Artificial Neural Networks for fMRI Data Analysis: A Survey

Ashwini S. Savanth[#1], Dr. P.A.Vijaya[*2]
[#] *Assistant Professor, Department of ECE, BNMIT, Bangalore, INDIA*
[*] *Professor & HoD, Department of ECE, BNMIT, Bangalore, INDIA*

***Abstract* — *fMRI is a valuable experimental and diagnostic tool for assessing the human body especially the brain. It has emerged as a successful tool in the investigation of cognitive functions. fMRI data was traditionally analysed with the univariate method, the popular one being Statistical Parametric Mapping based on the General Linear Model. But lately MVPA has been used to perform multivariate analysis of fMRI data. The multivariate approach originates from a field called as Machine Learning which is a branch of Artificial Intelligence. The Multivariate approaches have several advantages over the univariate approach, in that the Artificial Neural Networks (ANN) have outperformed some of the other classifiers such as Gaussian Naive Bayes, ICA and others. In this paper, an attempt is made to survey MVPA analysis of brain fMRI data using Artificial Neural Networks.*

Keywords — *Artificial Neural Networks, Feedforward, MVPA, Self Organizing Map, Univariate.*

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) is a valuable experimental and diagnostic tool for assessing the human body, especially the brain. It is a non-aggressive, non-radioactive, pain free and a non-invasive technique for studying brain activity and is quite popular among researchers and neuroscientists [1]. MRI is an extremely versatile imaging modality that can be used to study both the structure and function of the brain. Both structural and functional MRI images are acquired using the same scanner. Structural MRI helps in the study of brain structure whereas functional MRI (fMRI) is used in the study of cognitive and affective processes of the brain. fMRI has emerged as one of the most successful tool in the investigation of cognitive functions [2]. fMRI provides adequate spatial and temporal resolutions to measure the amplitude, location and timing of brain activity [3]. During the course of an fMRI experiment a series of brain images are acquired while the subject performs a set of tasks. The changes in the measured signal between individual images are used to make inferences regarding task related activations in the brain. The most common approach towards fMRI uses the Blood Oxygenation Level Dependent (BOLD) contrast. BOLD fMRI measures the ratio of oxygenated to deoxygenated haemoglobin in the blood. BOLD fMRI doesn't measure neuronal activity directly; instead it measures the metabolic demands of active neurons. fMRI data analysis is a massive data problem. Also, the relatively small changes in image intensity and existence of artifacts presents a challenge for accurately mapping task-related brain regions [4]. Several methods of fMRI data analysis have been reported in literature. The fMRI data analysis either uses the massive univariate approach or the multivariate approach.

II. UNIVARIATE VS MULTIVARIATE ANALYSIS

The univariate analysis is a single-voxel approach where each voxel is treated as a separate entity and statistical analysis is performed on that voxel. Statistical Parametric Mapping (SPM) based on the General Linear Model (GLM) performs voxel-by-voxel analysis which is massively univariate. It assumes a simple parametric linear model for signals with a specific noise structure and uses voxel-based linear regression analysis. It is widely used in the fMRI analysis mainly because of its simplicity of approach in principle and application. Usually in a SPM analysis linear convolution of an assumed hemodynamic response function (HRF) and the deterministic stimulus timing function is performed to construct reference functions. Certain factors like the modelling assumptions and the deterministic character assigned to the stimulus timing function could be too restrictive to capture the broad range of possible brain activation patterns in space and time and across subjects. Due to spatial coherence and temporal autocorrelation between brain voxels, a multivariate approach may be more suitable for fMRI analysis than the univariate approach [3]. The multivariate approach evaluates the covariance of the activated voxels across the regions of the brain. These results are easier to interpret as a signature of neural networks. On the other hand, univariate approaches are not capable of addressing the functional connectivity in the brain. The univariate approaches are forced to use more stringent and conservative, corrections for voxel-wise multiple comparisons. Hence the multivariate approach has a greater statistical power compared to univariate methods [5]. In MVPA the goal is to determine the model parameters that allow for the most accurate prediction of new observations. It seeks to create rules that can be used to categorize new observations.

In contrast, the GLM seeks to determine the model parameters that best fit the data at hand. GLM aims to see what to do with the data now, but MVPA sees what to do with the data now and in the future. MVPA has its own benefits. MVPA has increased sensitivity in detecting the presence of a particular mental representation in the brain. This makes the MVPA methods more feasible to measure the presence/ absence of cognitive states based on only a few seconds' worth of brain activity. Along with this the MVPA methods can be used to characterize how these cognitive states are represented in the brain [6]. It also focuses on the analysis and comparison of distributed patterns of activity hence helps to detect difference between conditions with higher sensitivity than the univariate analysis [7].

III. MULTI-VOXEL PATTERN ANALYSIS (MVPA)

The application of machine learning methods to fMRI data is referred to as Multi Voxel Pattern Analysis (MVPA). MVPA tools are referred to as classifiers or learning machines. MVPA focuses on multiple voxels instead of single voxels, and uses pattern classification algorithms on multiple voxels in order to decode the patterns of activity. In this approach, data from individual voxels within a region of interest are jointly analysed [7].

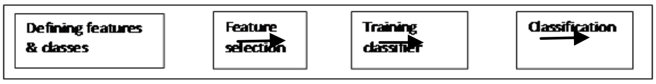

Fig 1: Steps involved in Multi Voxel Pattern Analysis

The steps involved in performing MVPA (Fig 1) includes the following: defining features and classes, feature selection, choosing a classifier, training and testing the classifier and examining results. For defining what information to be used as *features* there are many possible choices like the raw fMRI data over space and time, averaged fMRI data over a block, Beta values from a GLM analysis or average of several voxels in an ROI. The choice of the *class labels* depends upon the research question. In fMRI the number of features is usually larger than the number of observations. Hence it is beneficial to reduce the number of features through *feature selection*. This could involve using only voxels from a particular ROI or by dimensionality reduction technique such as SVD (Singular Value Decomposition) or PCA (Principle Component Analysis) or meta-analysis data. In literature, there are methods like Principal Feature Analysis [8], Gray Level Co-occurrence Matrix (GLCM) [1] which have been used for feature selection. The next step is selection of a suitable *classifier*. Classifiers can be either linear or non-linear. Most MVPA studies have used linear classifiers, including Correlation-based classifiers, Neural Networks without a hidden layer, Linear Discriminant Analysis, Linear Support Vector Machines (SVMs), and Gaussian Naive Bayes classifiers. Other MVPA analyses have used nonlinear classifiers; examples include Nonlinear Support Vector Machines and Neural Networks with hidden layers. Artificial Neural Networks have been used and compared with other methods by researchers. ANNs generally have outperformed some of these methods such as Gaussian Naive Bayes [9], ICA [3,16]. Hence Artificial Neural Networks is a good option to be used for fMRI data analysis.

IV. ARTIFICIAL NEURAL NETWORKS

Theoretical and computational neuroscience is the field concerned with the theoretical analysis and the computational modelling of biological neural systems. Since neural systems are intimately related to cognitive processes and behaviour, the field is closely related to cognitive and behavioural modelling. Artificial neural networks have been used by researchers to analyse brain fMRI data to a large extent. In that the Feedforward neural network and self organizing map are quite common and an attempt is made here to survey these methods proposed by several researchers [10].

A. Feedforward Neural Network

The feedforward neural network was the first and simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network. The simplest kind of neural network is a single-layer perceptron network, which consists of a single layer of output nodes; the inputs are fed directly to the outputs via a series of weights. Perceptrons can be trained by a simple learning algorithm that is usually called the delta rule. It calculates the errors between calculated output and sample output data, and uses this to create an adjustment to the weights, thus implementing a form of gradient descent. Multi-Layer Perceptron is a class of networks consisting of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications, the units of these networks apply a sigmoid function as an activation function. Multi-layer networks use a variety of learning techniques, the most popular being back-propagation [11].

A standard SLFN with N hidden units and C output units can be mathematically modelled as:

$$o = \sum_{i=1}^{N} \alpha_i f(w_i . x + b_i), x \in R^d \qquad ....(1)$$

where $f(\cdot)$ is the activation function of hidden units, o is the output vector, α_i is the weight vector connecting from the i-th hidden unit to the output units, b_i is the threshold of the i-th hidden unit, and $w_i=[w_{i1}, w_{i2}, ..., w_{id}]$ is the input weight vector connecting from the input units to the *i*-th hidden unit. $w_i \cdot x = \langle w_i, x \rangle$ is the inner product of w_i and x. The main goal of training process is to estimate the network weights w_i, α_i, and b_i so that they minimize the error function defined by:

$$E = \sum_{j=1}^{n}\left(o_j - t_j\right)^2 \qquad \text{..............(2)}$$

$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{N} \alpha_i\, f(w_i\,.\,x_j + b_i) - t_j\right)^2$$

Traditionally, the estimation of the network weights is based on the gradient-descent algorithms and a popular training algorithm based on this is the backpropagation (BP) in which the networks are turned based on gradient descent with error propagation from the output layer to the input layer. This algorithm has some problems such as local minima, overtraining, learning oscillation, etc. Huang et.al. proposed an efficient learning algorithm for SLFNs called as the extreme learning machine (ELM) in which the minimization process of error function was based on the linear model:

HA=**T**, ...(3)

where **H** is called as hidden layer output matrix of SLFN and defined by:

$$H = \begin{bmatrix} f(w_1\,.\,x_1 + b_1) & \cdots & f(w_N\,.\,x_1 + b_N) \\ \vdots & \ddots & \vdots \\ f(w_1\,.\,x_n + b_1) & \cdots & f(w_N\,.\,x_n + b_N) \end{bmatrix}$$

$$T = [\,t_1 \quad \dots \quad t_n\,]^T$$

$$A = [\,a_1 \quad \dots \quad a_C\,]$$

The SLFN along with RLS-ELM was used to decode the subject's cognitive states. The cognitive function considered in this study was a 'Picture versus Sentence' study. The classification performance of SLFNs trained by RLS-ELM was shown to be better than that of Gaussian Naive Bayes [9].

Jose Paulo Santos et. al. justified the use of Artificial Neural Networks to analyze fMRI data. The ANNs due to their complexity and computing load had been used only to limited parts of the brain. However Jose Paulo Santos et. al. proposed to use ANN to analyze whole brain fMRI data. Since the fMRI dataset is usually voluminous, dimensionality reduction was performed using probabilistic independent component analysis (PICA). The independent components then entered a simple backpropagation feedforward neural network, which, after training, was used to predict brands' assessments of a different set of subjects. A hidden layer with four nodes had produced best results. The selected activation function for the hidden nodes was "tansig", while for output neurons the function was "sigmoid". The conclusion was that ANNs can model complex cognitive processes, which could predict choices above chance level. Also the hidden nodes organize into separate and sounding cognitive processes. This opens the possibility to define cognition, not based on explicit task outcomes, but relying on implicit neural substrates [12].

J.A.Gutiérrez-Celaya et.al. built a supervised feedforward neural network-based classifier up on a classic Multi-Layer Perceptron (MLP) structure with three layers: input, hidden and output ones in order to explore the feasibility of automating the evaluation of stroke chronic patients' motor functions. The statistical pattern of brain activation corresponding to motor functionality captured by fMRI images was detected by artificial neural network based classifiers. The backpropagation algorithm was used for training the neural network and the gradient descent optimization technique to minimize the mean-squared error [13].

During a visual rivalry paradigm, Nicola Bertolino et.al. provided a method based on ANN to identify the different neural pattern of activity related to the processing of two classes of visual stimuli (houses and faces), applicable in the absence of behavioural indicators, indicating which stimulus is perceived by participant [14]. fMRI was studied as the subjects viewed binocular non-rivalry (BNR) and binocular rivalry(BR) tasks. First the BNR was used to identify brain areas involved in face and house decoding, then the ANN was trained on this data, and finally the trained ANN was employed to discriminate the pattern of activity in BR task analysis and verified the consistency of these results with the behavioural response. A one-layer Feed-Forward Neural Network with a Log-Sigmoid Transfer Function with hidden layer size of 65 neurons was employed. The Mean Square Error (MSE) relative to the difference between the target outputs (presented stimuli) and the values predicted by the model (network outputs) was used as a performance function. The main results showed that the trained ANN was able to generalize across the BNR and BR fMRI paradigms and identify with high accuracy the cognitive state of the participant during the BR condition. Fig 2 shows the bar plot of ANN percentage of successes for each subject.

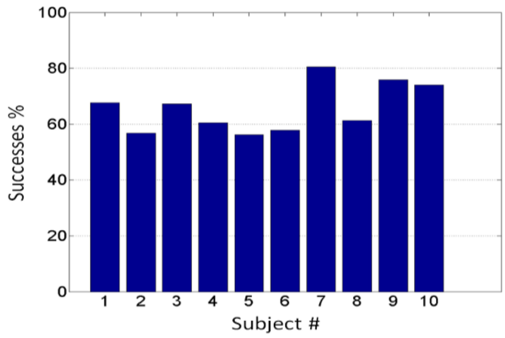

Fig 2: Bar plot of ANN percentage of successes for each subject

The single-subject analysis performed on the BNR data revealed that the participants showed activity for the contrast faces>houses in the posterior fusiform gyrus (i.e., FFA) and in the inferior occipital gyrus (i.e., OFA) (Fig. 3A), while for the contrast houses>faces in the parahippocampal gyrus (Fig. 3B).

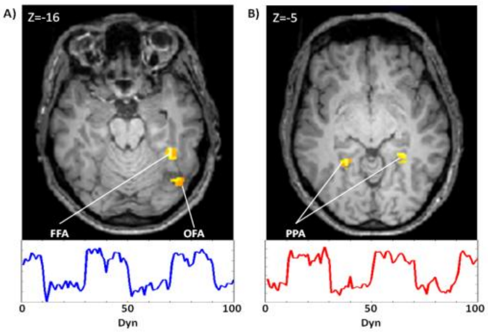

Fig 3(A,B): Example of resulting BOLD activity from GLM single subject analysis of BNR-localizer task.

B. Self-Organizing Map (SOM)

A Self-Organizing Map (SOM) is a kind of artificial neural network (ANN). It is trained using unsupervised learning. It is a method to do dimensionality reduction as it produces a low-dimensional (usually two-dimensional), discretized representation of the input space of the training samples, called a map. SOMs differ from other ANNs because they apply competitive learning as opposed to error-correction learning (such as backpropagation with gradient descent), and they use a neighbourhood function to preserve the topological properties of the input space. This makes SOMs useful for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling. Like most artificial neural networks, SOMs operate in two modes: training and mapping. "Training" builds the map using input examples (a competitive process, also called vector quantization), while "mapping" automatically classifies a new input vector. SOM has been used by many researchers in the analysis of fMRI data [15].

The SOM algorithm consists of two major steps: 1) determining the winner node; and 2) updating the weight vectors associated with the winner node and some of its neighbouring nodes. Prior to training, the weight vectors associated with each node of the map are suitably initialized. For a profitable initialization, the vectors are sampled evenly from the subspace spanned by the two largest principal component eigenvectors. The training expands over several iterations and is based on competitive learning. In each iteration, a vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T \in \mathrm{R}^n$ (where n is the length of fMRI data) from the input space is compared with the weight vectors of the nodes $\boldsymbol{m}_i = [m_{i1}, m_{i2}, \ldots, m_{in}]^T \in \mathrm{R}^n$ (where $i = 1, 2, \ldots, N$; N being the total number of nodes) to determine the winner node, often referred to as the best matching unit(BMU). The BMU refers to the node whose weight vector is the closest match of the input vector based upon a similarity metric. The most commonly used metric is the Euclidean metric:

$$\left| \boldsymbol{x} - \boldsymbol{m}c \right| = \min\{ \left| \boldsymbol{x} - \boldsymbol{m}_i \right| \} \quad i = 1, \qquad \ldots\ldots\ldots\ldots(4)$$

where $| \, |$ represents the Euclidean norm, $\boldsymbol{x}$ is the vector under consideration, $\boldsymbol{m}_i$ denotes the weight of the ith node on the map, and $\boldsymbol{m}_c$ represents the weight of the BMU. Once the BMU is determined, the weight vectors associated with the BMU and some of its neighbours in the map are updated to make them more similar to the input vector

$$\boldsymbol{m}_i(t+1) = \boldsymbol{m}_i(t) + h_{ci}(t)\,[\boldsymbol{x}(t) - \boldsymbol{m}_i(t)] \qquad \ldots\ldots\ldots(5)$$

where t is the current iteration number; $h_{ci}(t)$ is defined as the neighbourhood kernel that controls the number of neighbouring nodes to be updated and the rate of update in each iteration. The magnitude of this update decreases with time (iteration) and for nodes farther away from the BMU with a suitable kernel. In general, the neighbourhood kernel takes the form of a Gaussian function:

$$h_{ci}(t) = \alpha(t) \exp(- \left| r_i - r_c \right|^2 / 2\sigma^2(t)) \qquad \ldots\ldots\ldots(6)$$

where r_i and r_c are spatial coordinates of the ith node and the winner node, respectively, in the output space; σ is the fullwidth at half-maximum (FWHM) of the Gaussian kernel that determines the neighbouring nodes to be updated. α denotes the learning rate that controls how fast the weights get updated. Both σ and α decrease monotonically with the increase in the learning iteration t.

Meyer-Baese et.al. experimentally compared two exploratory data analysis methods for fMRI: the ICA techniques versus unsupervised clustering. One of the algorithms used in unsupervised clustering is SOM. From ROC analysis, it was observed that the clustering methods outperform the transformation-based methods and SOM was outperformed by topographical ICA [16].

Wellington P. dos Santos et.al. present a new approach for the detection of activated brain regions: the composition and analysis of synthetic multi and monospectral images using statistical methods and proposing non-parametrical methods based on Kohonen self-organized networks. SOM has an advantage of reducing a multispectral problem to a monospectral approach, eliminating the computational cost associated to the computing of the accumulated probabilistic distribution functions of the hypothesis tests [17].

Wei Liao et.al. proposed a method that integrates improved SOM and HC (Hierarchical clustering) in detecting and classifying brain activation. The validity of the algorithm was tested by a simulation study and real fMRI data, both of which included block-design and event-related experiments. The results show that the new integrated algorithm can identify activities that arise from different signal sources, other noise sources such as head motion, and different response patterns that arise from simultaneous stimulus tasks [18].

Fournel et. al. demonstrated that SOM algorithm is a good candidate for multi-voxel pattern analysis methods as it leads to good performance and allows to extract information about cognitive processes. Because SOM is designed for working with whole brain functional data, and due to the unsupervised nature of the algorithm, neuroimaging data can be analyzed without any prior assumptions. The results showed that, in the single-subject condition, the average classification performance on all conditions was 85.4% as shown in fig 4 and the average recognition performance on all condition was 92.72%. In the inter-subject classification, the average classification rate was 80.237% and the average prediction rate for the three experimental conditions was 83.33%. However, the study did not focus on projecting back the weights of an artificial neuron (i.e. a prototype of a cognitive state) into cerebral space [19].

Santosh B.Katwal et.al. proposed to use a graph-based visualization technique for SOM. The visualization scheme incorporated two metrics of SOM node connectivity

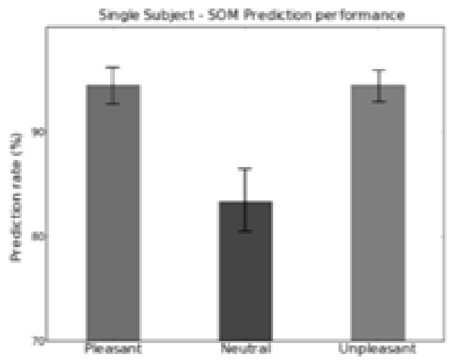

Fig 4: Average performance prediction of SOM algorithm for single subject classification

based on 1) local density distribution across SOM prototypes; and 2) local similarities (correlations) between the prototypes. The combined visualization effectively captured cluster boundaries and delineated detailed connectivity structures of the meaningful data [3]. The *Density-Based Connectivity Visualization, CONNDDvis was* realized by rendering of the connectivity matrix, CONNDD, over the SOM lattice. The existence of an edge between two prototypes $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$ on the graph indicates that they are neighbours in the input data space and the weight of the connection between them gives its connectivity strength:

$$\mathrm{CONNDD}(i,\ j) = |\mathrm{RF}_{ij}| + |\mathrm{RF}_{ji}| \quad i,\ j = 1,\ 2, \ldots\ldots(7)$$

where $|\mathrm{RF}_{ij}|$ denotes the number of input vectors in the receptive field of prototype $\boldsymbol{m}_i$ for which $\boldsymbol{m}_j$ is the second BMU($\boldsymbol{m}_i$ being the first BMU). The correlation coefficient matrix, CONNCC, which includes temporal similarities (correlation coefficients) of neighbouring prototypes, can be visualized graphically to display local similarities in the prototypes. The weight on the edges between two prototypes gives the measure of their similarity. The visualization obtained from CONNDDvis and CONNCCvis was merged to obtain a combined connectivity visualization that emphasizes delineation of connectivity structures of prototypes representing task-related signals.

$$\mathrm{CONNDDCC}(i, j)=\mathrm{CONNDD}(i, j) \times \mathrm{CONNCC}(i,j) \ldots\ldots\ldots..(8)$$

CONNDDCC denotes overall connectivity strength between m_i and m_j and includes both density-based connectivity and correlation-based connectivity between prototypes.

The performance of SOM applied in conjunction with the graph based visualization was compared with ICA and GLM. It was found that SOM outperformed ICA and GLM by providing highest sensitivity in classifying regions based on the timing of their responses. Fig 5 shows the voxels identified by SOM, ICA and GLM. There are variations to SOM which have been proposed like the Growing Self Organizing Map(GSOM) and Conscience Self-Organizing Map (CSOM).

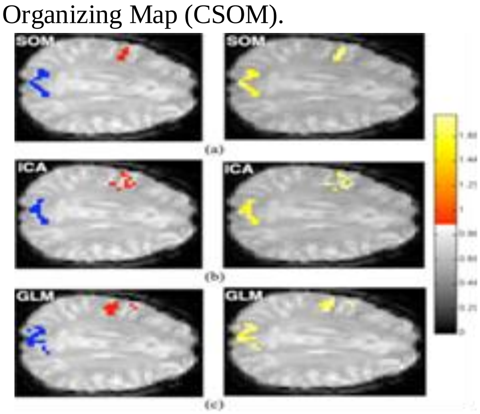

Fig 5: Voxels identified by SOM, ICA, GLM

Huang et.al. [21] demonstrated the potential of GSOM as a tool used for fMRI data analysis. When using SOM, users have to predefine the size of the map, i.e. the structure of SOM map is fixed whose capability of discovering data becomes limited. As an improvement version of the traditional SOM, Growing Self Organizing Map (GSOM) enables its map to grow dynamically based on the input data. A significant feature of GSOM is the spread factor (SF) parameter, which can be used to control the growth of the GSOM map. The value of spread factor (SF) is between 0 and 1. For lower SF, the lower level of spread of the map is displayed and vice versa, which means users can analyse data in hierarchical levels.

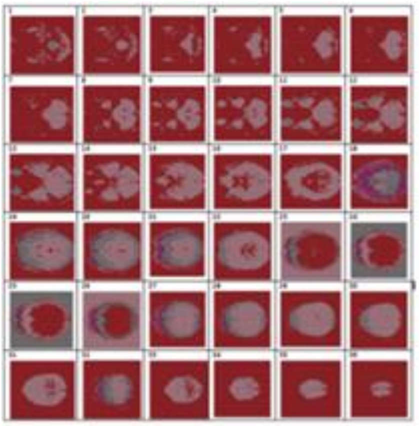

Fig.6 GSOM clustering images when subject was at rest

Huang et.al. used GSOM as a visualization tool to cluster fMRI finger tapping and non-tapping data. The finger tapping experiment during the fMRI scan is commonly conducted for clinical purpose, for instance, it enables researchers to profile Parkinson Disease characteristics because of finger tapping deficits already discovered in Parkinson patients. The GSOM starts with a minimal number of nodes (usually 4) and grows new nodes on the boundary based on a heuristic. By using the SF, the data analyst has the ability to control the growth of the GSOM. All the starting nodes of the GSOM are boundary nodes, i.e. each node has the freedom to grow in its own direction at the beginning. New Nodes are grown from the boundary nodes. Once a node is selected for growing all its free neighbouring positions will be grown new nodes. Due to the flexible structure and dynamic node adding capacity, the GSOM has shown to provide better visualization as well as faster processing speed compared to the SOM. A further key application has been the use of the SF parameter to develop GSOMs at different levels of spread, thus enabling the generation of hierarchies of clusters. The GSOM based analysis performed by Huang et.al. and the results are displayed as under. Fig.6 shows 36 GSOM clustering images corresponding to 36 brain horizontal slices when the subject was at rest. Fig.7 illustrates 36 GSOM clustering partitions images when the subject was tapping fingers.

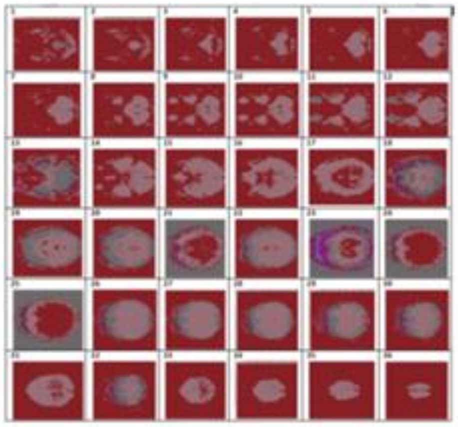

Fig.7 GSOM clustering images when subject was tapping fingers

Image 36 in Fig.6 is taken as an example, where in order to observe the detailed structure of the middle red area, hierarchical levels of spread factors are used as follows in Fig.8. Fig.8 indicates that GSOM is able to represent hierarchical levels visualization of areas of interest. Therefore, detailed insight of brain images were obtained by using higher level of spread factors.

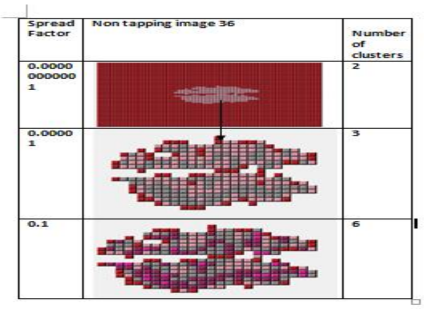

Fig.8 Sub region details by using higher spread factors

The scaled Euclidean similarity between images 35 in Fig.6 and Fig.7 is 1, which indicates that there are differences between these two images. However, with spread factor 0.00000000001, it was not possible to tell the difference by visualizing image 35 in Fig.5 and its counterpart in Fig.7. Therefore, a higher spread factor of 0.000001 was applied to the middle area of interest within these images. As shown in Fig.9, GSOM is able to distinguish tapping or non-tapping horizontal brain slice 35.

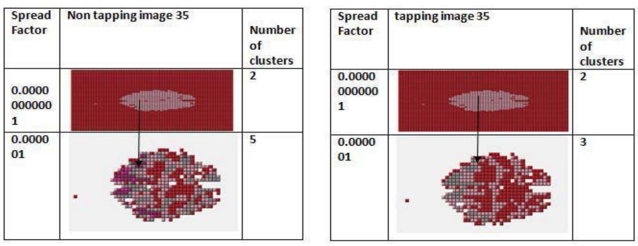

Fig.9 Comparison between images by using higher spread factors

Patrick O'Driscoll et.al. have used Conscience Self-Organizing Map (CSOM) as a clustering method for delineating functional sub-networks that are active when an individual performs an approach-avoidance paradigm [22]. The CSOM is a subsequently developed version of SOM because of its ability to enforce equi-probabilistic (maximum entropy) mapping of the data points $x \in M$ to the VQ prototypes. This facilitates optimal information transfer about the data distribution (with the given N prototypes). The CSOM learning algorithm is given here and also where it differs from the Kohonen SOM is indicated. After initialization of the prototypes w_i, learning consists of many cycles (indexed by t) through the following steps. *Competition*: For a random $x \in M$ find the closest (winner) prototype w_c:

$$c(x) = \mathrm{argmin}_i(\|x - w_i\| - bias_i), \qquad i = 1, \cdots, N \quad \text{..........(9)}$$

where the scalar quantity $bias_i$ is computed from the winning frequency F_i of w_i as $bias_i = \gamma(1/N - F_i)$, and the winning frequencies of all prototypes are updated after winner selection. (γ is a user-controlled parameter.) The $bias_i$ is the conscience, inducing infrequent winners to win more, frequent winners to win less data points. $\gamma = 0$ reduces eq. (9) to the winner selection of the Kohonen SOM. *Weight adaptation*: the winner w_c and its neighbours in the SOM lattice are moved closer to x.

$$w_i(t + 1) = w_i(t) + \alpha(t)\ h_{c,i}(t)\ (x - w_i) \quad \text{..............(10)}$$

The SOM lattice region influenced by the update is defined by the radially decreasing neighbourhood function $h_{c,i}(t)$ centered over the winner. For the Kohonen SOM, it is often a Gaussian, and initially must cover most of the SOM lattice. Both $h_{c,i}(t)$ and the learning rate $\alpha(t)$ must decrease with time t in order to achieve topologically correct ordering of the prototypes in the SOM grid. The CSOM has another advantage: it only needs to update the immediate neighbours in eq. (10) because cooperation across the lattice is ensured by the conscience mechanism. This leads to substantial savings in computation. Patrick O'Driscoll et.al. applied CSOM clustering using data from a single subject who was shown only three unpleasant faces. The results showed a representative image with several SOM clusters that cover known functional areas.

V. Conclusions

ANNs are suitable for BOLD fMRI signal analysis and perform better than the traditional GLM analysis, Gaussian Naive Bayes or the ICA. Initially due to the high complexity and computational load, the ANNs were applied only to a few parts of the brain or only to few ROIs. But this has been extended to the whole brain region along with dimensionality reduction methods and thereby giving good results. ANNs have been demonstrated to be suitable for modelling complex cognitive processes, detecting brain activity and functional connectivity. The SOM algorithm also proves to be a good candidate for MVPA analysis as it leads to good performance and allows for a coherent prototype projection in the standard space. The SOM algorithm also has a potential to be extended to determine the temporal sequence of brain processes and possibly reveal the dynamics of inter-regional influences in the brain. GSOM and CSOM are the different versions of SOM which are being investigated to be used as a potential tool for fMRI data analysis.

References

[1] Sahar Jafarpour, Zahra Sedghi, Mehdi Chehel Amirani, 'A Robust Brain MRI Classification with GLCM Features', International Journal of Computer Applications (0975 – 8887) Volume 37– No.12, January 2012

[2] Shantipriya Parida, Satchidananda Dehuri, 'Review of fMRI Data Analysis: A Special Focus on Classification', International Journal of E-Health and Medical Communications, 5(2), 1-26, April- June 2014

[3] Santosh B. Katwal, John C. Gore, Ren´e Marois, and Baxter P. Rogers,'Unsupervised Spatiotemporal Analysis of FMRI Data Using Graph-Based Visualizations of Self-Organizing Maps ', IEEE Transactions On Biomedical Engineering, VOL. 60, NO. 9, September 2013

[4] C.L. Chen, T.H. Wu, Y.T. Wu, Y.H. Huang, and J.S. Lee, 'Characterization of the univariate and multivariate techniques on the Analysis of Simulated and fMRI Datasets with Visual Task', Nuclear Science Symposium Conference Record, 2003 IEEE,

[5] Christian Habeck,Yaakov Stern,'Multivariate Data Analysis for Neuroimaging Data: Overview and application to Alzheimer's Disease', Cell Biochem Biophys. 2010 November ; 58(2): 53–67.

[6] Kenneth A. Norman, Sean M. Polyn, Greg J. Detre and James V. Haxby, 'Beyond mind-reading: multi-voxel pattern analysis of fMRI data', TRENDS in Cognitive Sciences (2006), doi:10.1016/j.tics.2006.07.005

[7] Martin A. Lindquist, 'The Statistical Analysis of fMRI Data', Statistical Science 2008, Vol. 23, No. 4, 439–464

[8] Lijun Wang, Yu Lei, Ying Zeng, Li Tong, and Bin Yan, "Principal Feature Analysis: A Multivariate Feature Selection Method for fMRI Data," Computational and Mathematical Methods in Medicine, vol. 2013, Article ID 645921, 7 pages, 2013. doi:10.1155/2013/645921

[9] Hieu Trung Huynh and Yonggwan Won, 'Decoding Cognitive States from fMRI Data Using Single Hidden-Layer Feedforward Neural Networks', Fourth International Conference on Networked Computing and Advanced Information Management 2008 IEEE

[10] https://en.wikipedia.org/wiki/Artificial_neural_network

[11] https://en.wikipedia.org/wiki/Feedforward_neural_network

[12] José Paulo Santos, Luiz Moutinho, 'Tackling the cognitive processes that underlie brands' assessments using artificial neural networks and whole brain fMRI acquisitions', IEEE International Workshop on Pattern Recognition in NeuroImaging 2011 IEEE Computer Society

[13] J. A. Gutiérrez-Celaya, R. Leder, R. Carrillo, A. Hawayek, J. Hernández and E. Sucar, ' fMRI-based Inverse Analysis of Stroke Patients' Motor Functions', March 28 - April 1, 2011, Rio De Janeiro, Brazil, 2011 IEEE

[14] Nicola Bertolino, Stefania Ferraro, Anna Nigri, Maria Grazia Bruzzone, Francesco Ghielmetti, 'A Neural Network Approach to fMRI Binocular Visual Rivalry Task Analysis', PLoS ONE 9(8):e105206. doi:10.1371/journal.pone.0105206
[15] https://en.wikipedia.org/wiki/Self-organizing_map
[16] A. Meyer-Baese, Axel Wismueller, and Oliver Lange, 'Comparison of Two Exploratory Data Analysis Methods for fMRI: Unsupervised Clustering Versus Independent Component Analysis', IEEE Transactions On Information Technology In Biomedicine, Vol. 8, No. 3, September 2004
[17] Wellington P. dos Santos, Ricardo E. de Souza, Ascendino F. D. e Silva, Ana Cristina T. Albuquerque, 'A Monospectral Approach for fMRI Analysis using Kohonen Self-Organized Networks', Seventh International Conference on Intelligent Systems Design and Applications, 2007 IEEE Computer Society
[18] Wei Liao, Huafu Chen, Qin Yang, and Xu Lei, 'Analysis of fMRI Data Using Improved Self- Organizing Mapping and Spatio-Temporal Metric Hierarchical Clustering', IEEE Transactions On Medical Imaging, Vol. 27, No. 10, October 2008
[19] Arnaud Fournel, Emanuelle Reynaud, 'Exploring Whole Brain fMRI Data with Unsupervised Artificial Neural Networks', IEEE International Workshop on Pattern Recognition in NeuroImaging 2011 IEEE Computer Society
[20] Arnaud Fournel, Emanuelle Reynaud, 'Unsupervised classification of whole-brain fMRI data with artificial neural networks', Neurosciences Computationnelles, "Neurocomp'10", Aug 2010, Lyon, France
[21] Pin Huang, Pubudu Pathirana, Damminda Alahakoon, Peter Brotchie, 'Application of Growing Self-Organizing Map to Distinguish between Finger Tapping and Non Tapping from Brain Images', Information and Automation for Sustainability (ICIAfS), 2012 IEEE 6th International Conference on 27-29 Sept. 2012
[22] Patrick O'Driscoll, Erzs´ebet Mer´enyi, Christof Karmonik, Robert Grossman, 'SOM and MCODE Methods of Defining Functional Clusters in MRI of the Brain', Conf Proc IEEE Eng Med Biol Soc. 2014;2014:734-7. doi: 10.1109/EMBC.2014.6943695