Comparison of classical EEG source analysis with deep learning

Jakob Winkler[1], Christian Uhl[1], Stefan Geißelsöder[1], Tim Erdbrügger[2], and Carsten Wolters[2]

[1] University of Applied Sciences Ansbach, Center for Signal Analysis of Complex Systems (CCS), Residenzstr. 8, 91522 Ansbach, Germany,
winkler20613@hs-ansbach.de
[2] University of Münster, Institute for Biomagnetism and Biosignalanalysis (IBB), Malmedyweg 15, 48149 Münster, Germany

Abstract. This paper discusses the challenges and methods for source reconstruction of evoked potentials using deep learning in the context of electroencephalography (EEG). We propose the use of deep learning to address known challenges and improve traditional approaches. We explain the creation of a suitable dataset for solving the inverse problem, including the simulation of neural activity and the use of lead field matrices for the forward solution. Furthermore, we undertake a comparative analysis of some initial deep learning models with similar classical methods.

Keywords: data simulation, deep learning, EEG, inverse problems, source localization

1 Introduction

Electroencephalography (EEG) is a widely used, non-invasive method for measuring neuronal activity. It is a particularly attractive method in the areas of clinical diagnosis and research due to its low cost and rapid implementation. The significant strength of EEG lies in its high temporal resolution [1].

In order to maintain the advantage of high temporal resolution while improving spatial resolution, attempts are made to localise the source of the neuronal signal in the brain. Mathematical models of the brain and skull based on measured signals are used for this purpose [2], [3]. However, it is important to note that all classical methods for EEG source localisation are always estimates, as the problem is inherently not uniquely solvable. Since only a small amount of measurement electrodes are attached to the surface of the head, the measured signal cannot be precisely assigned to one of thousands possible sources of activity, resulting in several possible activity patterns that generate the same measurement signal [4].

In numerous medical applications and other areas, deep learning has already proven its ability to significantly support human decision-making and provide valuable supplementary information for medical decisions [5], [6]. Consequently, deep learning could also offer new advantages in source localization or extend classical methods.

This paper discusses and evaluates some of the approaches to solving the source localisation of evoked potentials.

2 Methods

2.1 Data simulation

Before we can use machine learning or deep learning approaches, we need a suitable dataset for our problem. This dataset must contain x: the potential measured on the scalp surface and y: the underlying neuronal activity responsible for the measured potential. Since we cannot know the exact location of a neuronal activity in a realistic test environment, the data must be simulated, where the exact sources are known and can be freely chosen. In this work, we only use simulations with a single origin of source per sample.

There are various methods for generating such a dataset. In our work, we started the simulations with the MNE Python library, an open-source package designed for visualisation and analysis of human neurophysiological data [7]. We chose an average reference head model and computed a boundary element model (BEM) to solve the forward problem. Potential sources were reduced to points on the surface of the cortex [8].
The Institute for Biomagnetism and Biosignal Analysis (IBB) in Münster led by Carsten Wolters provided an anonymised dataset of 19 subjects with real EEG measurements, volumetric source spaces, as illustrated in Figure 1, and an individualised lead field matrix for the forward simulation:

The lead field matrices were calculated with DUNEuro using a six-layer finite element method (FEM) with geometrically adapted hexahedra [9]. With this approach, we can accurately track each step from the measurement conditions to the calculation of the individualised forward simulations and, for example, exchange the algorithm for the forward model. Another advantage is that we can directly compare the trained model with the measurements and the expected true source.

Initial studies using MNE's algorithms and models showed that it is challenging to track the results and compare them with real data. As a result, the second method of data generation, based on individualised lead field matrices for forward simulation, is used to train and evaluate the presented deep learning algorithms.

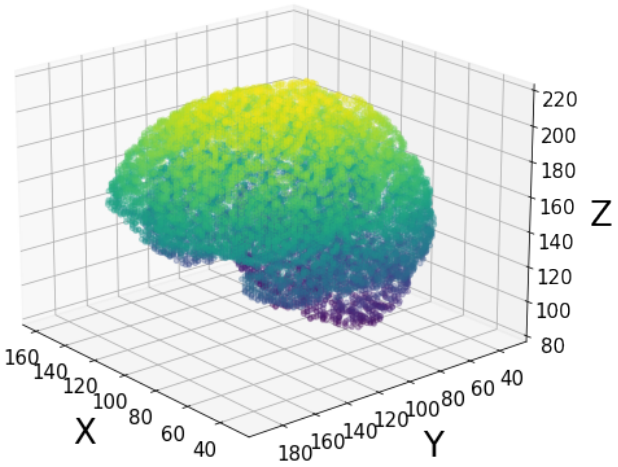

Fig. 1. Volumetric source space from proband 1. Each point represents a possible source of activity. For better visibility, the points are color graded in the z-direction.

To further improve the simulated data for a neural network, we added a Gaussian propagation of the neural activity. Instead of a single active source in the brain, a point is selected and the surrounding possible sources are also simulated with a portion of the original activation. This makes the data more realistic and improves the learning of the neural network, as the error function for training compares the predicted activity values at each possible source point with the actual values and calculates a difference. The greater the number of positive or active values, the more stable the error. This is particularly evident in the case of predicted activities that are close to the correct position. The amount of activation follows a Gaussian distribution and is dependent on the distance x to the original source point and is afterwards scaled:

$$\text{Gauss}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

In addition, we added some noise to the data to prepare the model for real data, which will inevitably contain some noise [10]. For this we use Gaussian white noise with a constant power density across all frequencies and zero mean [11]. To do this, we calculate the average power of the signal and then the required average power of the noise based on the average signal power to obtain a desired signal-to-noise (SNR) factor:

$$\text{SNR}_{\text{dB}} = 10 \log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \tag{2}$$

The noise is then randomly generated from a normal distribution with a mean of 0 and a standard deviation for each channel, and added to the original signal data. This process allows the creation of a suitable dataset for each desired signal-to-noise ratio.

A potential future direction of research is the use of a distinct, precisely calibrated noise for each subject, given that each subject exhibits a unique noise behaviour. This noise would then be calculated by the reference measurement without stimulus.

2.2 Neural network

Once the requisite dataset has been prepared, it can be used to train a model using deep learning algorithms. Two possible methods are fully-connected neural networks (FCNN) and convolutional neural networks (CNN), which have been implemented by other authors in previous studies [12], [13], [14]. In this paper, we have used a fully-connected neural network. The network's architecture comprises an input layer with 54 neurons, three hidden layers with 4096, 8192 and 12024 neurons, respectively, and an output layer with 14909 neurons, corresponding to the number of possible sources in the source space (see Table 1):

Table 1. Architecture of the neural network used

Layer	Number neurons
Input layer	54
1. hidden layer	4096
2. hidden layer	8192
3. hidden layer	12024
Output layer	14909

Each hidden layer is followed by a batch normalisation layer [15]. For hidden layers 1 and 3, a dropout factor of 0.2 was employed, while a dropout factor of 0.3 was applied to layer 2 [16]. The purpose of the batch normalisation layer is to enhance and stabilise the training process. The addition of dropout was intended to facilitate the network's ability to generalise and deliver enhanced performance, particularly in the presence of noise. The mean squared error (MSE) was employed as the loss function with the ground truth y_i and the predicted value $\hat{y}_i$:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3}$$

The adam optimizer, with a learning rate of 0.001, is used in this study [19]. The network architecture is not the primary focus of the work and has not yet been optimised to its full potential. Consequently, there is still considerable scope for improvement in performance.

2.3 Evaluation

In order to validate the method, it is necessary to compare it with a classical method with comparable prior information. For this purpose, a dipole scan is utilized. A dipole scan iterates over every possible source in the source space and calculates a forward solution at this point using the lead field matrix [17], [18]. The source that is being looked for is the one that best describes the measured signal data. This method was chosen for comparison because, like the neural network, it has the original lead field of the subject available. In the case of the neural network, the lead field is implicitly given because it learns from the simulation data that was constructed using the lead field. Consequently, both methods essentially have the same prior information, and the neural network has no advantage because it has learned the data simulation with great precision. In the neural network, the point with the highest estimated activity is taken as the prediction.

The evaluation involves the creation of 100 data points for each SNR, with the localisation errors of the methods averaged over these 100 samples. The Localization Error (LE) is defined as:

$$\mathrm{LE} = \|\mathbf{r}_{\mathrm{true}} - \mathbf{r}_{\mathrm{peak}}\|_2 \tag{4}$$

3 Results

The SNR values -10, -7.5, -5, -2.5, 0, 2.5, 5, 10, 15 and 20, as well as no noise at all ($SNR \to \infty$) were considered in the evaluation.

If only the dipole scan on the selected SNR range is considered, it can be seen that from about 20 dB the localisation error approaches zero (see Figure 2). By adding the Gaussian Propagation the prediction quality becomes slightly worse across all SNRs, but the rough curve remains the same (see Figure 3). This shows that the dipole scan cannot handle regional activities optimally, as it is optimised for individual dipole sources. Here, however, the dipole Scan produces perfect results with sufficient signal content.

For each of the SNRs, a neural network was trained on a dataset with the corresponding SNR. The dataset consisted of 50,000 samples each time and each model was trained for 50 epochs. We chose 50 epochs as a simple compromise between train loss and overfitting. The loss values of the last training iteration are shown in Table 2:

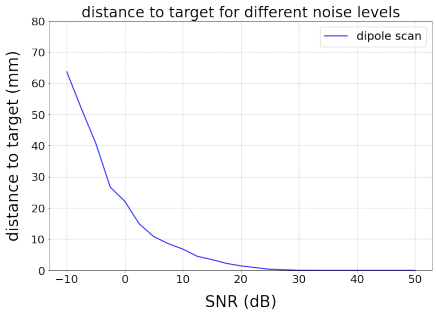

Fig. 2. Dipole scan localization error on different SNR´s without Gaussian propagation

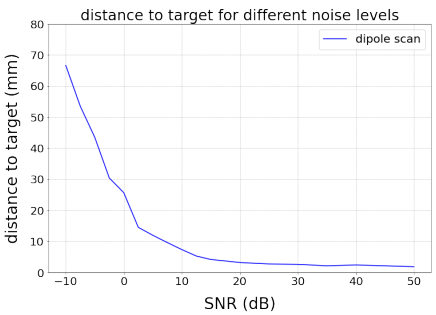

Fig. 3. Dipole scan localization error on different SNR´s with Gaussian propagation

Table 2. Loss values after 50 episodes of training

SNR	Train loss	Test loss
-10	16.93	1186582.14
-7.5	16.87	567281.55
-5	16.01	173864.21
-2.5	14.12	253.95
0	10.82	113.55
2.5	8.02	121818529.26
5	6.05	807080921.18
10	4.78	187600850.37
15	4.13	3813368.04
20	3.55	39060237.59
$\rightarrow \infty$	2.66	58650.25

We can see a stable decrease in the training error with increasing SNR. However, the error on test data is very inconsistent and high. When looking at the loss curves (Figure 4 and 5), it can be seen that the training error also decreases steadily within a training run. However, the test error increases in a random manner. The progression of the two loss curves indicates overfitting of the model to the training data. As model optimisation is not the focus of the work, we nevertheless use the models after 50 epochs for further evaluation. In future works this will be addresed with regularization methods to obtain more stable results.

In Figure 6 a single sample with no noise is shown. On the left there is the original ground truth data and on the right the predicted activity from the neural network. Looking at single samples, the prediction accuracy even on test data is quite precise.

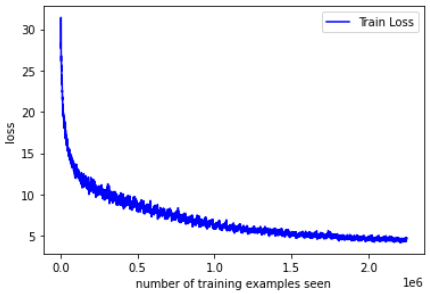

Fig. 4. Train loss of the 10dB SNR model during 50 epochs of training

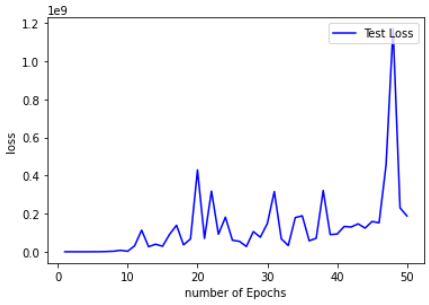

Fig. 5. Test loss of the 10dB SNR model during 50 epochs of training

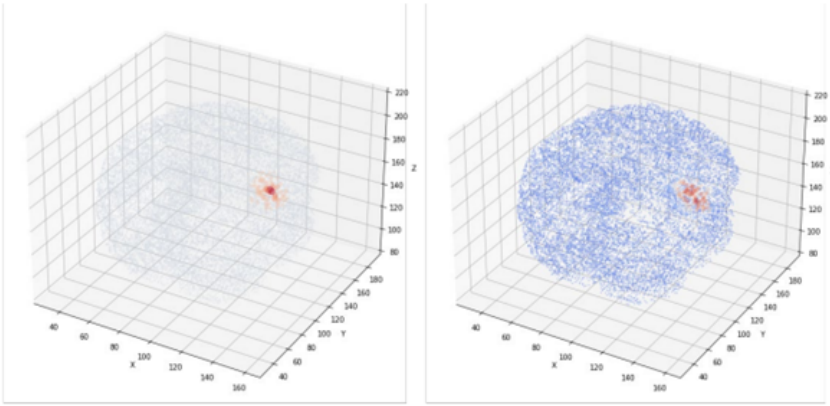

Fig. 6. Ground truth sample (left) and predicted activity from the neural network (right)

Figure 7 shows the localisation error at different SNR ratios. The blue graph shows a model that is trained with this noise level and the red graph shows the LE over all SNRs of a model that was trained completely without noise. It can be seen here that models that contain noise in the training data also deliver better results on noisy data without losing quality at high SNRs. Figure 8 shows in red again the model without noise compared to a model trained on 10 dB SNR. The model without noise is better at high SNR ratios, but the blue model is better in the 10 dB SNR range. Since we are always confronted with noise in real applications, we see the use of noise in the training data as an important step towards improving deep learning models for source localisation.

Figure 9 shows the localisation error in mm of the respective models compared to a dipole scan at all noise levels. The results are very similar with the neural networks performing slightly better across all SNRs.

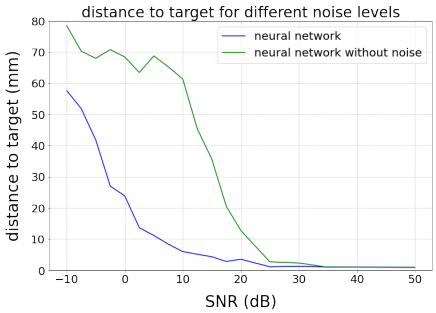

Fig. 7. LE of models trained on the specific SNR (blue) and LE of a model trained without noise (green)

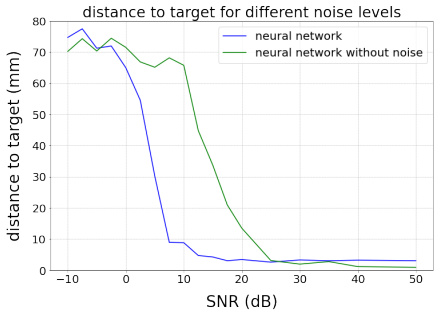

Fig. 8. Localization error of a model trained on 10 dB SNR (blue) and LE of a model trained without noise (green)

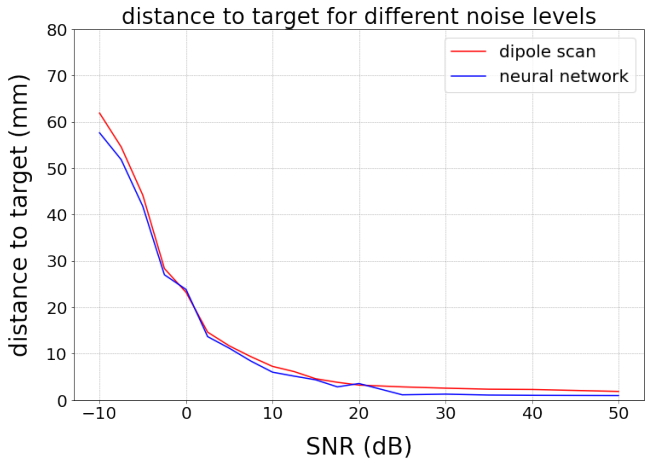

Fig. 9. Localization error of the neural networks (blue) and the LE of a dipole scan over different SNRs (red)

4 Discussion and outlook

The input of the neural network is fixed to the number of electrodes used, with a range of approximately 8 to 300. The number of outputs of the network correlates with the number of sources to be predicted in the source space, with a range of approximately 1,000 to 20,000. Consequently, the input dimension is always significantly smaller than the output dimension. In our example, this ratio is:

$$\mathbb{R}^{54} \rightarrow \mathbb{R}^{14909} \tag{5}$$

Potential solutions for this include the use of convolutional neural network (CNN) models or specifically trained autoencoders to reduce the dimensionality of the prediction. Furthermore, the incorporation of Gaussian propagation serves

to mitigate the dimensionality issue indirectly. By expanding the active region in the source space, the effective dimension of the predicted space is diminished to a slight degree. In this study, we solely employed Gaussian propagation to address the dimensionality challenge.

It is to be expected that the dipole Scan localises very well on datasets without noise. Without Gaussian propagation and without noise, this method has perfect accuracy as it also has all the necessary information. After adding Gaussian propagation, the accuracy decreases slightly. However, the advantage of the neural network over the dipole scan lies in the localisation of noisy data, as the network can learn to better deal with the noise.

Unfortunately a model trained for a specific noise level only produces very good results in the noise domain it is trained on. This could be solved by estimating the SNR of a measurement in reality and then using the model with the exact noise level for localisation. Another alternative would be to train a model with many different noise levels so that it performs well at all SNRs. However, it should be noted that the SNR in real measurements is always in a constant range around 15 dB and the models could then be optimised precisely for this level.

This work shows that deep learning models have the potential to achieve the same prediction quality as a dipole scan with a perfect lead field matrix, and possibly exceed it. The models were not particularly optimised for this work and it is to be expected that even better results can be achieved by optimising the model parameters. Despite the good performance of the deep learning models on the simulated data, validation on real data is still necessary to confirm the results.

5 Acknowledgement

This study was supported by ERA PerMed as project ERAPERMED2020-227 PerEpi (Bundesministerium für Gesundheit, project ZMI1-2521FSB006), by the Deutsche Forschungsgemeinschaft (DFG), projects WO1425/10-1, by the DAAD project 57663920 and by the German Federal Ministry of Education and Research (BMBF, Funding number: 05M20WBA)

References

1. Luck, Steven: An Introduction to The Event-Related Potential Technique. The MIT Press (2005).
2. Awada, K. A., Jackson, D. R., Williams, J. T., Wilton, D. R., Baumann, S. B., and Papanicolaou, A. C.: Computational aspects of finite element modeling in EEG source localization. IEEE transactions on biomedical engineering 44 (1997).
3. Buchner, H., Waberski, T. D., Fuchs, M., Wischmann, H. A., Wagner, M., and Drenckhahn, R.: Comparison of realistically shaped boundary-element and spherical

head models in source localization of early somatosensory evoked potentials. Brain topography 8 (1995).
4. Koles, Z. J.: Trends in EEG source localization. Electroencephalography and Clinical Neurophysiology 106 (1998).
5. Sandeep Kumar, E., Satya Jayadev, P. Deep Learning for Clinical Decision Support Systems: A Review from the Panorama of Smart Healthcare. In: Dash, S., Acharya, B., Mittal, M., Abraham, A., Kelemen, A. (eds) Deep Learning Techniques for Biomedical and Health Informatics. Studies in Big Data, vol 68. Springer, Cham. (2020). https://doi.org/10.1007/978-3-030-33966-1
6. Talaat, F.M., El-Sappagh, S., Alnowaiser, K. et al. Improved prostate cancer diagnosis using a modified ResNet50-based deep learning architecture. BMC Med Inform Decis Mak 24, 23 (2024). https://doi.org/10.1186/s12911-024-02419-0
7. MNE — MNE 1.5.0 documentation. Accessed 22 August 2023. https://mne.tools/stable/index.html
8. Hämäläinen, M.S., Sarvas, J. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Transactions on Biomedical Engineering, 36, 165-171. (1989).
9. M. Antonakakis, S. Schrader, Ü. Aydin, A. Khan, J. Gross, M. Zervakis, S. Rampp, C. H. Wolters: Inter-Subject Variability of Skull Conductivity and Thickness in Calibrated Realistic Head Models, NeuroImage, Volume 223 (2020). https://doi.org/10.1016/j.neuroimage.2020.117353
10. Sabine Leske, Sarang S. Dalal: Reducing power line noise in EEG and MEG data via spectrum interpolation. Neuroimage; 189: 763–776 (2019). 10.1016/j.neuroimage.2019.01.026
11. Norbert, Wiener, Cybernetics or Control and Communication in the Animal and the Machine. The MIT Press. (2019). https://doi.org/10.7551/mitpress/11810.001.0001
12. A. Delatolas, M. Antonakakis, C. H. Wolters, M. Zervakis.: EEG Source Analysis with a Finite-Element-based Convolutional Neural Network - feCNN. IEEE-EMBS (2022).
13. Hecker, L., Rupprecht, R., van Tebartz Elst, L., and Kornmeier, J.: ConvDip: A Convolutional Neural Network for Better EEG Source Imaging. Front. Neurosci. 15 (2021).
14. Roman Tankelevich: Inverse Problem's Solution Using Deep Learning: An EEG-based Study of Brain Activity. Part 1 - rel. 1.0. Researchgate (2019).
15. Sergey Ioffe, Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. ArXiv; 1502.03167. (2015).
16. Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research 15, 56, Page 1929-1958. (2014). http://jmlr.org/papers/v15/srivastava14a.html
17. C. C. Wood: Application of dipole localization methods to source identification of human evoked potentials, Ann. New York Acad. Sci., vol. 388, pp. 139-155, (1982).
18. M. Scherg: Fundamentals of dipole source potential analysis in Auditory Evoked Magnetic Fields and Potentials, Basel:Karger, vol. 6, (1989).
19. Diederik P. Kingma and Jimmy Ba: Adam: A Method for Stochastic Optimization. arXiv; 1412.6980 (2017). https://arxiv.org/abs/1412.6980