IEEE Transactions on Medical Imaging

EEG and MEG Inversion Using Convolutional and Recurrent Neural Networks

Journal:	*Transactions on Medical Imaging*
Manuscript ID	Draft
Manuscript Type:	Full Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Casanova, Joaquin; University of Florida, Electrical and Computer Engineering Stoecker-Sylvia, Zachary; Oceanit, Headquarters Miyamoto, Ryan; Oceanit, Headquarters Lin, Jenshan; University of Florida, Electrical and Computer Engineering
Keywords:	Electrophysical imaging < Imaging modalities, Brain < Object of interest, Machine learning < General methodology
Specialty/Area of Expertise:	

TRANSACTIONS ON MEDICAL IMAGING, VOL. XX, NO. YY

EEG and MEG Inversion Using Convolutional and Recurrent Neural Networks

Joaquin J. Casanova, *Member, IEEE,* Zachary D. Stoecker-Sylvia, Ryan Y. Miyamoto, *Member, IEEE,* and Jenshan Lin, *Fellow, IEEE*

***Abstract*—Real-time localization of neuronal activity has a number of uses, including brain-computer interfaces and medical diagnostics. Generally, this is done by taking measurements of the brain's magnetic and electric fields (magnetoencephalography [MEG] and electroencephalography [EEG]), and inverting the measurements. Most approaches are physics-based, and attempt to find the best estimate of the lead-field matrix, which relates the dipole activation to the field strength by minimizing the error of an estimate. To date, most approaches using the lead-field matrix are complicated and too slow in real time, aimed primarily at elucidating brain structural functional relationships from experimental data. We propose a new technique in which the location of peak neuronal current is estimated by treating EEG and MEG as a two-channel image, or time-series of images, which is processed by a neural network which returns the location of the dipole of peak magnitude. Four architectures are tested: 2-layer perceptron, convolutional neural network (CNN), recurrent neural network (RNN), and CNN feeding RNN. In the absence of true measures of neuronal activity, we used two publicly available MEG/EEG datasets, and treated the estimates of the traditional minimum-norm estimate (MNE) as true estimates. We test the four variations of the network architecture, and in the best case (CNN only), we achieve test dataset errors (RMSE of max dipole location) of 1.9 and 5.1 mm on two separate datasets of MEG/EEG time series from subjects with different types of stimulus.**

***Index Terms*—Electrophysical imaging, Brain, Machine learning.**

I. Introduction

THERE is a great need for interpretation of brain signals for both use in control of devices, for prosthetics, for example, or for disease diagnostics. Magnetoencephalography (MEG), the measurement of the brain's magnetic field, is typically done using cryogenic superconducting quantum interference devices (SQUID) on a dense grid surrounding the skull, and has been used in brain-computer interfaces [1]. Additionally, electroencephalography (EEG), is also used for BCI as well as medical diagnostics and study of brain stucture and function [2]. Particularly for the latter, there is great utility in using the external measurements of MEG and EEG to determine which parts of the brain are most electrically active, a problem known as localization or inversion. Functional magnetic resonance imaging (fMRI) can be used but instead measures blood flow, and is thus not a direct measurement of electrical activity [3].

J. Casanova and J. Lin are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, 32611 USA e-mail: jcasa@ufl.edu

R. Miyamoto and Z. Stoecker-Sylvia are with Oceanit.

Manuscript received

Classically, MEG/EEG inversion has followed one of a few approaches, with incremental enhancements as sophistication and computational power has increased. [4] provides an excellent review of techniques. [5] established the orignal framework of the dipole in a sphere model, in which a lead-field matrix based on magnetostatic and electrostatic equations, and the positions of dipoles, transforms a vector of dipole magnitudes into a vector of magnetic flux and electric potential at sensor locations. The problem of localization becomes a problem of finding the locations and magnitudes of dipoles that minimize the square error of predicted versus observed fields. Using this formalism and a finite-element model, [6] localized dipoles from MEG data. [7] used Kalman filtering to solve the dynamical inverse problem. The signal-space separation method transforms the data into basis functions in signal space to better filter background noise sources [8]. Beamformer techniques have also been applied for MEG/EEG inversion [9]. More recently, EEG has been inverted using artificial neural networks [10].

Neural networks have proven an important tool for accurate predictions from noisy and large datasets. A recent technique is convolutional neural networks, in which image inputs are convolved with weighted kernels [11]. To treat time-series or sequences, recurrent neural networks, incorporating temporal feedback, have been developed, such as the long-short-term memory RNN [12]. CNN and RNN have been combined in [13] to translate video sequences into text descriptions of the video.

Our proposed idea is to apply the combined CNN/RNN framework to sequences of MEG/EEG data, and predict the location of only the peak dipole. In this way we simplify the problem, and with approporiately trained networks, we could make quick predicts from data, possibly in real time. To demonstrate our technique, we use publicly available datasets, and use the dipole currents predicted by a well established technique (MNE) as the true value to which we compare our estimates and train our network [14]. Further, we investigate four variations of the network architecture (2-layer MLP, RNN, CNN, and CNN with RNN), to establish the advantages and disadvantages of each.

II. Methods

In this section, we describe our methodology, including datasets, preprocessing, and neural network architectures.

1) Datasets: To test our method, we used two multimodal datasets. Both were processed using the python-mne package [14].

One dataset [15] consisted of 19 subjects presented with stimulus of famous, unfamiliar, or scrambled faces. This data was obtained from the OpenfMRI database. Its accession number is ds000117. There were 102 magnetometers, 204 planar gradiometers, and 70 electrodes, with 3 electrodes used for ocular and cardiac arifacts. We excluded gradiometers, as our study is to develop an inversion technique for use with our own MEG system, which only has the radial magnetometers. We considered one subject only (sub007) as that subject's MRI was the only one to segment properly into skin, skull, and cortex layers (an important feature for the inversion step). We used the data processing pipeline provided with the dataset here: http://mne-tools.github.io/mne-biomag-group-demo/. The sampling rate was 1100 Hz, with 551 time points per trial, and oct6 (source spacing 4.9 mm) grid for the dipole mesh used in MNE. There were a total of 879 trials.

The second dataset [16] is included with MNE, and consists of trials where subjects received audio or visual stimulus: checkerboard patterns were presented into the left and right visual field, interspersed by tones to the left or right ear. The dataset includes, for each trial, one ocular channel (for artifact removal), 59 EEG channels 102, and MEG channels. The source grid used was oct5 (9.9 mm). The sampling rate was 150 Hz, for 106 time steps, with 239 trials.

A. Preprocessing

For both datasets, channels were preprocessed through low-pass filtering, followed by EOG/ECG artifact removal using direct measurements of EOG or ECG and/or independent component analysis. Data was epoched (cut into trials) but not averaged over each type of stimulus - each trial is considered as a separate datapoint. Further, prior to input to the neural network, the data was transformed by principle component analysis, to avoid numerical problems due to the small sensor units (fT and μV).

B. Description of Neural Networks

We considered four variations of the network structure, each of which was programmed using Google's Tensorflow API in Python (tensorflow.com). The simplest configuration was a multi-layer perceptron, where the PCA-transformed data is input as a vector at each timestep, and transformed by an input layer with ReLU activation functions, and then an output layer, with linear activation functions, to give the coordinate of the maximum current neuron at each timestep (Fig. 1). The next network type was a convolutional neural network, where the data at each timestep was interpolated to an 11 by 11 grid, such that the data forms a sequence of 2-channel (EEG and MEG) images. The grid size was chosen such that increasing the number of pixels had no marginal benefit on the model error. There were two convolutional layers, with kernel size 3 by 3. This was followed by a dense layer which is mapped to the output with a linear activation function. We also considered a recurrent neural network, with long short term memory (LSTM) cells, with an input layer, recurrent layer, and output layer, returning the dipole location of the maximum current neuron over the entire timeseries, not at each timestep. Finally, we considered a combination of the CNN and RNN networks (Fig. 2). Network parameters are given in Table I.

C. Training and testing

To train the networks, on each dataset the data was divided into randomly selected test, validation, and training sets, where the test set was 20% of the total dataset, validation was 20% of the remainder, and the remainder was divided into batches for training of 20% each. Cost was the RMSE of the location of the dipole, where the true value was taken as the Minimum Norm Estimate found using the MNE Python package. Validation error was logged every 100 steps, and each batch was trained for 1250 steps (giving 5000 train steps total). We used Adam optimization to train the network parameters [17] with a learning rate of 0.005. Five train/validation/test sets were randomly shuffled to ensure the results were reliable.

III. Results

Training and validation curves are given in Figs. 3 and 4, along with test dataset cost, for the auditory dataset and the faces dataset, respectively. Mean and standard deviation of test cost are given in Teble II. For both datasets, CNN and MLP networks outperform the networks including an RNN. While they are capable of achieving low training cost, they tend to overfit, and thus do poorly on the test set (test cost of around 45 mm for the faces dataset, 30 mm for the auditory dataset). For reference, the human head is about 75 mm in diameter. On the other hand, simpler networks did better, with the CNN only network performing best (cost of 5.1 mm for the auditory set, and 1.9 mm for the faces dataset). Fewer parameters allowed for better generalization, and the CNN layer allows us to preserve spatial information critical for inversion.

The poor generalization of the RNN networks led us to suspect that it may be too sensitive to noise in the location of the peak dipole. In other words, since we were using the MNE estimate of the dipole currents as truth, and these have some inherent error due to sensor noise and inversion error, the performance of the network to estimate the location may be poor. To test this, instead of comparing the neural network estimate to the peak dipole, we instead tried comparing it to the top 100 dipoles (which is approximately the top 5%-10%), and picking the closest one for calculating the cost. This has the effect of averaging out error in the dipoles estimated by MNE. Plots showing the training curves for this top-100 approach are shown in Figs. 5 and 6. Clearly, this reduces the test, training, and validation costs, but the performance ranking of the the four networks remains the same.

IV. Conclusion

We have presented a new technique for rapid inversion of MEG/EEG data, using neural networks to estimate the location of the peak dipole, rather than the current distribution over the cortex. In examining four possible network architectures, it was found that the CNN networks perform best in determining the location of the peak dipole at each timestep. More

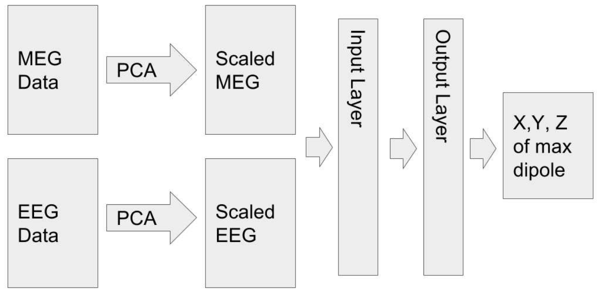

Fig. 1. Block diagram of MLP neural network.

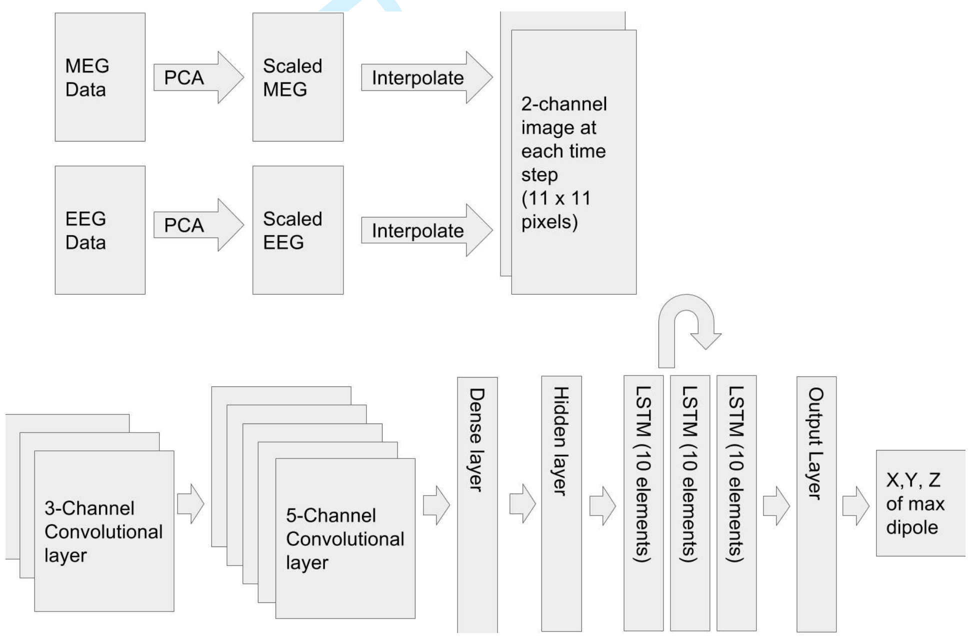

Fig. 2. Block diagram of CNN+RNN neural network.

Network	CNN	Hidden layer	RNN
MLP	N/A	10 units	N/A
CNN	2 layers (3, then 5 channels)	N/A	N/A
RNN	N/A	N/A	10 units, 3 layers
CNN+RNN	2 layers (3, then 5 channels)	N/A	10 units, 3 layers

TABLE I
NETWORK PARAMETERS.

complicated network architectures, or those without the CNN layers, performed worse, due to overfitting and loss of spatial information. Treating the EEG and MEG data as a two-channel image is an effective method for retaining patial field patterns, and is easily extensible to situations with more available sensors, such as NIRS. Future work may investigate using sequence-to-sequence neural machine translation to interpret timeseries.

Subject	MLP	CNN	RNN	CNN+RNN
Auditory (1 dipole)	5.387, 0.227	5.076, 0.134	35.270, 2.677	34.167, 2.373
Faces (1 dipole)	1.905, 0.023	1.895, 0.018	47.299, 2.249	42.576, 1.3027
Auditory (Best of 100 dipoles)	2.450, 0.404	2.138, 0.186	10.763, 0.645	9.752, 0.527
Faces (Best of 100 dipoles)	1.165, 0.012	1.121, 0.038	19.418, 1.600	17.820, 1.374

TABLE II
TEST SET RMSE FROM 5 CROSS-VALIDATION RUNS (MEAN, STANDARD DEVIATION)

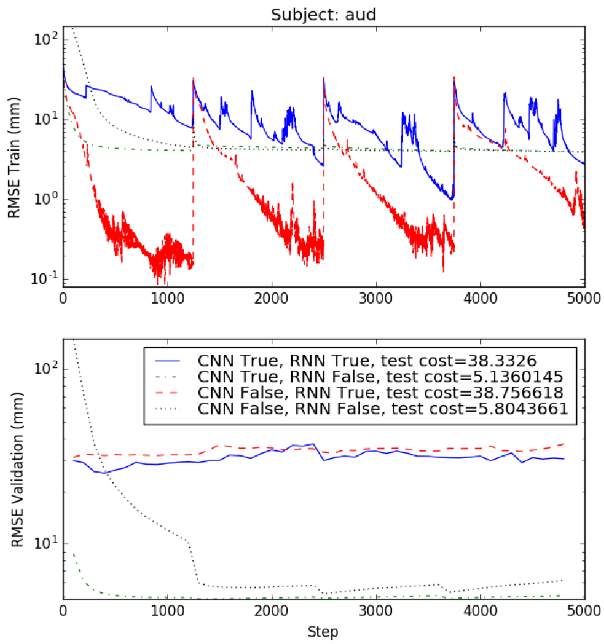

Fig. 3. Training/validation results for auditory stimulus dataset (1 dipole).

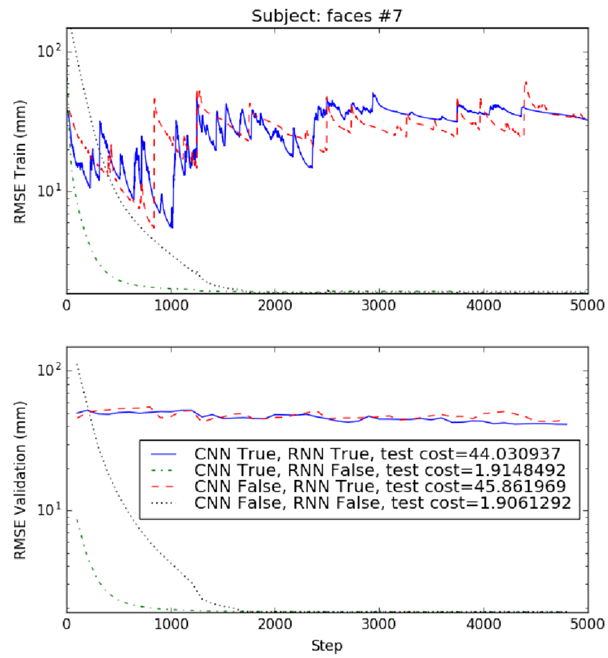

Fig. 4. Training/validation results for faces stimulus dataset (1 dipole).

ACKNOWLEDGMENT

The authors would like to thank DARPA for their support through contract W911NF-16-C-0057, as well as the MNE (https://martinos.org/mne/stable/index.html) project and OpenFMRI (https://openfmri.org/) project for software and datasets.

REFERENCES

[1] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, "An MEG-based brain–computer interface (BCI)," *Neuroimage*, vol. 36, no. 3, pp. 581–593, 2007.
[2] F. H. L. da Silva, "The impact of EEG/MEG signal processing and modeling in the diagnostic and management of epilepsy," *IEEE Reviews in Biomedical Engineering*, vol. 1, pp. 143–156, 2008.
[3] D. Sharon, M. S. Hämäläinen, R. B. Tootell, E. Halgren, and J. W. Belliveau, "The advantage of combining MEG and EEG: comparison to fMRI in focally stimulated visual cortex," *Neuroimage*, vol. 36, no. 4, pp. 1225–1235, 2007.
[4] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis," *Journal of neuroengineering and rehabilitation*, vol. 5, no. 1, p. 1, 2008.
[5] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatio-temporal meg data," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 6, pp. 541–557, 1992.
[6] H. Buchner, G. Knoll, M. Fuchs, A. Rienäcker, R. Beckmann, M. Wagner, J. Silny, and J. Pesch, "Inverse localization of electric dipole current sources in finite element models of the human head," *Electroencephalography and clinical Neurophysiology*, vol. 102, no. 4, pp. 267–278, 1997.
[7] A. Galka, O. Yamashita, T. Ozaki, R. Biscay, and P. Valdés-Sosa, "A solution to the dynamical inverse problem of EEG generation using spatiotemporal kalman filtering," *NeuroImage*, vol. 23, no. 2, pp. 435–453, 2004.
[8] S. Taulu, J. Simola, and M. Kajola, "Applications of the signal space separation method," *IEEE transactions on signal processing*, vol. 53, no. 9, pp. 3359–3372, 2005.
[9] K. Sekihara, S. S. Nagarajan, D. Poeppel, A. Marantz, and Y. Miyashita, "Reconstructing spatio-temporal activities of neural sources using an meg vector beamformer technique," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 7, pp. 760–771, 2001.
[10] G. Van Hoey, J. De Clercq, B. Vanrumste, R. Van de Walle, I. Lemahieu, M. D'Havé, and P. Boon, "EEG dipole source localization using artificial neural networks," *Physics in medicine and biology*, vol. 45, no. 4, p. 997, 2000.
[11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
[12] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[13] S. Venugopalan, H. Xu, J. Donahue, M. Rohrbach, R. Mooney, and K. Saenko, "Translating videos to natural language using deep recurrent neural networks," *arXiv preprint arXiv:1412.4729*, 2014.
[14] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen, "MNE software for processing MEG and EEG data," *Neuroimage*, vol. 86, pp. 446–460, 2014.
[15] D. G. Wakeman and R. N. Henson, "A multi-subject, multi-modal human neuroimaging dataset," *Scientific data*, vol. 2, 2015.
[16] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen *et al.*, "MEG

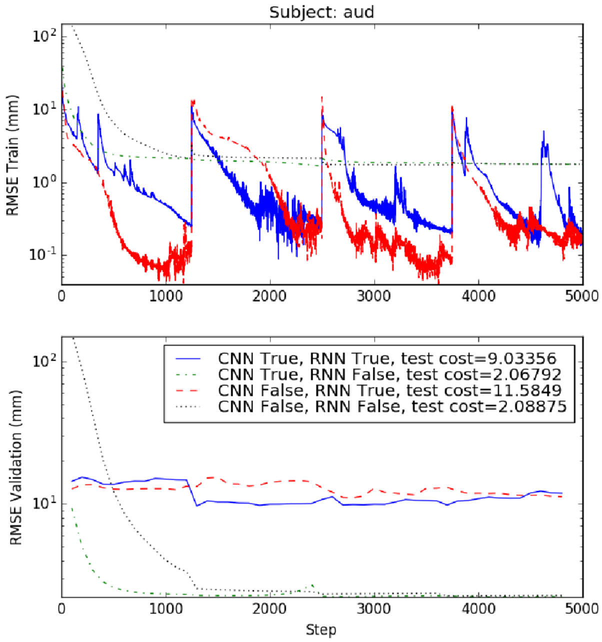

Fig. 5. Training/validation results for auditory stimulus dataset (Best of 100 dipoles).

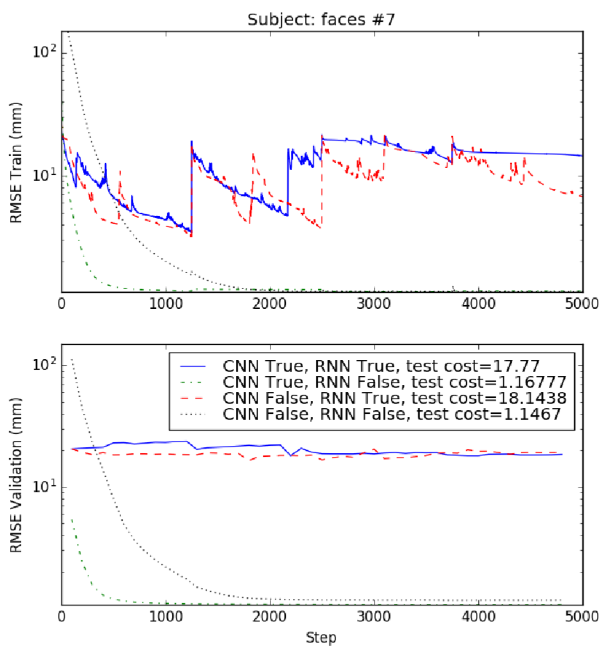

Fig. 6. Training/validation results for faces stimulus dataset (Best of 100 dipoles).

and EEG data analysis with MNE-Python," *Frontiers in neuroscience*, vol. 7, p. 267, 2013.

[17] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.