Analysis of Cortical Connectivity using Hopfield Neural Network

S.Dixit
Department of Neurology, University of Medicine and Dentistry of New Jersey,
Newark, NJ, USA
&
K.Mosier
Departments of Radiology and Surgery, Memorial Sloan-Kettering Cancer Center,
New York, NY, USA

Abstract: Functional Magnetic Resonance Imaging (fMRI) is increasingly recognized as a standard technique for brain mapping and determining the connectivity between cortical regions. Statistical approaches to determining cortical connectivity, e.g. Structural Equation Modeling between various regions of interest (ROI) in the active brain can be computationally inefficient. This study explores the utility of a Hopfield Neural Network to determine cortical connectivity in an fMRI data set.

Introduction:

fMRI has come to be recognized as a very important technique in analyzing functional brain activity. fMRI data are acquired as the subject performs some cognitive, sensory or motor task and the data displayed as a colorized map of active regions overlaid on MR images. An important step is to determine how these activated regions are functionally connected during performance of a particular task. Structural Equation Modeling (SEM) is now being widely used as statistical tool to find connectivity between various regions of the brain [1-3]. However, there are a number of limitations to SEM. SEM techniques are time consuming and tedious. SEM is based on multiple regression techniques to compare a connectivity model based on the observed data to a connectivity model based on hypothesized or known neuroanatomical models. Fitting models containing a large number of brain regions, for example, often requires multiple parameter changes, or changes to the model, to achieve a good fit. Moreover, as research extends into higher cognitive functions, neuroanatomical animal models are of limited value in guiding these decisions [4]. Neural networks may offer advantages to circumvent some of these limitations in SEM. In particular, neural network techniques may provide increased computational speed for larger

data sets, and the ability to efficaciously determine the effects of changing connectivity weights in one region of a large network. This paper presents the application of a neural network to simulate the way different areas in brain are connected while performing sensorimotor object exploration tasks.

Purpose:

Our purpose in this investigation was to simulate the cortical networks using Artificial Neural Networks during manipulation of objects of different geometry in the absence of visual input. The question we wished to address was whether the cortical connectivity model obtained using Artificial Neural Network would be consistent with the model obtained by Structural Equation Modeling.

Materials and Methods:

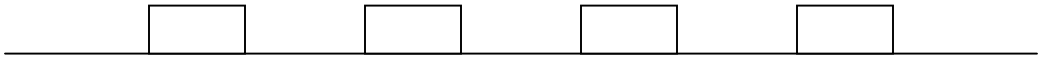
A total of four healthy adult human subjects were imaged in this study (3M, 1F; mean age = 33 years). The subjects were imaged on a GE 1.5T MR scanner using conventional Blood Oxygen Level Dependent (BOLD) techniques. The subjects were positioned supine in the bore with their heads centered in a standard "bird-cage" quadrature head coil. The subjects wore acoustically shielded headphones to attenuate gradient noise. In order to reduce potential head movement, their heads were restrained with cushions and tape. The functional MR data sets were acquired using a gradient echo echo-planar sequence with the following parameters: TR/TE = 4000/60 msec, 64x64 matrix, 24 cm field of view (FOV), 5 mm slice thickness and 90^0 flip angle. A total of 28 slices were acquired in the axial plane. During the same imaging session, high-resolution T1-weighted (TR/TE=450/16msec, 256x256, 24cm FOV, 5 mm slice thickness, 28 slices) images were obtained in the same axial planes for co-

registration with the functional data sets. Prior to the imaging session, the subject's hand dominance was determined using established survey criteria (Edinburgh Handedness Survey, Olfield, 1971). Institutionally approved informed consent was obtained for all subjects.

The task paradigm consisted of a mixed block and event-related design in which subjects freely explored 1 cm wooden objects of different geometry (square, triangle, rectangle, and cylinder) with the fingers of the dominant hand. Each imaging trial consisted of four blocks of object exploration for 20 seconds alternating with baseline or rest blocks of 30 seconds. In each exploration block, the subject explored a different object (Figure 1). The subject began the trial with dominant hand in the supine position with the arm in an extended but comfortable position. At the onset of each exploration block, an object was placed onto the palmar surface of the hand by the investigator. Subjects were instructed prior to the beginning of the experiment that when they felt the object in the palm, to begin exploration with the fingers. All imaging trials sessions were performed with the subject's eyes closed to avoid visual input, and subjects were monitored directly by the investigator to ensure compliance. All subjects were compliant with the task. Prior to imaging, the subjects were not informed what the objects would be, nor were they instructed to try to determine the nature or shape of the object.

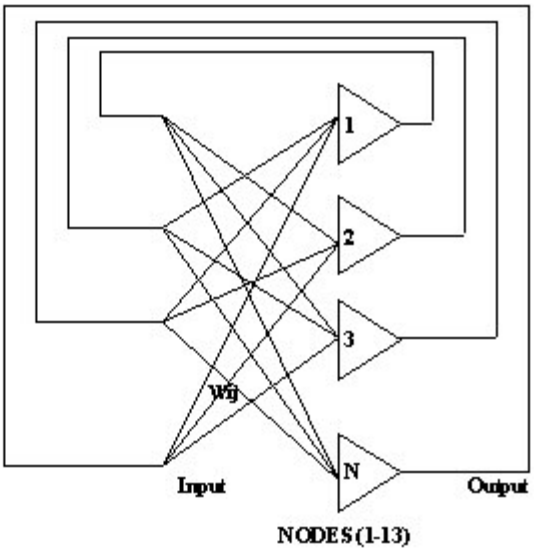
The raw echo-planar data were reconstructed and processed off-line using a custom algorithm incorporating routines written in IDL (Interactive Data Language, Boulder, CO) that run SPM99 (Statistical Parametric Mapping; Welcome Dept. of Neurology). In this study, a voxel-level significance of $p=0.001$ was achieved using a Z-threshold of 3.09. An automated region-of-interest (ROI) analysis was performed to detect all activated voxels within each brain gyrus, and the results tabulated with the following parameters: location

(gyrus or subcortical location: nuclei of the basal ganglia and thalamus), Z-score of the activated voxels, and the volume of the activated voxels. Activated voxels with a volume less than 2 were excluded from the data. The location of the activated voxels were designated according to MNI (Montreal Neurological Institute) coordinates and their localization confirmed by visual inspection of the activation maps overlaid to the T1-weighted images using conventional neuroanatomical landmarks. The product of the Z-score and the volume of activated voxels (STOT) for each activated brain region was used as the input for both the SEM and neural network analysis.



Neural Network analysis:

We employed an auto-associative neural network for this analysis (Figure 2). The auto-associative neural architecture used in our study was initially described by Hopfield [5]. The neural network consists of a set of nodes, which are connected to each other and which are arranged in layers. The network was a two-layer network consisting of an input layer and an output layer. The input layer was composed of a 13x1 vector of STOT values representing each of the activated brain areas.



Each activated brain area corresponded to a single node. The net synaptic input to the network was determined by $E_i = \sum_{j*i} w_{ji} \ (2* s_j - 1)$, where E_i = net Synaptic Weight, Sj = State of neuron (active or inactive), and w_{ji} = Weight of connection between neuron *j* and neuron *i.*

Training of the neural network was unsupervised; the system learns locally and automatically; the neurons adjusted their interconnectivity by adjusting their weights depending on their co-activation. Hebb's rule [6] was used for training of the network such that: dw = lr*a*p, where dw is the weight change for a given neuron, p = neuron's input, a= neuron's output and lr = learning rate. Hebb's rule is interpreted in a neurobiological sense in two parts: "(1) If two neurons on either side of a synapse (connection) are activated simultaneously (i.e. synchronously), then the strength of that synapse is selectively increased. (2) If two neurons on either side of a synapse are activated asynchronously, then that synapse is selectively weakened or eliminated." [7] The temporal resolution of the fMRI data collected in these experiments is considerably less than that for electrophysiological methods because the TR (time to repetition) was four seconds. Thus synchronous or asynchronous

behavior cannot be measured. The input vector used here is based on hemodynamic responses, rather than firing neurons, therefore, the connectivity reflects the covariance in activity between different brain areas over the time course of the image acquisition. Due to these temporal and hemodynamic constraints, the time delay between the different nodes in the network is assumed to be equal. The synaptic weights from one node to another can be either positive or negative. In the initial training set, all brain areas or nodes were assumed to be connected to all other areas.

The goal of ANN training was to obtain a stable weight matrix with Mean Square Error (MSE) as near zero as possible. Twenty iterations were required to reduce the MSE from 1.5 to zero. Once trained the number of epochs reduced considerably from twenty to ten or less epochs. The areas, which had inter-connection weights of more than 0.2, were retained in the model and were considered to be grouped together.

Structural Equation Modeling

Structural Equation Modeling (or Path Analysis) employs multiple regression techniques to determine connectivity between different brain regions in brain imaging data sets. In this study, a Pearson correlation was performed on the matrix of STOT values. The correlation matrix was decomposed with Principal Component Analysis in order to search for covariances in the time series suggestive of grouping or clustering among the different brain areas. A regression model is used to fit the data, $\eta = \beta n + \psi$, where η is the vector of variances, β = the matrix that defines the network, and ψ = vector of residual effects. The particular structural model used was $Y = (I-B)^{-1} \Phi (I-B')^{-1}$ (Systat v.10, SPSS, Chicago, ILL). The connectivity between regions in the structural model (the path coefficients) were defined either by the correlation coefficients from the Pearson correlation matrix, or by the inner

product of the loading factors from the PCA. Both the correlation matrix and the PCA loading factors were used to define the path coefficients in order to test different models of connectivity. The fit of the data to the model is determined by iterative fitting using a generalized least squares algorithm and the goodness of fit determined through chi-square analysis [8].

Results:

Training of the network as described above was achieved within twenty iterations to reach a desired MSE of zero (Figure 3).

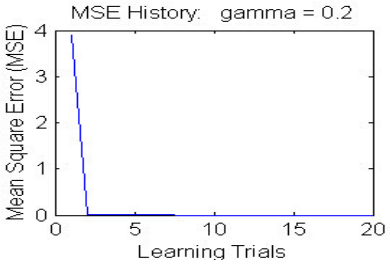


Figure 3. History of MSE as a function of epochs

The input matrix to the network consisted of 13 x 1 vector and the weight matrix of 13x13 was obtained after 20 iterations. The weight matrix obtained was used to determine connectivity between the different brain areas. Areas having interconnectivity weights of more than 0.2 were kept in the model. The areas showing high interconnection weights had the highest correlation and these areas were grouped together as shown in figure 4 (M1= primary motor cortex; SFG = superior frontal gyrus; CERBLM = cerebellum; S1 = primary sensory cortex; MFG = middle frontal gyrus; THAL = thalamus; SMG = supramarginal gyrus; SPL = superior parietal lobe; AG = angular gyrus; HIPPO = hippocampus; ITG= inferior temporal gyrus; IFGO = inferior frontal gryus; OCCIP = occipital.

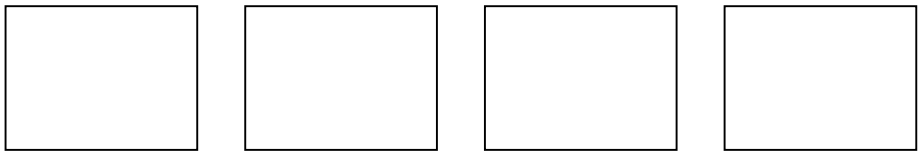
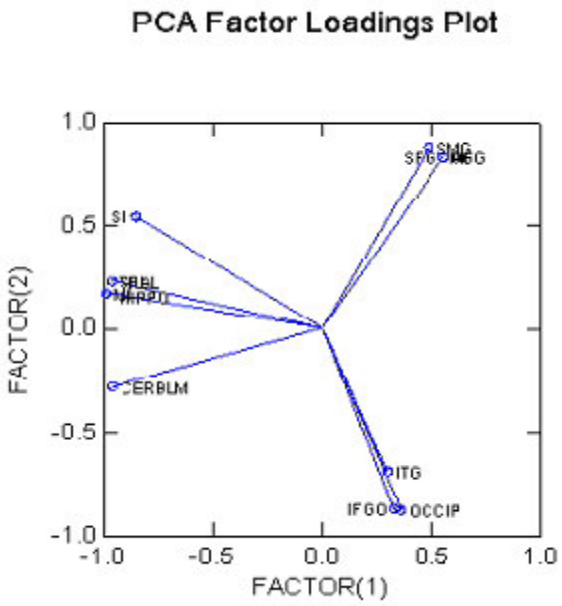


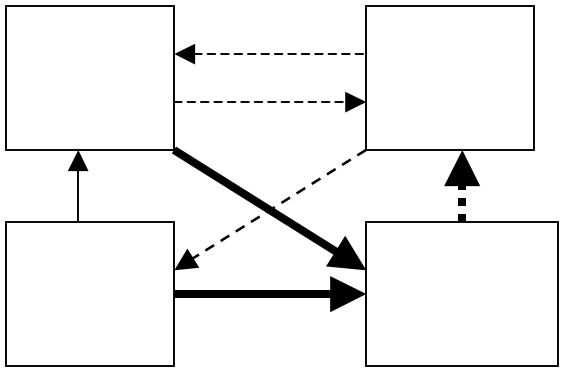
Figure 4 Groupings of the different brain areas obtained from the weight matrix

Principal components analysis of the input data yielded two main principal components accounting for 53% and 41% of the variance, respectively. The distribution of the different brain areas along the principal components is shown in Figure 5. A comparison



of Figures 4 and 5 shows that the grouping of brain region is consistent between that derived from the neural network analysis and that derived from the PCA. The connectivity model between the different groups of brain areas was determined through successive iterations of SEM fitting of the data to various hypothesized models. A distributed model of connectivity

in which each brain area was independent and reciprocally connected to every other brain area was tested and was not statistically significant, p= 0.000 (indicating a statistically significant difference between the hypothesized model and the data). The Pearson correlation matrix (cf. Methods, Structural Equation Modeling) was likewise used to determine a connectivity model, however this model did not fit the data p=0.000. The model obtained from the loading factors of the PCA was statistically significant, p =0.13 (no significant difference between the hypothesized model and the data). Figure 6 shows the connectivity model among the brain areas derived from the PCA.



In comparison, the weight matrix obtained through the ANN demonstrates connectivity between the different brain areas consistent with the model obtained through SEM (Figure 4).

Discussion:

The neural network used in this investigation attained a mean square error (MSE) in twenty iterations. This rapid training result was achieved due to the relatively small input vector. The weight matrix obtained after training showed correlations that were consistent

with those determined through SEM. Thus, the neural network extracted the covariance structure in the data that achieved the best fit of the connectivity model.

The results of this study demonstrate that a simple two-layer auto-associative neural network can effectively capture the essential features of a connectivity model of a whole brain cortical network. The input data to the neural network contains a covariance structure based on active brain areas. The goal of an auto-associative neural network is to reproduce the pattern of an input vector. Thus a Hopfield neural network is well suited to determine connectivity models derived from fMRI data.

References:

1. McKiernan, K.A., Conant, L.L., Chen, A., Binder, J.R. (2001, June). *Development and cross-validation of a model of linguistic processing using neural network and path analysis with FMRI data*. Poster presented at the annual meeting of Human Brain Mapping, Brighton, UK

2. Horwitz B, Tagments MA, McIntosh AR. Neural modeling, functional brain imaging and cognition. Trends in Cogn Sci 3(3): 91-98, 1999.

3. Horwitz B, Friston KJ, and Taylor JG. Neural Networks 13(8-9): 829-46, 2000.

4. Milner AD and Goodale MA. The visual brain in action. Oxford, Oxford Univ. Press, 1995.

5. Hopfield, J.J.Neural Network and physical systems with emergent collective computational abilities. Proceedings *of National Academy of Science 79(1982*). 2554-2558.

6. Hebb, D.O. Organization of behavior (New York, Wiley, 1949).

7. Haykin S. Neural Networks A Comprehensive Foundation. Prentice Hall, Upper Saddle River, NJ, 1999.

8. McIntosh AR and Gonzalez-Lima F. Structural Equation Modeling and It's Applications to Network Analysis in Functional Brain Imaging. Human Brain Mapping 2: 2-22, 1994.

Bio-sketch:

Shailja Dixit MD, M.S.

Shailja is the Director of the EEG laboratory at The University Hospital, University of Medicine and Dentistry of New Jersey and a Clinical Instructor in Department of Neurology at New Jersey Medical School. She is also Adjunct faculty in School of Public Health, NJMS. Her research work focuses on Analysis of fMRI data to study Cortical Connectivity using Artificial Neural Networks.

Kristine Mosier DMD, Ph.D.

Kristine is a an assistant attending (assistant professor) of radiology and surgery at Memorial Sloan Kettering Cancer Center and Adjunct Assistant Professor of Radiology at the New Jersey Medical School, Newark, NJ. Her laboratory is focused on how cortical networks are organized in the planning and execution of different sensorimotor behaviors.