Removing Noise from Extracellular Neural Recordings Using Fully Convolutional Denoising Autoencoders

Christodoulos Kechris[1]*, Alexandros Delitzas[1]*, Vasileios Matsoukas[1]*, Panagiotis C. Petrantonakis[2]

Abstract— Extracellular recordings are severely contaminated by a considerable amount of noise sources, rendering the denoising process an extremely challenging task that should be tackled for efficient spike sorting. To this end, we propose an end-to-end deep learning approach to the problem, utilizing a Fully Convolutional Denoising Autoencoder, which learns to produce a clean neuronal activity signal from a noisy multichannel input. The experimental results on simulated data show that our proposed method can improve significantly the quality of noise-corrupted neural signals, outperforming widely-used wavelet denoising techniques.

I. INTRODUCTION

Monitoring the activity of single or multiple neurons, by placing electrodes and registering extracellular recordings, is a key factor in understanding brain's mechanisms [1]. The desired goal when analyzing extracellular recordings is to identify which neurons activate by firing spikes, a technique known as Spike Sorting (SS) [2]. The SS process traditionally involves the following steps [3]. First, the recorded raw data are filtered since noise contamination, attributed mainly to nearby neural activity but also to hardware [1], can negatively influence the SS performance [4]. Second, a spike detection procedure is applied by manually or automatically setting an amplitude threshold on the filtered data. Third, a feature extraction step aims to separate the spike shapes. Finally, a clustering algorithm is utilized to group spikes with similar features into clusters, distinguishing the spike's source neuron.

For the filtering step, a simple bandpass filter in the range between 300 and 3000 Hz is commonly employed. The lower cutoff frequency removes the slow components related to Local Field Potential (LFP), while the upper prevents the visual distortion of spikes from high frequency noise. However, the use of standard bandpass filtering inserts distortions to the waveform shapes [5]. To overcome this limitation, wavelet denoising techniques have been widely adopted in neural signal processing. In these techniques, the signal is represented in the time-scale domain, thereby allowing a fine analysis of the signal's frequency components without loss of temporal information, in contrast to the traditional Fourier transform [6]. Two of the most prominent wavelet denoising methods for neural signals are the Discrete Wavelet

Transform (DWT) and the Stationary Wavelet Transform (SWT) [7], [6].

While traditional methods have performed fair enough, Deep Learning (DL) approaches have attracted attention because of their capacity to comprehend complex data representations. Biological systems have been intensively addressed by sophisticated DL approaches with impressive results including physiological signal denoising. In [8], a deep encoder-decoder algorithm is used for denoising 2D images of neuronal physiology, drastically improving the Signal-to-Noise Ratio (SNR) distribution. Convolutional Autoencoders have also showed promising potential in electoencephalogram (EEG) filtering, surpassing the traditional bandpass approach [9]. Moreover, in [10], Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks outperformed traditional wavelet methods in electrocardiogram (ECG) denoising. In [11] the authors experimentally show the superiority of the Fully Convolutional Denoising Autoencoder (FC-DAE) over the Dense and Convolutional Denoising Autoencoders in ECG noise reduction.

In our work, we propose a Fully Convolutional Denoising Autoencoder (FC-DAE) to eliminate noise distortion in extracellular neural activity recordings and retrieve a signal as close to ground truth as possible. We base our experiments on simulated datasets, evaluate our approach with respect to SNR and Root Mean Square Error (RMSE) metrics and compare with Wavelet Transform methods.

All the datasets and the code discussed in this paper are publicly available at Github[1].

II. MATERIALS AND METHODS

A. Proposed Autoencoder Network

An end-to-end Fully Convolutional Denoising Autoencoder was used in order to filter the noisy input signal. The FC-DAE receives as input the multichannel recording and produces a clean multichannel output containing only the neuronal activity signal. The network architecture is illustrated in Figure 1. In this work, we completely omit the Fully Connected layer that is found in CNNs, adopting the rationale proposed in [11]. This approach presents important benefits, mainly being computationally efficient and preserving the local information [12]. Additionally, we do not include any MaxPooling layers, since they may fail to preserve important structural details [11]. The network's encoder composes of five one-dimensional convolutional

*These authors contributed equally
[1]Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, GR 54124, Thessaloniki, Greece
[2]Information Technologies Institute, Centre for Research and Technology – Hellas (CERTH), GR 57001, Thessaloniki, Greece. (correspondence e-mail: ppetrant@iti.gr)

[1]https://github.com/AlexDelitzas/fcdae-neural-signal-denoising

layers while the decoder has a symmetric structure composed of five one-dimensional transposed convolutions.

The input recording is initially zero-meaned and then normalized in the $[-1, 1]$ range. Since the FC-DAE operates on batches of data, the recording is split into samples using non-overlapped windows of size L. The FC-DAE jointly processes the channels, taking advantage of their interrelations, and outputs the n-channel filtered sample of length L.

The window length was selected to be 512 samples, which corresponds to a duration of 51.2ms considering the frequency sampling of $10kHz$. The number of channels depends on the probe used for recording the extracellular action potentials (EAP).

The kernel size was set to 3 for all convolutional layers, since after experimentation it was observed that the small kernel achieved good results and ensured computational efficiency thanks to the small number of parameters. The ELU activation function was used for all the layers with the exception of the last one where the linear activation was used. The learning rate was set to 0.001 and the Mean Square Error loss function was utilized.

B. SEAset: Simulated Extracellular Action Potential Dataset

For the needs of this study a simulated dataset of extracellular recordings was generated. Since the supervised learning scheme is employed to train the FC-DAE, a ground truth clean recording is necessary. The patch-clamp approach has been used as a method to provide real-life EAPs together with the corresponding intracellular signal that can be considered as the ground truth. However, they are limited by the small number of cells that can be recorded simultaneously [13]. Thus, we opted for the simulation of the recordings to ensure control over the recording conditions.

The EAPs were simulated with the MEArec Python library [13]. Neuronal models of the L5 area were used to simulate the activity of three inhibitory and three excitatory neurons. To ensure the generation of realistic recordings the distance-correlated noise was added as described in [13] and the bursting and overlap settings were activated.

A simulated tetrode probe was selected to record the EAPs resulting in a four-channel input signal. The sampling frequency was set to $10kHz$. The ground truth clean signal was generated by aggregating the neuronal activity of all six cells into one single signal. Since the network's output is multichannel, the same aggregated signal was used as ground truth for all four channels.

To evaluate the robustness of our approach under different levels of noise, we generated signals contaminated with noise standard deviation ([13]) of 7, 9, 15 and 20 μV. For each noise level a set of ten 5s recordings was generated and for every recording, a randomly generated set of seeds was used for the MEArec's spiketrains, templates, convolution and noise seeds. For a better intuition on the SEAset recordings' noise contamination, we examined the observed SNR values. Let $\hat{X}^n_{i_{rec},l}$ be the $l-th$ noisy channel of the $i_{rec}-th$ recording contaminated with noise of level n μV, where $l \in [1,2,3,4]$, $i_{rec} \in [1,...,10]$ and $n \in [7,9,15,20]$ μV.

The corresponding ground truth clean signal is noted as $X^n_{i_{rec},l}$. We calculate the observed SNR of $\hat{X}^n_{i_{rec},l}$ as:

$$SNR^n_{i_{rec},l} = 10 \log_{10} \frac{\sum(X^n_{i_{rec},l})^2}{\sum(\hat{X}^n_{i_{rec},l} - X^n_{i_{rec},l})^2}$$

We then calculate the average $SNR^n_{l_{avg}}$ across all ten recordings for each channel l, $SNR^n_{l_{avg}} = \frac{1}{10} \sum_{i_{rec}} SNR^n_{i_{rec},l}$. Finally, for each noise level n μV we calculate the minimum and maximum SNR values across all four channels (Table I).

TABLE I

MIMIMUM AND MAXIMUM OBSERVED SNR FOR EACH NOISE LEVEL

Noise Level (μV)	Minimum SNR (dB)	Maximum SNR (dB)
7	0.60	2.87
9	-1.36	0.77
15	-5.29	-3.59
20	-7.02	-5.97

C. DWT and SWT filtering methods

Our method was evaluated in comparison with two wavelet denoising methods, the Discrete Wavelet Transform (DWT) and the Stationary Wavelet Transform (SWT). In general, all wavelet-based denoising techniques involve three steps: 1) Decomposition: Choosing a mother wavelet and a maximum decomposition level and then decomposing the signal to obtain the detail and approximation coefficients for each level, 2) Thresholding: Computing threshold values and applying soft or hard threshold to the detail coefficients for each level, 3) Reconstruction: Reconstructing the signal by using the modified coefficients.

The main difference between DWT and SWT lies upon the decimation operation. Specifically, the number of coefficients of each level is half that of the preceding level in DWT, while the number of coefficients is the same for each level in SWT. The retention of redundant data renders the SWT translation-invariant, which is a useful property for denoising neural recordings [14].

The effectiveness of wavelet denoising algorithms strongly depends on the implementation choices at the different steps, including the decomposition level, the mother wavelet, the threshold method and its type (i.e., how it is applied). For this reason, we explored several combinations of these parameters during the parameter tuning process. Inspired by [6], we explored five orthogonal and biorthogonal mother wavelets: Haar, Coiflet2 (Coif2), Daubechies4 (Db4), Biorthogonal6.8 (Bior6.8) and Symlet7 (Sym7). Regarding the threshold selection method, we explored four options [15], [16]: universal, minimax, SURE and heuristic SURE. The aforementioned threshold selection methods were combined with soft and hard type of thresholding. Furthermore, the signals were decomposed at four and five decomposition levels. Consequently, taking into account the possible values of the four implementation choices, our parameter space consisted of 80 different combinations.
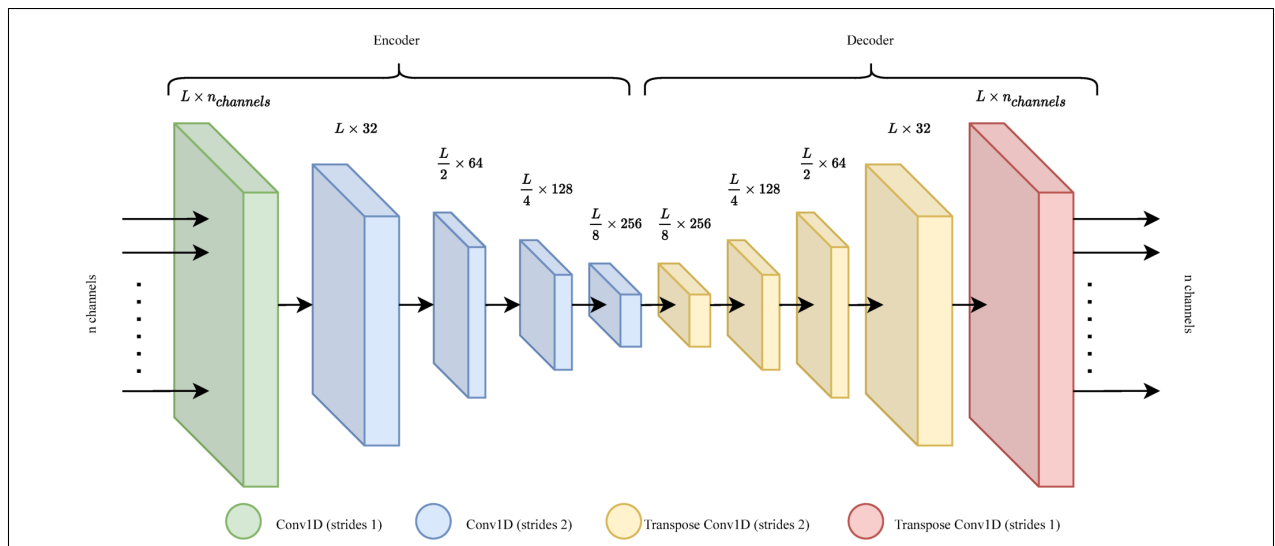
Fig. 1.   The Fully Convolutional Denoising Autoencoder architecture.

III. EXPERIMENTAL SETUP

We train and evaluate our model on the SEAset data employing the leave-one-out (LOO) strategy. This enables the evaluation of the model on previously unseen recordings and the examination of its generalization ability. For each noise level, the evaluation process was repeated ten times; each time, the training set consisted of nine 5s recordings, which provided a sufficient amount of samples for training, and the test set consisted of one 5s recording.

The same LOO evaluation process was followed for the parameter selection of the DWT and SWT wavelet denoising methods. Specifically, in each LOO iteration, the parameters were tuned on the training set by means of searching the parameter combination which minimizes the Mean Square Error between the denoised and the ground truth signal. Then, the performance of the methods was evaluated on the test set.

The per channel SNR improvement [17] and Root Mean Square Error (RMSE) are utilized as performance metrics. Each value occurs as the average of all ten iterations of the LOO process respectively. The SNR improvement provides an intuition on how well the filtering method removed the input signal's noise, while the RMSE evaluates the similarity between the method's output and the ground truth. This is important in spike sorting applications where the original action potential form can contain useful information about the source neuron and should be maintained by the preprocessing filtering method.

IV. EXPERIMENTAL RESULTS

Figure 2 presents an example of the denoising outputs for the FC-DAE, DWT and SWT and allows a qualitative examination of their performance. In the left column, a section of a test recording of noise 9 μV is processed by the three methods. The top plot displays the noisy input, while the second one the ground truth signal. The following three plots present the four-channel outputs of the FC-DAE, the

DWT and the SWT respectively. The DWT and SWT output signals are pretty similar and manage to isolate the spikes relatively well. The FC-DAE's output, however, is visibly cleaner. In areas of no spiking activity, the autoencoder manages to completely suppress the noise, in contrast to the wavelet denoising methods where some noise is still present. At the same time, the FC-DAE's spikes' shape is cleaner and quite close to the ground truth.

A close-up look at the signals' section between the samples 120 and 200 is plotted in the right column of Figure 2. Here, two consecutive EAPs have been recorded, with very little space between them. For legibility reasons, we only display the channel in which the DWT and SWT better performed, i.e. the second one. The FC-DAE output presents significantly more similarity to the ground truth, compared to the DWT and SWT signals, and did not fail to produce the second spike, unlike the wavelet methods.

The SNR improvement and RMSE for all four examined noise levels are summarized in Figure 3. Here, the initial comments made above are confirmed, for all noise levels. The FC-DAE manages to outperform the wavelet denoising methods both in terms of denoising and reconstructing the clean neuronal activity signal. Additionally, the FC-DAE presents significantly less variance in its estimations' RMSE across the four channels, in contrast to the DWT and SWT.

V. CONCLUSION

In this work, a Deep Fully Convolutional Denoising Autoencoder was developed for denoising EAP recordings. A realistic simulated dataset of neuronal recordings from the L5 area was created and then was used to train and evaluate the network. In the experiments conducted, our proposed method showed that it can achieve a substantial quality improvement of noisy recordings with various noise levels while simultaneously contributing to the effective reconstruction of the EAP. The proposed FC-DAE presented superior performance in comparison with the widely-used
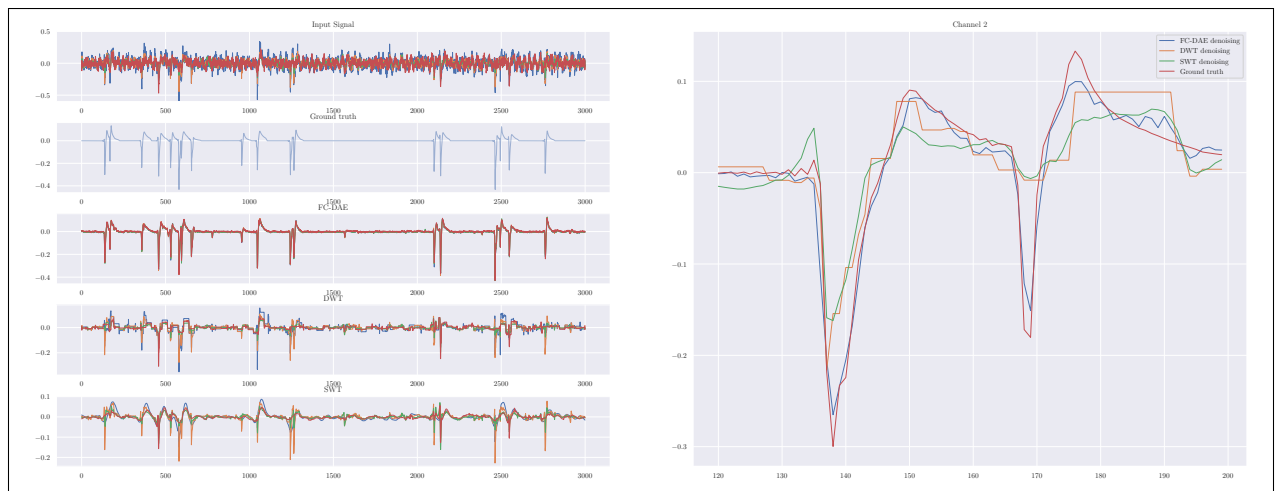
Fig. 2. Left: Example output of the FC-DAE, DWT and SWT methods for a section of recordings of noise level 9 μV. Right: Recording details from sample 120 to 200.
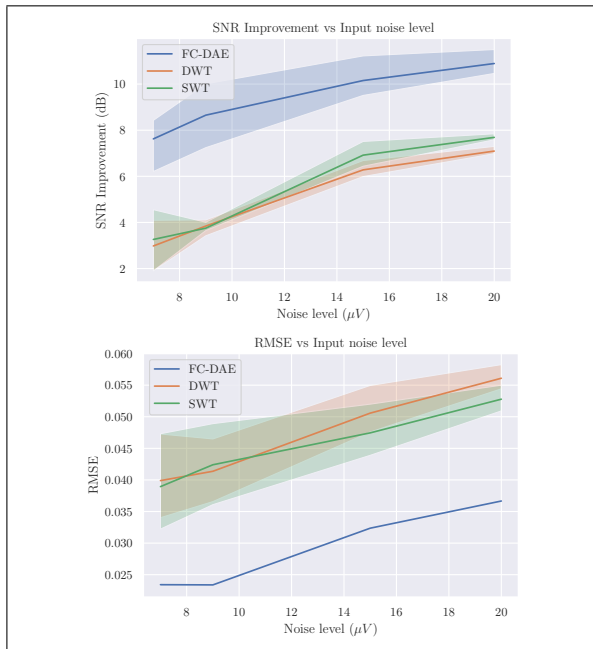

Fig. 3. Up: Average SNR Improvement across all channels for each method (bold line) and the minimum and maximum values (transparent range). Down: Similarly for each method's RMSE performance.

wavelet denoising methods of DWT and SWT. Future work will include denoising real recordings using the FC-DAE trained on simulated datasets. Further, evaluating the performance of spike sorting on FC-DAE pre-processed recordings will allow a more thorough examination of the denoising robustness. Finally, it should be investigated whether one model can handle EAPs from different brain areas with different spike shapes, or an area-specific model approach would be more appropriate.

REFERENCES

[1] S. Gibson, J. W. Judy, and D. Markovic, "Spike sorting: The first step in decoding the brain: The first step in decoding the brain," *IEEE Signal Processing Magazine*, vol. 29, no. 1, p. 124–143, Jan 2012.

[2] R. Quiroga, "Spike sorting," *Scholarpedia*, vol. 2, no. 12, p. 3583, 2007.

[3] H. G. Rey, C. Pedreira, and R. Quian Quiroga, "Past, present and future of spike sorting techniques," *Brain Research Bulletin*, vol. 119, p. 106–117, Oct 2015.

[4] M. S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials," *Network: Computation in Neural Systems*, vol. 9, no. 4, pp. R53–R78, 1998.

[5] A. Wiltschko, G. Gage, and J. Berke, "Wavelet filtering before spike detection preserves waveform shape and enhances single-unit discrimination," *Journal of Neuroscience Methods*, vol. 173, pp. 34–40, 08 2008.

[6] G. Baldazzi, G. Solinas, J. D. Valle, M. Barbaro, S. Micera, L. Raffo, and D. Pani, "Systematic analysis of wavelet denoising methods for neural signal processing," *Journal of Neural Engineering*, vol. 17, no. 6, p. 066016, dec 2020.

[7] R. J. Brychta, S. Tuntrakool, M. Appalsamy, N. R. Keller, D. Robertson, R. G. Shiavi, and A. Diedrich, "Wavelet methods for spike detection in mouse renal sympathetic nerve activity," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 1, pp. 82–93, 2007.

[8] J. Lecoq, M. Oliver, J. H. Siegle, N. Orlova, and C. Koch, *Removing independent noise in systems neuroscience data using DeepInterpolation*, Oct 2020.

[9] N. M. N. Leite, E. T. Pereira, E. C. Gurjão, and L. R. Veloso, "Deep convolutional autoencoder for eeg noise filtering," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2018, pp. 2605–2612.

[10] C. Arsene, "Complex deep learning models for denoising of human heart ecg signals," *ArXiv*, vol. abs/1908.10417, Jun 2020.

[11] 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. 60 806–60 813, 2019.

[12] S.-W. Fu, Y. Tsao, X. Lu, and H. Kawai, "Raw waveform-based speech enhancement by fully convolutional networks," in *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2017, pp. 006–012.

[13] A. P. Buccino and G. T. Einevoll, "Mearec: a fast and customizable testbench simulator for ground-truth extracellular spiking activity," *Neuroinformatics*, vol. 19, no. 1, pp. 185–204, 2021.

[14] D. Farina, W. Jensen, and M. Akay, *Wavelet Denoising and Conditioning of Neural Recordings*, 2013, pp. 173–182.

[15] B. Ergen, *Signal and Image Denoising Using Wavelet Transform*. InTech, 04 2012.

[16] I. M. Johnstone and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 59, no. 2, pp. 319–351, 1997.

[17] E. Fotiadou and R. Vullings, "Multi-channel fetal ecg denoising with deep convolutional neural networks," *Frontiers in Pediatrics*, vol. 8, 2020.