Enhancing Local Linear Models Using Functional Connectivity for Brain State Decoding

Orhan Fırat, Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

Mete Özay, Department of Computer Engineering, Middle East Technical University, Ankara, Turkey & Department of Computer Science, University of Birmingham, Birmingham, UK & Department of Electrical Engineering, Princeton University, Princeton, NJ, USA

İtir Önal, Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

Ilke Öztekin, Department of Psychology, Koç University, Istanbul, Turkey

Fatoş T. Yarman Vural, Department of Computer Engineering, Middle East Technical University, Ankara, Turkey

---

ABSTRACT

The authors propose a statistical learning model for classifying cognitive processes based on distributed patterns of neural activation in the brain, acquired via functional magnetic resonance imaging (fMRI). In the proposed learning machine, local meshes are formed around each voxel. The distance between voxels in the mesh is determined by using functional neighborhood concept. In order to define functional neighborhood, the similarities between the time series recorded for voxels are measured and functional connectivity matrices are constructed. Then, the local mesh for each voxel is formed by including the functionally closest neighboring voxels in the mesh. The relationship between the voxels within a mesh is estimated by using a linear regression model. These relationship vectors, called Functional Connectivity aware Mesh Arc Descriptors (FC-MAD) are then used to train a statistical learning machine. The proposed method was tested on a recognition memory experiment, including data pertaining to encoding and retrieval of words belonging to ten different semantic categories. Two popular classifiers, namely k-Nearest Neighbor and Support Vector Machine, are trained in order to predict the semantic category of the item being retrieved, based on activation patterns during encoding. The classification performance of the Functional Mesh Learning model, which range in 62-68% is superior to the classical multi-voxel pattern analysis (MVPA) methods, which range in 40-48%, for ten semantic categories.

Keywords: *Brain Decoding, Feature Extraction, Functional Magnetic Resonance Imaging (fMRI), Machine Learning, Multi Voxel Pattern Analysis (MVPA), Pattern Classification*

DOI: 10.4018/ijcini.2013070103

Several methods have been developed to understand how brain processes information. One in particular, aims to predict or decode the brain state, and/or the type of information associated with cognitive processes, based on distributed patterns of activation in the brain, acquired with functional magnetic resonance imaging (fMRI) using machine learning methods (Battle, Chichik, & Koller, 2006; Bressler & Tognoli, 2006; Haynes & Rees, 2006; Norman, Polyn, Detre, & Haxby, 2006; Richiardi, Eryilmaz, Schwartz, Vuilleumier, & Van De Ville, 2011; Shirer, Ryali, Rykhlevskaia, Menon, & Greicius, 2012; Wang, Hutchinson, & Mitchell, 2003). One of the major motivations of this study is to propose a model for pattern analysis of fMRI data pertaining to different cognitive states using statistical learning theory. This representation involves understanding, manipulating and predicting the behavior of the very complex nature of human brain. Massively coupled dynamic interactions of the brain at many scales cannot be fully understood by only employing the measurements recorded from the individual voxels. Therefore, there has been growing interest in using brain connectivity to reveal interactions between spatially distant regions. Brain connectivity describes neural processes as the outcomes of dynamic coordination among smaller elements (Bullmore & Sporns, 2009). Three main types of brain connectivity are reported in the literature: i) structural connectivity, basically reveals anatomic connections (pathways) of brain, such as physical links between neural elements, ii) functional connectivity, defined as statistical dependence between remote neural elements or regions across time, e.g. correlation and iii) effective connectivity, which analyzes brain connectivity using causal effects between neural elements, resulting in causal activation paths (Bressler & Menon, 2010; Friston & Friston, 1994).

Connectivity for decoding is mostly used for model selection and/or defining the neighborhood of seed neural elements, or regions (Li, Guo, Nie, Li, & Liu, 2009). For instance, in a study by (McIntosh, Nyberg, Bookstein, & Tulving, 1997), partial least squares for activation analysis were performed to construct a

cross block covariance matrix using PET data. Correlation based measures such as correlation/partial correlation, Granger causality, mutual-information or coherence are used for the selection of different functional interdependence functions. Ryali et al. (2012) measure sparse-partial correlation between multiple regions using elastic net penalty, which combines L1 and L2 norm regularization terms, in order to improve the sensitivity of the correlation measure. Patel, Bowman, and Rilling (2006) proposed a conditional dependence model which accounts for an imbalance between class conditional and posterior probabilities, to achieve at a measure of connectivity. Unlike correlation measures, Shirer et al. (2012) train a classifier to decode cognitive states after constructing functional connectivity matrices, analyzing increasing connectivity regions by subtracting connectivity matrices for each state. Richiardi et al. (2011) constructed functional connectivity matrices by using pair-wise Pearson correlation coefficients and employ graph matching to decode brain states.

In this study, we introduce an algorithm for modeling cognitive processes, based on the functional and structural connectivity in the brain. Structural connectivity is utilized for anatomic parcellation of the brain regions by clustering the voxel intensity values measured by fMRI. Next, functional connectivity is utilized within the clusters by different correlation measures. Functional connectivity matrices are formed to define functional neighborhood of a voxel. A local mesh is formed for each voxel (called the seed voxel) by including the -functionally closest neighbors (called the surrounding voxels) in the mesh. The relationships between the seed voxel and the surrounding voxels are modeled by estimating the arc weights of the mesh in a linear regression model. The arc weights, called *Functional Connectivity-aware Mesh Arc Descriptors* (FC-MAD) represent the relationship of each voxel to its functionally closest neighbors. Finally, the proposed FC-MAD features are used to train a classifier which recognizes type of information and/or cognitive state. In the current study, we particularly focused on classification of the type of information

being encoded and retrieved during memory operations. During the experiment, participants studied a list of words selected from one of ten pre-defined semantic categories and made recognition memory judgments while neural activation was recorded using fMRI (Oztekin & Badre, 2011; Oztekin, Curtis, & McElree, 2009). Accordingly, we tested whether is the proposed machine learning algorithm can successfully identify and differentiate of the type of information (i.e. the semantic category that the word belongs to) which is represented in the brain at a given time considering distributed patterns of brain activity associated with and during memory encoding and retrieval.

MESH LEARNING AND MESH ARC DESCRIPTORS (MAD)

Descriptive Analysis and Motivation for the Mesh Model

In this study, BOLD signals $v(t_i, \bar{s}_j)$ are measured at time instants t_i , $i = 1, 2, 3, \dots, N$, at voxel coordinates $\bar{s}_j$, $j = 1, 2, 3, \dots, M$, where N is the number of time samples, and M is the number of voxels. The data set $D = \{v(t_i, \bar{s}_j)\}$ consists of the voxels $v(t_i, \bar{s}_j)$, which are distributed in brain in three dimensions. Therefore, the position $\bar{s}_j = (x_j, y_j, z_j)$ of a voxel $v(t_i, \bar{s}_j)$ at time instant t_i is a three dimensional vector. At each time instant t_i , the participant is processing (either encoding or retrieving) a word belonging to a cognitive process. Therefore, the samples $v(t_i, \bar{s}_j)$ has an object label at each time instance. In Mesh Learning (Ozay, Öztekin, Öztekin, & Vural, 2012), the cognitive states are modeled by local meshes for each individual voxel, called seed voxel $v(t_i, \bar{s}_j)$, which is defined in a neighborhood system η_p (see Figure 1). In this mesh, voxel $v(t_i, \bar{s}_j)$ is

connected to p -nearest neighboring voxels $\{v(t_i, \bar{s}_k)\}_{k=1}^p$ by the arcs with weights $\{a_{i,j,k}\}_{k=1}^p$. Therefore, the relationship among the BOLD signals measured at each voxel, are represented by the arc weights. p -nearest neighbors, η_p , are defined as the spatially-nearest neighbors to the seed voxel, where the distances between the voxels are computed using Euclidean distances between the spatial coordinates $\bar{s}_j$ of the voxels in brain. The arc weights $a_{i,j,k}$ of the mesh are estimated by the following linear regression equation:

$$v(t_i, \bar{s}_j) = \sum_{\bar{s}_k \in \eta_p} a_{i,j,k} v(t_i, \bar{s}_k) + \varepsilon_{i,j}, \quad (1)$$

where $\varepsilon_{i,j}$ indicates the error of voxel $v(t_i, \bar{s}_j)$ at time instant t_i , which is minimized for estimating the arc weights $a_{i,j,k}$. This procedure is conducted by minimizing the expected square error defined as follows,

$$E(\mu_{i,j}^2) = E \left[\left(v(t_i, \bar{s}_j) - \sum_{\bar{s}_k \in \eta_p(\bar{s}_j)} a_{i,j,k} v(t_i, \bar{s}_k) \right)^2 \right], \quad (2)$$

where $\eta_p(\bar{s}_j)$ is the set of p -nearest neighbors of the j^{th} voxel at $\bar{s}_j$.

Minimizing Equation (2) with respect to $a_{i,j,k}$ is accomplished by employing Levinson-Durbin recursion (Vaidyanathan, 2007), where $E(\cdot)$ is the expectation operator. The arc weights $a_{i,j,k}$, which are computed for each seed voxel at each time instant t_i , is used to form the mesh arc vector $\bar{a}_{i,j} = [a_{i,j,1} \ a_{i,j,2} \ \dots \ a_{i,j,p}]$. Furthermore, a mesh arc matrix A_j is constructed by concatenating the mesh arc vectors at each time

Figure 1. The Star Mesh, which represents the voxel intensity values $v(t_i, \bar{s}_k)$ at the center and its 4-nearest neighbors. Blue node represents the center voxel and the orange nodes represent the surrounding voxels.

instant, $A_j = [\bar{a}_{1,j} \ \bar{a}_{2,j} \ \dots \ \bar{a}_{N,j}]^T$. Finally, feature matrix $F = [A_1 \ A_2 \ \dots \ A_M]$ which represents the Mesh Arc Descriptor (MAD), is constructed. The feature matrix, extracted during both memory encoding and retrieval stages is further used in training and testing phases in the classification of cognitive processes, respectively. For the details of the mesh learning algorithm see Firat et al. (2012) and Özay, Öztekin, Öztekin, and Yarman Vural (2011).

The motivation of representing voxels in the brain by local meshes can be validated by analyzing an individual voxels' intensity change and the change of the sum of squared difference of intensities

$$d_{\bar{s}_j, \eta_p(\bar{s}_j)} = \sum_{\bar{s}_k \in \eta_p(\bar{s}_j)} \left[v(t_i, \bar{s}_j) - v(t_i, \bar{s}_k) \right]^2 \text{ in the}$$

neighborhood of that voxel in time. Individual voxel intensity values, which are measured at each time instant, do not possess any discriminative information as illustrated in Figure 2 with red line. Note that the signal intensity

value for a voxel is almost constant for each time instant. Since the measurements along the time axis correspond to separate cognitive processes, in most of the problems, it is unlikely to discriminate them by using multi-voxel pattern analysis (MVPA) methods, which classify the voxel intensity values by a machine learning tool. On the contrary, there is a slight variation of the sum of squared distances of intensity values in differing neighbor sizes. The above observation shows that the relationships among voxels carry more information than individual voxel intensity values, at each time instant.

Descriptive Analysis and Motivation for the Mesh Model

The estimated MAD vectors, which represent relationships among the voxels in the same neighborhood system, have a high discriminative power compared to the individual voxel intensity values. As a result, the Mesh learning algorithm performs better than the well-known

Figure 2. Sum of squared difference, $d_{\bar{s}_j, n_p(\bar{s}_j)}$, of intensity values for a voxel and its N -nearest neighboring voxels over time in log space. The time axis indicates the fMRI measurements from 10 semantic categories.

individual voxel based algorithms (see also; Table 2). However, employing the Euclidean distance to form the neighborhood system may not fully represent the activation patterns in the brain, where the spatially distant neurons might exhibit functional connectivity. "Nearest" neighborhood in the mesh model implies spatial surroundings of the seed voxel when Euclidean distance is used, which may not be fully the case during cognitive processing. Additionally, it is well known that spatially close voxels are strongly coupled during cognitive processes (Sporns, 2011). Therefore, using Euclidean distance for defining neighborhoods for voxels may cause redundant meshes and mesh arc weights in a feature matrix. A partial improvement for this problem can be accomplished by the usage of functional connectivity. Selecting functional neighbors for each voxel and constructing the meshes based on the functional neighborhood results in a more discriminative feature matrix improving the classification performance.

Functional Connectivity

Given the time series of voxels $v(t, \bar{s}_i)$ and $v(t, \bar{s}_j)$, where $t = (t_1, t_2, \dots, t_N)$ is the time

vector and the variables are consecutive time instants, functional connectivity is defined as the measure of similarity between time series of these voxels. The voxels are considered to be functionally connected if they have similar functional properties. Therefore, the functional connectivity depends on the similarity measure. The similarity can be measured, for example, by estimating the correlation or covariance between pairs of time series. Functional connectivity is expected to capture patterns of deviations between distributed and often spatially remote regions in brain (Sporns, 2010) and constructed using an inter-regional analysis, where the regions are connected according to their functional properties (Friston & Friston, 1994).

Motivated by the benefits of functional connectivity in Mesh Learning; we further used functional connectivity to define the functional neighborhood, and to form the meshes for extracting MAD.

The estimated MAD vectors, which represent relationships among the voxels in the same neighborhood system, have a high discriminative power compared to the individual voxel intensity values. As a result, the Mesh learning

algorithm performs better than the well-known individual voxel based algorithms (see also; Table 2). However, employing the Euclidean distance to form the neighborhood system may not fully represent the activation patterns in the brain, where the spatially distant neurons might exhibit functional connectivity. "Nearest" neighborhood in the mesh model implies spatial surroundings of the seed voxel when Euclidean distance is used, which may not be fully the case during cognitive processing. Additionally, it is well known that spatially close voxels are strongly coupled during cognitive processes (Sporns, 2011). Therefore, using Euclidean distance for defining neighborhoods for voxels may cause redundant meshes and mesh arc weights in a feature matrix. A partial improvement for this problem can be accomplished by the usage of functional connectivity. Selecting functional neighbors for each voxel and constructing the meshes based on the functional neighborhood results in a more discriminative feature matrix improving the classification performance.

Functional Connectivity Graph

In order to represent functional connectivity in brain, we define a graph $G = (V, E)$. $G = (V, E)$, where $V = \{\vartheta_j\}_{j=1}^M$ is the set of nodes (vertices) and $E = \{e_{jk}\}_{j,k=1}^M$ is the set of edges. In this representation, a node ϑ_j corresponds to a time series, $v(\mathbf{t}, \bar{s}_j)$, which is measured at an individual voxel and an edge between ϑ_j and ϑ_k , is represented as $e_{jk} = \rho_{jk}(h)$, where $\rho_{jk}(h)$, is the functional connectivity coefficient, which is computed using a functional similarity measure time series of voxel signals $v(\mathbf{t}, \bar{s}_j)$, $j = 1, 2, 3, \dots, M$, at lag h using Equation (3).

In this study, edges in the functional connectivity graph are represented by symmetric dependence measures, in the time domain. It has been suggested that correlation based measures are well suited for functional connectivity analysis (Smith et al., 2011). Consequently,

we use cross-correlations to measure the functional similarity between time-series. The cross-correlation coefficient between two nodes, voxels ϑ_j and ϑ_k in our case, at lag h , $\rho_{jk}(h)$, is given by

$$\rho_{jk}(h) = \frac{\text{cov}_{jk}(v(\mathbf{t}, \bar{s}_j), v(\mathbf{t} + h, \bar{s}_k))}{\sqrt{\text{var}_j(v(\mathbf{t}, \bar{s}_j)) \cdot \text{var}_k(v(\mathbf{t}, \bar{s}_k))}}, \quad (3)$$

where cov_{jk} is the covariance of the signals measured at two voxels, and var_j is the variance of the signals measured at a voxel $v(\mathbf{t}, \bar{s}_j)$. If $\rho_{jk}(h) = \rho_{jk}(-h)$, $\rho_{jk}(h) \in [-1, 1]$ and $h = 0$, then we compute zero-order correlation using (3).

Local Patches

Constructing a functional connectivity graph by considering all voxels as individual nodes introduces scalability problems. In order to reduce the computational complexity, the voxels are first clustered with respect to their locations, where each cluster is called a local patch. Then, the functional connectivity graph is formed for the voxels in each local patch with approximately size π . This approach reduces the computational complexity from $O(M^2)$ to $O(C\pi^2)$ where M is the number of voxels, C is the number of local patches. Note that $\pi \ll M$, in practice.

The local patches are constructed by clustering the whole dataset $D = \{v(\mathbf{t}_i, \bar{s}_j)\}$, $i = 1, 2, 3, \dots, N$, $j = 1, 2, 3, \dots, M$, using Euclidean distance among spatial locations of voxels $\bar{s}_j = (x_j, y_j, z_j)$ in a self-tuning spectral clustering algorithm (Manor & Perona, 2004). After partitioning the whole dataset D into C clusters, functional connectivity is measured locally within these clusters. A cognitive process is then represented in a local patch (cluster) m ,

using a within cluster functional connectivity matrix FC_m , each of which forms the set of functional connectivity matrices $FC = \{FC_m\}_{m=1}^C$ which is employed in the model selection for mesh learning algorithm. Details of the within cluster functional connectivity matrix computation process are given in Algorithm 1. Figure 3, represents the local connectivity patterns in two generated clusters.

Functionally Connected Mesh

We define a local mesh around each voxel which consists of the set of functionally connected voxels. These meshes are then used to extract MAD features from the meshes which consist of functionally similar voxels. The suggested model is called Functional Mesh Learning and the extracted MAD features are called Functional Connectivity aware MeshArc Descriptors (FC-MAD).

Each element of functional connectivity matrix FC_m , represents a pair-wise correlation of two voxels, in a local patch. Since, the correlation between any two nodes lies in the interval $[-1, 1]$ interval, BOLD time-series of two nodes can either be positively correlated or negatively correlated.

Mathematically speaking, the functionally nearest neighbor of $v(t_i, \bar{s}_j)$ is defined as,

$$\eta_1^{fc}[v(t_i, \bar{s}_j)] = \left\{ v(t_i, \bar{s}_k) : \max(\rho_{jk}), \forall v(t_i, \bar{s}_k) \in FC_m(j', \cdot) \right\}. \quad (4)$$

Then, the p -functional neighborhood of a voxel $v(t_i, \bar{s}_j)$ is generated from the $(p-1)$ -functional neighborhood iteratively, selecting the functionally nearest neighbor of that voxel from $\eta_{p-1}^{fc}[v(t_i, \bar{s}_j)]^c$, where c indicates the complement set of η_{p-1}^{fc} . p -functionally nearest neighbors of voxel $v(t_i, \bar{s}_j)$ are obtained by adding the voxels in $\eta_{p-1}^{fc}[v(t_i, \bar{s}_j)]$ to the functionally nearest neighbor of η_p^{fc} , as follows;

$$\eta_p^{fc}[v(t_i, \bar{s}_j)] = \left\{ v(t_i, \bar{s}_k) \cup \eta_{p-1}^{fc}[v(t_i, \bar{s}_j)] : \max(\rho_{jk}), v(t_i, \bar{s}_k) \in \eta_{p-1}^{fc}[v(t_i, \bar{s}_j)]^c \right\}. \quad (5)$$

Figure 3. Sample functional connectivity matrices constructed for local patch 104 (a) and 54 (b) used in experiments. Each row of represents the correlation between a seed node (row index) and all other nodes in the local patch. The most positively correlated neighbor for voxel 5 in cluster 104 is voxel 2 and indicated with a circle (a), the most negatively correlated neighbor for voxel 9 in cluster 54 is voxel 35 and indicated with a circle (b).

Algorithm 1: Computation of within-cluster functional connectivity matrices;

Input: Dataset: $\{v(t, \bar{s}_j)\}$,
 $i = 1, 2, 3, \dots, N$, $j = 1, 2, 3, \dots, M$
Number of Clusters: C
Lag: h

Begin

$$FC = \{ \}$$

$$[c_1, c_2, \dots, c_C] = \text{cluster\_voxels\_by\_location}([\bar{s}_1, \bar{s}_2, \dots, \bar{s}_M])$$

1. **for** $m = 1$ **to** C
2.   **for** $j = 1$ **to** π_m
3.   **for each** $k \in \{c_m - j\}$
4.   Compute $FC_m(j, k) = \rho_{jk}(h)$ using equation (3);
5.   **endfor** (j)
6. $FC = FC \cup FC_m$
7.   **endfor** (m)

End

Output: The Set of Functional Connectivity Matrices FC

For a voxel v_j at location $\bar{s}_j$, the set of p -functionally nearest neighbors η_p^{fc} , consists of the p of the most strongly correlated voxels in row j of the functional connectivity matrix $FC_m(j, \cdot)$, which is computed in Algorithm 1, where m is the index of the cluster which voxel $v(t_i, \bar{s}_j)$ belongs to and j' is the translated index of the voxel in FC_m .

Equation (5) employs only positively correlated samples (ρ_{jk} values close to +1). Another definition for functional neighborhood can be given by using the negatively correlated samples (ρ_{jk} values close to -1). In this case, $\max(\cdot)$ operation of Equation (5) is replaced by $\min(\cdot)$ operation. In Figure 3, functionally nearest neighbor selection is illustrated by using most positively correlated (obtained by $\max(\cdot)$ operation) and most negatively correlated voxels (obtained by $\min(\cdot)$ operation). Note that, the order of FC-MAD cannot exceed the minimum number of voxels in all clusters, $p \leq \pi_m \forall m = 1, 2, 3 \dots C$.

Experiments for the fMRI Data Collection

In the experiment, a participant is shown lists of words selected from a pre-defined semantic category, while being scanned using fMRI, see (Oztekin & Badre, 2011; Oztekin et al., 2009). After the presentation of each study list, the participant solves math problems and following this delay period, decides whether a probe word matches one of the members of the study list ("old" or "new"). Employing a delay period (about 14 sec during which the participant solved math problems) allows independent assessment of encoding related (i.e. study list period) brain activation from retrieval related (i.e. during the test probe) activity patterns. With this approach, one can test whether it is possible to identify and differentiate semantic categories of information that is represented in the brain at a given time based on distributed patterns of brain activity associated with and during cognitive processing. A total of ten semantic categories were used in the study, which are 'animals', 'colors', 'furniture', 'body parts', 'fruits', 'herbs', 'clothes', 'chemical elements', 'vegetables' and 'tools'. We used the neural activation patterns collected during

Algorithm 2: Extract functional connectivity-aware mesh arc descriptors (FC-MAD); *fc-mad*

Input: Dataset: $D = \{v(t_i, \bar{s}_j)\}, i = 1, 2, 3, \dots, N, j = 1, 2, 3, \dots, M$

Order of FC-MAD: p
Functional Connectivity Matrices: FC

Begin

1. **for** $j=1$ to M
2. Compute p -**functional**-neighborhood $\eta_p^{fc}[v(t, \bar{s}_j)]$ of $v(t, \bar{s}_j)$ by analyzing FC ;
3. **for** $i=1$ to N
4. Compute $\bar{a}_{i,j}$ by minimizing (2) if $\eta_p^{fc}[v(t, \bar{s}_j)] \neq \emptyset$;
5. **endfor** (i)
6. Construct A_j using $\bar{a}_{i,j}$;
7. **endfor** (j)
8. Construct F using A_j ;

End

Output: Feature matrix F

encoding and retrieval phases, to train and test the classifier to predict the semantic categories.

Implementation of the Functional Mesh Learning Algorithm

Our dataset consists of 240 training samples from encoding phase and 239 test samples from the retrieval phase with 24 samples in each of 10 semantic categories. Our region of interest consisted of 8142 voxels covering the lateral temporal cortex. Results for FC-LRF were generated using k-nearest neighbor (k-nn) and Support Vector Machine (SVM) methods. The k value of k-nn and kernel parameters of SVM classifier are selected using cross validation in training set.

The number of clusters C in the proposed algorithm is a user specified parameter. Since the number of voxels in all clusters π_m is always much higher than FC-LRF order p , regardless of the cluster sizes, similar functionally connected meshes are formed. Therefore, it has practically no effect on the performance of the algorithm. This fact is illustrated in the performance results in Table 3. Three different correlation variants are employed to capture functional similarity between nodes; i)

cross-correlation which is given in (3), ii) peak correlation which captures the relationships between activation peaks and iii) scan correlation, which measures the correlations of waveforms at a specific scan of interest. The overall performance of the algorithm is improved only by 2-4% percent by employing improved correlation measures as peak correlation and scan correlation, due to limitations of space, and only the performance results obtained by cross-correlation are shown in Table 1 and 2.

In addition, we employ two different functionally-nearest neighbor selection approaches namely, selecting positively and negatively correlated neighbors. The performance results are illustrated in Table 1. *Functional Connectivity Toolbox* implementation (Zhou, Thompson, & Siegle, 2009) is used for the computation of the correlation measures.

The results in Table 2 show that the employment of functional connectivity in the mesh learning algorithm (Ozay et al., 2012) improves classification performances, considerably. When we classify the raw features of 8142 voxels (without LRF), we observe 48% and 40% performances. Note that Mesh Learning increases the performances to 58% and 45% and Functional Mesh Learning further increases the

Table 1. Performance results of the functional mesh learning algorithm, performance measures are indicated recall as *R*, precision as *P* and F-score as *F* in the header row

Class Label	<i>k</i> -NN						SVM					
	Positively Correlated Neighbors			Negatively Correlated Neighbors			Positively Correlated Neighbors			Negatively Correlated Neighbors		
1	63	58	60	63	54	58	54	54	54	57	54	55
2	79	75	77	75	75	75	61	71	65	63	79	70
3	79	76	78	75	71	73	76	79	78	69	75	72
4	74	68	71	78	61	68	64	75	69	58	61	60
5	58	68	63	58	59	59	58	46	51	55	50	52
6	63	72	67	63	81	71	75	63	68	81	54	65
7	67	75	71	75	71	73	62	67	64	65	71	68
8	71	64	67	63	61	62	67	58	62	68	54	60
9	67	56	61	67	65	66	55	67	60	57	71	63
10	54	65	59	67	73	70	52	50	51	63	50	56
Avg	67	68	68	68	67	68	62	63	63	64	62	63

Table 2. Classification performance comparison of proposed methods

Method Employed	Classification Performance (%)		
		<i>k</i> -NN	SVM
Without LRF – Raw Features		48	40
Mesh Learning		56	62
Functional Mesh Learning using Positive Correlation		68	63
Functional Mesh Learning using Negative Correlation		67	62

Table 3. Classification performances for varying number of local patches (clusters)

Number of Local Patches	32	64	128	256	Standard Dev.
Recall	66,97	66,56	67,81	67,39	0,54
Precision	68,44	67,71	67,84	67,77	0,33

performances up to 68% and 63% using *k*-NN and SVM methods, respectively. The main issue which increases the performance is basically the selection of nearest neighbors by using functional connectivity of the voxels in brain.

CONCLUSION

In this study, we propose a new machine learning approach, called Functional Mesh Learning in order to classify cognitive process, based on distributed patterns of neural activation patterns

in the brain. In the current data set, the model was tested during memory process and performed successfully. The proposed method employs functional connectivity in order to define local meshes to represent the relationships between the voxels and their *p-functionally nearest neighbors*.

Our goal is to be able to model cognitive processes based on neural activation patterns in the brain. The present set of results indicate that the suggested Functional Mesh Learning model can be used to classify cognitive states and types of information represented during these cognitive operations based on distributed patterns of brain activity. In the current study, we only focused on modeling memory encoding and retrieval processes. Future research extending these findings to a wider range of cognitive operations would bring additional insight into the generality of the success of the proposed algorithm for modeling brain during cognitive processing and improving Functional Mesh Learning algorithm by eliminating drawbacks such as the linearity of the mesh model, selecting the optimal FC-MAD order value p , lack of modeling brain hierarchy and brain pathways.

REFERENCES

- Battle, A., Chechik, G., & Koller, D. (2006). *Temporal and cross-subject probabilistic models for fmri prediction tasks*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.116.745>
- Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Sciences*, 14(6), 277–290. doi:10.1016/j.tics.2010.04.004 PMID:20493761
- Bressler, S. L., & Tognoli, E. (2006). Operational principles of neurocognitive networks. *International journal of psychophysiology. Official Journal of the International Organization of Psychophysiology*, 60(2), 139–148. doi:10.1016/j.ijpsycho.2005.12.008
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews. Neuroscience*, 10(3), 186–198. doi:10.1038/nrn2575 PMID:19190637
- Firat, O., Özay, M., Önal, I., Karagöz, B., Öztekin, İ., & Yarman Vural, F. T. (2012). A mesh learning approach for brain data modeling. In *Proceedings of the IEEE 20th Conference on Signal Processing and Communications Applications (SIU)*.
- Friston, K. J., & Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *HUMAN BRAIN MAPPING*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.160.256>
- Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience*, 7(7), 523–534. doi:10.1038/nrn1931 PMID:16791142
- Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, 33(2), 131–139. doi:10.1016/j.compmedimag.2008.10.011 PMID:19111443
- Manor, L., & Perona, P. (2004). Self-tuning spectral clustering. *Advances in Neural Information Processing Systems*, 17, 1601–1608. Retrieved from <http://www.citeulike.org/user/dgrianna/article/3382909>
- McIntosh, A., Nyberg, L., Bookstein, F., & Tulving, E. (1997). Differential functional connectivity of prefrontal and medial temporal cortices during episodic. *Human Brain Mapping*. Retrieved from <http://www.mendeley.com/research/differential-functional-connectivity-prefrontal-medial-temporal-cortices-during-episodic/>
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430. doi:10.1016/j.tics.2006.07.005 PMID:16899397
- Ozay, M., Öztekin, İ., Öztekin, U., & Vural, F. T. Y. (2012). Mesh learning for classifying cognitive processes. *Arxiv: 1205.2382. Neural and Evolutionary Computing; Artificial Intelligence; Computer Vision and Pattern Recognition; Machine Learning*. Retrieved from <http://arxiv.org/abs/1205.2382>
- Özay, M., Öztekin, İ., Öztekin, U., & Yarman Vural, F. T. (2011). Neuroinformatics 2011: Modeling cognitive states using machine learning techniques. In *Proceedings of the annual meeting of Neuroinformatics Congress*, Boston, MA. Retrieved March 13, 2012, from <http://www.neuroinformatics2011.org/abstracts/modeling-cognitive-states-using-machine-learning-techniques>

- Oztekin, I., & Badre, D. (2011). Distributed patterns of brain activity that lead to forgetting. *Frontiers in human neuroscience*, 5, 86. doi:10.3389/fnhum.2011.00086
- Oztekin, I., Curtis, C. E., & McElree, B. (2009). The medial temporal lobe and the left inferior prefrontal cortex jointly support interference resolution in verbal working memory. *Journal of Cognitive Neuroscience*, 21(10), 1967–1979. doi:10.1162/jocn.2008.21146 PMID:18855551
- Patel, R. S., Bowman, F. D., & Rilling, J. K. (2006). A Bayesian approach to determining connectivity of the human brain. *Human Brain Mapping*, 27(3), 267–276. doi:10.1002/hbm.20182 PMID:16092131
- Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P., & Van De Ville, D. (2011). Decoding brain states from fMRI connectivity graphs. *NeuroImage*, 56(2), 616–626. doi:10.1016/j.neuroimage.2010.05.081 PMID:20541019
- Ryali, S., Chen, T., Supekar, K., & Menon, V. (2012). Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty. *NeuroImage*, 59(4), 3852–3861. doi:10.1016/j.neuroimage.2011.11.054 PMID:22155039
- Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*, 22(1), 158–165. doi:10.1093/cercor/bhr099 PMID:21616982
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., & Nichols, T. E. et al. (2011). Network modelling methods for FMRI. *NeuroImage*, 54(2), 875–891. doi:10.1016/j.neuroimage.2010.08.063 PMID:20817103
- Sporns, O. (2010). Networks of the brain. *MIT Press*. Retrieved from <http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=12288>
- Sporns, O. (2011). The human connectome: A complex network. *Annals of the New York Academy of Sciences*, 1224, 109–125. doi:10.1111/j.1749-6632.2010.05888.x PMID:21251014
- Vaidyanathan, P. P. (2007). The theory of linear prediction. *Synthesis Lectures on Signal Processing*, 2(1), 1–184. doi:10.2200/S00086ED-1V01Y200712SPR003
- Wang, X., Hutchinson, R., & Mitchell, T. M. (2003). Training fMRI classifiers to detect cognitive states across multiple human subjects. *IN NIPS03*, 16. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.15.7797>
- Zhou, D., Thompson, W. K., & Siegle, G. (2009). MATLAB toolbox for functional connectivity. *NeuroImage*, 47(4), 1590–1607. doi:10.1016/j.neuroimage.2009.05.089 PMID:19520177