

Article

MEG Source Localization via Deep Learning

Dimitrios Pantazis ^{1,†} and Amir Adler ^{1,2,*,†}

¹ McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; pantazis@mit.edu

² Electrical Engineering Department, Braude College of Engineering, Karmiel 2161002, Israel

* Correspondence: adleram@braude.ac.il

† These authors contributed equally to this work.

Abstract: We present a deep learning solution to the problem of localization of magnetoencephalography (MEG) brain signals. The proposed deep model architectures are tuned to single and multiple time point MEG data, and can estimate varying numbers of dipole sources. Results from simulated MEG data on the cortical surface of a real human subject demonstrated improvements against the popular RAP-MUSIC localization algorithm in specific scenarios with varying SNR levels, inter-source correlation values, and number of sources. Importantly, the deep learning models had robust performance to forward model errors resulting from head translation and rotation and a significant reduction in computation time, to a fraction of 1 ms, paving the way to real-time MEG source localization.

Keywords: magnetoencephalography; deep learning; source localization; inverse problems



Citation: Pantazis, D.; Adler, A. MEG Source Localization Via Deep Learning. *Sensors* **2021**, *21*, 4278. <https://doi.org/10.3390/s21134278>

Academic Editors: Gianluca Esposito and Andrea Bizzego

Received: 29 April 2021

Accepted: 17 June 2021

Published: 22 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Accurate localization of functional brain activity holds promise to enable novel treatments and assistive technologies that are in critical need by our aging society. The ageing of the world population has increased the prevalence of age-related health problems, such as physical injuries, mental disorders, and stroke, leading to severe consequences for patients, families, and the health care system. Emerging technologies can improve the quality of life of patients by (i) providing effective neurorehabilitation, and (ii) enabling independence in everyday tasks. The first challenge may be addressed by designing neuromodulatory interfacing systems that can enhance specific cognitive functions or treat specific psychiatric/neurological pathologies. Such systems could be driven by real-time brain activity to selectively modulate specific neurodynamics using approaches such as transcranial magnetic stimulation [1,2] or focused ultrasound [3,4]. The second challenge may be addressed by designing effective brain-machine interfaces (BMI). Common BMI control signals rely on primary sensory- or motor-related activation. However, these signals only reflect a limited range of cognitive processes. Higher-order cognitive signals, and specifically those from prefrontal cortex that encode goal-oriented tasks, could lead to more robust and intuitive BMI [5,6].

Both neurorehabilitation and BMI approaches necessitate an effective and accurate way of measuring and localizing functional brain activity in real time. This can be achieved by electroencephalography (EEG) [7,8] and MEG [9–11], two non-invasive electrophysiological techniques. EEG uses an array of electrodes placed on the scalp to record voltage fluctuations, whereas MEG uses sensitive magnetic detectors called superconducting quantum interference devices (SQUIDs) [12] to measure the same primary electrical currents that generate the electric potential distributions recorded in EEG. Since EEG and MEG capture the electromagnetic fields produced by neuronal currents, they provide a fast and direct index of neuronal activity. However, existing MEG/EEG source localization methods offer limited spatial resolution, confounding the origin of signals that could be used for neurorehabilitation or BMI, or are too slow to compute in real time.

Deep learning (DL) [13] offers a promising new approach to significantly improve source localization in real time. A growing number of works successfully employ DL to

achieve state-of-the-art image quality for inverse imaging problems, such as X-ray computed tomography (CT) [14–16], magnetic resonance imaging (MRI) [17–19], positron emission tomography (PET) [20,21], image super-resolution [22–24], photoacoustic tomography [25], synthetic aperture radar (SAR) image reconstruction [26,27] and seismic tomography [28]. In MEG and EEG, artificial neural networks have been used in the past two decades to predict the location of single dipoles [29–31] or two dipoles [32], but generally DL methods have received little attention with the exception of a few recent studies. Cui et al. (2019) [33] used long-short term memory networks (LSTM) to identify the location and time course of a single source; Ding et. al (2019) [34] used an LSTM network to refine dynamic statistical parametric mapping solutions; and Hecker et al. (2020) [35] used feedforward neural networks to construct distributed cortical solutions. These studies are limited by the number of dipoles, or aim to address the ill-posed nature of distributed solutions.

Here, we develop and present a novel DL solution to localize neural sources comprising *multiple* dipoles, and assess its accuracy and robustness with simulated MEG data. We use the sensor geometry of the whole-head Elekta Triux MEG system and define as source space the cortical surface extracted from a structural MRI scan of a real human subject. While we focus on MEG, the same approaches are directly extendable to EEG, enabling a portable and affordable solution to source localization.

2. Background on MEG Source Localization

Two noninvasive techniques, MEG and EEG, measure the electromagnetic signals emitted by the human brain and can provide a fast and direct index of neural activity suitable for real-time applications. The primary source of these electromagnetic signals is widely believed to be the electrical currents flowing through the apical dendrites of pyramidal neurons in the cerebral cortex. Clusters of thousands of synchronously activated pyramidal cortical neurons can be modeled as an equivalent current dipole (ECD). The current dipole is therefore the basic element used to represent neural activation in MEG and EEG localization methods.

In this section we briefly review the notations used to describe measurement data, forward matrix, and sources, and formulate the problem of estimating current dipoles. Consider an array of M MEG or EEG sensors that measures data from a finite number Q of equivalent current dipole (ECD) sources emitting signals $\{s_q(t)\}_{q=1}^Q$ at locations $\{\mathbf{p}_q\}_{q=1}^Q$. Under these assumptions, the $M \times 1$ vector of the received signals by the array is given by:

$$\mathbf{y}(t) = \sum_{q=1}^Q \mathbf{l}(\mathbf{p}_q) s_q(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{l}(\mathbf{p}_q)$ is the topography of the dipole at location \mathbf{p}_q and $\mathbf{n}(t)$ is the additive noise. The topography $\mathbf{l}(\mathbf{p}_q)$, is given by:

$$\mathbf{l}(\mathbf{p}_q) = \mathbf{L}(\mathbf{p}_q) \mathbf{q}, \quad (2)$$

where $\mathbf{L}(\mathbf{p}_q)$ is the $M \times 3$ forward matrix at location \mathbf{p}_q and \mathbf{q} is the 3×1 vector of the orientation of the ECD source. Depending on the problem, the orientation \mathbf{q} may be known, referred to as *fixed-oriented* dipole, or it may be unknown, referred to as *freely-oriented* dipole.

Assuming that the array is sampled N times at t_1, \dots, t_N , the matrix \mathbf{Y} of the sampled signals can be expressed as:

$$\mathbf{Y} = \mathbf{A}(\mathbf{P})\mathbf{S} + \mathbf{N}, \quad (3)$$

where \mathbf{Y} is the $M \times N$ matrix of the received signals:

$$\mathbf{Y} = [\mathbf{y}(t_1), \dots, \mathbf{y}(t_N)], \quad (4)$$

$\mathbf{A}(\mathbf{P})$ is the $M \times Q$ mixing matrix of the topography vectors at the Q locations $\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_Q]$:

$$\mathbf{A}(\mathbf{P}) = [\mathbf{l}(\mathbf{p}_1), \dots, \mathbf{l}(\mathbf{p}_Q)], \quad (5)$$

\mathbf{S} is the $Q \times N$ matrix of the sources:

$$\mathbf{S} = [\mathbf{s}(t_1), \dots, \mathbf{s}(t_N)], \quad (6)$$

with $\mathbf{s}(t) = [s_1(t), \dots, s_Q(t)]^T$, and \mathbf{N} is the $M \times N$ matrix of noise:

$$\mathbf{N} = [\mathbf{n}(t_1), \dots, \mathbf{n}(t_N)]. \quad (7)$$

Mathematically, the localization problem can be cast as an optimization problem of computing the location and moment parameters of the set of dipoles whose field best matches the MEG/EEG measurements in a least-squares (LS) sense [36]. In this paper we focus on solutions that solve for a small parsimonious set of dipoles and avoid the ill-posedness associated with imaging methods that yield distributed solutions, such as minimum-norm [9] (Figure 1). Solutions that estimate a small set of dipoles include the *dipole fitting* and *scanning* methods. Dipole fitting methods solve the optimization problem directly using techniques that include gradient descent, Nelder–Meade simplex algorithm, multistart, genetic algorithm, and simulated annealing [37–40]. However, these techniques remain unpopular because they converge to a suboptimal local optimum or are too computationally expensive.

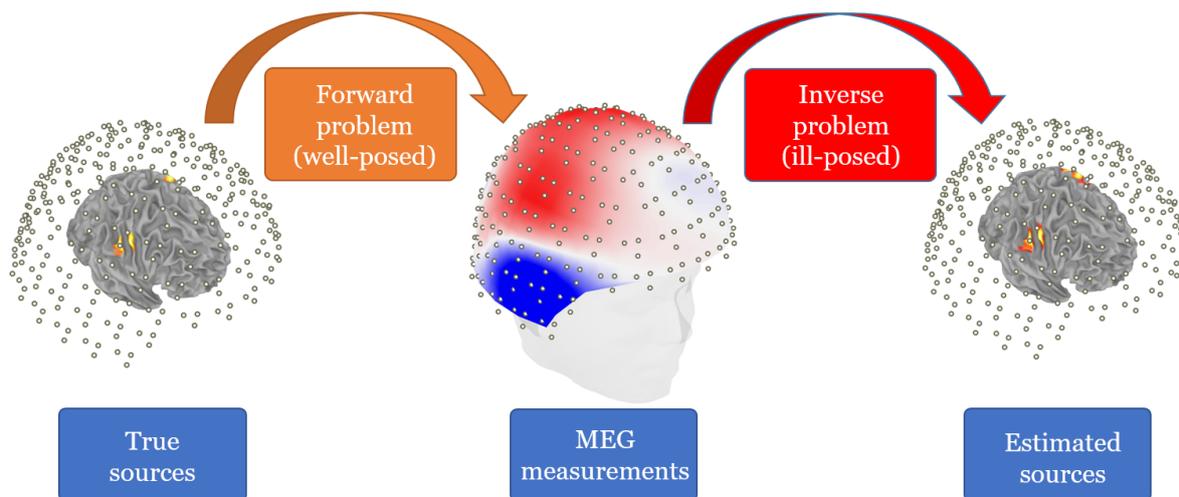


Figure 1. MEG forward and inverse problems. In the forward problem, a well-posed model maps the true sources activation to the MEG measurement vector. In the inverse (and ill-posed) problem, an inverse operator maps the measurement vector to the estimated sources activation.

An alternative approach is scanning methods, which use adaptive spatial filters to search for optimal dipole positions throughout a discrete grid representing the source space [11]. Source locations are then determined as those for which a metric (localizer) exceeds a given threshold. While these approaches do not lead to true least squares solutions, they can be used to initialize a local least squares search. The most common scanning methods are beamformers [41,42] and MUSIC [36], both widely used for bioelectromagnetic source localization, but they assume uncorrelated sources. When correlations are significant, they result in partial or complete cancellation of correlated (also referred to as synchronous) sources and distort the estimated time courses. Several multi-source extensions have been proposed for synchronous sources [43–51]; however, they require some a-priori information on the location of the synchronous sources, are limited to the localization of pairs of synchronous sources, or are limited in their performance.

One important division of the scanning methods is whether they are *non-recursive*

or *recursive*. The original Beamformer [41,42] and MUSIC [36] methods are non-recursive and require the identification of the largest local maxima in the localizer function to find multiple dipoles. Some multi-source variants are also non-recursive (e.g., [44–47]), and as a result they use brute-force optimization, assume that the approximate locations of the neuronal sources have been identified a priori, or still require the identification of the largest local maxima in the localizer function. To overcome these limitations, non-recursive methods have recursive counterparts, such as RAP MUSIC [52], TRAP MUSIC [53], Recursive Double-Scanning MUSIC [54], and RAP Beamformer [7]. The idea behind the recursive execution is that one finds the source locations iteratively at each step, projecting out the topographies of the previously found dipoles before forming the localizer for the current step [7,52]. In this way, one replaces the task of finding several local maxima with the easier task of finding the global maximum of the localizer at each iteration step. While recursive methods generally perform better than their non-recursive counterparts, they still suffer from several limitations, including limited performance, the need for high signal-to-noise ratio (SNR), non-linear optimization of source orientation angles and source amplitudes, or inaccurate estimation as correlation values increase. They are also computationally expensive due to the recursive estimation of sources.

3. Background on Deep Learning

Inverse problems in signal and image processing were traditionally solved using analytical methods; however, recent DL [13] solutions, as exemplified in [17,55], provide state-of-the-art results for numerous problems including X-ray computed tomography, magnetic resonance image reconstruction, natural image restoration (denoising, super-resolution, deblurring), synthetic aperture radar image reconstruction and hyper-spectral unmixing, among others. In the following, we review DL principles, which form the basis for the DL-based solutions to MEG source localization presented in the next section, including concepts such as network layers and activation functions, empirical risk minimization, gradient-based learning, and regularization.

DL is a powerful class of data-driven machine learning algorithms for supervised, unsupervised, reinforcement and generative tasks. DL algorithms are built using Deep Neural Networks (DNNs), which are formed by a hierarchical composition of non-linear functions (layers). The main reason for the success of DL is the ability to train very high capacity (i.e., hypothesis space) networks using very large datasets, often leading to robust *representation learning* [56] and good *generalization* capabilities in numerous problem domains. Generalization is defined as the ability of an algorithm to perform well on unseen examples. In statistical learning terms, an algorithm $\mathcal{A} : \mathcal{X} \rightarrow \mathcal{Y}$ is learned using a training dataset $\mathcal{S} = \{(x_1, y_1), \dots, (x_N, y_N)\}$ of size N , where $x_i \in \mathcal{X}$ is a data sample and $y_i \in \mathcal{Y}$ is the corresponding label (for example, source location coordinates). Let $\mathcal{P}(\mathcal{X}, \mathcal{Y})$ be the true distribution of the data, then the expected risk is defined by:

$$\mathcal{R}(\mathcal{A}) = E_{x,y \sim \mathcal{P}(\mathcal{X}, \mathcal{Y})}[\mathcal{L}(\mathcal{A}(x), y)], \quad (8)$$

where \mathcal{L} is a loss function that measures the misfit between the algorithm output and the data label. The goal of DL is to find an algorithm \mathcal{A} within a given capacity (i.e., function space) that minimizes the expected risk; however, the expected risk cannot be computed since the true distribution is unavailable. Therefore, the empirical risk is minimized instead:

$$\mathcal{R}_{\mathcal{E}}(\mathcal{A}) = \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\mathcal{A}(x_i), y_i), \quad (9)$$

which approximates the statistical expectation with an empirical mean computed using the training dataset. The *generalization gap* is defined as the difference between the expected risk to the empirical risk: $\mathcal{R}(\mathcal{A}) - \mathcal{R}_{\mathcal{E}}(\mathcal{A})$. By using large training datasets and high capacity algorithms, DL has been shown to achieve a low generalization gap, where an

approximation of the expected risk is computed using the learned algorithm and a held-out testing dataset $\mathcal{T} = \{(x_1, y_1), \dots, (x_M, y_M)\}$ of size M , such that $\mathcal{S} \cap \mathcal{T} = \emptyset$.

In the following subsections we describe the main building blocks of DNNs, including multi-layer perceptron and convolutional neural networks.

3.1. Multi-Layer Perceptron (MLP)

The elementary building block of the MLP is the *Perceptron*, which computes a non-linear scalar function, termed *activation*, of an input $\mathbf{x} \in \mathbf{R}^n$, as follows

$$y = f(\mathbf{w}^T \mathbf{x} + b), \quad (10)$$

where \mathbf{w} is a vector of weights and b is a scalar bias. A common activation function is the Sigmoid [13], defined as

$$f(z) = \frac{1}{1 + e^{-z}},$$

and in this case the perceptron output is computed as follows

$$y = \frac{1}{1 + e^{-(\mathbf{w}^T \mathbf{x} + b)}}.$$

A single layer of perceptrons is composed of multiple perceptrons, all connected to the same input vector \mathbf{x} , with a unique weight vector and bias, per perceptron. A single layer of perceptrons can be formulated in matrix form, as follows:

$$\mathbf{y} = f(\mathbf{W}\mathbf{x} + \mathbf{b}), \quad (11)$$

where each row of the matrix \mathbf{W} corresponds to the weights of one perceptron, and each element of the vector \mathbf{b} corresponds to the bias of one perceptron. The MLP is composed of multiple layers of perceptrons, such that the output of each layer becomes the input to the next layer. Such hierarchical composition of k non-linear functions is formulated as follows:

$$\mathbf{F}(\mathbf{x}; \Theta) = f_k(f_{k-1}(\dots f_2(f_1(\mathbf{x}; \theta_1); \theta_2); \theta_{k-1}); \theta_k), \quad (12)$$

where $\theta_i = [\mathbf{W}_i, \mathbf{b}_i]$ are the parameters (i.e., weights and biases) of the i -th layer and $\Theta = [\theta_1, \theta_2, \dots, \theta_k]$ is the set of all network parameters.

In our source localization solution for a single MEG snapshot, detailed in Section 4.1, we utilize a MLP architecture that accepts the MEG measurement vector as an input and the final layer provides the sources locations.

In the supervised learning framework, the parameters Θ are learned by minimizing the empirical risk, computed over the training dataset \mathcal{S} . The empirical risk can be regularized in order to improve DNN generalization, by mitigating over-fitting of the learned parameters to the training data. The empirical risk is defined by

$$\mathbf{J}(\Theta) = \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\mathbf{F}(\mathbf{x}_i; \Theta), y_i), \quad (13)$$

and the optimal set of parameters Θ^* are obtained by solving

$$\Theta^* = \arg \min_{\Theta} \mathbf{J}(\Theta). \quad (14)$$

The minimization of the empirical risk is typically performed by iterative gradient-based algorithms, such as the stochastic gradient descent (SGD)

$$\hat{\Theta}_{t+1} = \hat{\Theta}_t - \lambda \nabla_{\Theta} \mathbf{J}(\Theta), \quad (15)$$

where $\hat{\Theta}_t$ is the estimate of Θ^* at the t -th iteration, $\lambda > 0$ is the learning rate, and the approximated gradient $\nabla_{\Theta} J(\Theta)$ is computed by the *back-propagation* algorithm using a small random subset of examples from the training set S .

3.2. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) were originally developed for processing input images, using the weight sharing principle of a convolutional kernel that is convolved with input data. The main motivation is to reduce significantly the number of required learnable parameters, as compared to processing a full image by perceptrons, namely, allocating one weight per pixel for each perceptron. A CNN is composed of one or more convolutional layers, where each layer is composed of one or more learnable kernels. For a 2D input $I(i, j)$, a convolutional layer performs the convolution (some DL libraries implement the cross-correlation operation) between the input to the kernel(s)

$$C(i, j) = (K * I)(i, j) = \sum_{m, n} W(m, n) I(i - m, j - n), \quad (16)$$

where $W(m, n)$ is the kernel and $C(i, j)$ is the convolution result. A bias b is further added to each convolution results, and an activation function $f()$ is applied, to obtain the *feature map* $F(i, j)$ given by

$$F(i, j) = f(C(i, j) + b). \quad (17)$$

A convolutional layer with K kernels produces K feature maps, where kernels of 1D, 2D or 3D are commonly used. Convolutional layers are often immediately followed by sub-sampling layers, such as *MaxPooling* that decimates information by picking the maximum value within a given array of values, or *AveragePooling* that replaces a given array of values by their mean. CNN networks are typically composed by a cascade of convolutional layers, optionally followed by fully-connected (FC) layers, depending on the required task.

In our source localization solution using multiple MEG snapshots, detailed in Section 4.2, we utilize 1D convolutions for spatio-temporal feature extractors, which are followed by FC layers for computing the estimated sources locations.

4. Deep Learning for MEG Source Localization

In this section we present the proposed deep neural network (DNN) architectures and training data generation workflow.

4.1. MLP for Single-Snapshot Source Localization

MEG source localization is computed from sensor measurements using either a single snapshot (i.e., a single time sample) or multiple snapshots. The single snapshot case is highly challenging for popular MEG localization algorithms, such as MUSIC [36], RAP-MUSIC [52], and RAP-Beamformer [7], all of which rely on the data covariance matrix. A single snapshot estimation of the covariance matrix is often insufficient for good localization accuracy of multiple simultaneously active sources, especially in low and medium signal-to-noise ratios (SNR). Since the input in this case is a single measurement MEG vector, we implemented four layer MLP-based architectures, where the input FC layer maps the M -dimensional snapshot vector to a higher dimensional vector and the output layer computes the source(s) coordinates in 3D, as illustrated in Figure 2. We refer to this model as *DeepMEG-MLP*. We implemented three DeepMEG-MLP models, corresponding to $Q = 1, 2$, and 3 sources, as summarized in Table 1.

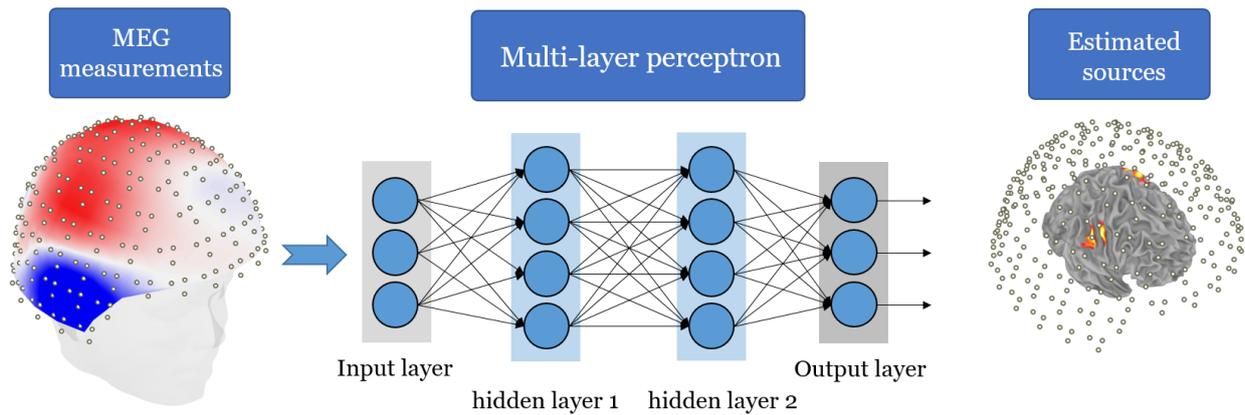


Figure 2. Illustration of the MLP-based MEG source localization solution. The end-to-end inversion operator performs mapping from the MEG measurement space to the source locations space.

Table 1. Evaluated DeepMEG Architectures. Fully-connected layers are denoted by FC(N , 'activation'), where N is the number of perdepstrons and the activation function is Sigmoid for all layers excluding output layers, which compute sources locations and do not employ an activation function. Convolutional layers are denoted by Conv1D(L , T), where L is the number of 1D kernels, and T is the length of each 1D kernel.

Input	Network	1st Layer	2nd Layer	3rd Layer	4th Layer	5th Layer	Parameters
Single MEG Snapshot	MLP-1	FC (3000, 'sigmoid')	FC (2500, 'sigmoid')	FC (1200, 'sigmoid')	FC (3, 'none')	-	11,428,303
	MLP-2	FC (3000, 'sigmoid')	FC (2500, 'sigmoid')	FC (1200, 'sigmoid')	FC (6, 'none')	-	11,431,906
	MLP-3	FC (3000, 'sigmoid')	FC (2500, 'sigmoid')	FC (1200, 'sigmoid')	FC (9, 'none')	-	11,435,509
MEG Time-Series	CNN-1	Conv1D ($L = 32$, $T = 5$)	FC (3000, 'sigmoid')	FC (2500, 'sigmoid')	FC (1200, 'sigmoid')	FC (3, 'none')	11,711,295
	CNN-2	Conv1D ($L = 32$, $T = 5$)	FC (3000, 'sigmoid')	FC (2500, 'sigmoid')	FC (1200, 'sigmoid')	FC (6, 'none')	11,714,898
	CNN-3	Conv1D ($L = 32$, $T = 5$)	FC (3000, 'sigmoid')	FC (2500, 'sigmoid')	FC (1200, 'sigmoid')	FC (9, 'none')	11,718,501

4.2. CNN for Multiple-Snapshot Source Localization

For multiple consecutive MEG snapshots we implemented a CNN-based architecture with five layers, in which the first layer performs 1D convolutions on the input data and the resulting 1D feature maps are processed by three FC layers with sigmoid activation, and an output FC layer which computes the source locations. We refer to this model as *DeepMEG-CNN*, as illustrated in Figure 3. The 1D convolutional layer forms a bank of $L = 32$ space-time filters (which can also be interpreted as beamformers [57]). Each 1D temporal filter spans $T = 5$ time samples. A different 1D filter is applied to the time course of each of the M sensors with uniquely learned coefficients. We implemented three DeepMEG-CNN models, corresponding to $Q = 1, 2$, and 3 sources, as summarized in Table 1.

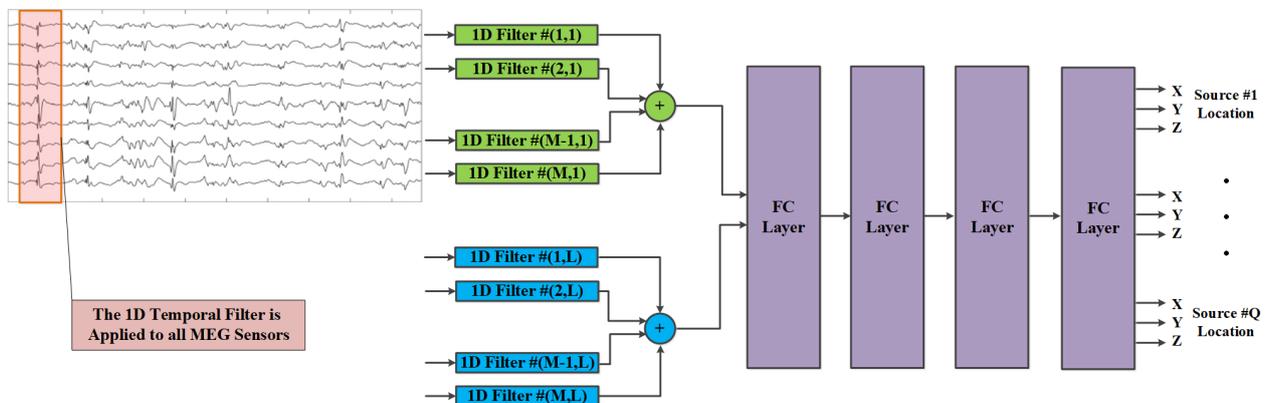


Figure 3. Illustration of the CNN-based MEG source localization solution. The end-to-end inversion operator performs mapping from the MEG time-series measurement space, using a bank of L space-time filters and four fully connected (FC) layers, to the source locations space.

4.3. Data Generation Workflow

To train the deep network models and evaluate their performance on source localization, we need to know the ground truth of the underlying neural sources generating MEG data. Since this information (i.e., the true locations of the sources) is unavailable in real MEG measurements of human participants, we performed simulations with an actual MEG sensor array and a realistic brain anatomy and source configurations. Specifically, the sensor array was based on the whole-head Elekta Triux MEG system with 306-channel probe unit including 204 planar gradiometer sensors and 102 magnetometer sensors (Figure 4a). The geometry of the MEG source space was modeled with the cortical manifold extracted from a T1-weighted MRI structural scan from a real adult human subject using Freesurfer [58]. This source configuration is consistent with the arrangement of pyramidal neurons, the principal source of MEG signals, in the cerebral cortex. Sources were restricted to 15,002 grid points over the cortex (Figure 4b). The lead field matrix, which represents the forward mapping from the activated sources to the sensor array, was estimated using BrainStorm [59] based on an overlapping spheres head model [60]. This model has been shown to have accuracy similar to boundary element methods in MEG data but is orders of magnitude faster to compute [60].

Simulated MEG sensor data was generated by first activating a few sources randomly selected on the cortical manifold with activation time courses $s_i(t)$. The time courses were modeled with 16 time points sampled as mixtures of sinusoidal signals of equal amplitude with correlations in the range of 0 to 0.9 and frequencies in the range of 10 to 90 Hz. The corresponding sensor measurements were then obtained by multiplying each source with its respective topography vector $l(p_i)$ (Figure 4c). Finally, Gaussian white noise was generated and added to the MEG sensors to model instrumentation noise at specific SNR levels. The SNR was defined as the ratio of the Frobenius norm of the signal-magnetic-field spatiotemporal matrix to that of the noise matrix for each trial as in [61].

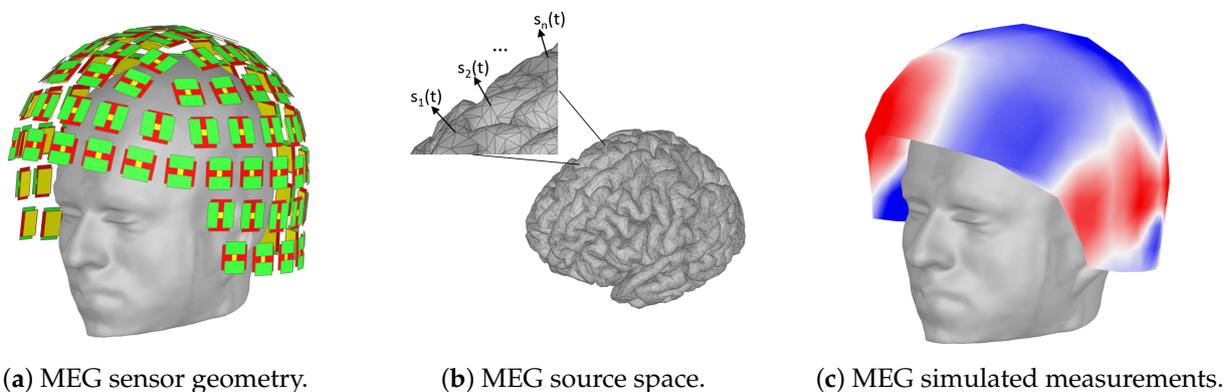


Figure 4. Simulation of MEG data for deep learning localization. (a) Simulations used the anatomy of an adult human subject and a whole-head MEG sensor array from an Elekta Triux device. (b) Cortical sources with time course $s_i(t)$ were simulated at different cortical locations. (c) The activated cortical sources yielded MEG measurements on the cortex that, combined with additive Gaussian noise, comprised the input to the deep learning model.

5. Performance Evaluation

5.1. Deep Network Training

The DeepMEG models were implemented in TensorFlow [62] and trained by minimizing the mean-squared error (MSE) between the true (i.e., MEG data labels) and estimated sources locations, using the SGD algorithm (15) with a learning rate $\lambda = 0.001$ and batch size of 32. The DeepMLP networks were trained with datasets of 1 million simulated snapshots, generated at a fixed SNR level, yet, as discussed in the following these trained models operate well in a wide range of SNR levels. The DeepCNN networks were trained using data generation on the fly and a total of 9.6 million multiple-snapshot signals per network. Data generation on the fly was utilized in order to mitigate the demanding memory

requirements of offline data generation in the case of multiple-snapshots training set. The DeepCNN network models were trained with MEG sensor data at a fixed SNR-level and random inter-source correlations, thus, learning to localize sources with a wide range of inter-source correlation levels.

5.2. Localization Experiments

We assessed the performance of the deepMEG models using simulated data as described in the data generation workflow. To assess localization accuracy in different realistic scenarios, we conducted simulations with different SNR levels and inter-source correlation values. We also varied the number of active sources to validate that localization is accurate even for multiple concurrently active sources.

During inference, we compared the performance of the deep learning model against the popular scanning localization solution RAP-MUSIC [52]. All experiment were conducted with 1000 Monte-Carlo repetitions per each SNR and inter-source correlation value. In each scenario, we used the deepMEG and RAP-MUSIC with the corresponding number of sources, which is assumed known by both methods. Estimation of the number of sources can be conducted with the Akaike information criterion (AIC), Bayesian information criterion (BIC), or cross-validation and is beyond the scope of this work.

5.2.1. Experiment 1: Performance of the DeepMEG-MLP Model with Single-Snapshot Data

We assessed the localization accuracy of the DeepMEG-MLP model against the RAP-MUSIC method for the case of two simultaneously active dipole sources. The DeepMEG-MLP model was trained with 10 dB SNR data, but inference used different SNR levels ranging from -10 dB to 20 dB (Figure 5). The DeepMEG model outperformed the RAP-MUSIC method at high SNR values, but had worse localization results in low SNR values (<5 dB). As expected, both methods consistently improved their localization performance with increasing SNR values.

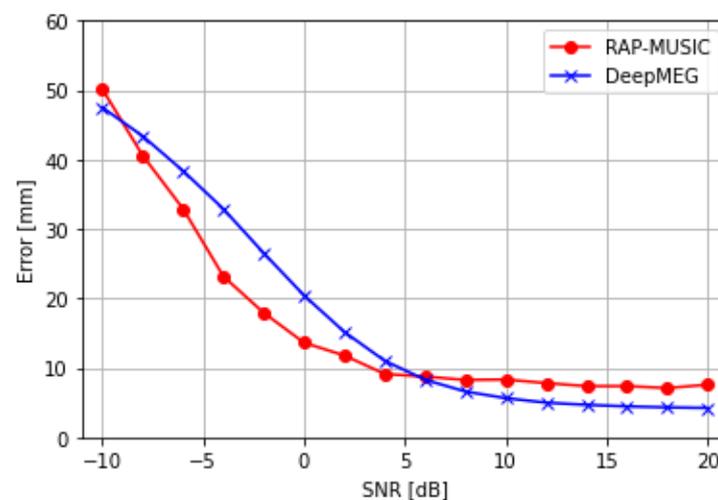


Figure 5. Localization accuracy of the DeepMEG-MLP model at different SNR levels for the cases of two dipole sources.

5.2.2. Experiment 2: Performance of the DeepMEG-CNN Model with Multiple-Snapshot Data

We extended the above experiment for the case of multiple snapshot data with $T = 16$ time samples and two or three sources with different inter-source correlation values. The DeepMEG-CNN model was trained with -15 dB SNR data, and inference used -15 dB, -12.5 dB, and -10 dB SNR. In the low SNR case (-15 dB), the DeepMEG-CNN consistently outperformed the RAP-MUSIC method with the exception of high (0.9) correlation values where the RAP-MUSIC had a slightly better accuracy (Figure 6a,b). As SNR increased to

−12.5 dB, the Deep MEG model remained overall better or had comparable performance to RAP-MUSIC (Figure 6c,d). This advantage was lost at −10 dB SNR, where RAP-MUSIC had an advantage (Figure 6e,f).

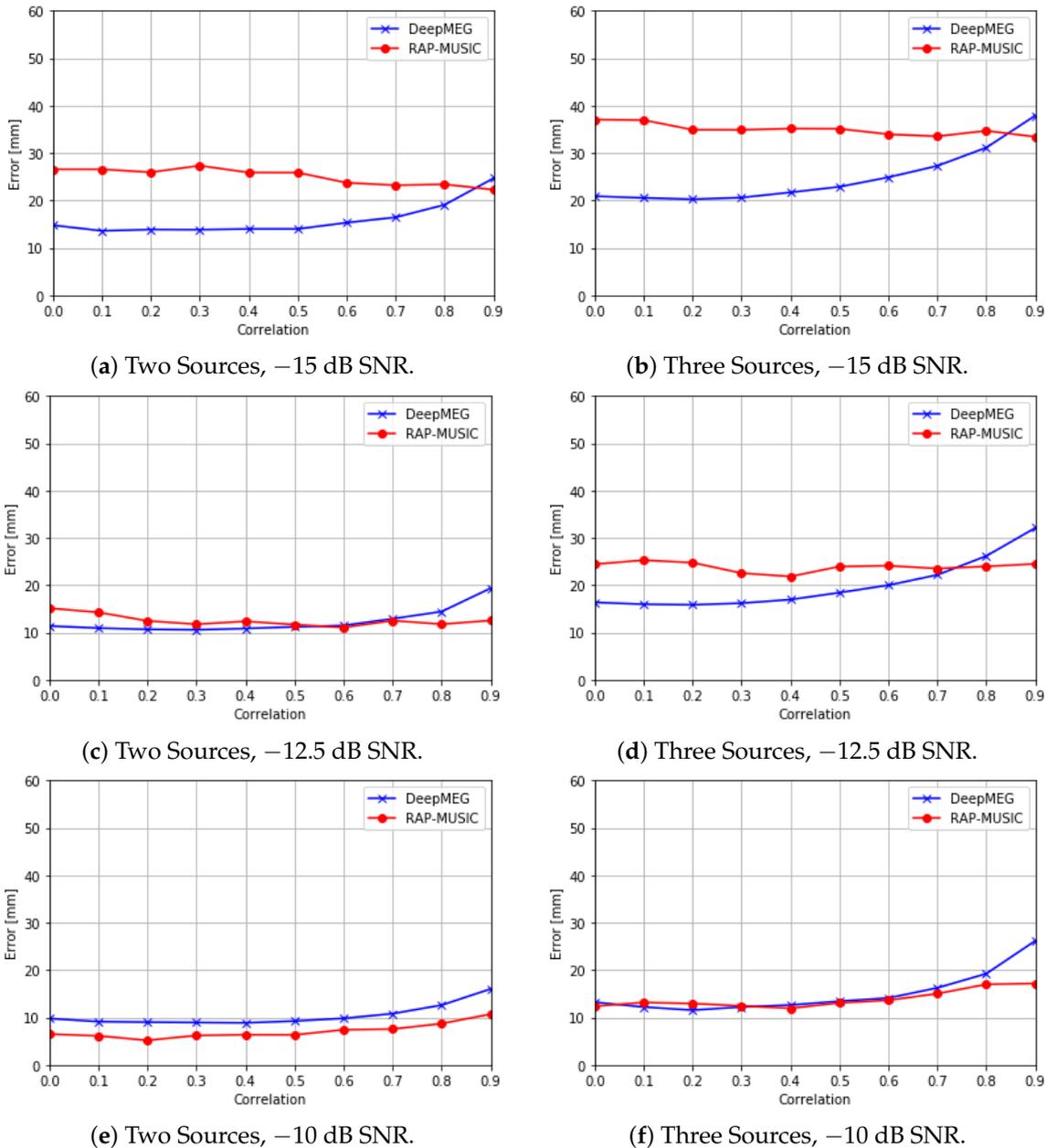


Figure 6. Localization accuracy of the DeepMEG-CNN model with $T = 16$ time samples at different inter-source correlation values for the cases of two and three sources with −15 dB, −12.5 dB and −10 dB SNR levels.

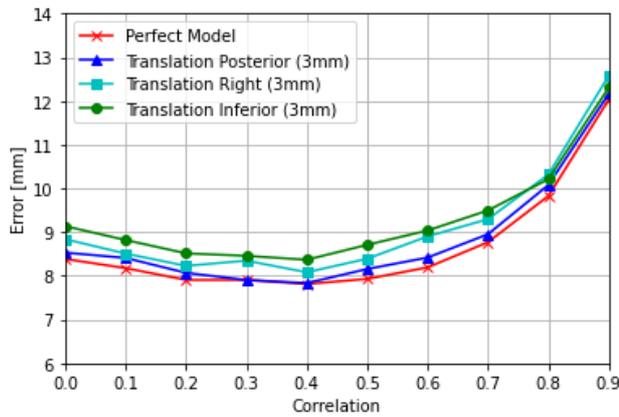
5.2.3. Experiment 3: Robustness of DeepMEG to Forward Model Errors

Here we assumed that the actual MEG forward model is different from the ideal forward model that was used for building the training set of the DeepMEG solution. The actual model was defined as follows:

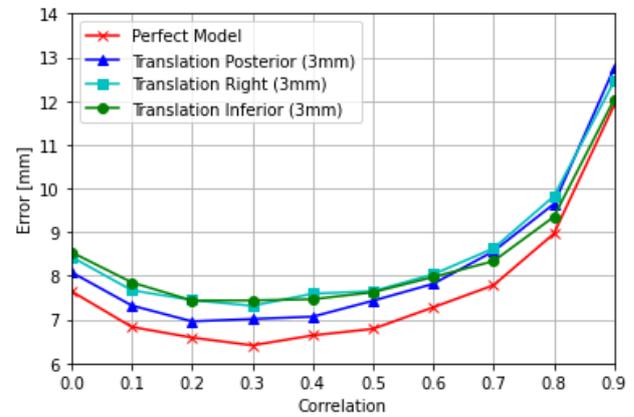
$$\tilde{\mathbf{A}}(\mathbf{P}) = \mathbf{A}(\mathbf{P}) + \Delta\mathbf{A}(\mathbf{P}), \quad (18)$$

where $\Delta\mathbf{A}(\mathbf{P})$ denotes the *model error* matrix, which can be *arbitrary*. In this experiment we considered the model error matrix to represent inaccuracies due to imprecise estimation of the position of the head caused by translation or rotation registration errors. Figure 7

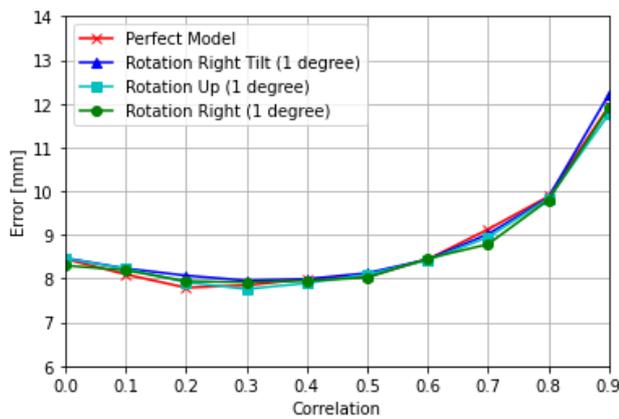
presents the DeepMEG-CNN localization accuracy with model registration errors equal to 3 mm translation and 1 degree rotation across different axes. The DeepMEG-CNN achieved robust localization accuracy in all cases: up to 1 mm degradation for head translations and up to 0.25 mm degradation for head rotations.



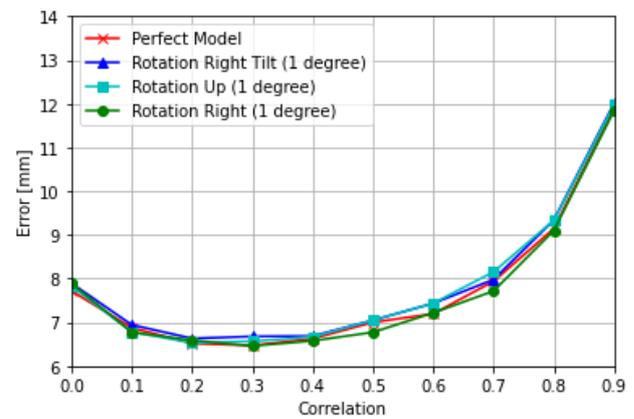
(a) Two sources, -5 dB SNR, Head Model Translation.



(b) Three sources, 0 dB SNR, Head Model Translation.



(c) Two sources, -5 dB SNR, Head Model Rotation.



(d) Three sources, 0 dB SNR, Head Model Rotation.

Figure 7. Robustness of DeepMEG-CNN localization accuracy to forward model errors due to head model translations by 3 mm (a,b), and head model rotations by 1 degree (c,d).

5.3. Real-Time Source Localization

An important advantage of the DL approaches is that they have significantly reduced computational time, paving the way to real-time MEG source localization solutions. We conducted a computation time comparison (compute hardware: CPU 6-Core Intel i5-9400F@3.9 Ghz, CPU RAM 64 GB, GPU Nvidia GeForce RTX 2080), detailed in Table 2, for each of the proposed DL architectures and the RAP-MUSIC algorithm. The comparison reveals 4 orders of magnitude faster computation of DeepMEG models as compared to the RAP-MUSIC, which requires $O(Q) M \times M$ matrix inversions and $O(Q \times N_{dipoles})$ matrix-vector multiplications, with each multiplication having complexity of $O(M^2)$. Here, the total number of dipoles $N_{dipoles} = 15,002$ and the number of sensors $M = 306$.

Table 2. Computation Time Comparison.

Sources	Time Samples	Algorithm	Time [ms]
1	1	RAP-MUSIC	135.47
1	1	DeepMEG MLP-1	0.19
2	1	RAP-MUSIC	452.17
2	1	DeepMEG MLP-2	0.19
3	1	RAP-MUSIC	736.76
3	1	DeepMEG MLP-3	0.19
1	16	RAP-MUSIC	136.59
1	16	DeepMEG CNN-1	0.25
2	16	RAP-MUSIC	478.23
2	16	DeepMEG CNN-2	0.27
3	16	RAP-MUSIC	741.51
3	16	DeepMEG CNN-3	0.27

6. Conclusions

Fast and accurate solutions to MEG source localization are crucial for real-time brain imaging, and hold the potential to enable novel applications in neurorehabilitation and BMI. Current methods for multiple dipole fitting and scanning do not achieve precise source localization because it is typically computationally intractable to find the global maximum in the case of multiple dipoles. These methods also limit the number of dipoles and temporal rate of source localization due to high computational demands. In this article, we reviewed existing MEG source localization solutions and fundamental DL tools. Motivated by the recent success of DL in a growing number of inverse imaging problems, we proposed two DL architectures for the solution of the MEG inverse problem, the DeepMEG-MLP for single time point localization, and the DeepMEG-CNN for multiple time point localization.

We compared the performance of DeepMEG against the popular RAP-MUSIC localization algorithm and showed improvements in localization accuracy in a range of scenarios with variable SNR levels, inter-source correlation values, and number of sources. Importantly, the DeepMEG inference was estimable in less than a millisecond and thus was orders of magnitude faster than RAP-MUSIC. Fast computation was possible due to the high optimization of modern DL tools, and even allows the rapid estimation of dipoles at near the 1 kHz sampling rate speed of existing MEG devices. This could facilitate the search for optimal indices of brain activity in neurofeedback and BCI tasks.

A key property of DeepMEG was its robustness to forward model errors. The localization performance of the model remained relatively stable even when introducing model errors caused by imprecise estimation of the position of the head due to translations or rotations. This is critical for real-time applications where the forward matrix is not precisely known, or movement of the subject introduces time-varying inaccuracies.

To apply the DeepMEG models in real MEG experiments, the networks must first be trained separately for each subject using the forward matrix derived from individual anatomy. Training should take place offline using simulated data, and then the trained model can be applied for inference on real-data. The amount of simulated data for training can be easily increased if necessary. To improve the validity of the predictions, the distributional assumptions of the simulated data should follow those of the experimental data as closely as possible, including the number of sources, and realistic noise distribution and shape of source time courses. Future work is needed to determine how to best specify and evaluate the impact of different distributional assumptions in the quality of the DL model predictions in different experimental settings. Once a model achieves the desirable performance, extending to new subjects could be achieved with transfer learning to fine tune the parameters with reduced computational cost.

While the DeepMEG-MLP and DeepMEG-CNN architectures yielded promising localization results, more research is needed to explore different architectures, regularizations, loss functions, and other DL parameters that may further improve MEG source localization. It is also critical to assess whether DL models remain robust to model errors under more cases of realistic perturbations, beyond the imprecise head position estimation assessed here. This includes inaccurate modeling of the source space (e.g., cortical segmentation errors); imprecise localization of sensors within the sensor array; and inaccurate estimation of the forward model (e.g., unknown tissue conductivities, analytical approximations or numerical errors) [63,64].

As both theory and simulations suggest, when the true number of sources is known, RAP-MUSIC and TRAP-MUSIC (a generalization of RAP-MUSIC) localize sources equally well [53]. Thus our localization comparisons should also extend to TRAP-MUSIC. Of course, this holds under the critical assumption that another method has assessed the true number of sources, as required by the DeepMEG-MLP and DeepMEG-CNN models. This can be done with AIC, BIC, cross-validation, or other criteria [65]. However, as extension of our work, different DL architectures could be developed to estimate both the number of sources and their location simultaneously, obviating the need to separately estimate the true number of sources.

Author Contributions: Conceptualization, D.P. and A.A.; Methodology, D.P. and A.A.; Software, D.P. and A.A.; Validation, D.P. and A.A.; Formal analysis, D.P. and A.A.; Investigation, D.P. and A.A.; Resources, D.P. and A.A.; Writing—original draft preparation, D.P. and A.A.; Writing—review and editing, D.P. and A.A.; Visualization, D.P. and A.A.; Supervision, D.P. and A.A.; Project administration, D.P. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The second author wishes to thank Braude College of Engineering for supporting this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Valero-Cabré, A.; Amengual, J.L.; Stengel, C.; Pascual-Leone, A.; Coubard, O.A. Transcranial magnetic stimulation in basic and clinical neuroscience: A comprehensive review of fundamental principles and novel insights. *Neurosci. Biobehav. Rev.* **2017**, *83*, 381–404. [[CrossRef](#)]
2. Kohl, S.; Hannah, R.; Rocchi, L.; Nord, C.L.; Rothwell, J.; Voon, V. Cortical Paired Associative Stimulation Influences Response Inhibition: Cortico-cortical and Cortico-subcortical Networks. *Biol. Psychiatry* **2019**, *85*, 355–363. [[CrossRef](#)]
3. Folloni, D.; Verhagen, L.; Mars, R.B.; Fouragnan, E.; Constans, C.; Aubry, J.F.; Rushworth, M.F.; Sallet, J. Manipulation of Subcortical and Deep Cortical Activity in the Primate Brain Using Transcranial Focused Ultrasound Stimulation. *Neuron* **2019**, *101*, 1109–1116.e5. [[CrossRef](#)]
4. Darrow, D. Focused Ultrasound for Neuromodulation. *Neurotherapeutics* **2019**, *16*, 88–99. [[CrossRef](#)]
5. Min, B.K.; Chavarriaga, R.; del Millán, R.J. Harnessing Prefrontal Cognitive Signals for Brain–Machine Interfaces. *Trends Biotechnol.* **2017**, *35*, 585–597. [[CrossRef](#)] [[PubMed](#)]
6. Pichiorri, F.; Mattia, D. Brain-computer interfaces in neurologic rehabilitation practice. In *Handbook of Clinical Neurology*; Elsevier: Amsterdam, The Netherlands, 2020; Volume 168, pp. 101–116. [[CrossRef](#)]
7. Ilmoniemi, R.; Sarvas, J. *Brain Signals: Physics and Mathematics of MEG and EEG*; MIT Press: Cambridge, MA, USA, 2019.
8. Niedermeyer, E.; Schomer, D.L.; Lopes da Silva, F.H. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2012.
9. Hämmäläinen, M.; Hari, R.; Ilmoniemi, R.J.; Knuutila, J.; Lounasmaa, O.V. Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* **1993**, *65*, 413–497. [[CrossRef](#)]
10. Baillet, S. Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* **2017**, *20*, 327–339. [[CrossRef](#)] [[PubMed](#)]

11. Darvas, F.; Pantazis, D.; Kucukaltun-Yildirim, E.; Leahy, R. Mapping human brain function with MEG and EEG: Methods and validation. *NeuroImage* **2004**, *23*, S289–S299. [[CrossRef](#)] [[PubMed](#)]
12. Kleiner, R.; Koelle, D.; Ludwig, F.; Clarke, J. Superconducting quantum interference devices: State of the art and applications. *Proc. IEEE* **2004**, *92*, 1534–1548. [[CrossRef](#)]
13. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
14. Gupta, H.; Jin, K.H.; Nguyen, H.Q.; McCann, M.T.; Unser, M. CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction. *IEEE Trans. Med. Imaging* **2018**, *37*, 1440–1453. [[CrossRef](#)] [[PubMed](#)]
15. Jin, K.H.; McCann, M.T.; Froustey, E.; Unser, M. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Trans. Image Process.* **2017**, *26*, 4509–4522. [[CrossRef](#)] [[PubMed](#)]
16. Gjestebj, L.; Yang, Q.; Xi, Y.; Zhou, Y.; Zhang, J.; Wang, G. Deep Learning Methods to Guide CT Image Reconstruction and Reduce Metal Artifacts. In Proceedings of the SPIE Medical Imaging, Orlando, FL, USA, 11–16 February 2017. [[CrossRef](#)]
17. Liang, D.; Cheng, J.; Ke, Z.; Ying, L. Deep Magnetic Resonance Image Reconstruction: Inverse Problems Meet Neural Networks. *IEEE Signal Process. Mag.* **2020**, *37*, 141–151. [[CrossRef](#)]
18. Wang, S.; Su, Z.; Ying, L.; Peng, X.; Zhu, S.; Liang, F.; Feng, D.; Liang, D. Accelerating magnetic resonance imaging via deep learning. In Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016.
19. Schlemper, J.; Caballero, J.; Hajnal, J.V.; Price, A.N.; Rueckert, D. A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. *IEEE Trans. Med. Imaging* **2018**, *37*, 491–503. [[CrossRef](#)] [[PubMed](#)]
20. Gong, K.; Catana, C.; Qi, J.; Li, Q. PET Image Reconstruction Using Deep Image Prior. *IEEE Trans. Med. Imaging* **2019**, *38*, 1655–1665. [[CrossRef](#)] [[PubMed](#)]
21. Kim, K.; Wu, D.; Gong, K.; Dutta, J.; Kim, J.H.; Son, Y.D.; Kim, H.K.; El Fakhri, G.; Li, Q. Penalized PET Reconstruction Using Deep Learning Prior and Local Linear Fitting. *IEEE Trans. Med. Imaging* **2018**, *37*, 1478–1487. [[CrossRef](#)] [[PubMed](#)]
22. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)]
23. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1646–1654. [[CrossRef](#)]
24. Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1132–1140. [[CrossRef](#)]
25. Hauptmann, A.; Lucka, F.; Betcke, M.; Huynh, N.; Adler, J.; Cox, B.; Beard, P.; Ourselin, S.; Arridge, S. Model-Based Learning for Accelerated, Limited-View 3-D Photoacoustic Tomography. *IEEE Trans. Med. Imaging* **2018**, *37*, 1382–1393. [[CrossRef](#)]
26. Yonel, B.; Mason, E.; Yazici, B. Deep Learning for Passive Synthetic Aperture Radar. *IEEE J. Sel. Top. Signal Process.* **2018**, *12*, 90–103. [[CrossRef](#)]
27. Budillon, A.; Johnsy, A.C.; Schirinzi, G.; Vitale, S. Sar Tomography Based on Deep Learning. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 3625–3628. [[CrossRef](#)]
28. Araya-Polo, M.; Jennings, J.; Adler, A.; Dahlke, T. Deep-learning tomography. *Lead. Edge* **2018**, *37*, 58–66. [[CrossRef](#)]
29. Zhang, Q.; Yuasa, M.; Nagashino, H.; Kinouchi, Y. Single dipole source localization from conventional EEG using BP neural networks. In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong, China, 1 November 1998; Volume 4, pp. 2163–2166. [[CrossRef](#)]
30. Abeyratne, U.R.; Zhang, G.; Saratchandran, P. EEG source localization: A comparative study of classical and neural network methods. *Int. J. Neural Syst.* **2001**, *11*, 349–359. [[CrossRef](#)]
31. Sciabassi, R.J.; Sonmez, M.; Sun, M. EEG source localization: a neural network approach. *Neurol. Res.* **2001**, *23*, 457–464. [[CrossRef](#)] [[PubMed](#)]
32. Yuasa, M.; Zhang, Q.; Nagashino, H.; Kinouchi, Y. EEG source localization for two dipoles by neural networks. In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong, China, 1 November 1998; Volume 4, pp. 2190–2192.
33. Cui, S.; Duan, L.; Gong, B.; Qiao, Y.; Xu, F.; Chen, J.; Wang, C. EEG source localization using spatio-temporal neural network. *China Commun.* **2019**, *16*, 131–143. [[CrossRef](#)]
34. Dinh, C.; Samuelsson, J.G.; Hunold, A.; Hämäläinen, M.S.; Khan, S. Contextual Minimum-Norm Estimates (CMNE): A Deep Learning Method for Source Estimation in Neuronal Networks. *arXiv* **2019**, arXiv:1909.02636.
35. Hecker, L.; Rupperecht, R.; Tebartz van Elst, L.; Kornmeier, J. ConvDip: A convolutional neural network for better M/EEG Source Imaging. *Front. Neurosci.* **2021**, *15*, 533. [[CrossRef](#)]
36. Mosher, J.; Lewis, P.; Leahy, R. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Trans. Biomed. Eng.* **1992**, *39*, 541–557. [[CrossRef](#)]
37. Huang, M.; Aine, C.; Supek, S.; Best, E.; Ranken, D.; Flynn, E. Multi-start downhill simplex method for spatio-temporal source localization in magnetoencephalography. *Electroencephalogr. Clin. Neurophysiol. Potentials Sect.* **1998**, *108*, 32–44. [[CrossRef](#)]

38. Uutela, K.; Hamalainen, M.; Salmelin, R. Global optimization in the localization of neuromagnetic sources. *IEEE Trans. Biomed. Eng.* **1998**, *45*, 716–723. [[CrossRef](#)]
39. Khosla, D.; Singh, M.; Don, M. Spatio-temporal EEG source localization using simulated annealing. *IEEE Trans. Biomed. Eng.* **1997**, *44*, 1075–1091. [[CrossRef](#)]
40. Jiang, T.; Luo, A.; Li, X.; Kruggel, F. A Comparative Study Of Global Optimization Approaches To Meg Source Localization. *Int. J. Comput. Math.* **2003**, *80*, 305–324. [[CrossRef](#)]
41. van Veen, B.D.; Drongelen, W.v.; Yuchtman, M.; Suzuki, A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* **1997**, *44*, 867–880. [[CrossRef](#)]
42. Vrba, J.; Robinson, S.E. Signal Processing in Magnetoencephalography. *Methods* **2001**, *25*, 249–271. [[CrossRef](#)]
43. Dalal, S.S.; Sekihara, K.; Nagarajan, S.S. Modified beamformers for coherent source region suppression. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 1357–1363. [[CrossRef](#)]
44. Brookes, M.J.; Stevenson, C.M.; Barnes, G.R.; Hillebrand, A.; Simpson, M.I.; Francis, S.T.; Morris, P.G. Beamformer reconstruction of correlated sources using a modified source. *NeuroImage* **2007**, *34*, 1454–1465. [[CrossRef](#)]
45. Hui, H.B.; Pantazis, D.; Bressler, S.L.; Leahy, R.M. Identifying true cortical interactions in MEG using the nulling beamformer. *NeuroImage* **2010**, *49*, 3161–3174. [[CrossRef](#)]
46. Diwakar, M.; Huang, M.X.; Srinivasan, R.; Harrington, D.L.; Robb, A.; Angeles, A.; Muzzatti, L.; Pakdaman, R.; Song, T.; Theilmann, R.J.; et al. Dual-Core Beamformer for obtaining highly correlated neuronal networks in MEG. *NeuroImage* **2011**, *54*, 253–263. [[CrossRef](#)]
47. Diwakar, M.; Tal, O.; Liu, T.; Harrington, D.; Srinivasan, R.; Muzzatti, L.; Song, T.; Theilmann, R.; Lee, R.; Huang, M. Accurate reconstruction of temporal correlation for neuronal sources using the enhanced dual-core MEG beamformer. *NeuroImage* **2011**, *56*, 1918–1928. [[CrossRef](#)] [[PubMed](#)]
48. Moiseev, A.; Gaspar, J.M.; Schneider, J.A.; Herdman, A.T. Application of multi-source minimum variance beamformers for reconstruction of correlated neural activity. *NeuroImage* **2011**, *58*, 481–496. [[CrossRef](#)] [[PubMed](#)]
49. Moiseev, A.; Herdman, A.T. Multi-core beamformers: Derivation, limitations and improvements. *NeuroImage* **2013**, *71*, 135–146. [[CrossRef](#)]
50. Hesheng, L.; Schimpf, P.H. Efficient localization of synchronous EEG source activities using a modified RAP-MUSIC algorithm. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 652–661.
51. Ewald, A.; Avarvand, F.S.; Nolte, G. Wedge MUSIC: A novel approach to examine experimental differences of brain source connectivity patterns from EEG/MEG data. *NeuroImage* **2014**, *101*, 610–624. [[CrossRef](#)] [[PubMed](#)]
52. Mosher, J.C.; Leahy, R.M. Source localization using recursively applied and projected (RAP) MUSIC. *IEEE Trans. Signal Process.* **1999**, *47*, 332–340. [[CrossRef](#)]
53. Mäkelä, N.; Stenroos, M.; Sarvas, J.; Ilmoniemi, R.J. Truncated RAP-MUSIC (TRAP-MUSIC) for MEG and EEG source localization. *NeuroImage* **2018**, *167*, 73–83. [[CrossRef](#)] [[PubMed](#)]
54. Mäkelä, N.; Stenroos, M.; Sarvas, J.; Ilmoniemi, R.J. Locating highly correlated sources from MEG with (recursive) (R)DS-MUSIC. *Neuroscience* **2017**, preprint. [[CrossRef](#)]
55. Lucas, A.; Iliadis, M.; Molina, R.; Katsaggelos, A.K. Using Deep Neural Networks for Inverse Problems in Imaging: Beyond Analytical Methods. *IEEE Signal Process. Mag.* **2018**, *35*, 20–36. [[CrossRef](#)]
56. Bengio, Y.; Courville, A.; Vincent, P. Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [[CrossRef](#)]
57. Sainath, T.N.; Weiss, R.J.; Wilson, K.W.; Li, B.; Narayanan, A.; Variiani, E.; Bacchiani, M.; Shafran, I.; Senior, A.; Chin, K.; et al. Multichannel Signal Processing With Deep Neural Networks for Automatic Speech Recognition. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2017**, *25*, 965–979. [[CrossRef](#)]
58. Fischl, B.; Salat, D.H.; van der Kouwe, A.J.; Makris, N.; Ségonne, F.; Quinn, B.T.; Dale, A.M. Sequence-independent segmentation of magnetic resonance images. *NeuroImage* **2004**, *23*, S69–S84. [[CrossRef](#)] [[PubMed](#)]
59. Tadel, F.; Baillet, S.; Mosher, J.C.; Pantazis, D.; Leahy, R.M. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* **2011**, *2011*, 1–13. [[CrossRef](#)]
60. Huang, M.X.; Mosher, J.C.; Leahy, R.M. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Phys. Med. Biol.* **1999**, *44*, 423–440. [[CrossRef](#)]
61. Sekihara, K.; Nagarajan, S.S.; Poeppel, D.; Marantz, A.; Miyashita, Y. Reconstructing spatio-temporal activities of neural sources using an MEG vector beamformer technique. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 760–771. [[CrossRef](#)]
62. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, H.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. Available online: Available online: [tensorflow.org](https://www.tensorflow.org) (accessed on 22 June 2021).
63. Cho, J.H.; Vorwerk, J.; Wolters, C.H.; Knösche, T.R. Influence of the head model on EEG and MEG source connectivity analyses. *Neuroimage* **2015**, *110*, 60–77. [[CrossRef](#)] [[PubMed](#)]
64. Gençer, N.G.; Acar, C.E. Sensitivity of EEG and MEG measurements to tissue conductivity. *Phys. Med. Biol.* **2004**, *49*, 701. [[CrossRef](#)] [[PubMed](#)]
65. Wax, M.; Adler, A. Detection of the Number of Signals by Signal Subspace Matching. *IEEE Trans. Signal Process.* **2021**, *69*, 973–985. [[CrossRef](#)]