denoised segment are presented in blue, orange and black, respectively. The selected portions of Figure 3a provide the exact moment of artifact occurrence: in Fp2 channel there is an eye blink, in F7 channel there is an eye saccade and in T7 channel there is muscle activity. These portions are zoomed in in Figures 3b, 3c and 3d. The PSD of the EEG time series, in the selected portions, are provided in Figures 3e, 3f and 3g.

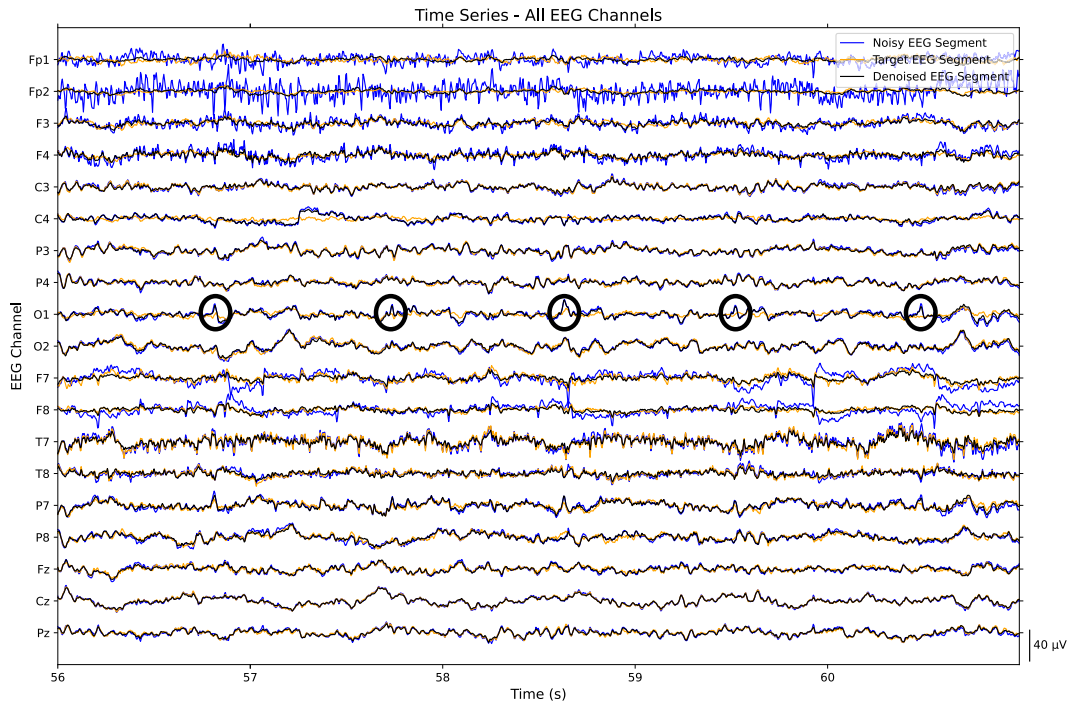


FIGURE 4. Five seconds of all channels of an example EEG segment. The noisy segment, target segment and denoised segment are presented in blue, orange and black, respectively. The selected portions provide the exact moments when the cardiac artifacts occur.

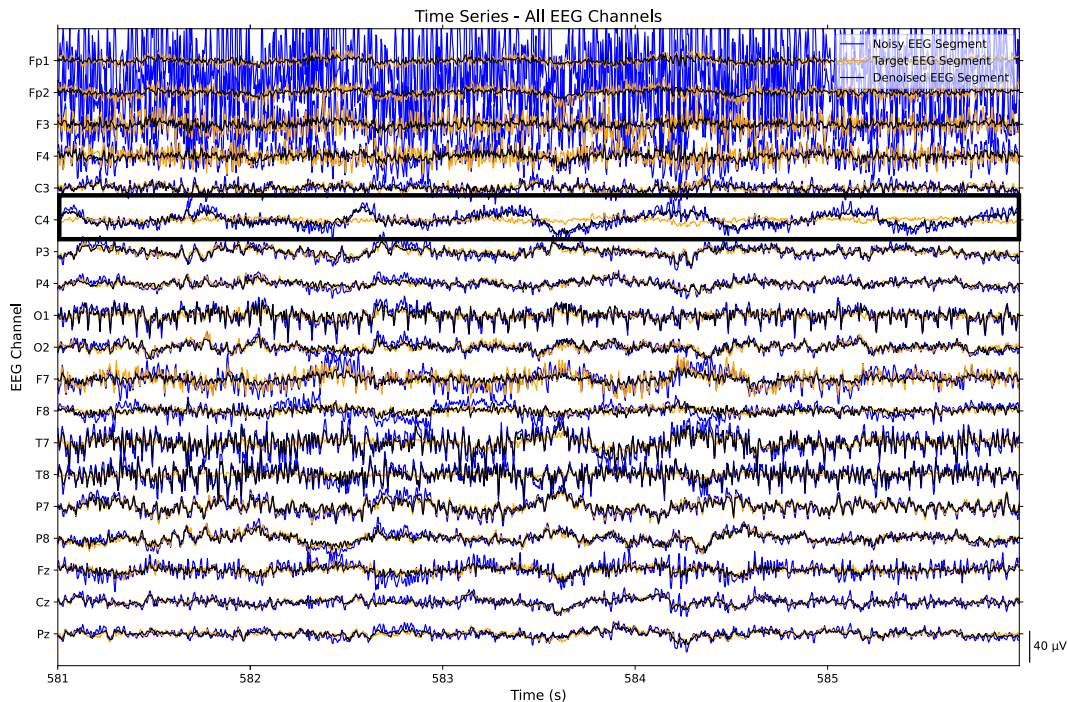


FIGURE 5. Five seconds of all channels of an example EEG segment. The noisy segment, target segment and denoised segment are presented in blue, orange and black, respectively. The selected portion evidences the EEG channel where the pulse artifacts occur.

electrodes which leads to a change of impedance. In this case, the model was able to remove this interference from the noisy segment.

Figure 7 evidences that the model was not able to extract brain information from time intervals when there were electrode connection errors. We would expect that this type of

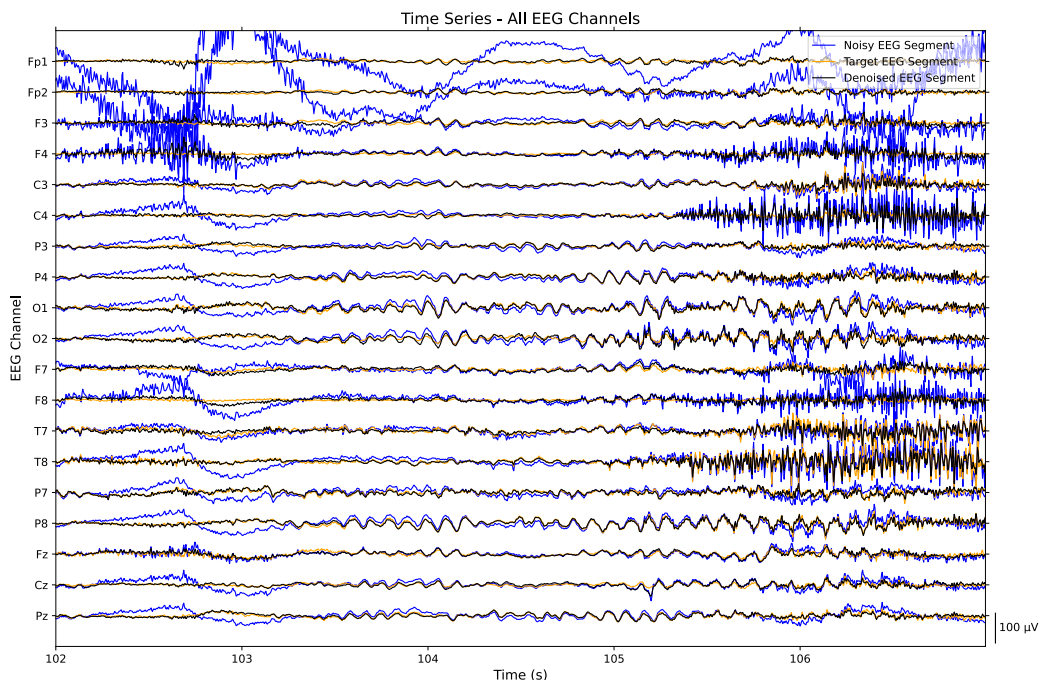


FIGURE 6. Five seconds of all channels of an example EEG segment containing electrode movements. The noisy segment, target segment and denoised segment are presented in blue, orange and black, respectively.

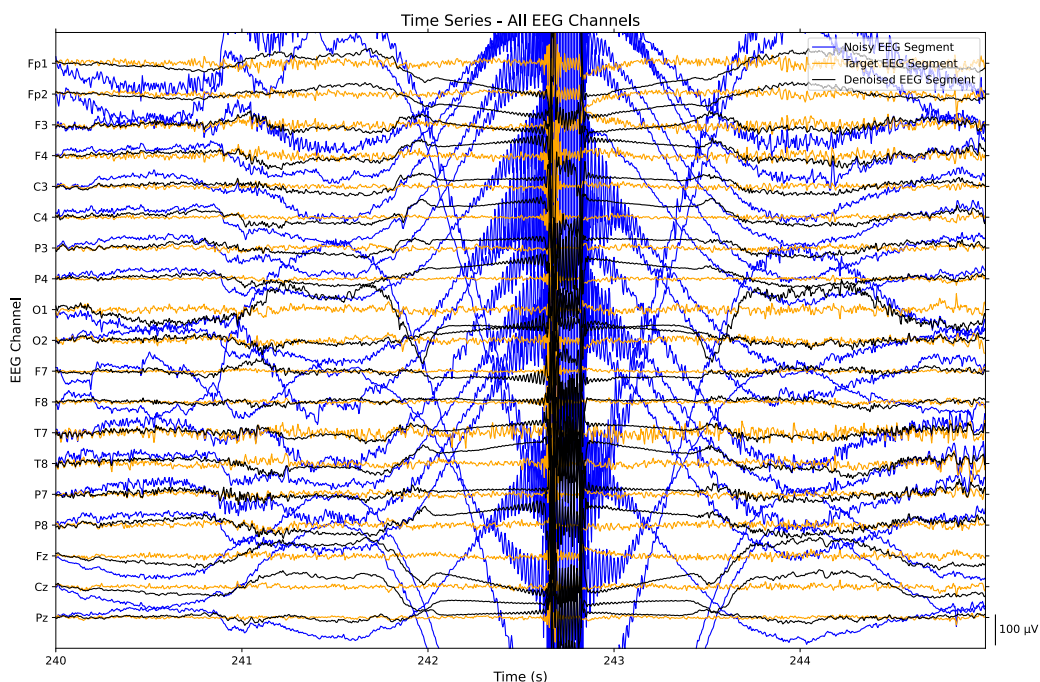


FIGURE 7. Five seconds of all channels of an example EEG segment containing experimental errors which were not removed in the first step of the EEG preprocessing algorithm. The noisy segment, target segment and denoised segment are presented in blue, orange and black, respectively.

artifact would be removed in the first stage of the EEG preprocessing algorithm. The algorithm was designed to remove portions, with an amplitude greater than 5 mV or lower than -5 mV, when the connection error occurred on several channels, simultaneously. Therefore, it is possible that some

portions, with connection problems, still remained after the initial EEG preprocessing.

Figure 8 shows that the model has learned not to make considerable transformations when noise is not present in the EEG segments. However, it is seen that there was an

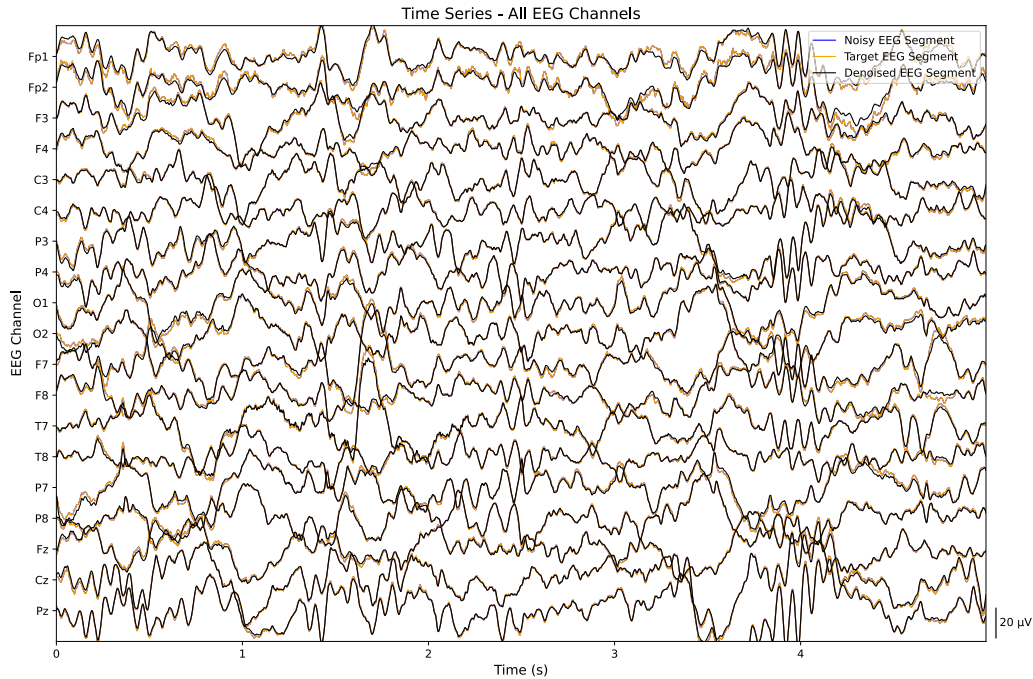


FIGURE 8. Five seconds of all channels of an example EEG segment that does not contain any noisy artifact. The noisy and target segments are equal as there are no artifacts in the EEG data. These segments are presented in orange whereas the denoised segment is presented in black.

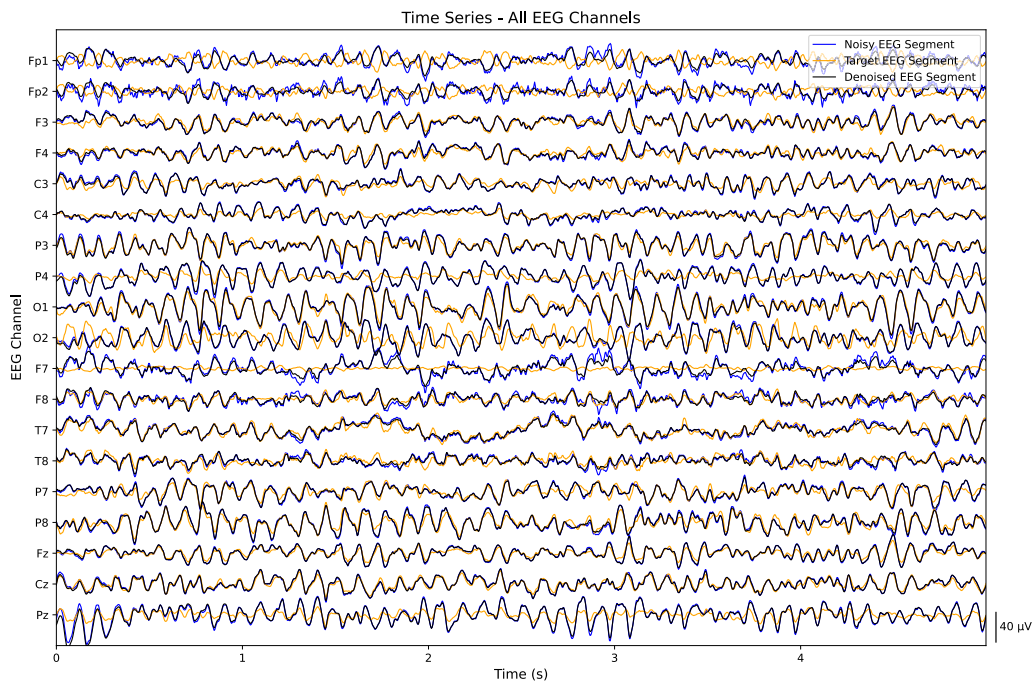


FIGURE 9. Five seconds of all channels of an example EEG segment which had some brain information removed by visual inspection that was not removed by the EEG artifact removal model. The noisy segment, target segment and denoised segment are presented in blue, orange and black, respectively.

attenuation of the high frequency waves in EEG channels, specially, where high amplitude artifacts usually appear such as Fp1, Fp2, F7 and F8. This means that the model focused excessively on removing artifacts on these channels containing low frequency artifacts, and thus, failed to learn high frequency details.

Figure 9 presents a very important behaviour of the developed model. This figure shows a portion of a segment, with a connection error, that resulted in removing more than half of the ICs by visual inspection. In this case, the model learned to analyse small windows of the entire noisy segment and to keep the data that did not have any noisy artifact. Before

developing our models, we removed every EEG segment with more than 50% noisy ICs. Therefore, the models may not have learned to remove excessive information in situations as this one. This is an important advantage comparing to independent component analysis (ICA) approaches, because it preserved brain information while attenuating the influence of artifacts.

In summary, we may conclude that our model could attenuate artifacts such as eye blinks, eye saccades, muscle activity, and channel movements present in Figures 3b, 3c, 3d, and 6, respectively. Furthermore, it could perform reasonable reconstructions when no artifacts were present on the EEG data (see Figures 8 and 9). However, the model could not handle rare EEG artifacts such as cardiac artifacts, pulse artifacts, and saturated segments present in Figures 4, 5, and 7, respectively.

C. EVALUATION METRICS

The developed EEG artifact removal model was assessed using the evaluation metrics presented in Subsection II-E. Thus, we computed the metrics for all independent test samples. The metrics were calculated for each EEG channel, independently. Therefore, for each EEG channel, we obtained 910 values for root mean squared error (RMSE), relative root mean squared error (RRMSE), and Pearson correlation coefficient (PCC) for noisy and denoised segments and 875 values for signal-to-noise ratio (SNR) differences.

When inspecting the results for RMSE, RRMSE, PCC, and SNR difference, we observed skewed distributions explained by the existence of some outliers in these metrics. These outliers result from some experimental errors that were not removed in the initial preprocessing pipeline. Therefore, instead of using the common central tendency statistics, mean and standard deviation, we utilised the median and interquartile range (see Figure 10). Mean and standard deviation are available in Supplementary Material.

Figures 10a and 10b present the median and interquartile range values of the RMSE and RRMSE, respectively. As stated in Section II-E, RMSE evaluates the reconstruction error whereas RRMSE measures the normalised reconstruction error. The lower these metrics are, the closer the obtained denoised data are to the target one. In general, these values decreased when using the DCNN model, which suggests that the model learned to approximate the noisy segments to the target ones.

According to Figures 10a and 10b, Fp1, Fp2, F7, and F8 were the EEG channels associated with a larger decrease in RMSE and RRMSE. This occurred as a result of the removal of ocular artifacts, which typically have an amplitude higher than the brain data. Channels F3, F4, T7, and T8 also evidenced a large reduction of these metrics. Although all EEG channels contained some muscle activity at a certain period of time, F3, F4, T7, and T8 were usually contaminated with this artifact over several segments. Therefore, results suggest that the developed model was able to reduce the presence of these artifacts, even in highly-corrupted channels. Comparing RMSE and RRMSE for the denoised segments, it is seen that

although the channels O1, O2, P7 and P8 evidence RMSE values similar to the those obtained for channels C3, C4, P3 and P4, these channels present RRMSE values among the lowest of all channels. This means that the expected root mean squared (RMS) of channels O1, O2, P7 and P8 were greater than the expected RMS of the channels C3, C4, P3 and P4. Therefore, we conclude that the same error value has a lower relevance for those channels.

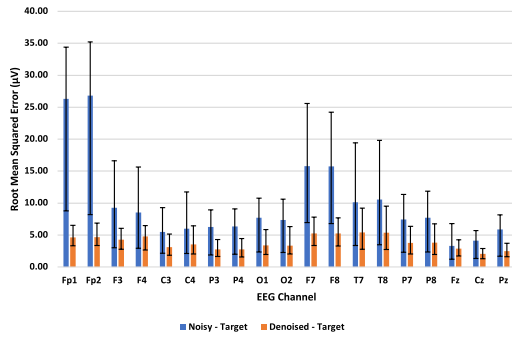
Figure 10c shows the median and interquartile range values of the PCC values. As mentioned in Section II-E, PCC measures the linear correlation between two time series. Therefore, the higher this metric is, the greater is the linear correlation between the obtained denoised segment and the target one. In general, PCC increased after using the DCNN model, which suggests that the noisy segments became more linearly correlated with the target ones after using it. As already verified in Figures 10a and 10b, the larger PCC increase can be seen for the EEG channels containing ocular artifacts (Fp1, Fp2, F7 and F8).

Figure 10d shows the median and interquartile range values of the SNR difference values. As mentioned in Section II-E, SNR difference measures the improvement of the SNR after using the DCNN model. Positive values suggest a SNR increase, after using the model, whereas negative values suggest a SNR decrease and, therefore, a lower success in denoising the EEG. Although the overall results evidence the improvement of the SNR for all channels, the Fz channel presents an interquartile range that contains the zero value. As seen in Figures 10b and 10c, the interquartile range of the results for this channel, before using the DCNNs, contains almost optimal values. This indicates that for our test dataset, this channel was less corrupted by artifacts. Therefore, it was practically unchanged by the model.

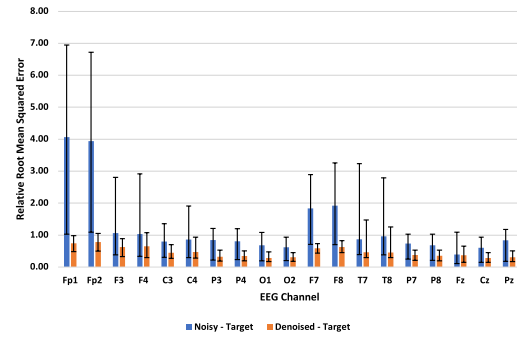
D. COMPARISON WITH DIFFERENT ARTIFACT REMOVAL MODELS

For each artifact removal model, we computed the RMSE, RRMSE, PCC, and SNR difference for all independent test samples. Furthermore, we obtained the computation times. Contrary to the previous section, to simplify the comparison, we computed these metrics using all channels. Table 3 presents the RMSE, the RRMSE and the PCC for the original data and denoised data reconstructed by our model, 1D-ResCNN, and extended Infomax ICA-MARA. Furthermore, it contains the SNR difference and computation times for each artifact removal model.

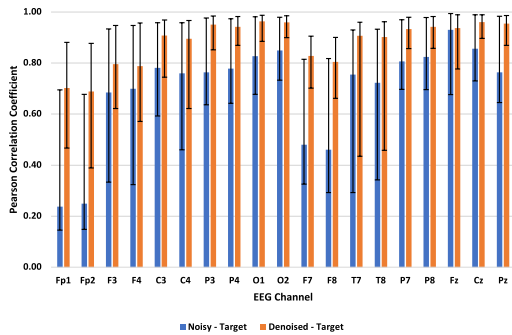
We also performed pairwise comparisons, using non-parametric tests (Kruskal-Wallis [46] and Dunn-Šidák [47]), between all approaches, using all statistical metrics, to study whether there are statistical differences between them. To compare the performances, we used a significance level of 0.05. Figure 11 presents those pairwise comparisons. For RMSE, RRMSE and prediction time, lower values are related with lower noise levels whereas for PCC and SNR difference, higher values are preferred.



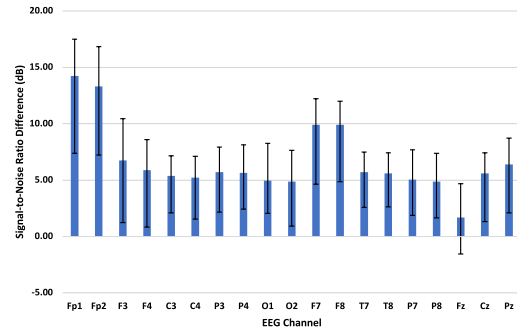
(a) RMSE of the EEG segments, for each EEG channel, before and after applying DCNN model.



(b) RRMSE of the EEG segments, for each EEG channel, before and after applying DCNN model.



(c) PCC of the EEG segments, for each EEG channel, before and after applying DCNN model.



(d) SNR difference of the EEG segments, for each EEG channel.

FIGURE 10. Performance results using the standard evaluation metrics reported in EEG-based studies. Each bar corresponds to the median value and the interquartil range.

TABLE 3. Statistical metrics used to compare the EEG artifact removal models. The values are presented in the format median (first quartile - third quartile).

	Noisy	Denoised (Our Model)	Denoised (1D-ResCNN)	Denoised (Extended Infomax ICA - MARA)
RMSE (Microvolts)	13.88 (4.53 - 19.04)	4.83 (2.92 - 6.90)	9.83 (6.93 - 12.70)	5.28 (3.32 - 7.62)
RRMSE	1.57 (0.50 - 2.29)	0.52 (0.32 - 0.77)	1.03 (0.61 - 1.51)	0.63 (0.38 - 0.81)
PCC	0.54 (0.40 - 0.90)	0.86 (0.70 - 0.95)	0.53 (0.37 - 0.83)	0.78 (0.60 - 0.93)
SNR Difference	-	8.81 (4.83 - 10.05)	3.12 (0.11 - 3.78)	7.58 (3.08 - 10.37)
Computation Time (Seconds)	-	0.29 (0.29 - 0.30)	3.37 (3.35 - 3.38)	384.52 (375.43 - 391.49)

Results presented in Table 3 evidence that our model obtained considerably lower reconstruction errors and higher PCCs and SNR differences, compared to the 1D-ResCNN. Furthermore, our model is faster than 1D-ResCNN. Figure 11 shows that the differences between both methods were statistical significant (p-value (RMSE) ≈ 0 ; p-value (RRMSE) ≈ 0 ; p-value (PCC) ≈ 0 ; p-value (SNR Diff) ≈ 0 ; p-value (Prediction Time) ≈ 0).

Comparing our model with the extended Infomax ICA-MARA, results provided in Table 3 evidence that our model obtained lower median reconstruction error and higher median PCCs and SNR differences. Figure 11 also evidences that our model obtained lower RRMSE, higher PCC and SNR difference with significant statistical differences (p-value (RRMSE) = 0.015; p-value (PCC) ≈ 0 ; p-value (SNR Diff) = 0.003). However, it shows that both models did not

obtain significant statistical differences for RMSE (p-value = 0.087). Additionally, results demonstrated that our approach is considerably faster than the extended Infomax ICA-MARA which lasted around 6 minutes on average compared to less than a second (p-value ≈ 0) in our method.

IV. DISCUSSION

The automatic electroencephalogram (EEG) artifact removal approach, presented in this article, is based on deep convolutional neural networks (DCNNs). It was designed to automatically remove several artifacts, commonly observed in long-term EEG signals, such as ocular artifacts, muscle activity, cardiac activity, pulse artifacts, and electrode connection issues, in a similar way to that performed by experts. Studies cited in the literature review, which developed deep learning models that whether only remove one

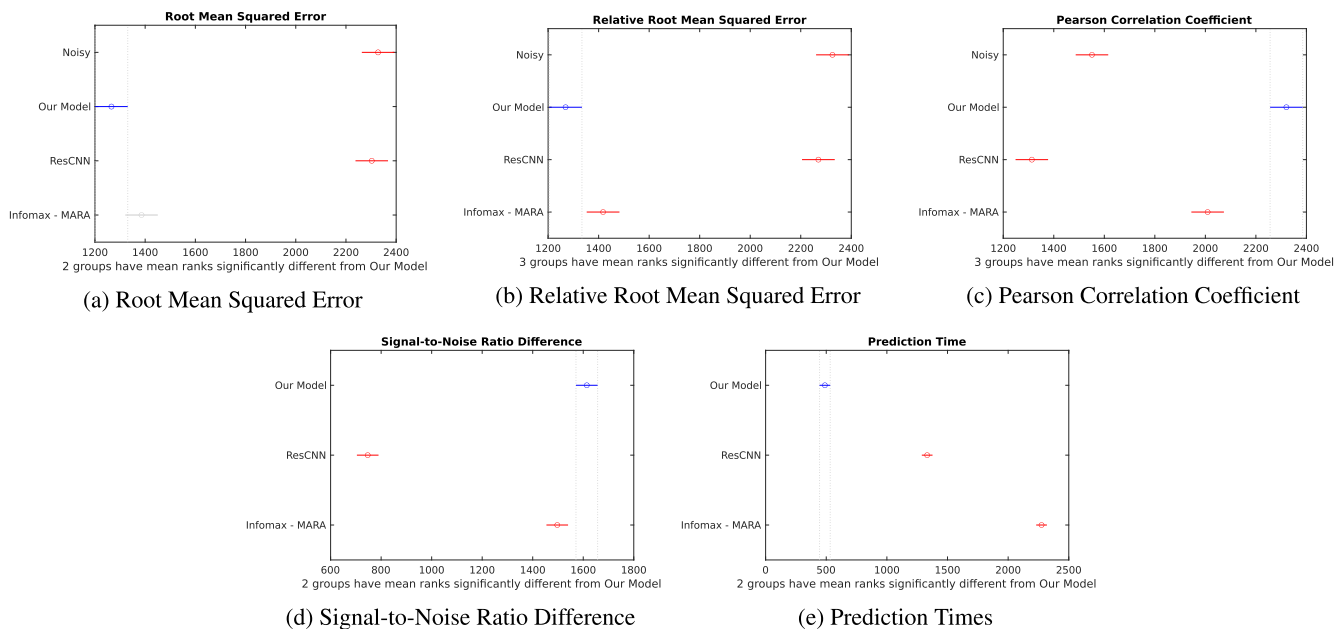


FIGURE 11. Pairwise comparison plots comparing the statistical metrics of the noisy data and the denoised data reconstructed by our model, by 1D-ResCNN and by extended Infomax ICA-MARA. The presented values represent the mean ranks and the standard errors obtained from the statistical test. Overlapping confidence intervals mean that there are no significant statistical differences between the results under analysis. These figures were obtained using the *multcompare* function from MATLAB.

type of artifact [22], [23] or were trained with artificially generated noisy EEG data [24], [26], do not fully simulate real application scenarios such as clinical EEG evaluation. Our approach is a step forward because it was able to remove several artifacts present in long-term signals collected from epileptic patients, in pre-surgical monitoring.

Our approach was developed using EEG data previously processed using independent component analysis (ICA). ICA is a linear decomposition method. Therefore, the reconstruction of EEG segments, after removing noisy independent components (ICs), is performed using linear equations. However, our model must not only automate the linear reconstruction of the EEG segments without artifacts but also the nonlinear decisions performed by experts when classifying the ICs.

The developed approach was able to attenuate the influence of artifacts while preserving brain information. Additionally, it was able to recognise artifacts within a segment and minimise information related to these ones keeping the remainder data. Thus, it could preserve more information than ICA [23], [26]. These results could be related to the removal of training data with more than half of their ICs classified as noise, i.e., the model did not learn to excessively remove data. Finally, the model removed artifacts from signals which were not used in training, which means that it did not overfit to signals from patients used in training, and, therefore, may be used in EEG signals from new subjects.

We found out that the model had difficulty in preserving small details of the EEG signals in channels where high amplitude artifacts were common, e.g., Fp1 and Fp2 channels. This behaviour was also noticed by Yang *et al.* [23].

They reported that gamma bands (more than 30 Hz) were not perfectly reconstructed when ocular artifacts were present, which means their model also lost high frequency details when signals were corrupted with this type of artifacts. Note that loss functions, which aim at reducing the reconstruction error, firstly find out how to decrease the larger errors, and secondly learn the smaller details. Therefore, as these channels presented high amplitude artifacts, the model learned to remove their influence before learning how to reconstruct low amplitude data. In the EEG signals, the frequency is, in most cases, inversely proportional to the amplitude. Thus, although our model could preserve low frequency EEG data, it may require a different training setup, e.g., increasing the training set, searching for the optimal deep learning architecture, using longer training times or replacing the utilised loss function by another one, in order to improve its high frequency detail reconstruction.

Results evidenced that our model obtained the greatest performance among the tested artifact removal models. Compared to the 1D-ResCNN, this could be explained by the fact that the latter was developed using simulated noisy EEG data which could not precisely mimic real noisy EEG segments. Considering the extended Infomax-MARA, our approach presented a minor loss of information because instead of removing the entire source related with the noise, it focuses on removing just the time interval when the artifact occurs. Furthermore, our model is faster than the other evaluated approaches. Therefore, it may be used to remove artifacts from signals in real-time scenarios. Our approach could be, for example, deployed into the IBM’s TrueNorth Neurosynaptic System [48], [49], which is a power-efficient

neuromorphic chip that can be adapted to implement deep convolutional neural networks [50], to remove artifacts from EEG signals before epileptic seizure prediction.

As our approach was developed with EEG segments acquired using a sampling rate of 256 Hz, it is restricted to acquisition systems using the same number of samples per second. However, this may not be seen as a strong limitation because studies often consider scalp EEG that was either obtained using this sampling rate [51]–[56] or using higher sampling rates which were subsequently downsampled for further analysis [57]–[60]. Moreover, as it was trained using multi-channel EEG segments, it is also restricted to the same channel placement over scalp. Furthermore, it is limited to segments lasting up to 10 minutes, i.e., signals with longer duration must be segmented before being processed.

The model was developed with 19-channel EEG segments previously processed using ICA. ICA can only find a number of independent sources at most equal to the number of used channels. Although some authors state that 19 may be considered as a high number of EEG channels, ICA usually performs more accurately with EEG data with at least 64 channels [61], [62]. Therefore, the reconstruction capability of our model may be limited by the performance of the ICA decomposition. However, as our main goal was to develop a model that would be able to work in real long-term scenarios, we were restricted to low-density EEG data that were available in the database.

We developed our model using EEG data from epileptic patients under pre-surgical monitoring conditions. These data were acquired without conditioning patients' activities. This means that data contain several artifacts which most probably are present in EEG signals acquired for other research purposes. Therefore, although it was developed using epileptic EEG data, it may be used for denoising other types of EEG signals.

V. CONCLUSION

This work demonstrates the potential of deep learning architectures in the development of models that can automatically remove artifacts from electroencephalograms (EEGs) in less than a second.

Removing artifacts present in real long-term EEG signals, by visual inspecting the independent sources of the signals, is a time consuming task since it requires the examination of several hours of data. Therefore, we developed a deep learning approach to automatically and quickly remove artifacts, such as eye blinks, eye movements, muscle activity, cardiac activity, and electrode connection interferences. In this way, we could use it later to automatically eliminate noise from EEG signals from other patients, available in the EPILEPSIAE database, or for removing noise in real-time scenarios.

Experimental results suggested that the developed model was able to attenuate the influence of the artifacts in the EEG signals. Furthermore, compared to other approaches, our model could combine a minor reconstruction error with a fast computation, making it suitable to be used to preprocess

real-time long-term EEG signals. This demonstrates that EEG artifact removal models, based on deep neural networks, developed using real EEG signals, should be taken into consideration when noisy artifacts are present in the EEG data.

Following this study, we plan to develop deep convolutional neural network models using each EEG channel individually and compare them with the model presented in this article. In this way, if the new approaches achieve similar or better performance, they could be used to remove artifacts from noisy segments acquired with any type of acquisition system, as long as one provides the same sampling rate.

REFERENCES

- [1] P. L. Nunez and R. Srinivasan "The physics—EEG interface," in *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd ed. New York, NY, USA: Oxford Univ. Press, 2006, pp. 3–55.
- [2] S. Sanei and J. A. Chambers, "Introduction to EEG," in *EEG Signal Processing*. Hoboken, NJ, USA: Wiley, 2007, ch. 1, pp. 1–34.
- [3] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems," *Neuroimage*, vol. 34, no. 4, pp. 1600–1611, 2007.
- [4] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—Practices and possibilities," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 488–500, May 2012.
- [5] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: A review," *Clin. Neurophysiol.*, vol. 46, nos. 4–5, pp. 287–305, 2016.
- [6] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—State-of-the-art and guidelines," *J. Neural Eng.*, vol. 12, no. 3, pp. 1–23, 2015.
- [7] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: Spectral and topographical characteristics," *Clin. Neurophysiol.*, vol. 114, no. 9, pp. 1580–1593, 2003.
- [8] S. Boudet, L. Peyrodie, P. Gallois, and C. Vasseur, "A global approach for automatic artifact removal for standard EEG record," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 5719–5722.
- [9] R. N. Roy, S. Charbonnier, and S. Bonnet, "Eye blink characterization from frontal EEG electrodes using source separation and pattern recognition algorithms," *Biomed. Signal Process. Control*, vol. 14, pp. 256–264, Nov. 2014.
- [10] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: A review," *Sensors*, vol. 19, no. 5, pp. 1–18, 2019.
- [11] P. J. Karoly, H. Ung, D. B. Grayden, L. Kuhlmann, K. Leyde, M. J. Cook, and D. R. Freestone, "The circadian profile of epilepsy improves seizure forecasting," *Brain*, vol. 140, no. 8, pp. 2169–2182, Aug. 2017.
- [12] M. O. Baud, J. K. Kleen, E. A. Mirro, J. C. Andrechak, D. King-Stephens, E. F. Chang, and V. R. Rao, "Multi-day rhythms modulate seizure risk in epilepsy," *Nature Commun.*, vol. 9, no. 1, p. 88, Dec. 2018.
- [13] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig, "Independent EEG sources are dipolar," *PLoS ONE*, vol. 7, no. 2, pp. 1–14, 2012.
- [14] S. Choi, A. Cichocki, H. M. Park, and S.-Y. Lee, "Blind source separation and independent component analysis: A review," *Neural Inf. Process. Lett. Rev.*, vol. 6, no. 1, pp. 1–57, 2005.
- [15] H. Nolan, R. Whelan, and R. B. Reilly, "FASTER: Fully automated statistical thresholding for EEG artifact rejection," *J. Neurosci. Methods*, vol. 192, no. 1, pp. 152–162, 2010.
- [16] I. Winkler, S. Haufe, and M. Tangermann, "Automatic classification of artifactual ICA-components for artifact removal in EEG signals," *Behav. Brain Functions*, vol. 7, Aug. 2011, Art. no. 30.
- [17] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, "ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features," *Psychophysiology*, vol. 48, no. 2, pp. 229–240, 2011.
- [18] T. Radüntz, J. Scouten, O. Hochmuth, and B. Meffert, "EEG artifact elimination by extraction of ICA-component features using image processing algorithms," *J. Neurosci. Methods*, vol. 243, pp. 84–93, Mar. 2015.
- [19] T. Radüntz, J. Scouten, O. Hochmuth, and B. Meffert, "Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features," *J. Neural Eng.*, vol. 14, no. 4, 2017, Art. no. 031001.

- [20] P. Croce, F. Zappasodi, L. Marzetti, A. Merla, V. Pizzella, and A. M. Chiarelli, "Deep convolutional neural networks for feature-less automatic classification of independent components in multi-channel electrophysiological brain recordings," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 8, pp. 2372–2380, Aug. 2019.
- [21] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "ICLabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, Sep. 2019.
- [22] R. Ghosh, N. Sinha, and S. K. Biswas, "Automated eye blink artefact removal from EEG using support vector machine and autoencoder," *IET Signal Process.*, vol. 13, no. 2, pp. 141–148, Apr. 2019.
- [23] B. Yang, K. Duan, C. Fan, C. Hu, and J. Wang, "Automatic ocular artifacts removal in EEG using deep learning," *Biomed. Signal Process. Control*, vol. 43, pp. 148–158, May 2018.
- [24] N. M. N. Leite, E. T. Pereira, E. C. Gurjao, and L. R. Veloso, "Deep convolutional autoencoder for EEG noise filtering," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Dec. 2018, pp. 2605–2612.
- [25] H. Zhang, C. Wei, M. Zhao, Q. Liu, and H. Wu, "A novel convolutional neural network model to remove muscle artifacts from EEG," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 1265–1269.
- [26] W. Sun, Y. Su, X. Wu, and X. Wu, "A novel end-to-end 1D-ResCNN model to remove artifact from EEG signals," *Neurocomputing*, vol. 404, pp. 108–121, Sep. 2020.
- [27] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu, "EEGdenoiseNet: A benchmark dataset for end-to-end deep learning solutions of EEG denoising," 2020, *arXiv:2009.11662*.
- [28] M. Ihle, H. Feldwisch-Drentrup, C. A. Teixeira, A. Witon, B. Schelter, J. Timmer, and A. Schulze-Bonhage, "EPILEPSIAE—A European epilepsy database," *Comput. Methods Programs Biomed.*, vol. 106, no. 3, pp. 127–138, 2012.
- [29] J. Klatt, H. Feldwisch-Drentrup, M. Ihle, V. Navarro, M. Neufang, C. Teixeira, C. Adam, M. Valderrama, C. Alvarado-Rojas, A. Witon, M. L. Van Quyen, F. Sales, A. Dourado, J. Timmer, A. Schulze-Bonhage, and B. Schelter, "The EPILEPSIAE database: An extensive electroencephalography database of epilepsy patients," *Epilepsia*, vol. 53, no. 9, pp. 1669–1676, 2012.
- [30] M. D'Alessandro, G. Vachtsevanos, R. Esteller, J. Echaz, S. Cranstoun, G. Worrell, L. Parish, and B. Litt, "A multi-feature and multi-channel univariate selection process for seizure prediction," *Clin. Neurophysiol.*, vol. 116, no. 3, pp. 506–516, Mar. 2005.
- [31] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.
- [32] C. A. Teixeira, B. Direito, M. Bandarabadi, M. L. Van Quyen, M. Valderrama, B. Schelter, A. Schulze-Bonhage, V. Navarro, F. Sales, and A. Dourado, "Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients," *Comput. Methods Programs Biomed.*, vol. 114, no. 3, pp. 324–336, 2014.
- [33] M. F. Pinto, A. Leal, F. Lopes, A. Dourado, P. Martins, and C. A. Teixeira, "A personalized and evolutionary algorithm for interpretable EEG epilepsy seizure prediction," *Sci. Rep.*, vol. 11, no. 1, p. 3415, Dec. 2021.
- [34] F. Perrin, J. Pernier, O. Bertrand, and J. F. Echallier, "Spherical splines for scalp potential and current density mapping," *Electroencephalogr. Clin. Neurophysiol.*, vol. 72, no. 2, pp. 184–187, 1989.
- [35] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [36] T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources," *Neural Comput.*, vol. 11, no. 2, pp. 417–441, 1999.
- [37] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
- [38] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. 30th Int. Conf. Mach. Learn.*, vol. 28, Atlanta, GA, USA: Citeseer, 2013, pp. 1–6.
- [39] G. Zhong, X. Ling, and L. Wang, "From shallow feature learning to deep learning: Benefits from the width and depth of deep architectures," *WIREs Data Mining Knowl. Discovery*, vol. 9, no. 1, Jan. 2019, Art. no. e1255.
- [40] K. Kawaguchi, J. Huang, and L. P. Kaelbling, "Effect of depth and width on local minima in deep learning," *Neural Comput.*, vol. 31, no. 7, pp. 1462–1498, Jul. 2019.
- [41] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
- [42] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," 2015, *arXiv:1505.00853*.
- [43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [44] M. M. N. Mannan, M. A. Kamran, and M. Y. Jeong, "Identification and removal of physiological artifacts from electroencephalogram signals: A review," *IEEE Access*, vol. 6, pp. 30630–30652, 2018.
- [45] H.-T. Chiang, Y.-Y. Hsieh, S.-W. Fu, K.-H. Hung, Y. Tsao, and S.-Y. Chien, "Noise reduction in ECG signals using fully convolutional denoising autoencoders," *IEEE Access*, vol. 7, pp. 60806–60813, 2019.
- [46] P. E. McKight and J. Najab, "Kruskal–Wallis test," in *The Corsini Encyclopedia of Psychology*. Hoboken, NJ, USA: Wiley, 2010, p. 1.
- [47] A. Dinno, "Nonparametric pairwise multiple comparisons in independent groups using Dunn's test," *Stata J., Promoting Commun. Statist. Stata*, vol. 15, no. 1, pp. 292–300, Apr. 2015.
- [48] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, B. Brezzo, I. Vo, S. K. Esser, R. Appuswamy, B. Taba, A. Amir, M. D. Flickner, W. P. Risk, R. Manohar, and D. S. Modha, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014.
- [49] I. Kiral-Kornek, S. Roy, E. Nurse, B. Mashford, P. Karoly, T. Carroll, D. Payne, S. Saha, S. Baldassano, and T. O'Brien, "Epileptic seizure prediction using big data and deep learning: Toward a mobile system," *EBioMedicine*, vol. 27, pp. 103–111, Jan. 2018.
- [50] S. K. Esser, P. A. Merolla, J. V. Arthur, A. S. Cassidy, R. Appuswamy, A. Andreopoulos, D. J. Berg, J. L. McKinstry, T. Melano, D. R. Barch, C. D. Nolfo, P. Datta, A. Amir, B. Taba, M. D. Flickner, and D. S. Modha, "Convolutional networks for fast, energy-efficient neuromorphic computing," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 41, pp. 11441–11446, 2016.
- [51] G. Borghini, P. Aricò, G. Di Flumeri, N. Sciaraffa, A. Colosimo, M.-T. Herrero, A. Bezerianos, N. V. Thakor, and F. Babiloni, "A new perspective for the training assessment: Machine learning-based neurometric for augmented User's evaluation," *Frontiers Neurosci.*, vol. 11, p. 325, Jun. 2017.
- [52] G. Di Flumeri, G. Borghini, P. Aricò, N. Sciaraffa, P. Lanzi, S. Pozzi, V. Vignali, C. Lantieri, A. Bichicchi, A. Simone, and F. Babiloni, "EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings," *Frontiers Hum. Neurosci.*, vol. 12, p. 509, Dec. 2018.
- [53] N. Friedman, T. Fekete, K. Gal, and O. Shriki, "EEG-based prediction of cognitive load in intelligence tests," *Frontiers Hum. Neurosci.*, vol. 13, p. 191, Jun. 2019.
- [54] P. Aricò, M. Reynal, G. Di Flumeri, G. Borghini, N. Sciaraffa, J.-P. Imbert, C. Hurter, M. Terenzi, A. Ferreira, S. Pozzi, V. Betti, M. Marucci, A. C. Telea, and F. Babiloni, "How neurophysiological measures can be used to enhance the evaluation of remote tower solutions," *Frontiers Hum. Neurosci.*, vol. 13, p. 303, Sep. 2019.
- [55] H. Chu, C. K. Chung, W. Jeong, and K.-H. Cho, "Predicting epileptic seizures from scalp EEG based on attractor state analysis," *Comput. Methods Programs Biomed.*, vol. 143, pp. 75–87, May 2017.
- [56] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals," *Comput. Biol. Med.*, vol. 99, pp. 24–37, Aug. 2018.
- [57] K. M. Puk, K. C. Gandy, S. Wang, and H. Park, "Pattern classification and analysis of memory processing in depression using EEG signals," in *Proc. Int. Conf. Brain Informat.* London, U.K.: Springer, 2016, pp. 124–137.
- [58] K. B. Mikkelsen, P. Kidmose, and L. K. Hansen, "On the keyhole hypothesis: High mutual information between ear and scalp EEG," *Frontiers Hum. Neurosci.*, vol. 11, p. 341, Jun. 2017.
- [59] N. Zink, A.-K. Stock, A. Vahid, and C. Beste, "On the neurophysiological mechanisms underlying the adaptability to varying cognitive control demands," *Frontiers Hum. Neurosci.*, vol. 12, p. 411, Oct. 2018.
- [60] S. Kohli and A. J. Casson, "Machine learning validation of EEG+ tACS artefact removal," *J. Neural Eng.*, vol. 17, no. 1, 2020, Art. no. 16034.
- [61] I. Rejer and P. Górski, "MAICA: An ICA-based method for source separation in a low-channel EEG recording," *J. Neural Eng.*, vol. 16, no. 5, Sep. 2019, Art. no. 056025.
- [62] M. Klug and K. Gramann, "Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments," *Eur. J. Neurosci.*, pp. 1–15, Oct. 2020.



tion, and deep learning. His research interests include natural language processing, signal processing, machine learning, and deep learning applied to the improvement of health applications.

FÁBIO LOPES received the M.Sc. degree in biomedical engineering from the University of Coimbra, in 2019. He is currently pursuing the Ph.D. degree with the Department of Informatics Engineering, University of Coimbra, and the Department of Epilepsy, University Medical Center, Freiburg, Germany. His M.Sc. thesis was in the field of natural language processing in clinical domain. His Ph.D. research is in the field of electroencephalogram, epileptic seizure prediction, and deep learning. His research interests include natural language processing, signal processing, machine learning, and deep learning applied to the improvement of health applications.



gram, accelerometry signal, and heart and pulmonary sound signals, and more recently, inserted in the scope of her Ph.D. degree, electroencephalogram and electrocardiogram. She is working towards the improvement of epileptic seizure anticipation using neuro-cardiovascular information fusion and dynamic classification.

ADRIANA LEAL received the M.Sc. degree in biomedical engineering from the University of Coimbra, in 2015, where she is currently pursuing the Ph.D. degree with the Department of Informatics Engineering, Faculty of Sciences and Technology. Her research interests include signal processing and pattern recognition (data mining and machine learning) techniques. She has studied and applied these approaches to analyze biosignals, including surface electromyogram, accelerometry signal, and heart and pulmonary sound signals, and more recently, inserted in the scope of her Ph.D. degree, electroencephalogram and electrocardiogram. She is working towards the improvement of epileptic seizure anticipation using neuro-cardiovascular information fusion and dynamic classification.



JÚLIO MEDEIROS received the M.Sc. degree in biomedical engineering from the University of Coimbra, in 2019, where he is currently pursuing the Ph.D. degree with the Department of Informatics Engineering. His M.Sc. thesis topic was in the field of bio-signal processing and software engineering. His research interests include signal processing and biofeedback and machine learning applied to understanding and mitigating human error in software development tasks.



MAURO F. PINTO received the M.Sc. degree in biomedical engineering from the University of Coimbra, in 2018, where he is currently pursuing the Ph.D. degree. He is an Invited Assistant Professor at the Department of Informatics Engineering, University of Coimbra. His research interests include the prediction of EEG epilepsy seizures and multiple sclerosis disease progression and machine learning interpretability.



ANTÓNIO DOURADO (Member, IEEE) received the B.Sc. (five years) degree in electrical engineering from the University of Coimbra, Portugal, in 1977, the M.Sc. (DEA) and Ph.D. (DI) degrees in automatique from LAAS du CNRS, Université Paul Sabatier, Toulouse, France, in 1981 and 1983, respectively, and the Habilitation degree in informatics engineering from the University of Coimbra, in 1997. Since 2002, he has been a Full Professor with the Faculty of Sciences and

Technology of the University of Coimbra (FCTUC). He was a Co-Founder of the research Center for Informatics and Systems of the University of Coimbra (CISUC) and the Founder of the Adaptive Computation Group that he directed, until 2019. Before, since 1989, he has been an Assistant, an Auxiliary Professor, and an Associate Professor in the Informatics Engineering and Electrical Engineering Departments, University of Coimbra. He has been and is involved in national projects and international projects at European level (included some Networks of Excellence). He has been the Coordinator of the FP7 EU Project EPILEPSIAE. He is author or coauthor of more than 250 international publications in refereed journals, book chapters, and conferences. His specialization and main research interests include computational intelligence, neural networks, fuzzy systems, signal processing, data mining for medical and industrial applications, intelligent control, algorithms for EEG-ECG processing for epileptic seizures prediction and neurological diseases diagnosis and prevention, and information and communications technologies for personalized health care. He has been a member of more than 200 international scientific program committees, a reviewer of numerous international journals, and an evaluator of several international projects. He is a member of IEEE Computational Intelligence Society, IEEE System Man and Cybernetic Society, IEEE Engineering in Medicine and Biology Society, and ACM. He has been a Co-Founder of European Control Association and the President and a Co-Founder of Portuguese Association of Automatic Control (IFAC National Member). He is a member of the Ordem dos Engenheiros (Portuguese Engineering Professional Association).



MATTHIAS DÜMPELMANN received the Ph.D. degree in electrical engineering from Ruhr-Universität-Bochum, Germany. He worked in academic research at the Department of Epileptology, University Hospital, Bonn, and the Department of Eletrônica e Sistemas, Federal University of Pernambuco, Recife, Brazil. He has industrial experience at a SME in medical device development. His current position is at the Epilepsy Center, University Medical Center, Freiburg, Germany. His research interests include the registration and analysis of biosignals and medical images with emphasis in EEG, wearables, and brain imaging. He is an Affiliate of the IEEE Engineering in Medicine and Biology Society.



CÉSAR TEIXEIRA graduated in systems and computation engineering from the University of the Algarve, Portugal, in 2003. He received the Ph.D. degree in electronics engineering and informatics from the University of the Algarve, in 2008. He is currently an Associate Professor at the Departments of Informatics Engineering, Faculty of Sciences and Technology, University of Coimbra, Portugal. His expertise is on bio-signal processing, classification, and modeling. More precisely on EEG/ECG-based epileptic seizure prediction and ultrasound imaging. From 2009 to 2012, he took a scientific leadership role in the European project EPILEPSIAE, which have as main objective the development of a transportable device for epileptic seizure prediction. He is the Principal Investigator of the Portuguese national project RECOD (PTDC/EEI-EEE/5788/2020). He has participated in several other national, including HeartSafe (PTDC EEI-PRO 2857 2012), iCIS (CENTRO-07-0224-FEDER-002003), and Base (BASE-POCI-01-0145-FEDER-031581). He has participated in international projects, including EPILEPSIAE (FP7-ICT-2007-211713), Welcome (FP7-611223), Link (H2020-692023), and Welmo (H2020-ICT-2018-2). His publication record includes over 100 papers in refereed national and international journals and conferences. He served as the Vice Director for CISUC. He is an Associate Editor of *Research on Biomedical Engineering*.

...