

ANKARA YILDIRIM BEYAZIT UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES



**CLUSTERING FUNCTIONAL MRI DATA USING A ROBUST
UNSUPERVISED LEARNING ALGORITHM**

Ph.D. Thesis by

Hadeel Kassim ALJOBOURI

Electrical and Electronics Engineering Department

May, 2017

ANKARA

CLUSTERING FUNCTIONAL MRI DATA USING A ROBUST UNSUPERVISED LEARNING ALGORITHM

A Thesis Submitted to

The Graduate School of Natural and Applied Sciences of

Ankara Yıldırım Beyazıt University

**In Partial Fulfillment of the Requirements for the Degree of Doctor of
Philosophy in Electrical and Electronics Engineering Department**

by

Hadeel Kassim ALJOBOURI

May, 2017

ANKARA

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled “**CLUSTERING FUNCTIONAL MRI DATA USING A ROBUST UNSUPERVISED LEARNING ALGORITHM**” completed by **Hadeel Kassim ALJOBOURI** under the supervision of **Assoc. Prof. Dr. İlyas ÇANKAYA** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of **Doctor of Philosophy in Electrical and Electronics Engineering**.

Assoc. Prof. Dr. İlyas ÇANKAYA

Supervisor

Prof. Dr. Fatih V. ÇELEBİ

Jury member

Assist. Prof. Dr. Ömer KARAL

Jury Member

Prof. Dr. Recep DEMİRCİ

Jury member

Assoc. Prof. Dr. Orhan Murat KOÇAK

Jury Member

Prof. Dr. Fatih V. ÇELEBİ

Director

Graduate School of Natural and Applied Sciences

ETHICAL DECLARATION

I hereby declare that, in this thesis which has been prepared in accordance with the Thesis Writing Manual of Graduate School of Natural and Applied Sciences,

- All data, information and documents are obtained in the framework of academic and ethical rules,
- All information, documents and assessments are presented in accordance with scientific ethics and morals,
- All the materials that have been utilized are fully cited and referenced,
- No change has been made on the utilized materials,
- All the works presented are original,

and in any contrary case of above statements I accept to renounce all my legal rights.

Date: 19.06.2017

Hadeel ALJOBOURI

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor Dr. İlyas ÇANKAYA for his inspiring encouragement and prompt guidance. Without his help I would not have had the opportunity to focus on this subject. I would also like to thank Dr. Orhan Murat KOÇAK for his advice.

Additionally, I want to express my thanks to all those who provided valuable assistance during the past five years in the Graduate School of Natural and Applied Sciences/ Ankara Yildirim Beyazit University.

My deepest love and thanks go to my dear mother and the soul of my dearest father. Lastly but certainly not least, I would like to thank my sisters, brother and my two kids. It is their love that leads me through many difficulties.

I would especially like to thank my husband Hussain A. JABER for all the patience, strength and unconditional support over these last years. Without his encouragement and support I would not have been able to concentrate and complete this thesis.

2017, 25 May

Hadeel Kassim ALJOBURI

CLUSTERING FUNCTIONAL MRI DATA USING A ROBUST UNSUPERVISED LEARNING ALGORITHM

ABSTRACT

Functional Magnetic Resonance Imaging (fMRI) has provided neuroscientists with a powerful tool to examine brain activity by calculating the levels of oxygen in the blood and generates a sequence of 3-D images. Clustering approach is a model-free analysis; it has the ability of defining the active zones in the brain without the need of prior knowledge about activation patterns or experiment as the classical and statistical General Linear Model (GLM) method.

The goal of this proposal is to find a solution for choosing an appropriate clustering approach to obtain the best performance for whole brain functional connectivity by means of data analysis. In this work, a novel and robust unsupervised learning approach is proposed; it relies on using a Robust Growing Neural Gas (RGNG) algorithm into a real auditory fMRI dataset. The main contribution of this work is running the RGNG algorithm into fMRI dataset with a comparison to NG and GNG algorithms, which is not used for the purpose yet or any other applications also. Another comparison has been done with the model-based data analysis approach using a Statistical Parametric Mapping (SPM) package which is based on GLM. The output results demonstrate that the presented RGNG approach is clearly superior to other approaches as revealed by their performance measured by Minimum Description Length (MDL) and Receiver Operating Characteristic (ROC) analysis.

A MATLAB-based graphical user interface (GUI) tool is designed and implemented as a software package for fMRI data analysis. A new model for neuroscience data analysis is developed in this work which is easily accessible by researchers. The proposed work could help the neurologist and psychologist for better interpretation of fMRI dataset.

Keywords: Clustering techniques, Data mining, Graphical User Interface (GUI), Growing Neural Gas (GNG), Neural Gas (NG), Robust Growing Neural Gas (RGNG)

FONKSİYONEL MRI VERİLERİNİN GÜRBÜZ DENETİMSİZ ALGORİTMASI İLE KÜMELENDİRİLMESİ

ÖZ

Fonksiyonel Manyetik Rezonans Görüntüleme (fMRI), kandaki oksijen seviyelerini hesaplayarak ve bir dizi 3 – boyutlu görüntü oluşturarak beyin aktivitesini incelemede sinir bilimciler için güçlü bir araç olmuştur. Kümeleme yaklaşımı model içermeyen bir analizdir; bu yaklaşım, klasik ve istatistiksel Genel Doğrusal Model (GLM) olarak, aktivasyon desenleri veya deney hakkında önceden bilgi sahibi olmaksızın beyindeki aktif bölgeleri tanımlama özelliğine sahiptir.

Bu araştırmanın amacı, tüm beyinin işlevsel bağlantı veri analizinde en iyi performansı elde etmeye yönelik uygun bir kümeleme yöntemi seçmek için çözüm bulmaktır. Bu çalışmada, gerçek bir işitsel fMRI veri setine ve Güçlü Artan Sinirsel Gaz (RGNG) algoritması kullanımına dayanan yeni ve güçlü bir denetimsiz öğrenme yöntemi önerilmiştir. Bu çalışmanın başlıca katkısı, NG ve GNG algoritmalarıyla mukayese ederek, daha önce bu amaç için veya başka herhangi bir uygulamada kullanılmayan RGNG algoritmasının fMRI veri setleri ile iç içe kullanılmasıdır. GLM'ne dayalı İstatistiksel Parametrik Eşleme (SPM) paketi kullanarak model tabanlı veri analizi yöntemi ile diğer bir karşılaştırma daha yapılmıştır. Çıktı sonuçları, önerilen RGNG yönteminin, Minimum Tanımlama Uzunluğu (MDL) ve Alıcı İşletimsel Özellikler (ROC) analizi ile ölçülen performanslarıyla ortaya konduğu üzere, diğer yöntemlerden açık ara daha üstün olduğunu göstermektedir.

fMRI veri analizi için MATLAB tabanlı bir Grafiksel Kullanıcı Arayüzü (GUI) yazılım paketi olarak tasarlanmış ve uygulanmıştır. Bu çalışmada, sinirbilimi araştırmacılarının verilerini analiz edebilecekleri ve kolayca erişebilecekleri yeni bir model geliştirilmiştir. Önerilen bu model, fMRI veri setinin daha iyi yorumlanması için nörolog ve psikologlara yardımcı olabilecektir.

Anahtar Kelimeler: Kümeleme teknikleri, Veri madenciliği, Grafiksel Kullanıcı Arayüzü (GUI), Artan Sinir Gazı (GNG), Sinirsel Gaz (NG), Güçlü Artan Sinirsel Gaz (RGNG)

CONTENTS

Ph.D. THESIS EXAMINATION RESULT FORM	ii
ETHICAL DECLARATION.....	iii
ACKNOWLEDGEMENTS.....	iv
ABSTRACT	v
ÖZ.....	vi
NOMENCLATURE	x
LIST OF TABLES.....	xii
LIST OF FIGURES	xiii
CHAPTER 1 - INTRODUCTION.....	1
1.1 Magnetic Resonance Imaging (MRI) against fMRI Technique	2
1.2 Basic MRI Concepts.....	6
1.3 FMRI Image Formation.....	7
1.4 Outline of Thesis	9
CHAPTER 2 - MEDICAL BACKGROUND	11
2.1 Brain Medical Background	11
2.2 FMRI Clinical Uses.....	12
2.3 FMRI Complexity and Noises.....	13
2.4 FMRI Paradigm Design.....	15
2.5 Aim and Contribution.....	17
2.6 Summary	17
CHAPTER 3 - BRAIN CLUSTERING AND PARCELLATION	19
3.1 FMRI Brain Parcellation	19
3.2 FMRI Data Analysis Packages.....	20
3.3. FMRI Data Analysis Techniques.....	23
3.4 Literature Review	24
3.4.1 Model-Driven Methods	25
3.4.2 Data-Driven Methods	26
3.4.2.1 Decomposition Exploratory Approaches.....	27
3.4.2.2 Clustering Exploratory Approaches.....	28
3.5 Summary	31

CHAPTER 4 - FMRI DATA ANALYSIS AND PREPROCESSING.....	32
4.1 FMRI Dataset Types	32
4.2 Image Spatial Preprocessing	35
4.2.1 Realignment	37
4.2.2 Coregistration	38
4.2.3 Segmentation	41
4.2.4 Normalization	41
4.2.5 Smoothing.....	44
4.3 General Linear Model (GLM).....	47
4.4 Discussion	54
CHAPTER 5 - PROPOSED ALGORITHMS WITH FMRI DATA ANALYSIS...	55
5.1 FMRI Data-Driven Techniques	55
5.2 Clustering Approaches with fMRI	57
5.3 Neural Gas (NG) Algorithm.....	59
5.4 Growing Neural Gas (GNG) Algorithm.....	64
5.5 Robust Growing Neural Gas (RGNG) Algorithm.....	70
5.6 Comparison among Three Clustering Approaches.....	77
5.7 Summary	80
CHAPTER 6 - ALGORITHMS IMPLEMENTATION AND RESULTS.....	81
6.1 Case of Studies	82
6.1.1 Classification Rate	82
6.1.2 Partition Quality.....	82
6.1.3 Minimum Cluster Number.....	83
6.1.4 Mean Square Error.....	84
6.2 Experimental Results with Synthetic Data.....	84
6.3 Prototype-Based Clustering Package	90
6.4 Experimental Results with the Real FMRI Data.....	94
6.5 Comparison of RGNG and SPM with Auditory Data.....	98
6.6 Visual fMRI Data Mining Package.....	100
6.7 Discussion	108
CHAPTER 7 - CONCLUSIONS AND FUTURE WORK.....	110

7.1 Conclusions	110
7.2 Future Directions	112
REFERENCES	114
RESUME (CV)	128



NOMENCLATURE

Acronyms

3D	Three dimensional
4D	Four dimensional
.hdr	header information file
.img	image data file
AFNI	Analysis of Functional NeuroImages
BOLD	Blood Oxygenation Level Dependent
BSS	Blind Source Separation
CR	Classification Rate
CSF	Cerebro Spinal Fluid
CURE	Clustering Using Representatives
DICOM	Digital Imaging and Communications in Medicine
EEG	Electroencephalogram
EPI	Echo Planar Imaging
FIL	Functional Imaging Laboratory
fM	functional image
fMRI	functional Magnetic Resonance Imaging
FCM	Fuzzy C-Means
FPR	False Positive Ratio
FSL	FMRIB Software Library
FWHM	Full Width at Half Maximum
GLM	General Linear Model
GNG	Growing Neural Gas
GUI	Graphical User Interface
hardcl	Hard Competitive Learning
HRF	hemodynamic response function
HDR	hemodynamic response
ICA	Independent Component Analysis
ICs	Independent Components
LBG	Linde-Buzo-Gray
MDL	Minimum Description Length

MEG	Magneto Encephalogram
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
NG	Neural Gas
NIFTI	Neuroimaging Informatics Technology Initiative
MCN	Minimum Cluster Number
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PLS	Partial Least Squares
PQ	Partition Quality
RF	Radio Frequency
RGNG	Robust Growing Neural Gas
ROC	Receiver Operating Characteristic
ROI	Region of Interest
rs-fMRIs	resting-state functional Magnetic Resonance Images
SNR	Signal-to-Noise Ratio
SOM	Self-Organizing Map
SPECT	Single Photon Emission Computed Tomography
SPM	Statistical Parametric Mapping
TC	Time Course
TE	Echo Time
TPR	True Positive ratio
TR	Repetition Time

LIST OF TABLES

Table 1.1 The Major Functional Brain Measurement Techniques.....	2
Table 3.1 An overview of major fMRI software packages	21
Table 3.2 Comparison among statistical, transformation and clustering based approache	24
Table 4.1 Overview of some medical image data formats	35
Table 5.1 Comparison among Three Artificial Neural Network Approaches Based on Unsupervised Clustering for fMRI Analysis	79
Table 6.1 Clustering results of the synthetic data	87



LIST OF FIGURES

Figure 1.1 MRI system block diagram.....	3
Figure 1.2 MRI against fMRI Image.....	4
Figure 1.3 MR Field directions (a) longitudinal and transverse relaxation; (b) B_0 and B_1 direction; (c) protons alignment inside magnetic field	5
Figure 1.4 TR and TE representation	7
Figure 1.5 MR Image formation steps	7
Figure 1.6 Schematic representation of HDR	9
Figure 2.1 Major parts of the brain	11
Figure 2.2 Major cerebrum lobes	12
Figure 2.3 FMRI scans showing schizophrenia and autism patients compared to healthy controls	13
Figure 2.4 (a) Orbitofrontal region, which has dropout bordered by the blue box; (b) Dropout near nasal sinuses and ear canals	14
Figure 2.5 The effect of head motion (motion is seen as activation along one side of the brain)	15
Figure 2.6 The proposed data mining system architecture.....	18
Figure 3.1 An example page of SPM package's windows.....	22
Figure 3.2 An example page of FSL package's windows.....	22
Figure 3.3 FMRI data analysis in literature	25
Figure 4.1 Original fMRI 4D data and MR image sequences with 5.5 mm spacing between slice	33
Figure 4.2 The SPM menu window	37
Figure 4.3 Realignment Result for 96 fM free auditory data	38
Figure 4.4 Mutual information coregistration output result (mean fM4 and sM).....	39
Figure 4.5 Segmentation results, (a) original image (sM); (b) segmented grey matter (c1); (c) white matter (c2); (d) CSF (c3).....	40
Figure 4.6 The Brodmann areas 3D	41
Figure 4.7 MNI versus Talairach brain atlas.....	42
Figure 4.8 Normalization result (fM20), (a) after normalization; (b) before normalization.....	43
Figure 4.9 Gaussian kernel.....	44
Figure 4.10 Smoothing result (fM20), (a) after smoothing; (b): before smoothing..	45

Figure 4.11 Whole brain fMRI image (a) image with preprocessing; (b) image without preprocessing	46
Figure 4.12 GLM analyses in terms of an fMRI experiment	48
Figure 4.13 A boxcar function	49
Figure 4.14 The HRF	50
Figure 4.15 Box car model function convolved with HRF.....	50
Figure 4.16 The design matrix	51
Figure 4.17 A hypothesis BOLD data.....	52
Figure 4.18 Predicted BOLD response TC	52
Figure 4.19 Experimental paradigm "silence" and "talk"	53
Figure 5.1 The flowchart of NG Algorithm	63
Figure 5.2 The flowchart of GNG Algorithm	69
Figure 5.3 The flowchart of RGNG Algorithm.....	78
Figure 6.1 Clustering with screw synthetic dataset for N=7, by running NG, GNG and RGNG techniques.....	85
Figure 6.2 Clustering with set5 synthetic dataset for N=10, by running NG, GNG and RGNG techniques.....	85
Figure 6.3 Clustering with snail synthetic dataset for N=12, by running NG, GNG and RGNG techniques.....	86
Figure 6.4 MDL values versus the number of clusters running NG, GNG and RGNG techniques on synthetic data, for: (a) N=7; (b) N=10; (c) N=12.....	90
Figure 6.5 Main window of the Prototype-Based Clustering software package.....	91
Figure 6.6 Different datasets with their information, (a) Snail data; (b) Screw data; (c) Ring data; (d) Set5 data	92
Figure 6.7 RGNG clustering with Set3 data (N=14).....	93
Figure 6.8 Comparison of MDL values for N=16.....	94
Figure 6.9 Active areas in the brain auditory cortex area using SPM package.....	95
Figure 6.10 Results of the clustering techniques for a transparent brain image	96
Figure 6.11 Clusters overlaid into the anatomical image.....	97
Figure 6.12 Activation zones of whole brain running: (a) RGNG; (b) SPM package	98
Figure 6.13 ROC analyses, (a) Good classifier; (b) Bad classifier	99
Figure 6.14 ROC curves of the auditory fMRI dataset	100
Figure 6.15 FMRI Clustering and Analysis software package	101
Figure 6.16 FMRI dataset for whole brain before the preprocessing.....	102

Figure 6.17 FMRI dataset for whole brain after the preprocessing	102
Figure 6.18 FMRI brain slices (a) slice no20; (b) slice no46.....	104
Figure 6.19 RGNG clusters on a transparent image.....	105
Figure 6.20 NG clusters overlaid in the anatomical image.....	106
Figure 6.21 GNG clusters overlaid in the anatomical image	106
Figure 6.22 RGNG clusters overlaid in the anatomical image	107



CHAPTER 1

INTRODUCTION

Brain is the most complex, fascinating and mysterious organ in the human body; it is responsible for controlling of many voluntary and involuntary activities of the body as sensory perception, memory, seeing, consciousness, attention, thinking, language, movement, emotions and etc. For the last few years, many mysteries are becoming clear with the development of cognitive neuroscience and the advancement of functional brain imaging techniques. Table 1.1 shows the major functional brain measurement techniques which was developed and used for understanding the brain functions and brain mappings such as Electroencephalogram (EEG), Magneto Encephalogram (MEG), Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT) and functional Magnetic Resonance Imaging (fMRI).

EEG system records the electrical activity of the brain over a period of time, using multiple electrodes placed on the scalp. MEG scan, records small magnetic fields produced by electrical currents occurring naturally in the brain, using very sensitive magnetometers called super conducting detectors and amplifiers (SQUIDS). Moreover, PET and SPECT are nuclear medicine tomographic imaging technique using gamma rays. PET scan uses a radioactive substance called a tracer, which is introduced into the body to search for disease in the body. SPECT scan by using radiopharmaceuticals administered intravenously or by inhalation to evaluate function in the human brain. While, fMRI measure the brain activity by detecting changes associated with blood flow.

EEG and MEG define the neuronal functions in real time for 10-100 msec, but provide poor spatial resolution (mm-cm). So fMRI has better spatial resolution (provide spatial information about the activation regions of the human brain) rather than EEG with visualized functional images, and better temporal resolution rather than PET or SPECT. Moreover, it is non-invasive; it can be adapted to many types of

experimental paradigms, has high sensitivity and gives non-ionizing radiation unlike X-Ray or Computer Tomography (CT) scan. So fMRI is a powerful tool to provide important information about the brain.

Table 1.1 The Major Functional Brain Measurement Techniques.

Techniques	Technique Types	Invasiveness	Brain Property Used	Temporal Resolution	Spatial Resolution
EEG	Electrophysiological	Non-Invasive	Electrical	Millisecond	Poor
MEG	Electrophysiological	Non-Invasive	Magnetic	Millisecond	Potentially Good
PET, SPECT	Neuroimaging	Invasive	Hemodynamic	Minutes	Millimeters
fMRI	Neuroimaging	Non-Invasive	Hemodynamic	Seconds	Millimeters

After this short introduction about the major functional brain measurement techniques and especially fMRI, the organization of the rest of this chapter is as follows: an introduction to Magnetic Resonance Imaging (MRI) against fMRI technique is presented in Section 1.1, and Section 1.2 presents basic MRI concepts. Then in Section 1.3, the information about the fMRI image formation is introduced. Finally a brief overview about all the chapters is given in Section 1.4.

1.1 Magnetic Resonance Imaging (MRI) against fMRI Technique

MRI consists of three main parts (Figure 1.1): Huge Magnet, Gradient Coils and Radio Frequency (RF) coils to produce three types of magnetic fields that are required in MRI scanners. So it uses a powerful magnetic field (1.5 - 7.0 Tesla), RF pulses and a computer for evaluating various parts of the body [1]. MRI image depends on the structure of water molecules in hydrogen nuclei and their

measurements in the body. The hydrogen atoms absorb the RF pulses generated by the transmitter RF coil and re-emit it as MR which is distinguished as a small voltage in a receive RF coil. Hydrogen nuclei return back to a relaxed state by T1 and T2 relaxation mechanisms. Different body tissues have different relaxation times that depend on the hydrogen molecule size and binding to other molecules [2].

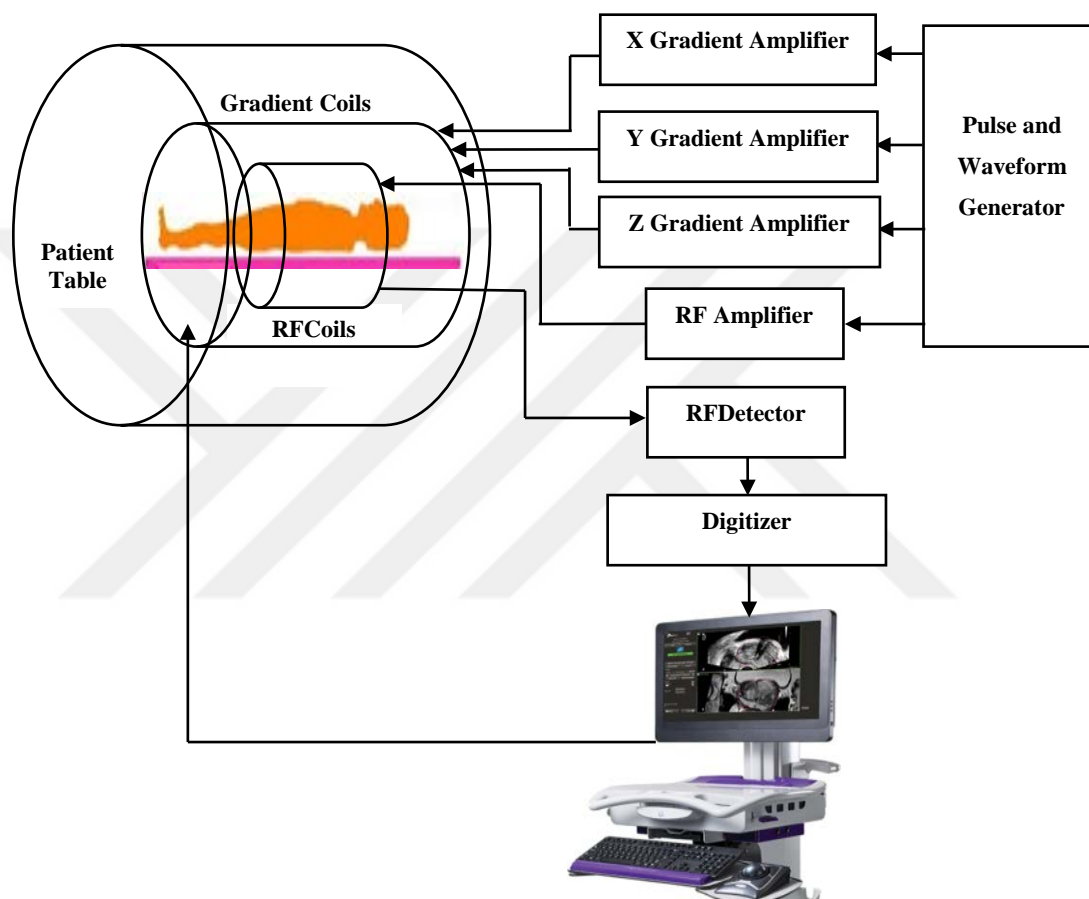


Figure 1.1 MRI system block diagram

The same as MRI hardware is fMRI, while fMRI uses MRI to calculate the levels of oxygen in the blood or Blood Oxygenation Level Dependent (BOLD) that measures the ratio of oxygenated to deoxygenated hemoglobin in the blood and is closely related to the neural activity. MRI produces the anatomical of organs, soft tissues, bone and all other internal body structures which allow physicians to evaluate them and detect the presence of certain diseases, while the fMRI views the metabolic function for measuring brain activity because it measures the hemodynamic response

function (HRF) or metabolic demands (oxygen consumption) of functional or active neurons in the brain or spinal cord and doesn't measure neuronal activity directly.

So MRI views anatomical or structural image with high resolution and produces the differences between tissue types with respect to space, while fMRI views the functional image with low resolution and produces the differences between tissue types with respect to time (Figure 1.2).

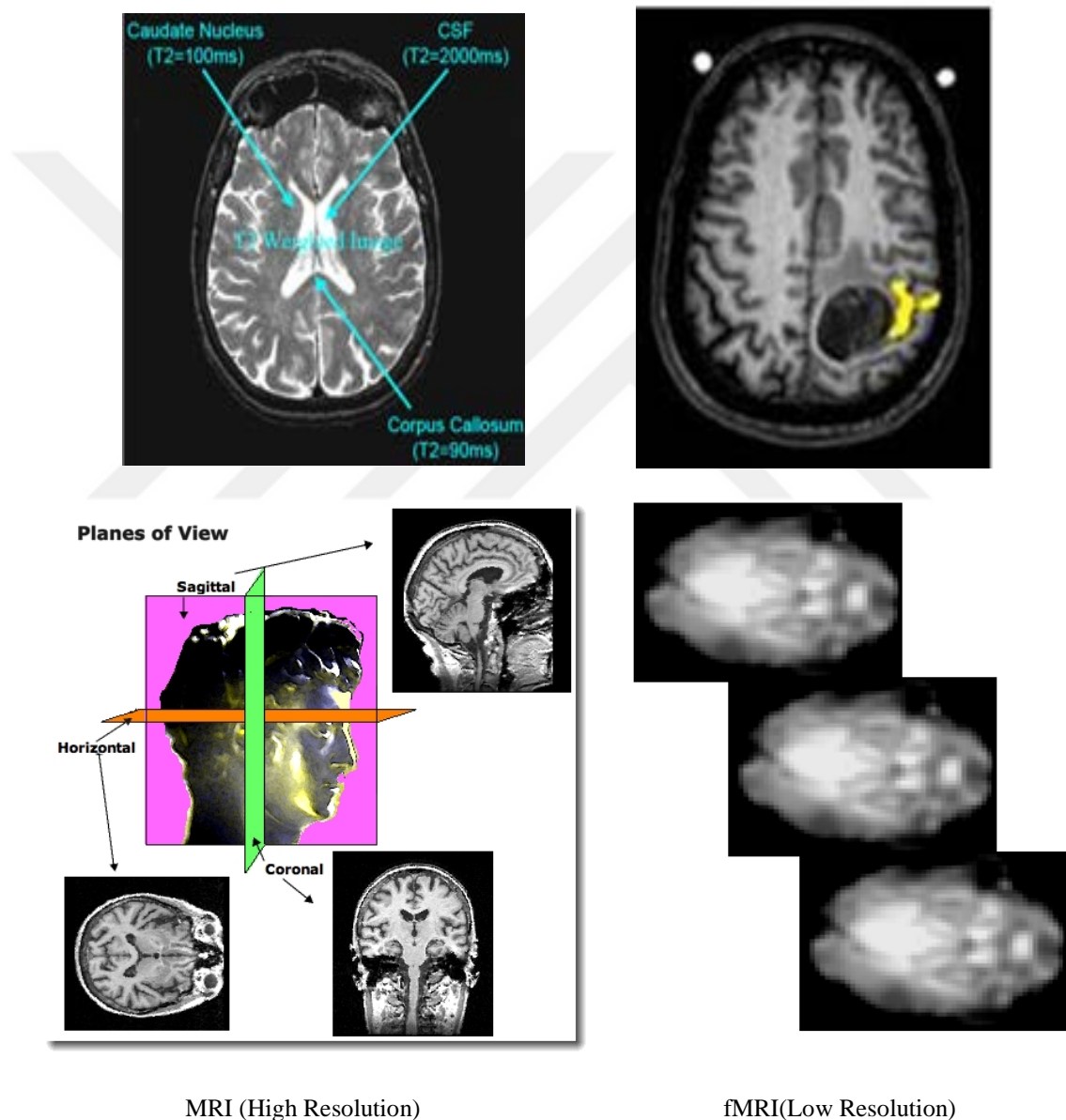
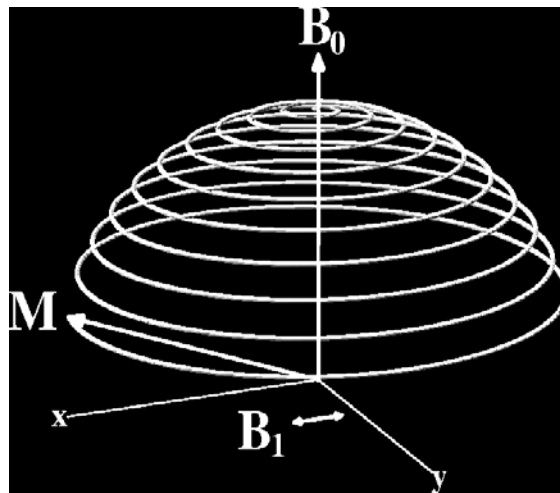
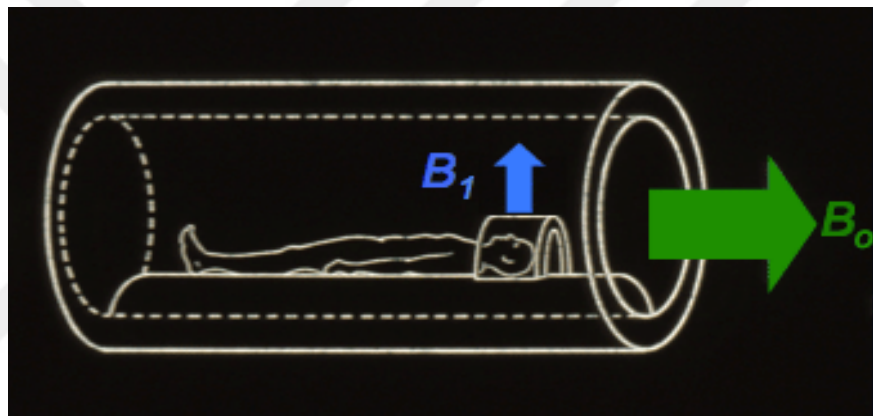


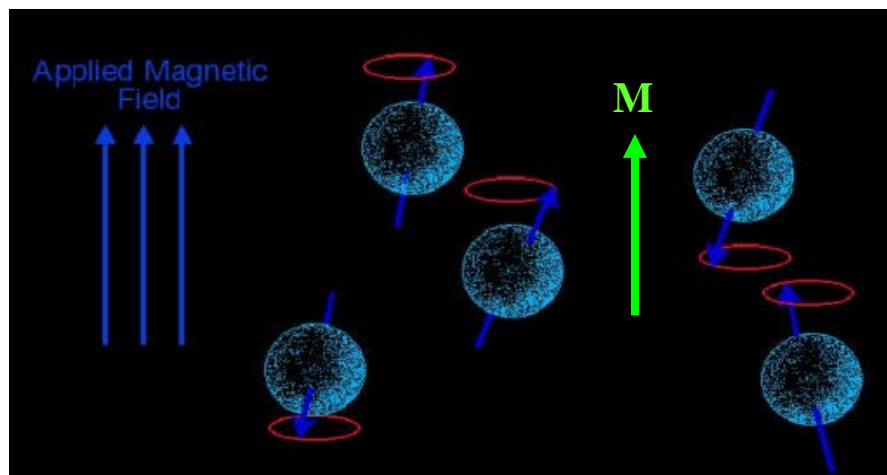
Figure 1.2 MRI against fMRI Image.



(a) Longitudinal relaxation (along B_0) transverse relaxation (along B_1).



(b) B_0 , the applied magnetic field; B_1 , RF pulse.



(c): protons aligned with or against B_0 inside magnetic field; M = net magnetization.

Figure 1.3 MR Field directions (a) longitudinal and transverse relaxation; (b) B_0 and B_1 direction; (c) protons alignment inside magnetic field.

The net magnetization vector M has two components (Figure 1.3):

M_Z — is the longitudinal component aligned with the applied magnetic field B_0 , and M_{xy} — is the transverse component in the plane orthogonal to B_0 .

Before applying the RF pulse B_1 , M is in the equilibrium state where M_Z is maximum and M_{xy} is zero.

After applying RF pulse B_1 , M_Z becomes small and M_{xy} becomes large. When B_1 is switched off, M will return to the equilibrium state. The recovery of M_Z to the initial magnetization M_0 after the RF pulse (longitudinal relaxation) is characterized by the relaxation time constant T_1 . The decay of M_{xy} after the RF pulse (transverse relaxation) is characterized by the relaxation time constant T_2 .

1.2 Basic MRI Concepts

There are some basic concepts, glossary related to MRI and fMRI; so it is important to understand their meanings:

- Repetition Time (TR) of the RF pulse: represents the length of time between consecutive points on a repeating series of pulses and echoes, smaller TR is faster imaging (Figure 1.4).
- Echo Time (TE): represents the time from the center of the RF pulse to the center of the echo (Figure 1.4).
- GRE: Gradient echo sequences.
- Slices: 3 dimensional (or 4 dimensional) images are usually formed by collecting thin slices together; so 30 slices with 3 mm thickness are needed to cover the whole brain which takes about 50-100 ms to collect the data for one slice image only.
- Voxel: represents the smallest 3 dimensional unit in the image.
- T1 relaxation: represents the rate at which the object goes from a non-magnetized state to a magnetized state measuring the longitudinal relaxation along B_0 and typically used to form anatomical images.
- T2 relaxation: represents the rate of decay of the MR signal after the RF pulse is excited measuring the transverse relaxation along B_1 .

- $T2^*$: represents the overall decay forming the functional image in fMRI.

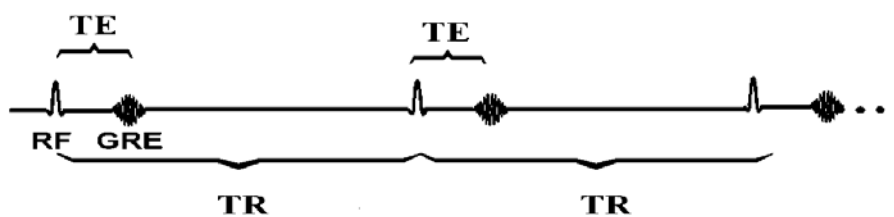


Figure 1.4 TR and TE representation.

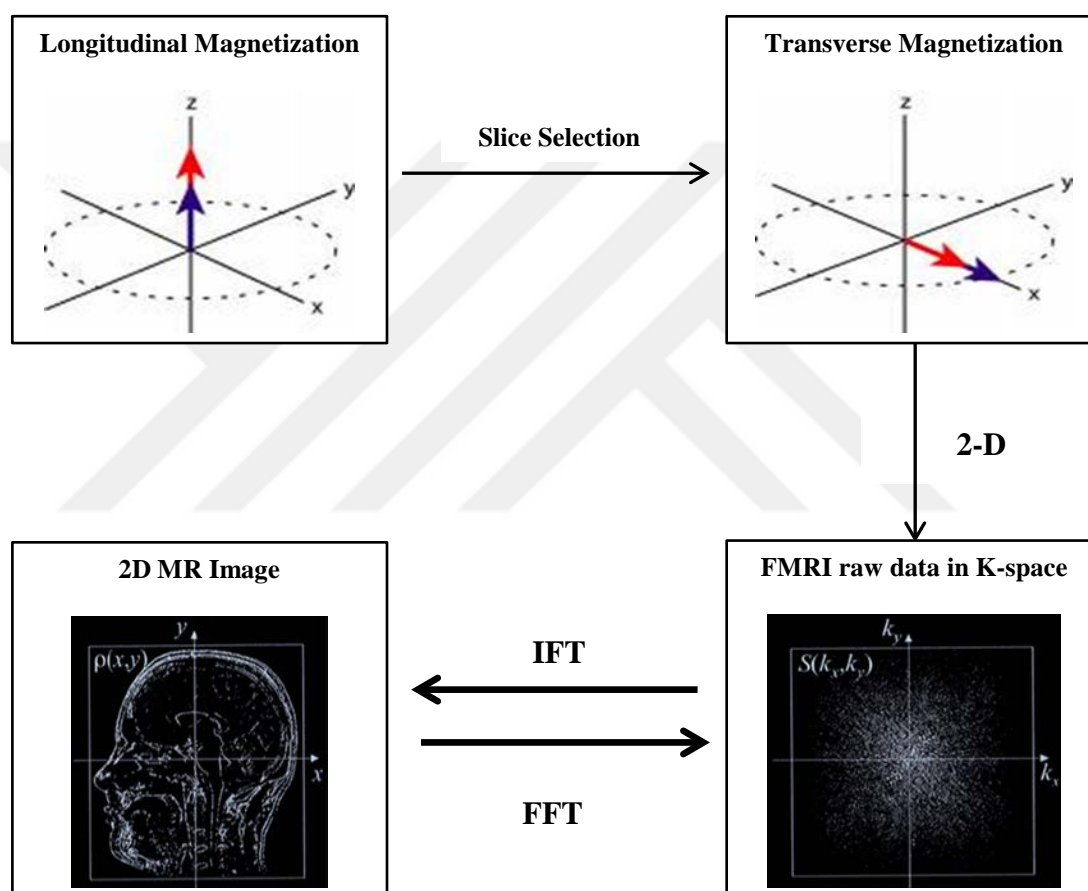


Figure 1.5 MR Image formation steps.

1.3 fMRI Image Formation

The fMRI image formation consists of three steps (Figure 1.5):

1. Slice selection
2. Spatial encoding
3. Image reconstruction

Slice selection step occurs when the particular slice is excited by the spins that generates a T1weighted image (Anatomical image) at which the object goes from a non-magnetized state to a magnetized state measuring longitudinal relaxation, a T2 weighted image is the rate of decay of the MR signal after the RF pulse is excited measuring the transverse relaxation and overall decay is T2* producing a functional image. Then the spatial distribution related to the spins is coded by the two-dimensional gradient impulse that gives the MR signal in k-space which holds raw data in the spatial frequency domain. 2D MR images are finally reconstructed from the raw data in the k-space by use of a Fourier Transform (FT).

T2* images triggered by the change of MR signal in the neuronal activity which is known as the Hemodynamic Response (HDR). HDR obtained by the reduction in the amount of deoxygenated blood and represents temporal properties of the brain. HDR shape consists of three phases as shown in Figure 1.6, which varies according to the rate and duration of the neural activity.

From this figure it can be recognized that the amplitude increases when the rate of neural activity increases and the width also increases when the duration of the neuronal activity increases [1]. The HDR has three phases shown in the previous figure can be summarized as follows [3]:

- *Initial dip*: occurs due to the transient increase in deoxyhemoglobin during the oxygen consumption after beginning of neural activity, which causes MR signal decreases below baseline.
- *Overcompensation*: occurs when MR signal increases above baseline due to increase of neuronal activity. Increase of neuronal activity causes oxygenated blood supplied more increased than the oxygenated blood extracted as well as decreasing in deoxygenated hemoglobin.
- *Undershoot*: occurs due to neuronal activity end, which causes the MR signal amplitude gradually decreases below the baseline level and then returns back to the baseline level due to a combination.

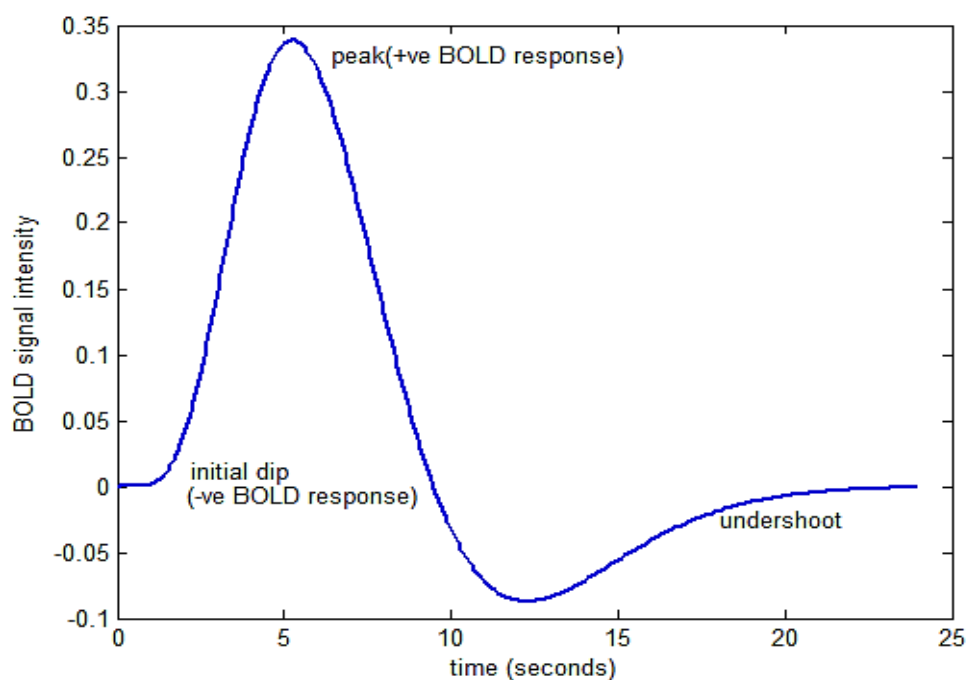


Figure 1.6 Schematic representation of HDR.

1.4 Outline of Thesis

The following chapters present the theoretical and practical research done through the development of this dissertation. The research presents robust and efficient algorithms for detecting the active areas in the brain. This thesis is structured in seven chapters that are briefly explained here:

Chapter two: is a presentation of the brain medical background and fMRI clinical uses. Thereafter complexity and noises of fMRI signal as well as paradigm design will be presented. This chapter introduces short review about the aim and the contribution of the thesis.

Chapter three: is a presentation of fMRI brain clustering and parcellation technique principle, followed by the packages and techniques used in fMRI data analysis. Thereafter literature review will be introduced about fMRI researches starting from the first of using BOLD signals until to using parcellation and clustering approaches.

Chapter four: information about the fMRI dataset type is introduced in this chapter, followed by preprocessing steps using free Statistical Parametric Mapping (SPM)

software package. Thereafter the conventional General Linear Model (GLM) method that used by SPM will be presented with its applications.

Chapter five: this chapter aims to introduce the proposed model which has different and important features in comparison with others clustering approaches using a novel Robust Growing Neural Gas (RGNG) approach, and apply this algorithm into fMRI dataset to detect the active areas in the brain with a comparison to the GNG algorithm which is not used for that purpose yet. A complete explanation of the three algorithms Neural Gas (NG), Growing Neural Gas (GNG) and RGNG is presented in this chapter. Also flowcharts will be designed for each of the three algorithms and presented; as well as a comparison among the three artificial neural network approaches that based on unsupervised clustering for fMRI Analysis is proposed and will be presented in a table.

Chapter six: shows the output results were obtained by feeding the proposed learning algorithms to synthetic and real fMRI dataset. These results were formulated on a visual software package using Graphical User Interface (GUI). The validity of the obtained results was proofed using statistical analysis.

Finally, the conclusions and future work are presented in chapter 7.

CHAPTER 2

MEDICAL BACKGROUND

fMRI provides spatial information about the activation regions of brain that is activated due to various cognitive and/or motor functions repeatedly, and generates a sequence of three dimensional (3D) images. There are a large number of artifacts can effects on fMRI dataset. These artifacts can roughly be divided into the scanner-induced and physiological artifacts. So, there are some software packages available for analyzing fMRI data. All this information as well as some introduction about the brain and fMRI medical using will be explained in details in the next sections.

2.1 Brain Medical Background

Brain is the center of the nervous system; Figure 2.1 shows the major parts of the brain which consists of the cerebrum, cerebellum and brain stem. This figure also shows the three main types of brain “matter” which are Gray matter, white matter and Cerebro Spinal Fluid (CSF). The top part of the head is the cerebrum, which is the biggest parts and most people think of as the brain. The smaller is the cerebellum, while the part that connects the brain to the spinal cord is the brain stem.

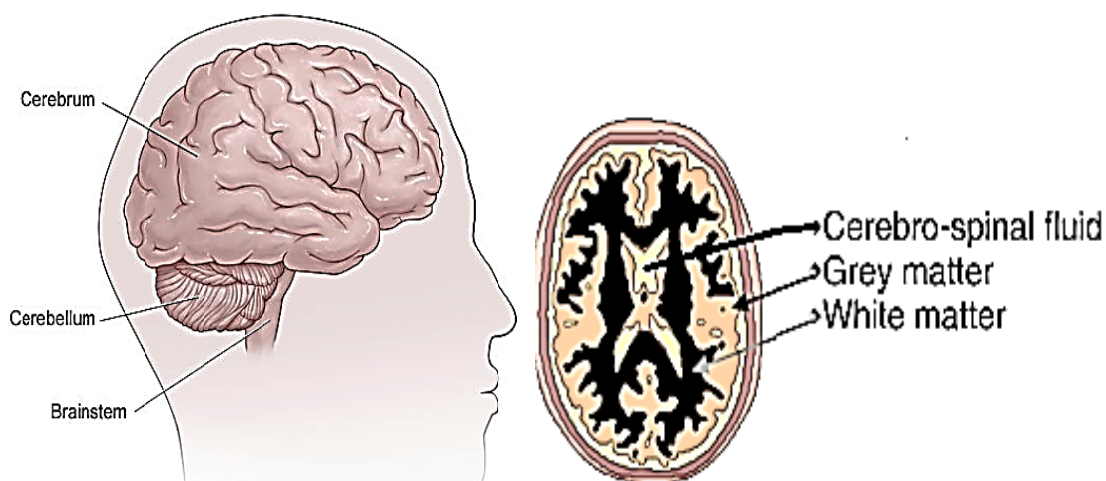


Figure 2.1 Major parts of the brain.

Mainly the brain is divided into two left and right hemispheres where the functional areas are differentiated according to the cerebrum lobes that split into four “lobes” or areas as shown in Figure 2.2:

- Frontal lobe is responsible for personality, working memory, decision making, controlling inhibition, movement, etc.
- Parietal lobe is responsible for sensation, maths, music, humor, spatial tasks, etc.
- Temporal lobe is responsible for hearing, memory, language, emotions, object recognition, etc.
- Occipital lobe is responsible for vision only.

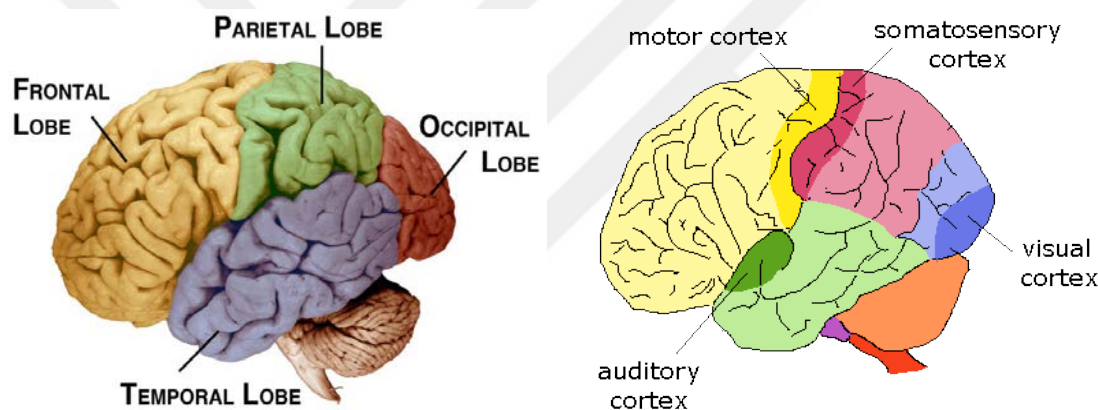


Figure 2.2 Major cerebrum lobes.

2.2 FMRI Clinical Uses

FMRI is used to understand neuronal mechanisms behind the disorders, such as Bipolar Disorder, Parkinson Disease, Autism Spectrum Disorders, Schizophrenia, and Alzheimer Disease.

- *Bipolar Disorder (BD)*: is a brain disorder with four basic types. Bipolar disorder causes clear changes in mood, energy, and activity levels [4].
- *Parkinson's Disease (PD)*: is a gradual disorder of the nervous system which develops progressively and affecting the human movement, it started with a simple tremor in one hand only then may cause the movement slow or hard [5].

- *Autism Spectrum Disorder (ASD)*: is characterized by weakness in social reactions and inability to communicate with the environment, it occurs in more than 1% of children and the parents can recognize the autism signs in the first two years of the child's life [6].
- *Schizophrenia (SCZ)*: are characterized by progressive impairment in cognitive and social functioning, but the severe phases of the two disorders are various [7].
- *Alzheimer's disease (AD)*: is progressive slowly over time, which causes trouble with memory, thinking and behavior as a type of dementia that lead to the inability to communicate with daily tasks [8].

Figure 2.3 shows fMRI scans for schizophrenia and autism patients compared to healthy controls.

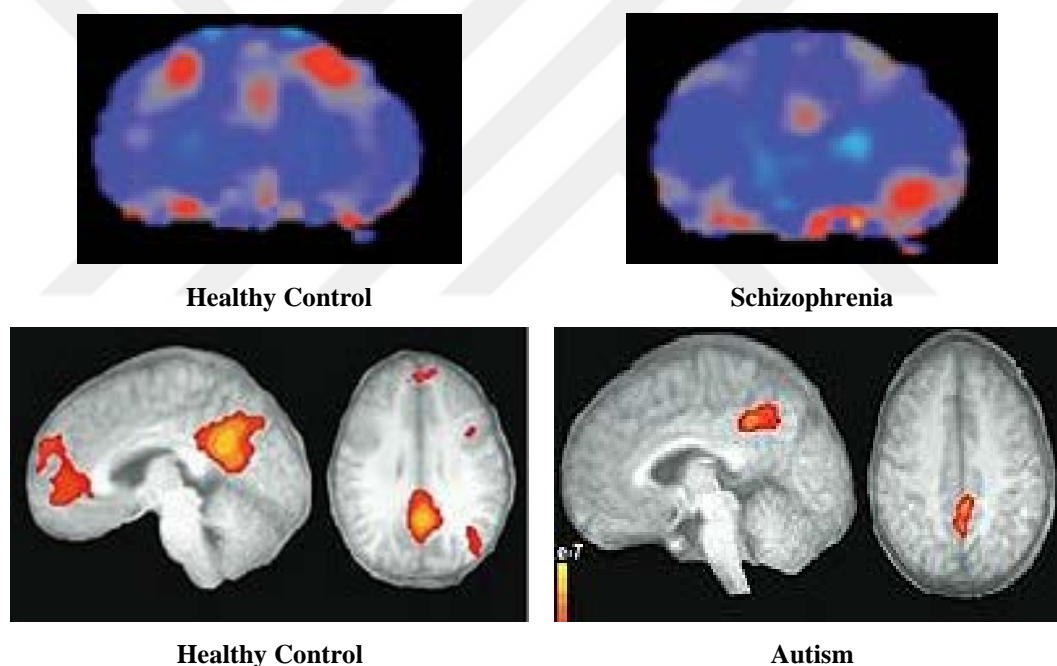


Figure 2.3 FMRI scans showing schizophrenia and autism patients compared to healthy controls.

2.3 FMRI Complexity and Noises

The complexity of fMRI data analysis is related to the huge dimensions of the data as well as having a lot of artifacts or noises which caused mainly due to various reasons such that related to hardware system (MRI scanner itself), individuals themselves (e.g. head motion) or physiological effects. Noises related to hardware system as B_0 Inhomogeneity and signal noise. Signal noise occurs due to bad RF shielding,

operation of the RF coil as well as the presence of metal in the patient that can lead also for B_0 inhomogeneity and B_0 field distortion [9]. Nevertheless, the artifact due to the thermal motion of free electrons in the fMRI hardware system (eddy currents or scanner heating) that causes heterogeneity of the magnetic field, there is another type of noise originated from the fMRI hardware when the image has "dropout" or dark in the brain regions adjacent to the air-tissue interfaces [10]. The air-tissue interfaces regions like nasal sinuses and ear canals that have dropout in the medial frontal lobe (orbitofrontal cortex) and dropout in lateral temporal lobes respectively as shown in Figure 2.4.

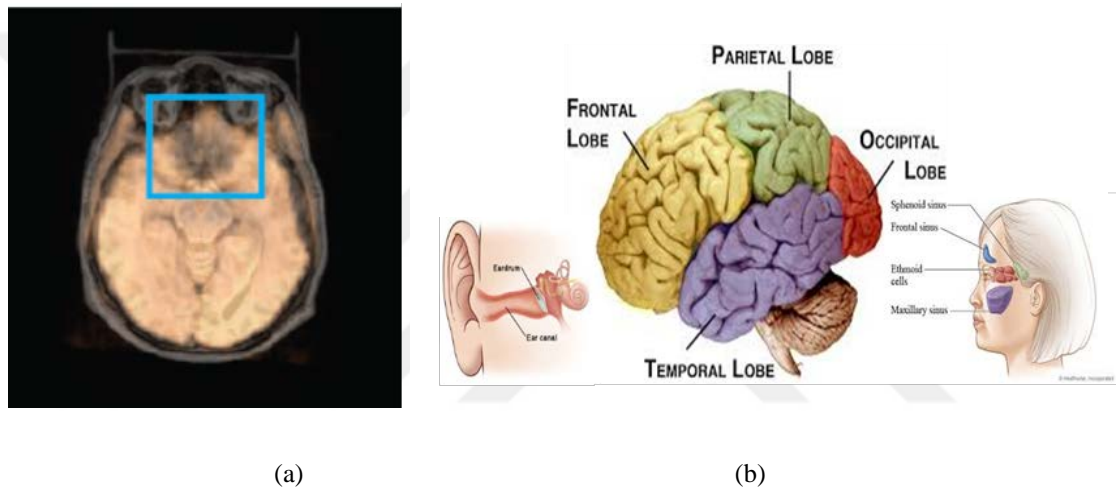


Figure 2.4 (a) Orbitofrontal region, which has dropout bordered by the blue box; (b) Dropout near nasal sinuses and ear canals.

Another part of the fMRI artifact sources are originated from the subjects themselves because of the fact that the patient is never completely motionless during experiments. Thermal noises of examining tissues, voluntary or involuntary movements of the patient as head movements (Figure 2.5) and eye movements are examples of subjects' movement [11]. While the subject's heartbeat, respiration and internal movement related to the blood flow pulsation are examples of the physiological effects.

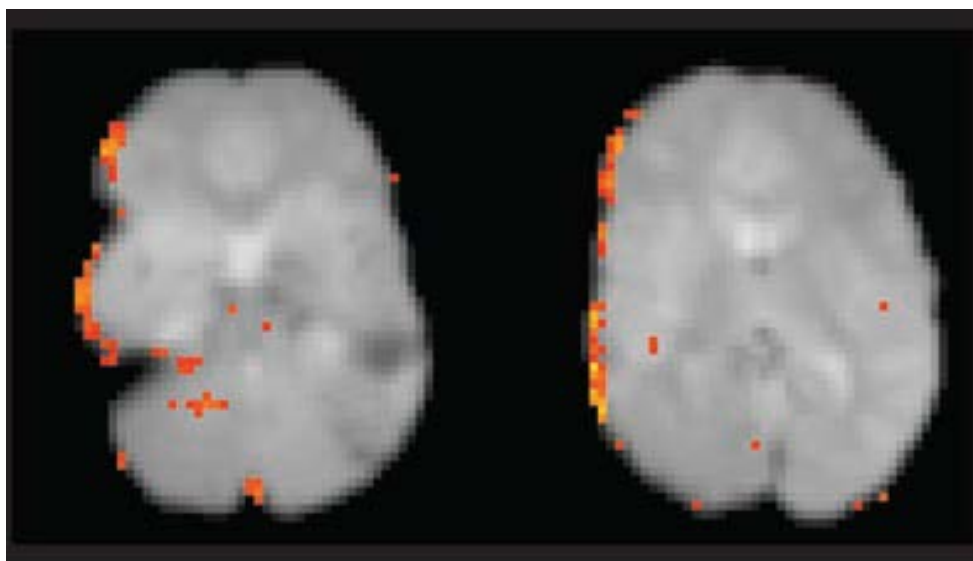


Figure 2.5 The effect of head motion (motion is seen as activation along one side of the brain).

There are also different types of noises related to signal processing which causes a variety of artifacts due to partial volume, Gibbs phenomenon and chemical shift. Partial volume occurs due to the presence of large voxel size. If there are two tissues with different resonance frequencies under an external magnetic field, chemical shift occurs due to the mis-registration between the positions of these tissues. While at high spatial frequencies, Gibbs phenomenon appears when images with bright or dark lines adjacent to sharp boundaries [12].

2.4 fMRI Paradigm Design

Task-based BOLD/fMRI studies which employ a certain (cognitive) paradigm can be divided into three main types of the paradigm designs:

1. block designs
2. event-related designs
3. mixed designs

Before analyzing the fMRI dataset, there are some important factors must be defined using the paradigm design during an fMRI experiment, as the temporal structure, construction and behavioral predictions of paradigm tasks executed by the subject [13]. During the experiments, the subject lying in the MRI scanner, BOLD response is defined during the cognitive task which defines a set of brain functions. The neural

responses or stimuli may be a receptive nature during the visual or auditory tasks, or reactive nature as memorize images during a certain period of time or respond to a stimulus (tap of a finger as an example).

Block designs are the oldest functional imaging paradigms; it is widely used in PET studies and dominated the first years of fMRI experimentation. Task periods or epochs are alternated with periods of rest that represent different cognitive states. In the simplest form, the experiment is composed of two states defined by different conditions as finger tapping stimulation and resting stimulation. Another example is the auditory dataset which used in this work; this experiment is composed of two states as auditory stimulation using bi-syllabic words (e.g. “mother”, “house”, “weather”, “movie” etc.) and rest stimulation. The main advantage of block designs is the highest signal-to-noise ratio (SNR), statistical power and maximal time efficiency. The main disadvantage of the block designs is their limited use, because in some cases it could be difficult to preserve a cognitive task for long periods of time, so information about the HDR and fMRI signal timing are difficult to measure. In other hand, the block design is not suitable as the presentation of an unpredictable stimulus.

In these cases event-related designs can be applied using single or multiple tasks and stimuli to take place at short and variable time intervals. This type of experimental design emerged in the mid of 90th providing the high degree of flexibility required for sophisticated neuropsychological experiments. This experiment can be used with the faster image acquisition of fMRI which enabled the detection of small changes in the HDR. Disadvantages of event-related designs include lower SNR and statistical power, with longer acquisition times and more trials per subject required. Analysis of the data is significantly more complex and dependent on accurate modeling of the HRF.

Mixed paradigms designs combine the features of blocked and event-related designs. An example of the mixed paradigms is semi randomized conditions with rest periods in between which takes place during the task blocks. The main advantage of mixed

designs is related to maintaining the preferable SNR characteristics of block designs with the flexibility of even-related ones.

2.5 Aim and Contribution

The aim of this research is to analyze fMRI dataset by running new and robust prototype-based unsupervised clustering algorithm. This approach and in comparison with other approaches is tested with the fMRI real auditory dataset. Figure 2.6 shows the data flow of the proposed data mining system architecture oriented toward brain function exploration which should be able cope with the problem of model-based method.

The process in the block is composed of five stages:

1. preprocessing of the raw data
2. clustering voxels together based on the similarity of their intensity profile in Time Course (TC) of fMRI image
3. overlay with the structural image
4. visual fMRI image
5. validation

Each part of the introduced block diagram will be explained in details in the next chapters.

2.6 Summary

fMRI quite used between imaging methods for studying brain functions in humans. In this chapter, information about brain medical background, types of artifact effect on fMRI as well as paradigm design is introduced.

The framework of the proposed work is introduced in the form of a block diagram. This block diagram shows the process of fMRI data analysis which is performed in the thesis; which includes: fMRI preprocessing steps, clustering voxels using the proposed algorithm, overlaying these clusters with the brain anatomical image, visualization and validation. The details of each part of those who introduced in the block will be expanded in the next chapters.

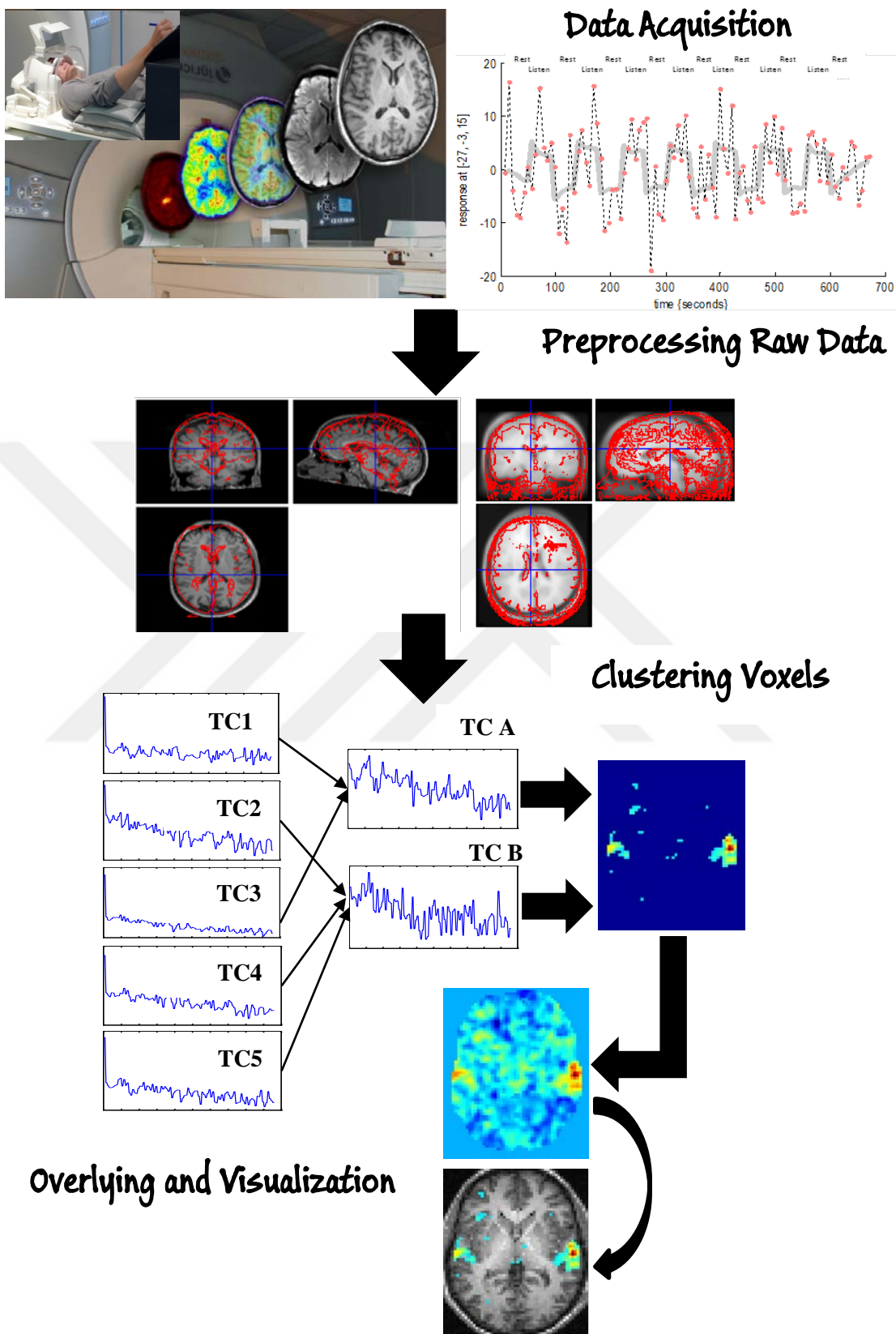


Figure 2.6 The proposed data mining system architecture.

CHAPTER 3

BRAIN CLUSTERING AND PARCELLATION

The fMRI technique scans whole or part of the brain repeatedly and generates a sequence of 3-D images. The voxels of the brain that represent the real activity is very difficult to be detected because of a weak SNR, the presence of artifacts and nonlinear properties. Due to these difficulties, using of data mining is an important. Data mining is used as a complement or replacement of the classical methods when it is difficult to predict what will occur during the acquisition of detection. FMRI dataset mining using a data-driven unsupervised clustering approach is proposed in this thesis.

An explanation of fMRI data analysis techniques as well as the presence of packages used in the same purpose is presented in this chapter. Thereafter, a literature review about fMRI researches is presented; starting from the first of using BOLD signals until to using parcellation and clustering approaches.

3.1 FMRI Brain Parcellation

To obtain the best performance for whole brain functional connectivity data analysis, the brain must be divided into many region of interest (ROI) to be used as network nodes. The structures of ROI's are normally at the level of many voxels constituting which is a possibly small brain region, and rarely at the level of a specific voxel. Several methods were proposed for defining ROIs and study function beyond the voxel description, which include using three strategies: (1) Randomly splitting the brain into anatomical or functional ROI's, (2) Anatomical brain atlas, (3) Brain parcellations using data-driven or clustering functional data.

For randomly splitting the brain into ROI's, the selection of these regions is depended on background and long experiments because of the cancellation problem. So, any signal lies outside the ROI will be ignored as a consequence and the final

results will not fit perfectly the new data. Therefore, this is regarded as a limitation by using the first strategy for defining ROI's [14].

The anatomical brain atlas provides a set of ROI's that cover all the brain volume [15- 17] or structures (anatomically, functionally or based on connectivity). There are two limitations of using the brain atlases: (1) All brain atlases are inconsistent between them [18], (2) Each atlas may not perfectly fit the data.

Brain parcellations are either anatomical or functional parcellations. The parcels in anatomical parcellations must be performed with the most appropriate atlas. The functional parcellations can be derived either from resting-state functional Magnetic Resonance Images (rs-fMRIs), activation data or other analyses. Functional parcellations use data-driven or clustering functional data.

Parcellation approaches use brain activity and clustering approaches to divide the brain into many parcels or regions with some degree of homogeneous characteristics. So, the brain is divided into regions with some degree of signal homogeneity. Parcellation approaches help for analysis and interpretation of neuroimaging data as well as mining these data because of the amount of fMRI data are huge.

3.2 FMRI Data Analysis Packages

The size of fMRI data is huge and the detection of voxels in the raw data that represent the real activity is very complex because of a weak SNR and of the existence of artifacts. So, powerful techniques from signal, image processing and statistics are fundamental to get a successful interpretation of fMRI experiments in order to go from the raw data to the finished image production. For processing and analysis of fMRI data, several software packages are now available some of them are freely available as: SPM, FMRIB software library (FSL), analysis of functional neuroimages (AFNI) and etc.; while others not as BrainVoyager. Table 3.1 introduces examples of these packages, Figures 3.1 and 3.2 show the most of famous two packages SPM [19] and FSL [20]. In this thesis SPM was used with fMRI dataset, by applying preprocessing steps which including: realignment, coregistration, normalization, segmentation and smoothing before starting with fMRI data analysis.

These software packages focused on testing a brain behavior by the statistical methods and detected the more functional voxels in the brain area under comparable conditions. So using these packages are limited to the prefixed model and cannot be able to give the output results without the expecting about what will have.

In this thesis, a novel approach relies on the RGNG algorithm to detect the active areas in the brain is presented which is not used for that purpose yet. Before starting with feeding our learning algorithm to the dataset, a number of preprocessing operations steps are required by using SPM as explained in the following chapters.

Table 3.1 An overview of major fMRI software packages.

Package	Developer	Platforms	Licensing	Properties
SPM	University College London	MATLAB	Open source	<ul style="list-style-type: none"> • Very active development community • Emphasis on fast processing • Command-line and GUI
FSL	Oxford University	UNIX	Open source	<ul style="list-style-type: none"> • Emphasis on time series modeling and cool new techniques • Command-line and GUI
AFNI	NIMH	UNIX	Open source	<ul style="list-style-type: none"> • Part command-line, part GUI
BrainVoyager	Brain Innovation	Windows, Mac, Linux	Commercial (closed source)	<ul style="list-style-type: none"> • GUI interface • Good surface inflation visualization • Fairly large set of tools, but smaller community
3D Slicer	BSD-style	Windows, Mac, Linux	Open source	<ul style="list-style-type: none"> • Tools for visualization, segmentation and quantification
FreeSurfer	FreeSurfer Software	Mac, Linux	Open source	<ul style="list-style-type: none"> • Automated tools for reconstruction of the brain's cortical surface

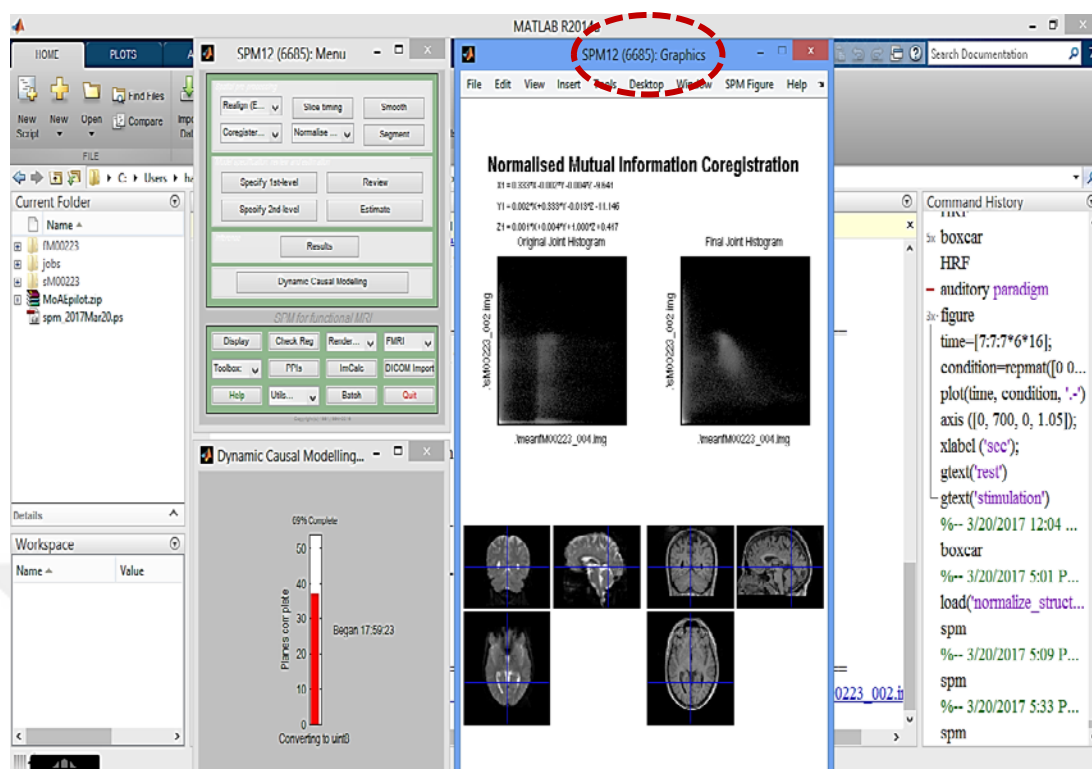


Figure 3.1 An example page of SPM package's windows.

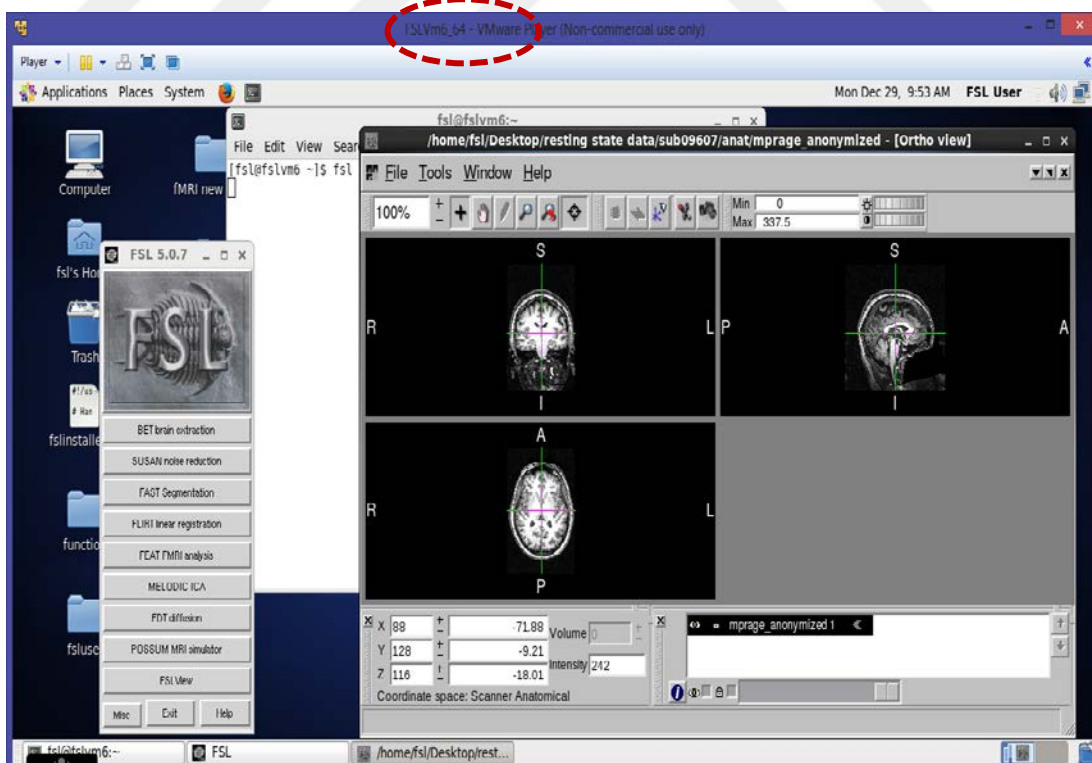


Figure 3.2 An example page of FSL package's windows.

3.3 FMRI Data Analysis Techniques

FMRI data analysis methods could be divided mainly into two categories: model-driven or model-based (hypothesis) and data-driven or model-free (exploratory) approaches. The methods based on model-driven deal with a definite activation pattern, response functions or experiment. So firstly they need a prior knowledge about them, and then test the analyzed data statistically about the presence or absence of the response. The methods related to this category differ from, either by the statistical method or the signal estimation procedure for performing the activation. The commonly used GLM method belongs to this category and is the most fundamental and basic approach used for fMRI data analysis, but it needs a prior knowledge about the activation patterns or experiment.

The methods based on data-driven count all the voxels simultaneously. Model-free approach has the ability to define the active zones and find structures in the brain and fMRI data competently without need of prior knowledge about activation patterns or experiment paradigm. The methods related to data-driven approach can be divided mainly into two groups: blind source separation (BSS) and clustering approach.

BSS tries to find unobserved signals or ‘sources’ from several observed mixtures, and generate a model of the data. There are different methods for blind signal separation: principal component analysis (PCA) [21, 22], independent component analysis (ICA) [23- 26] and canonical correlation analysis (CCA) [27]. These methods are used to separate the mixtures to obtain the source signals. FSL package used melodic ICA, which is

a model-free approach, but it is insufficient for most fMRI datasets because ICA has some limitations. ICA is attempting to find maximally independent maps and split the wide activation areas into a number of maps which have a strong correlation between TCs of different components. Also the independent components (ICs) from ICA decomposition are not ordered i.e., it is associated with the model order selection for linear model-based region extraction is still an open problem. So, it is difficult to identify any ICs are non-linear activation correlated or not.

Clustering [28, 29] analysis is based on group voxels according to their TCs signals into a similar HDR over time. As well as all the advantages presented previously for our proposed clustering algorithm RGNG and that will be explained carefully next; it is mainly a data-driven or model free (exploratory) approach.

Clustering techniques are considerable model-free or exploratory data analysis approaches and have the ability to define the active zones and find structures in the brain and fMRI data competently without need of prior knowledge about activation patterns or experiment paradigm. Table 3.2 compared statistical, transformation and clustering methods.

Table 3.2 Comparison among statistical, transformation and clustering based approaches.

Based Approach	FMRI Analysis Methods	Approach Properties
Statistical	Model-Driven/ Model-Based/ Hypothesis Used with SPM package and based on GLM	The most fundamental, basic and commonly used approach for fMRI data analysis, but it needs previous knowledge about activation patterns or experiments.
Transformation	Data-Driven/ Model Free/ Exploratory Used with FSL package and based on melodic IC analysis	It is based on linear mixing and is unordered. Thus, it must deal with independent data.
Clustering	Data-Driven/ Model-Free/ Exploratory RGNG is an example which were used in this thesis	It can define active zones and identify structures in the brain and fMRI data competently without the need for previous knowledge about activation patterns or experiments.

3.4 Literature Review

In this section a brief of literature review is introduced about several researchers who worked in fMRI researches starting from the first of using BOLD signals until to

using parcellation and clustering approaches. The following literature review can be summarized in a simple flowchart as shown in Figure 3.3.

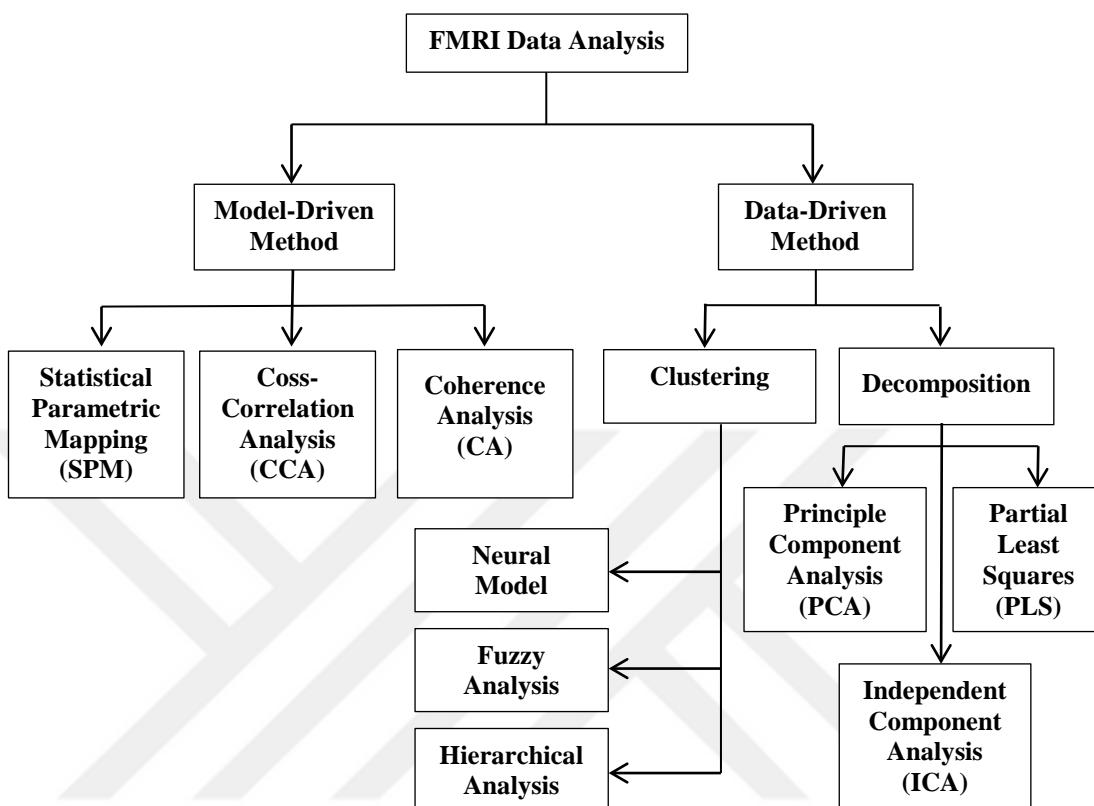


Figure 3.3 FMRI data analysis in literature.

3.4.1 Model-Driven Methods

FMRI has provided neuroscientists with a powerful tool to examine brain activity by measuring the levels of oxygen in the blood or BOLD signals which are regarded as an indirect approach for neural activity measurement. There are three groups published the first BOLD fMRI studies in 1992. Kwong et al. (1992) introduced study related to the activity in the human primary visual (V1) and motor (M1) cortex using 1.5 T MRI. Brain activation was obtained by visual stimulation and hand squeezing; the results showed that with both areas changes with MR signal agree with the corresponding stimulation [30]. Ogawa et al. (1992) introduced a similar experiment to evaluate changes in gradient-echo planar imaging (EPI) resulting from longer visual stimuli using 4T MRI [31]. The results founded the BOLD signal production by T2* effects through using different image-acquisition TE. Bandettini et

al. (1992) introduced study related to a motor task by asking the subjects to touch each finger to thumb repetitively; the results showed that the local signal increase of $4.3 \pm 0.3\%$ in the human primary motor cortex [32].

Due to the complication in fMRI acquisition and a low SNR associated with the BOLD signals, then a preprocessing step is needed for fMRI data analysis. Several projects can be initiated relating to fMRI data analysis. Friston K. j. et al. (1994) introduced one of the most popular model-driven methods for time-series fMRI data analysis using GLM method [33]. This method used for detecting correlations between sensory input and the brain's physiological response measured with fMRI. Aguirre et al., (1998) [34] and Handwerker et al. (2004) [35] proved that the BOLD responses of the human brain can vary across subjects, different regions in the brain, trials, days, and even their effects on statistical analyses. Cao and Worsley (1999) proposed two new types of random field using the cross correlation field [36]. The cross correlation field is the usual sample correlation coefficient for a set of pairs of Gaussian random fields. The output results are derived from the geometry of the deflection set which are used to detect the regions of high correlation human brain activity.

Due to the variability of BOLD signal, several methods have been proposed to overcome these influence in model-driven study. Woolrich et al. (2004) proposed different constrained linear basis sets for the HRF modeling. The HRF basis set has been defined and modeled the BOLD signals from different subjects with different HRFs from the predefined basis set [37].

3.4.2 Data-Driven Methods

There are other fMRI data analyses methods do not depend on the shapes of BOLD signal. Data-driven analysis is widely used for fMRI data processing in which the detection of brain activation is obtained from the information of the fMRI signal only.

3.4.2.1 Decomposition Exploratory Approaches

Backfrieder et al. (1996) used the decomposition PCA with visual and motor stimulation experiments for fMRI data analysis; they showed that their method yields accurate absolute quantification of in vivo brain activity [38]. Goutte et al. (1999) [39] and Gao and Yee (2003) [40] used Temporal Clustering Analysis (TCA) in fMRI time series and successfully extracted the main components of the responses and brain activation maps when the timing and location of the activation are completely unknown. McIntosh et al. (2004) proposed an effective multivariate analytic tool Partial Least Squares (PLS) for the brain activity detection [41]. The results showed that this method provided a robust statistical assessment by using event-related fMRI data without making assumptions about the shape of the HRFs. Beckmann and Smith (2004) [42], Yi-Ou et al. (2007) [43] and Wang and Peterson (2008) [44] used the decomposition ICA technique for fMRI time series and they successfully extracted the main components of the responses.

Parcellation method is another approach that used for fMRI data analysis; it is used to overcome the mis-registration problem and dealing with the limitation of spatial normalization. As it is explained previously that there are either anatomical or functional parcellation, the most important area of neuroscience is the functionally parcellation of the human cerebral cortex. Parcellation of the human brain has been done by Brodmann in early 20th century (Figure 4.6).

Flandin et al. (2002) used a brain parcellation technique to overcome the shortcomings of spatial normalization for model-driven fMRI data analysis [45]. By using the GLM parameters and group analysis with anatomical T1 MR Slice, they parcellate the brain of each subject into about 1000 homogenous parcels functionally. Thirion et al. (2006) used a multi-subject whole brain parcellation technique to overcome the shortcomings of spatial normalization of fMRI data set [46]. Using the GLM parameter analysis, they parcellate the whole brain into a certain number of

parcels. They collected voxels from all subjects together, and then derived parcel prototypes by using C-means clustering algorithm on GLM parameters.

Data-driven analysis is widely used for fMRI data processing to overcome the limitation associated with the shape of the HRF and task-related signal changes that must be introduced. As well as the assumption about the shapes of BOLD model, there is another limitation related to the subject behavior during the task. So data-driven analysis is used with parcellation technique for fMRI data processing, in which the detection of brain activation is obtained from the information of the fMRI signal only.

Yongnan Ji et al. (2009) introduced a parcellation approach for fMRI dataset based on ICA and PLS instead of the GLM, and they used a spectral clustering of the PLS latent variables to parcellate all subjects data [47]. Thomas et al. (2013) proposed a novel computational strategy to divide the cerebral cortex into disjoint, spatially neighboring and functionally homogeneous parcels using hierarchical clustering parcellation of the brain with rs-fMRI [48]. Thirion B. et al. (2014) studied the criteria accuracy of fit and reproducibility of the parcellation across boot strap samples on both simulated and two tasks-based fMRI dataset for the Ward, spectral and k-means clustering techniques [49]. They addressed the question of which clustering technique is appropriate and how to optimize the corresponding model. The experimental results showed that Ward's clustering performance was the best among the alternative clustering methods. ICA and PCA techniques regarded as fine methods to separate the fMRI signals into a group of the defined components, but each faced difficulty to predict what occurs during acquisition and must deal with the limitations of their independence and orthogonality respectively [50].

3.4.2.2 Clustering Exploratory Approaches

In fMRI researches, clustering techniques are considerable as a model-free or exploratory data analysis approaches and have the ability to define the active zones and find structures in the brain and fMRI data competently without need of prior

knowledge about activation patterns or experiment. But till now the problem of choosing appropriate clustering techniques is still existent.

Evgenia et al. (2003) compared NG method to its corresponding winner-takes-all method (Hard Competitive Learning (hardcl)) and winner-takes-all batch method (k-means); k-means was used widely in fMRI clustering techniques where the update of only the winning center is done before any adaptations [51]. The obtained results showed that the NG performance was the best in comparison to its competitor. Meyer-Baese et al. (2004) compared two exploratory methods that widely used for fMRI data analysis: the ICA techniques versus unsupervised clustering [52]. A comparative quantitative evaluation between three methods of unsupervised clustering (SOM, NG network and fuzzy clustering) versus three methods of ICA techniques (FastICA, Infomax and topographic ICA). The experimental results showed that the unsupervised clustering techniques have the ability to extract features for a small number of ICs fully but are restricted to the linear mixture assumption. Also, they performed better than ICA but require more time for processing rather than the ICA methods. Wismuller et al. (2004) compared NG method with Kohonen's self-organizing map (SOM) and a fuzzy clustering scheme based on deterministic annealing when applied to fMRI studies [53]. The experimental results showed that both NG network and fuzzy clustering technique based on deterministic are better than SOM in terms of identifying signal components with high correlation to the fMRI stimulus (found component TC, quantization error and activation maps). Also the results showed the NG network performance is the best among the two other clustering techniques about the quantization error and the identification of signal components with high correlation to the stimulus function. Dimitriadou et al. (2004) compared the efficiency and power of several cluster analysis techniques as crisp (NG, SOM, k-means, hardcl, maximum distance and CLARA), hierarchical and fuzzy (c-means, fuzzy competitive learning) techniques for fully artificial (mathematical) and synthesized (hybrid) fMRI data set [54]. They compared these methods using two performance measures, correlation coefficient and the weighted Jaccard Coefficient (wJC). The experimental results showed that NG and the k-means technique performance was the best among all other methods. Ana et al.

(2006) found experimentally that One-Pass realizations of SOM and NG provided good results with respect to distortion or even improved over the slower realizations [55]. According to the experimental results, they proposed that the appropriate setting for the convergence analysis of SOM, NG and similar competitive clustering algorithms is the field of Graduated Nonconvexity algorithms and they can be set in this framework easily.

Lachiche et al. (2005) introduced a new interactive data mining approach to fMRI images which was not been used for that purpose at that time and showed that the GNG has successfully recognized the active areas in brain fMRI images [56]. The idea of defining a distance between voxels of fMRI images was argued and suggested that this distance should be based on the signal only. Korczak (2007) introduced a new interactive data mining technique to fMRI images for cerebral activity observation which based on data-driven approach [57]. Different unsupervised clustering techniques were presented, developed and tested on sequences of fMRI images. Five clustering techniques applied to synthetic and real data are GNG, Kohonen's SOM, Linde-Buzo-Gray (LBG), K-means and Clustering Using Representatives (CURE). The experimental results showed that the GNG technique performance was the best among all other clustering methods with acceptable robustness. Heydar et al. (2009) developed the technique of the GNG network, which can run the optimal number of clusters automatically [58]. The experimental results used artificial and real fMRI dataset with the proposed algorithm which is an improved version of the GNG algorithm. They compared the Jaccard coefficient of the proposed technique with some well-known clustering techniques such as K-means, NG, GNG and Fuzzy C-Means (FCM); the results showed that the performance of the proposed technique is better than others.

From all the presented different techniques previously, it is evaluated the GNG technique as the best clustering performance and give an acceptable robustness [50], but also has some limitations associated with the sensitivity for initialization, the order of input vectors and existence of many outliers [59]. Different from the works

above, this thesis focuses on using the RGNG algorithm into fMRI data to detect the active areas in the brain.

This work related to using RGNG with fMRI dataset will be the first time in the literature, so there is no such study has been done according to the researcher knowledge. The proposed technique will be tested on real and free auditory fMRI dataset which is available for education and evaluation purposes. The real auditory fMRI activation data was taken from the London research institute “Welcome Trust Centre for Neuroimaging at University College of London (UCL)”. These data were used by some works as Lachiche et al. (2005), Korczak (2007) and Heydar et al. (2009) as show in the literatures previously.

This proposal differentiates itself from the literature on the following major fronts. While the GNG is originating from the NG algorithm by Fritzke (1995, 1997), the RGNG algorithm was introduced by Qin and Suganthan (2004) within the GNG structure. RGNG network possesses better robustness properties than GNG, by succession the properties of the original GNG algorithm that overcoming the robustness issues associated with it and incorporating it with several robust strategies, such as outlier resistant scheme, adaptive modulation of learning rates and cluster repulsion method.

3.5 Summary

In summary, this chapter introduced fMRI data analysis and inference, and the most important packages used in fMRI data analysis. The human brain parcellation technique principle was presented as well as literature review about fMRI researches starting from the first of using BOLD signals until to using parcellation and clustering approaches.

The research was started with investigating the available articles concerns on fMRI data mining and analysis, which were published in scientific journals. This section is important to give a background and review of the previous work related to fMRI, who analyzed the algorithms for clustering and processing different types of data as well as different packages used for fMRI preprocessing steps.

CHAPTER 4

FMRI DATA ANALYSIS AND PREPROCESSING

Before starting with fMRI data analysis using the proposed methods in this work, fMRI dataset must be preprocessed using one of the open source packages. These steps are important for removing unwanted data or noises from the dataset, improving the fMRI mapping and preparing it for the analysis.

In this chapter, the preprocessing steps are presented using free SPM software package. The output results from each of the preprocessing steps are introduced by using real auditory fMRI dataset. Thereafter, an introduction of the conventional GLM method that is used by SPM is presented.

4.1 FMRI Dataset Types

MRI data are usually stored in a binary data file as either 8- or 16-bit integers. The size of the data file on disk is the product of the data size and the dimensions of the image. For example, a standard MRI image (with dimensions of $64 \times 64 \times 32$ voxels) requires 256 kilobytes of disk space when stored as a 16-bit integer, but 1,024 kilobytes (one megabyte) when stored as a 64-bit floating point value.

Two-dimensional structural MR images after the reconstruction from the raw data are stored and manipulated as a slice. Each slice represents a grid or matrix of points, and these matrices represent the pixels with M rows and N columns. Several slices are required to produce a 3D volume image for creating a three-dimensional matrix with size $M \times N \times K$.

fMRI generates four dimensional (4D) image datasets which are sequences of 3D volumetric images obtained over a period of time. 4D medical images represent changes of 3D medical images over the fourth dimension (time) as shown in Figure 4.1.

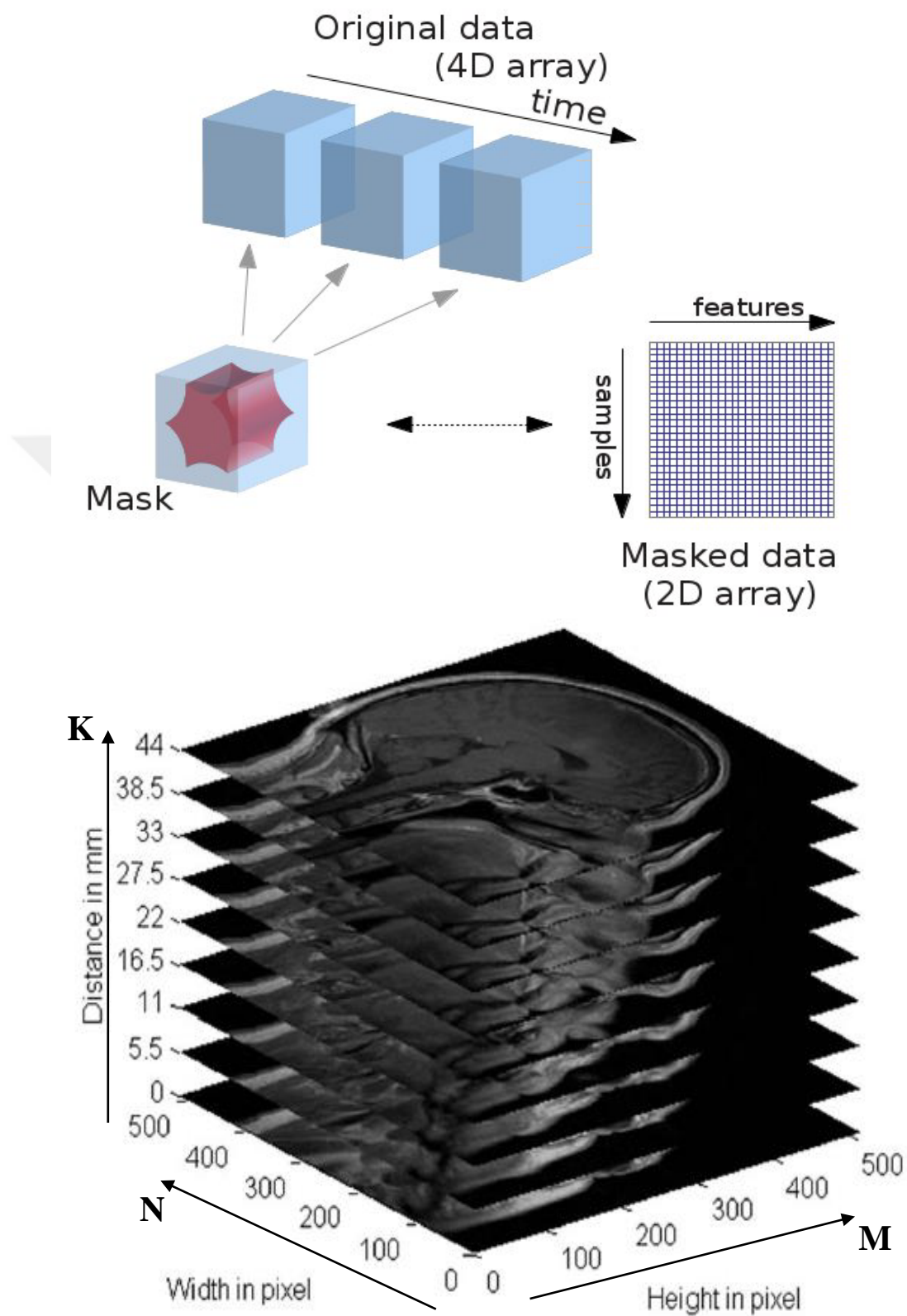


Figure 4.1 Original fMRI 4D data and MR image sequences with 5.5 mm spacing between slice. [10]

Generally MRI images are stored as 3D data files, but fMRI data are collected as a series of images, so they can be stored either as a set of 3D files or as a single 4D file where the fourth dimension is the time. By using four-dimensional file, the total files number is minimized, so it is preferred to store data as four-dimensional if possible. The problem in using 4D files is the impossibility of analyzing easily, and not all analysis packages can handle with 4D files.

In neuroimaging, there is a different large number of medical image formats and the most important three of them are described in Table 4.1. The most typical output data format for MRI scanners is the digital imaging and communications in medicine (DICOM)-files format where every pixel corresponds to a voxel and their information is codified in the digital image format. DICOM is considered as a standard of medical images and it is a useful format that stores MR image as well as the information related to the position and orientation of patient and MR scan image.

Another output MRI format is ANALYZE format, which consists of two types of files with extensions image data file (.img) and header information file (.hdr). An image file contains either a set of cross sectional images or uncompressed pixel data, while the header file contains a history and dimensions of the data as well as information about the subject, type of image, imaging parameters, image dimensions and so on. Neuroimaging Informatics Technology Initiative (NIFTI) consists of one or two types of files with extensions single file (.nii) or dual file (.img and .hdr).

The typical fMRI dataset can be saved in different file formats (e.g., DICOM, ANALYZE, NIFTI) according to the scanner type and the software package used for data analysis. Each of the file format used for storing MRI data and the data can be converted from one format to another using different conversion techniques which are available in most analysis software.

Table 4.1 Overview of some medical image data formats.

Format Name	File Extension	Origin
DICOM	None	ACR/NEMA consortium
Analyze	.img and .hdr	Analyze software, Mayo Clinic
NIfTI	.nii or .img and .hdr	NIH Neuroimaging Informatics Tools Initiative

4.2 Image Spatial Preprocessing

With different experimental conditions, the acquired fMRI data are formed as a combination of the BOLD signal changes and noises or artifact. These artifacts are caused mainly due to some reasons such that related to hardware system (MRI scanner itself), individuals themselves (e.g. head motion) or physiological effects. So and in order to prepare fMRI raw data for analysis, a number of preprocessing operations steps are required before extracting the data of interest from it. The aim of these preprocessing operations steps is to detect and repair artifacts as well as preparing data for the later statistical functions assumptions.

As it was explained previously, there are several software packages (SPM, FSL, AFNI, Brain Voyager) used for processing and analysis of fMRI data. Some freely available are SPM and FSL, while others are not. SPM is one of a free and open source package that runs within the MATLAB environment. It was developed at the functional imaging laboratory (FIL) at University College London, by Karl Friston who is a team leader [19].

The presented work use SPM for analyzing a free auditory fMRI data [60] as an example for introducing the spatial pre-processing stages. Auditory fMRI data set that used in this work is comprises whole brain BOLD/EPI images acquired on a modified 2T Siemens MAGNETOM Vision system [61]. Each acquisition consisted

of 64 contiguous slices (64X64X64 3X3X3mm voxels). Acquisition took 6.05s, with the scan to scan TR set arbitrarily to 7s. 96 acquisitions were made (TR=7s) from a single subject, in blocks of 6 scans (acquired during the same condition as a stimulation or rest), giving 16 blocks and each block for 42s. The condition for successive blocks is alternated between rest and auditory stimulation, starting with rest. The functional data starts at acquisition 4, functional image (fM4). Auditory stimulation was bi-syllabic words (e.g. “mother”, “house”, “weather”, “movie” etc.) presented binaurally at a rate of 60 per minute. Due to T1 effects it is advisable to discard the first few scans (there were no “dummy” lead in scans).

FMRI dataset is preprocessed by applying the following steps:

- Realignment
- Coregistration
- Segmentation
- Normalise
- Smoothing

The most important tools used with the preprocessing steps are shown in the SPM menu window in Figure 4.2 which described most of its stages in the following sections. Results output from the preprocessing steps are described in each section below.

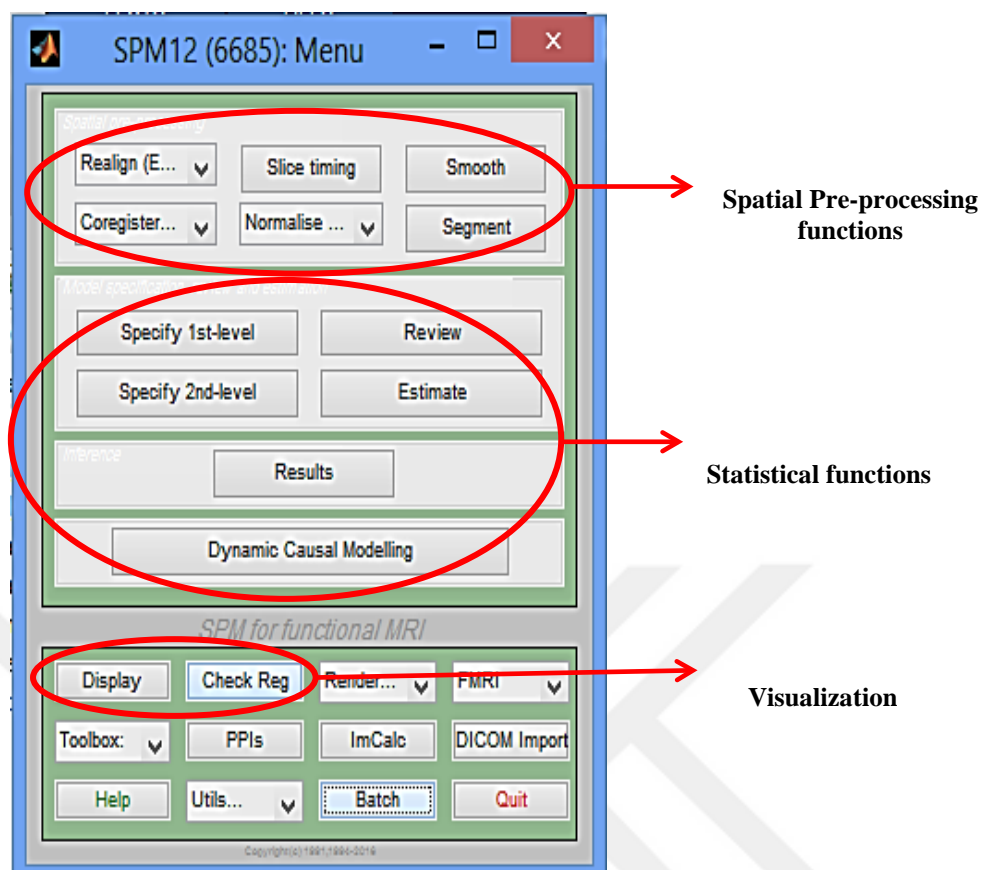


Figure 4.2 The SPM menu window.

4.2.1 Realignment

In the previous chapter, an example of a head motion was given. The problem of head movement during scan time can be reduced by minimizing the misalignment between images in an fMRI time series using a rigid body transformation [11] and realigning the images to a single reference image. The realignment procedure includes basically motion estimation (trial and error) and correction through the translations or moving the image in X, Y, or Z direction and rotations over the X, Y, and Z axis. Realignment Result for the auditory raw data with 96 functional images (fM) is shown in Figure 4.3. Top panels plot translation between timepoints and the bottom panels plot rotation after motion correction. These plots reflect the parameters of the rigid body transformation that are estimated for each timepoint in comparison to the reference image. The zero parameters refer to the matching of the reference image to itself exactly.

Image realignment

```

1 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
2 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
3 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
4 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
5 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
6 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
7 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
8 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
9 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory sr
10 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory s
11 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory s
12 C:\Users\hadeel\Desktop\PhD. 10th semester 2-2017\Thesis Work\auditory s
..... etc

```

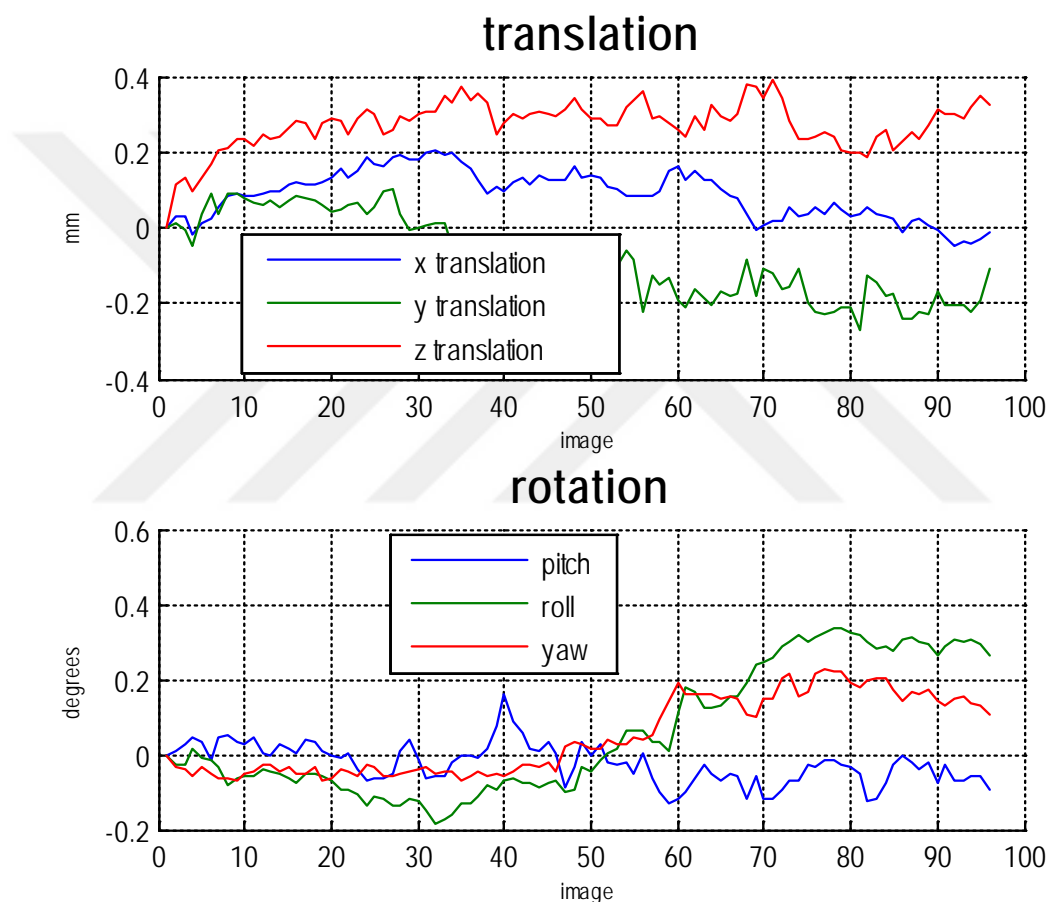


Figure 4.3 Realignment Result for 96 fM free auditory data.

4.2.2 Coregistration

Coregistration is used to align functional images (T2* weighted) with the anatomical (structural) MRI images (T1 or T2 weighted) of the same subject. This process is based on using different cost function which is called Mutual Information [61]. SPM

implements the coregistration by maximizing the mutual information. Figure 4.4 shows the coregistration between the reference functional mean image created from the realignment running and the structural image.

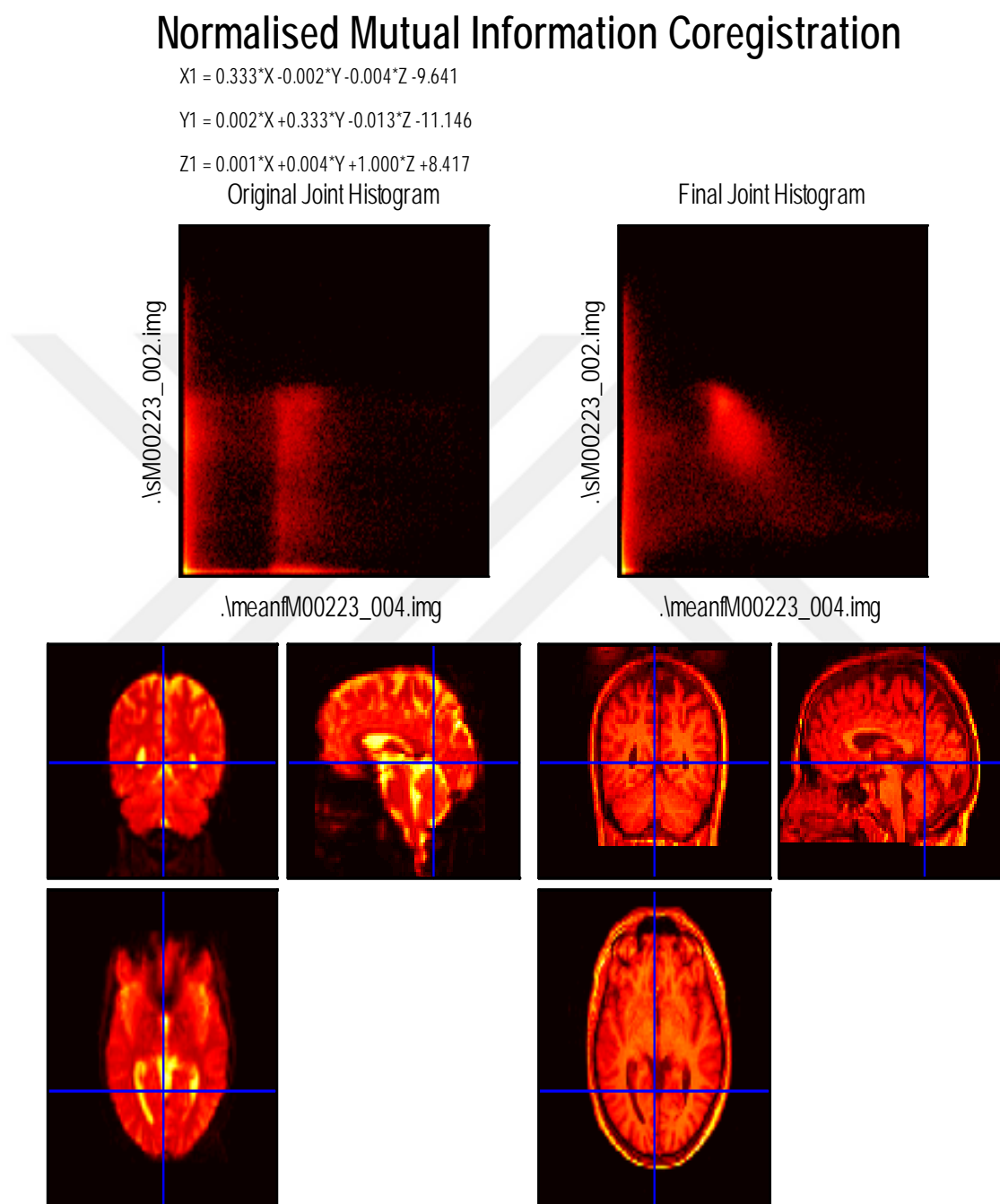


Figure 4.4 Mutual information coregistration output result (mean fM4 and sM).

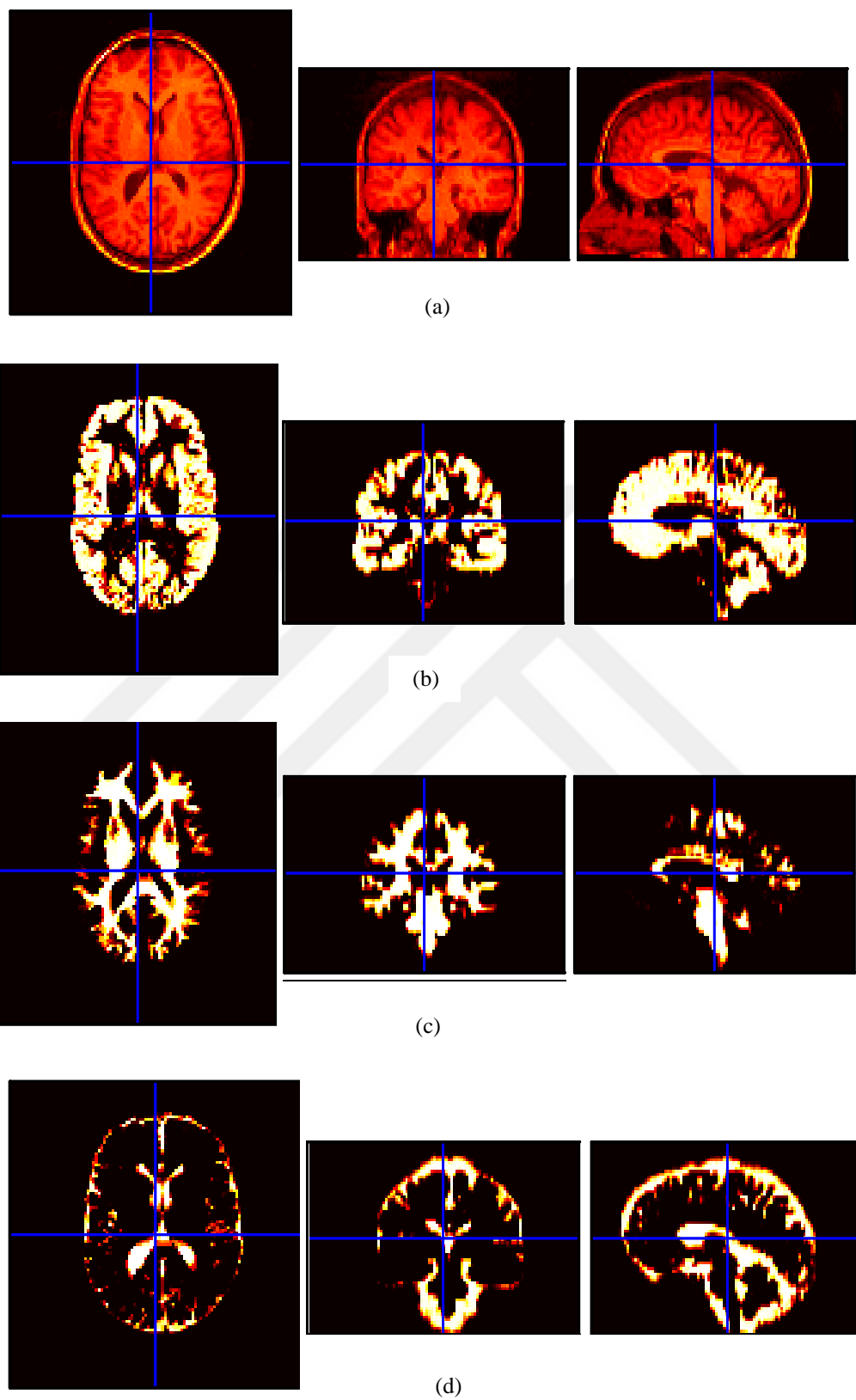


Figure 4.5 Segmentation results, (a) original image (sM); (b) segmented grey matter (c1); (c) white matter (c2); (d) CSF (c3).

4.2.3 Segmentation

Segmentation process is performed to view grey matter, white matter, or CSF separately for each scan. From the original image, segmenting scan results produce three types of used and useful images for doctors and researchers. Figure 4.5 shows the segmentation results of the structural image. This figure shows the original sM in part (a), while part (b) shows the segmented grey matter (c1) result, part (c) shows the segmented white matter (c2) result and part (d) shows the CSF (c3) segmentation result.

4.2.4 Normalization

Normalization is a procedure used to put functional data into a coordinate system or standardized Montreal Neurological Institute (MNI) space. MNI templates are standard brains by using a large series of MRI scans on normal controls; which are developed at the Montreal Neurological Institute. MNI-152 is defined from the average of 152 subjects as an alternative method of a single brain Talairach atlas which is still common to use in functional brain imaging studies. Talairach brain atlas is a 3-dimensional coordinate system of the brain anatomy. The atlas has Brodmann's areas classification in a rather approximate way of the brain. Parcellation of the human brain has been done by Brodmann in early 20th century, based on the brain cytoarchitecture and divided it into 52 different fields as shown in Figure 4.6.



Figure 4.6 The Brodmann areas 3D.

Figure 4.7 shows MNI versus Talairach brain atlas. This picture shows the top of the brain is higher in MNI, and the temporal lobes are so lower and larger than for the Talairach brain.

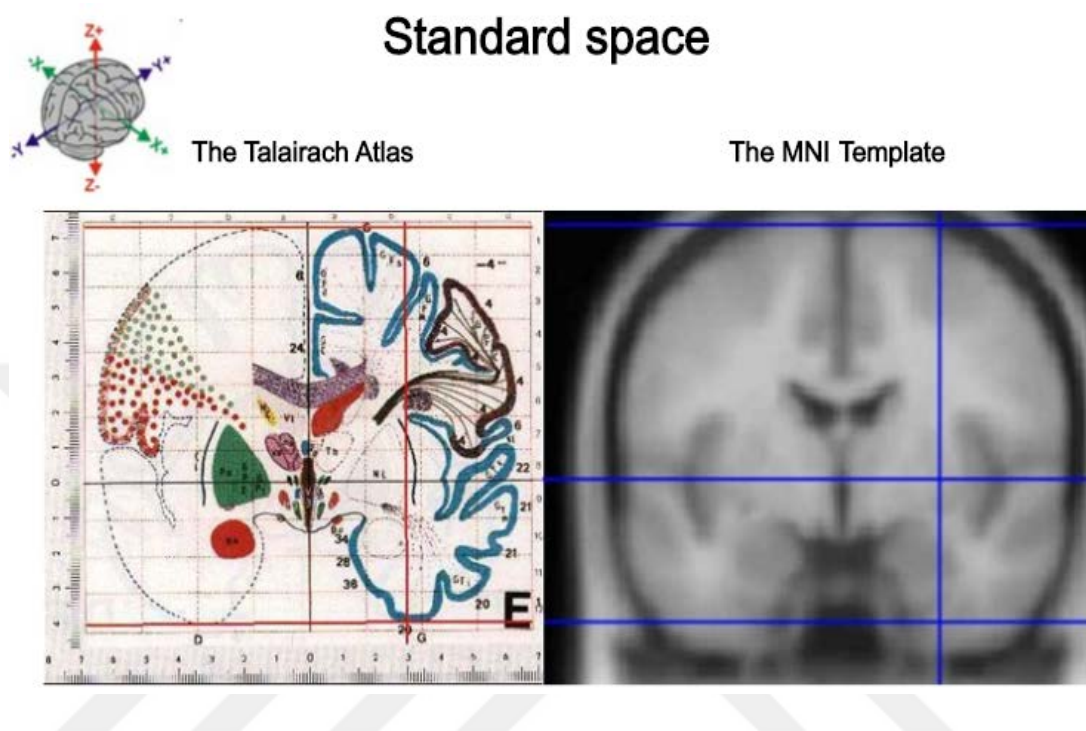
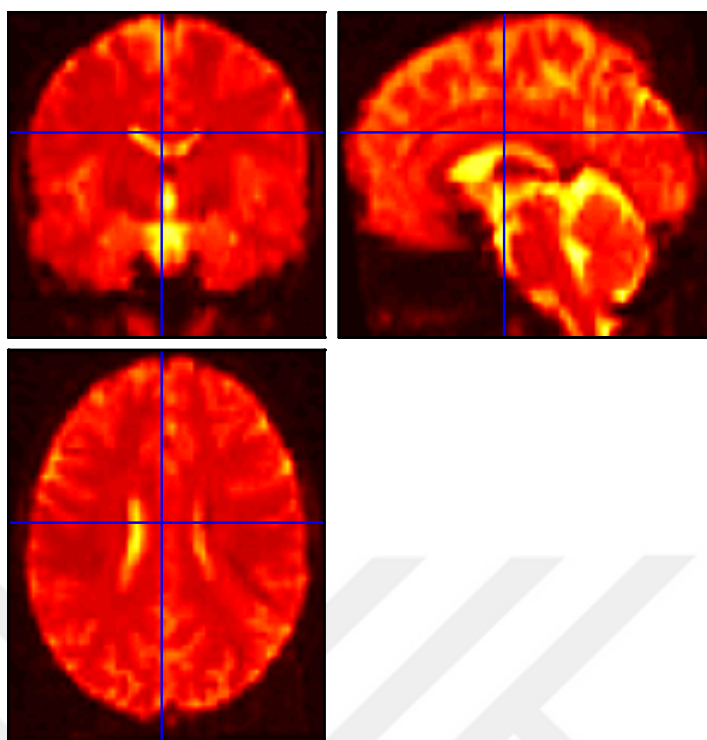
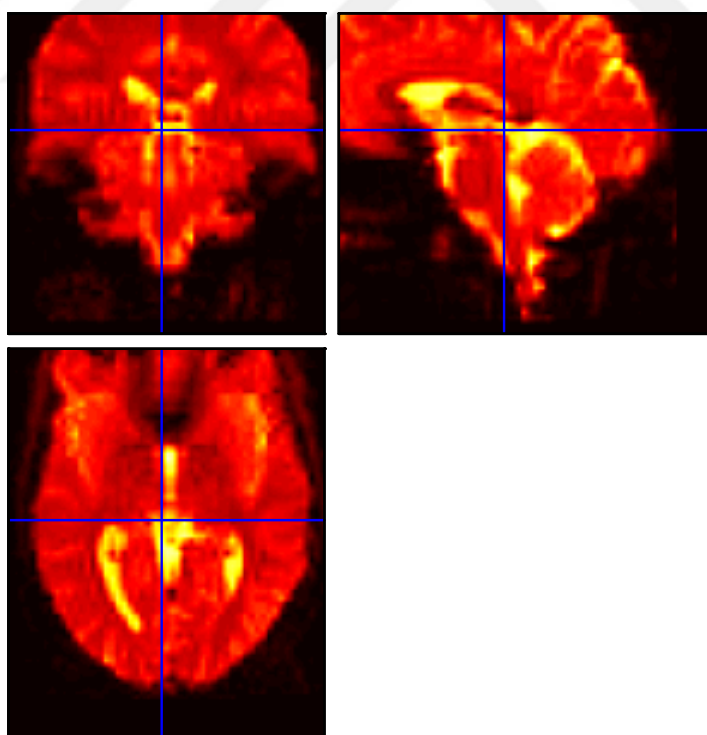


Figure 4.7 MNI versus Talairach brain atlas

SPM uses MNI template images, which are the most common templates used for fMRI spatial normalization. Normalize function can be used to put any scan registered with the anatomical that was segmented into template images which averaged scans of multiple subjects. Figure 4.8 shows the normalization output result for the functional slice image number 20 as an example (fM20).



(a)



(b)

Figure 4.8 Normalization result (fM20), (a) after normalization; (b) before normalization.

4.2.5 Smoothing

The fMRI images are used smooth function as a final step in spatial pre-processing to correct the results of brain mapping and to increase the SNR by improving the few remaining of functional/ anatomical differences between subjects. Smoothing is achieved by averaging every voxel with a weighted sum of its neighbors (across adjacent voxels) to blur the functional images. The weighting is defined by a Gaussian kernel (Figure 4.9) with full width at half maximum (FWHM) size.

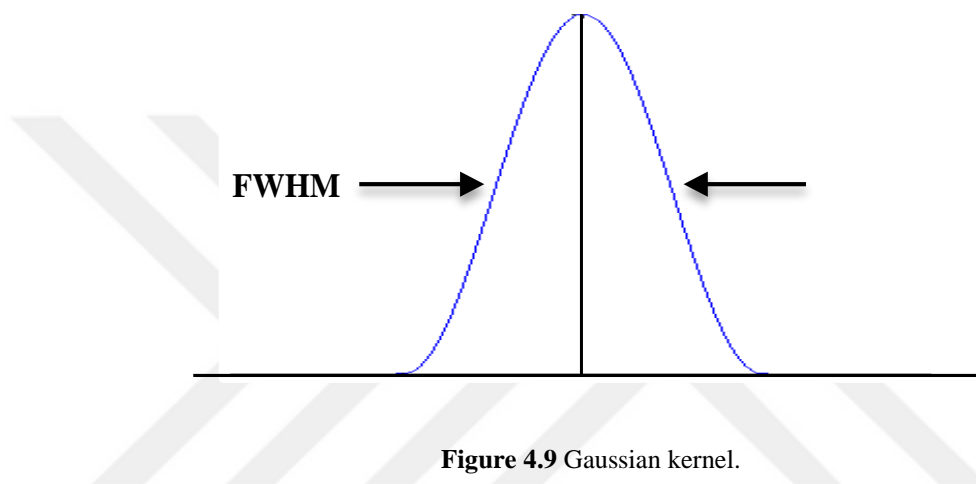
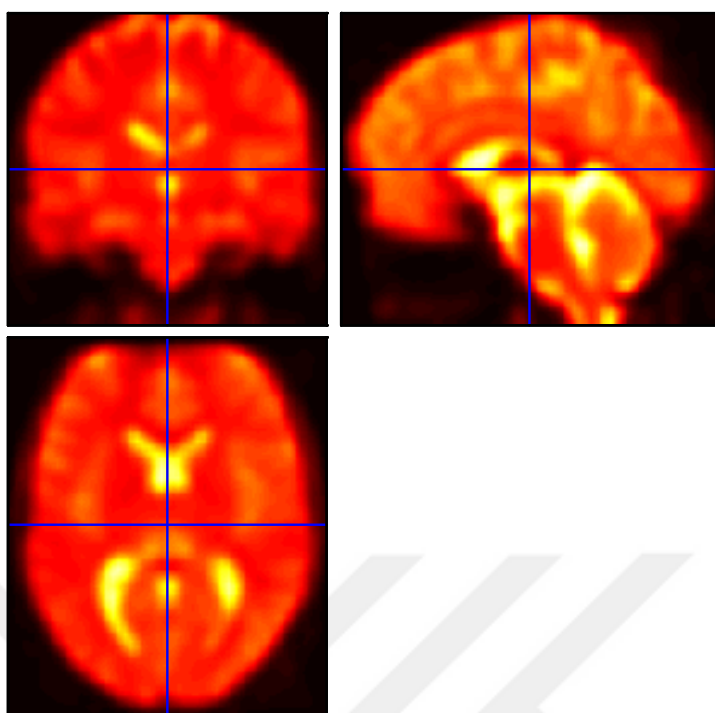


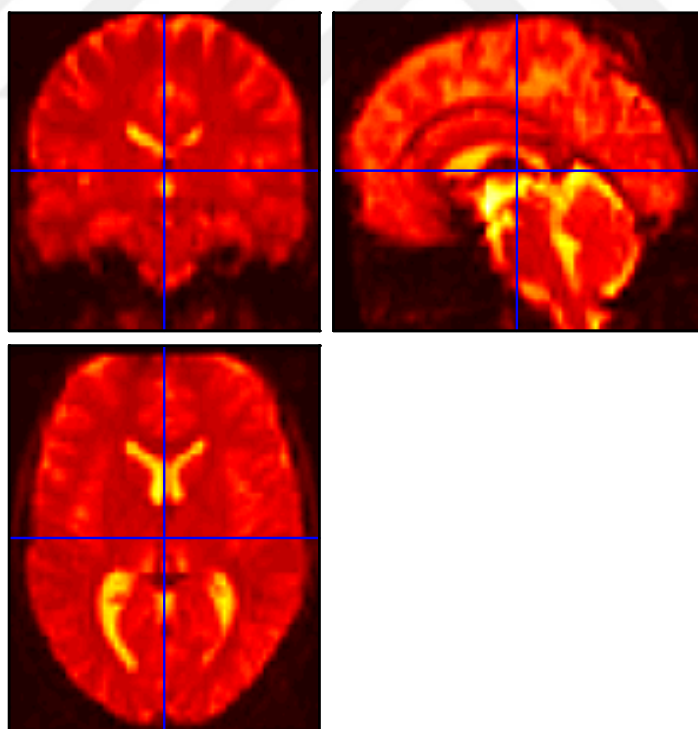
Figure 4.9 Gaussian kernel.

The FWHM is the size of the Gaussian that is given by its full width at half maximum. The larger the FWHM, the more smoothing get and vice versa. Figure 4.10 shows the output smoothing result using a Gaussian smoothing kernel of 6 for the functional slice image number 20 as an example (fM20). The above image in this figure shows the final step in the preprocessing steps with respect to the functional image without preprocessing shown down.

Another example showing the importance of the preprocessing steps is seen in Figure 4.11. This figure shows the whole brain fMRI images with clusters with and without preprocessing. So, these steps will effect on finding the true active zones with the same paradigm in the specified area of the brain. The result of fMRI image without preprocessing leads to the presence of a lot of unreal active areas in the shape of the image. These areas looked activated, but in fact are just an artifact and they are scattered all around the image and even outside the image.

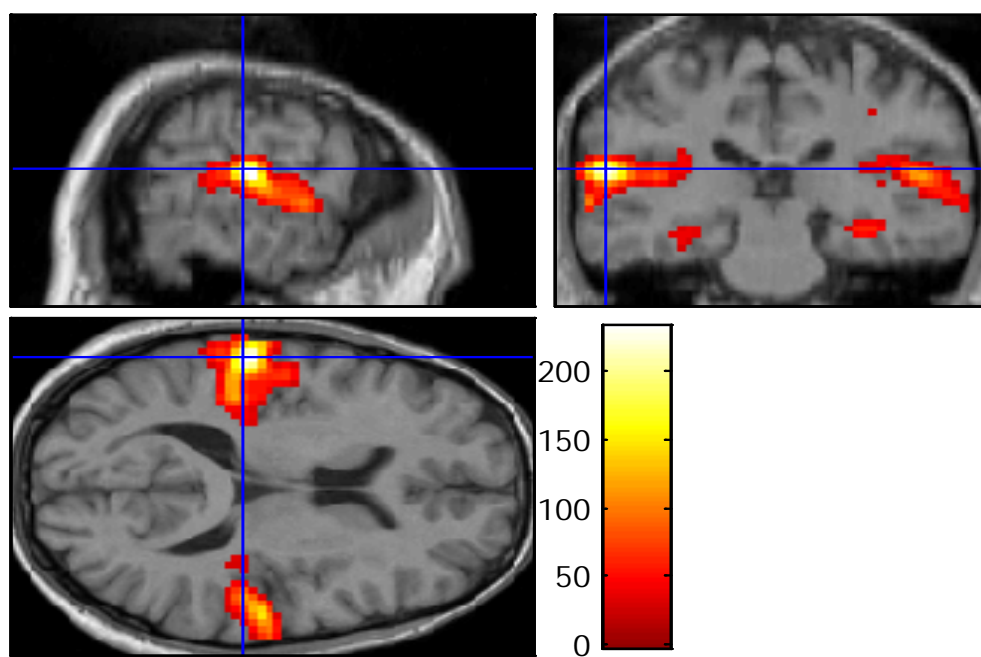


(a)

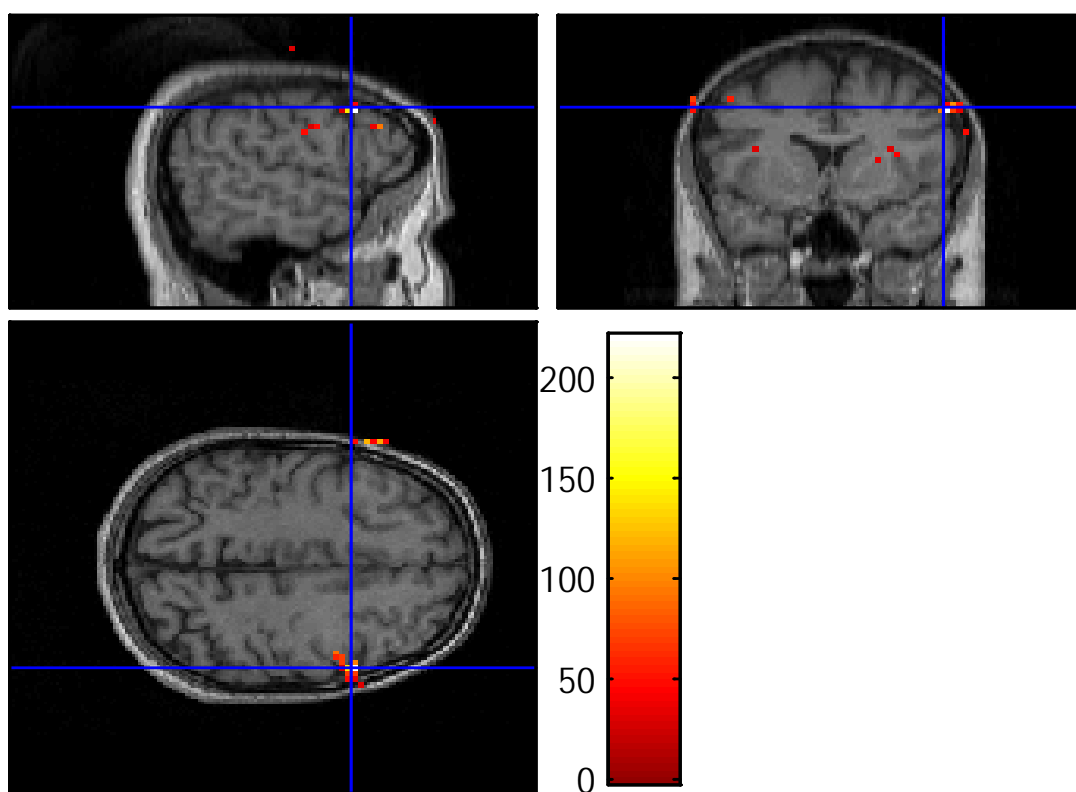


(b)

Figure 4.10 Smoothing result (fM20), (a) after smoothing; (b): before smoothing.



(a)



(b)

Figure 4.11 Whole brain fMRI image (a) image with preprocessing; (b) image without preprocessing.

In sum, the functional images were reoriented to MNI space, which is standard brain formed by using a large series of MRI scans on normal controls developed at the Montreal Neurological Institute. Then the functional raw data were realigned to correct for the head movements. The high-resolution anatomical T1 images were coregistered with the realigned functional images to strong the functional activation detection. Segmentation process is not mandatory. SPM12 uses MNI template image, which are the most common templates used for fMRI spatial normalization. In this step, the anatomical and functional images were spatially normalized into MNI space. Finally, the functional raw data were spatially smoothed with a Gaussian smoothing kernel of 6.

4.3 General Linear Model (GLM)

The GLM is one of the most important, fundamental, basic and common model-driven methods in fMRI data analysis by performing the statistical tests on each voxel. The general name in the GLM model means that this model can be used for many different statistical analysis types, as correlations, one-sample *t*-tests, two-sample *t*-tests, *F*-tests, Analysis of Variance (ANOVA), and Analysis of Covariance (ANCOVA) [11]. GLM can remove the effects that may confound with the analysis using a suitable model as a comparison to the common *t*-test and correlation analysis.

GLM is a regression method, so firstly constitute a statistical model to describe the BOLD signals corresponding to the stimulation, and then apply the model to fMRI data. GLM statistical analysis is implemented on each voxel providing a statistical measure of the brain structure corresponding to that voxel is activated during the stimulation. While GLM statistical analysis using null hypothesis is performed when the model does not match the TC of that voxel. The term "null hypothesis" usually refers to a general statement or default position that there is no relationship between two measured phenomena, or no difference among groups [62].

In GLM-based analysis, the BOLD responses' shapes are assumed to be the same for all subjects and voxels and any contrast of the BOLD responses is neglected. GLM analysis in terms of an fMRI experiment can be expressed as:

$$Y = X\beta + \varepsilon \quad \varepsilon \sim N(0, V) \quad (4.1)$$

where:

Y : measured response in one voxel of the fMRI data set, i.e. TC of the voxel we want to analyze (voxel time series)

X : design matrix (expected BOLD response)

β : vector parameters (unknown)

ε : residual errors

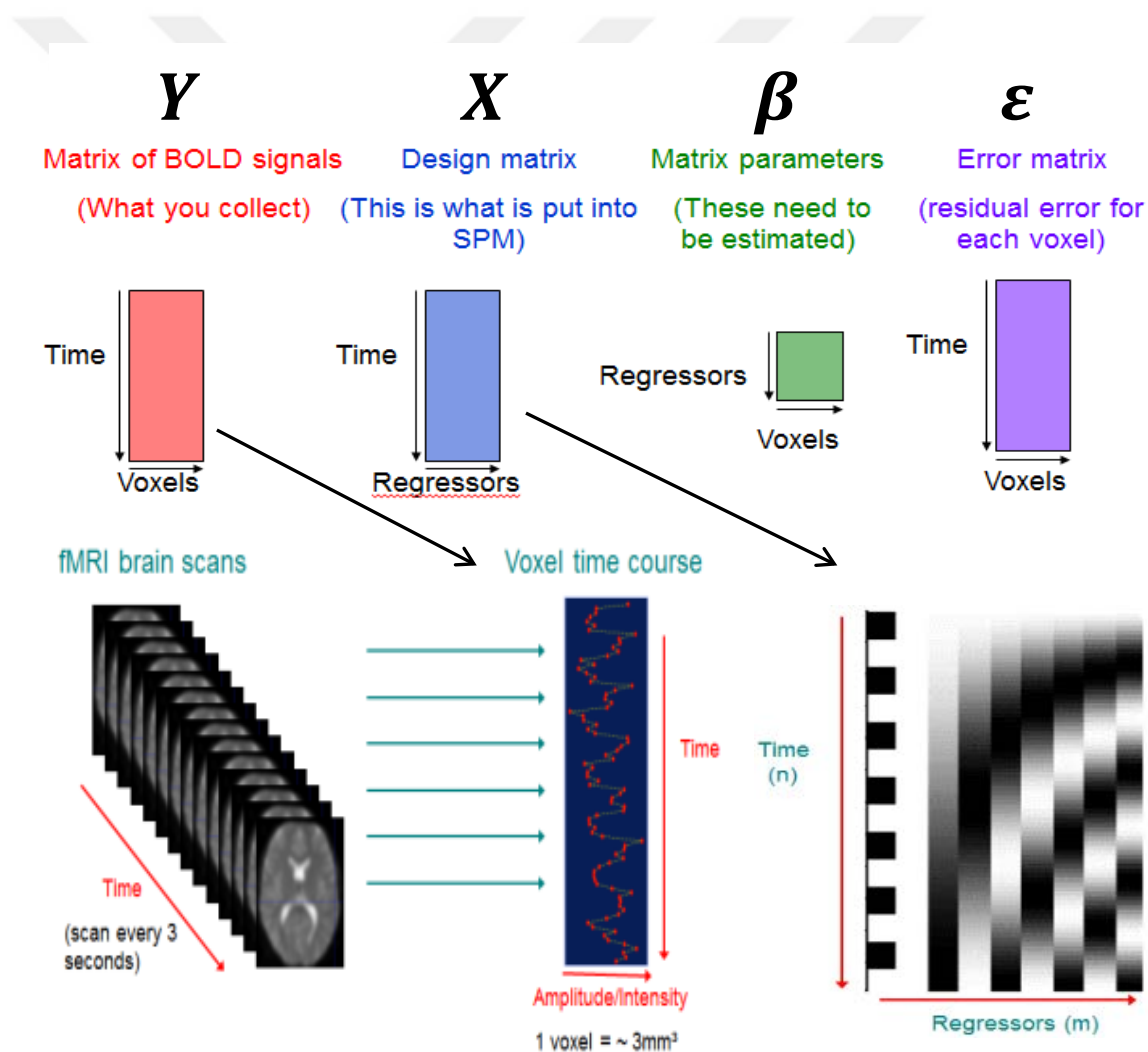


Figure 4.12 GLM analyses in terms of an fMRI experiment.

Figure 4.12 shows the above equation in more details. The design matrix X of the BOLD response contains the predictor variables that represent the experimental conditions under which the observations were made. Each column of the design matrix represents the effect of the experience or an effect that may confound the results, while each row of the design matrix represents a different fMRI scan.

There are different models for the BOLD response for obtaining the explanatory variables or predictors using: Boxcar model, Convolutional model, Balloon model and Subspace model. A boxcar model is used for the GLM analysis in SPM package.

Box-car model is the simplest BOLD response model by using a square-wave function only, to represent the observed BOLD response to be applicable for the block design, and ignores how the system (brain and MR-scanner) modulates the applied experimental paradigm. Box car function shown in Figure 4.13 as an example of box-car model of the neural activation that could be expected from a task-responsive voxel in a block design experiment with 250sec stimulation blocks and 250sec rest blocks.

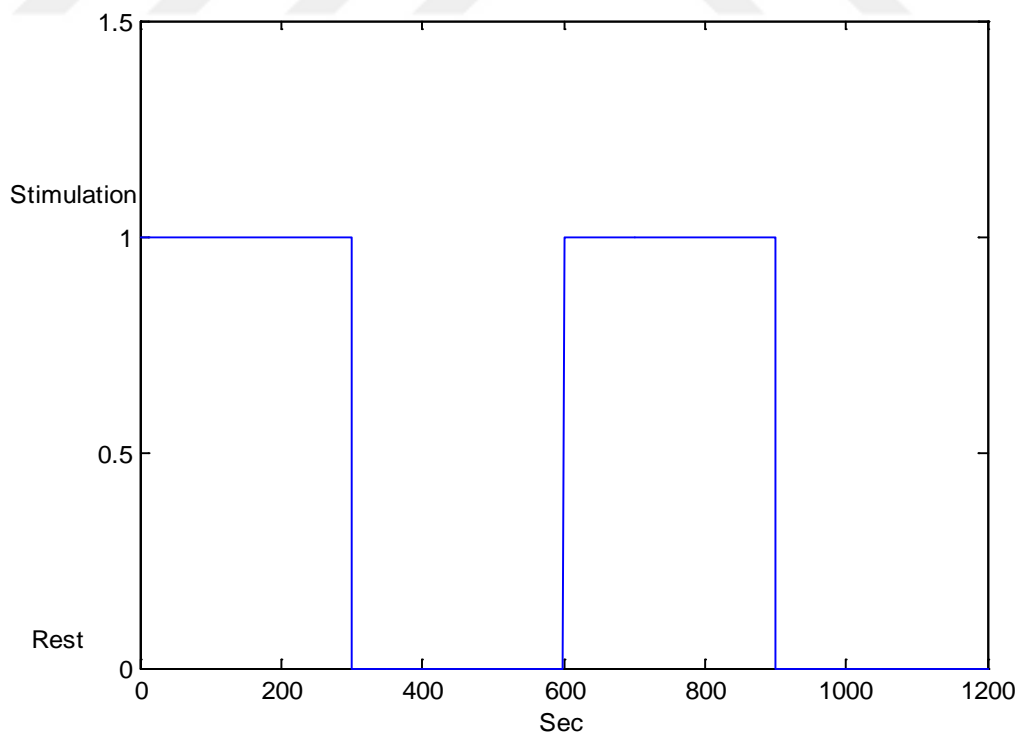


Figure 4.13 A boxcar function.

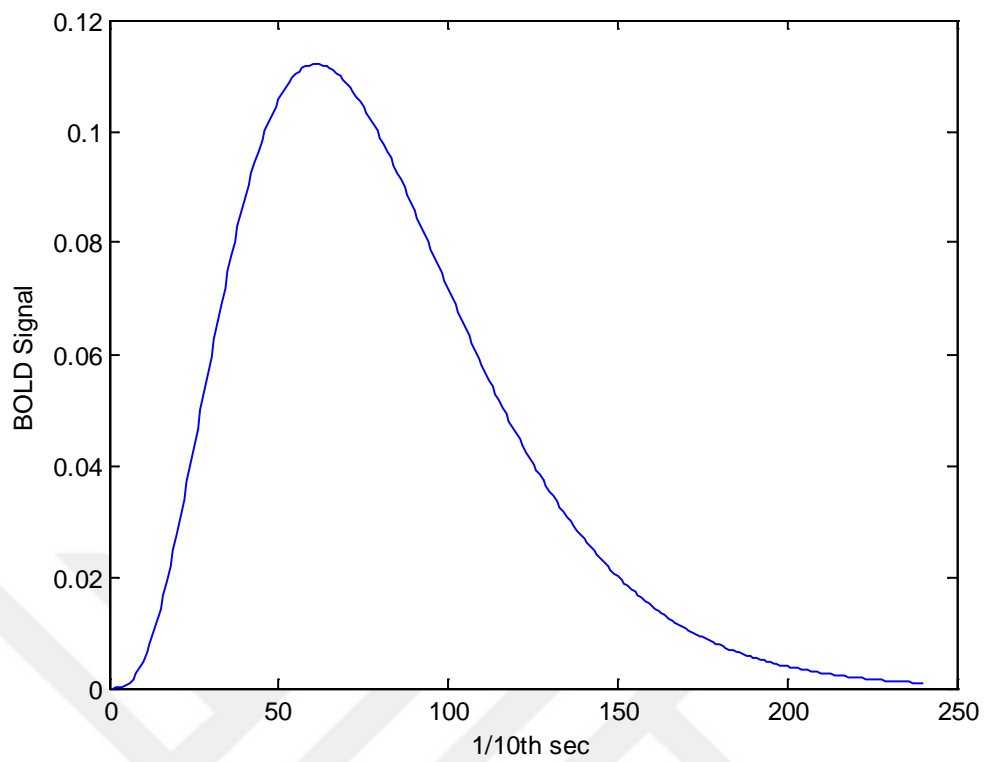


Figure 4.14 The HRF

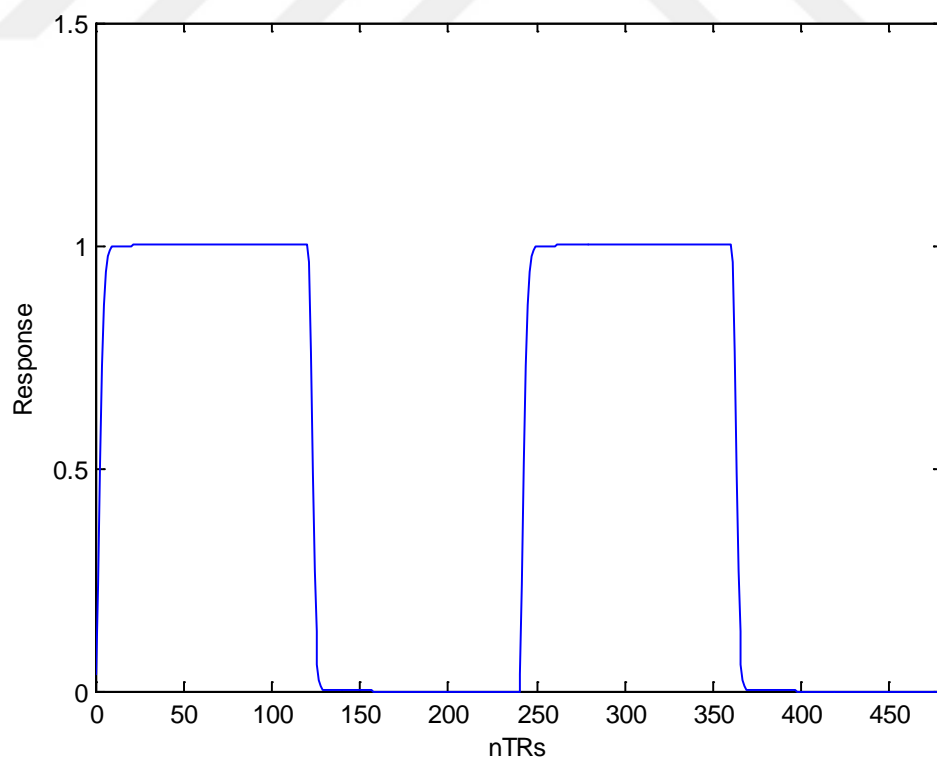


Figure 4.15 Box car model function convolved with HRF.

Figure 4.14 shows an example of the BOLD HRF or gamma function; x -axis defines in 1/10th sec for smooth plot. The box car model function is convolved with HRF to create a convolved regressor as shown in Figure 4.15; number of time points in the fMRI TC (nTRs) is 480sec. These regressors represent and match the rise and fall in BOLD signal (greyscale).

Figure 4.16 shows the generated design matrix with three columns. First column represents the on/off stimulus result of convolution the HRF and boxcar. Second column represents the baseline condition (constant or rest term); While third column represents linear drift (typical scan artifact).

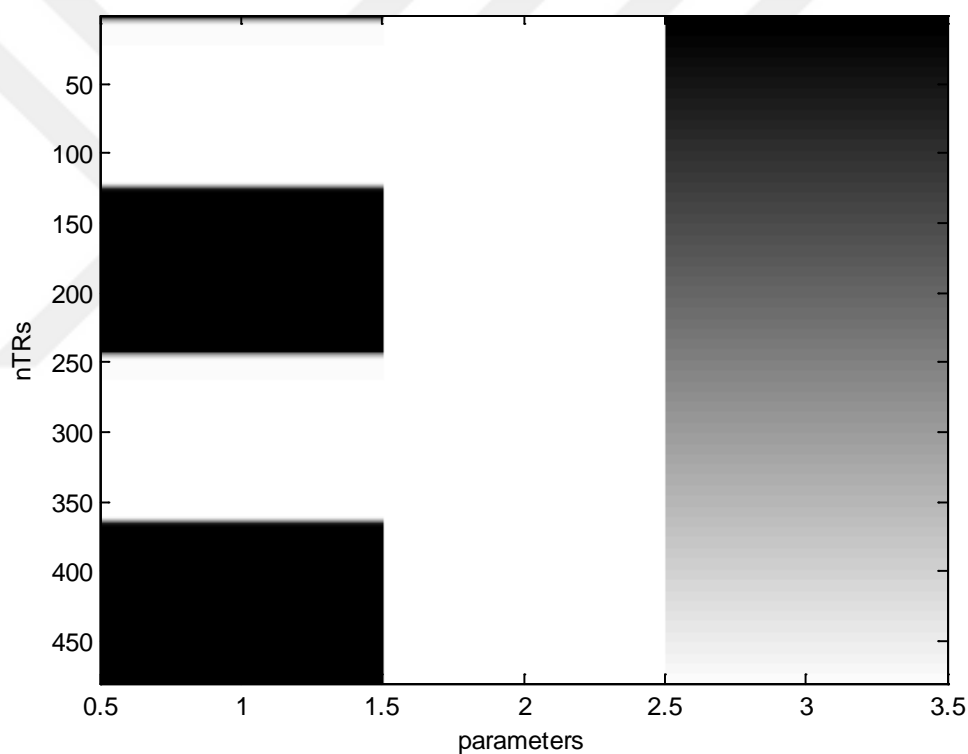


Figure 4.16 The design matrix.

A hypothesis BOLD data can be generated by adding baseline activation, linear scanner drift and noise to the output results obtained in Figure 4.15 which represents the convolution of a boxcar function with HRF, as shown in Figure 4.17. Finally the predicted BOLD signal by the GLM is plotted in Figure 4.18 as a near square with respect to the observed BOLD response. Regarding to auditory fMRI dataset used in

this work, Figure 4.19 shows the experimental paradigm for 96 acquisitions were made (TR=7s) in blocks of 6, giving 16 * 42s block.

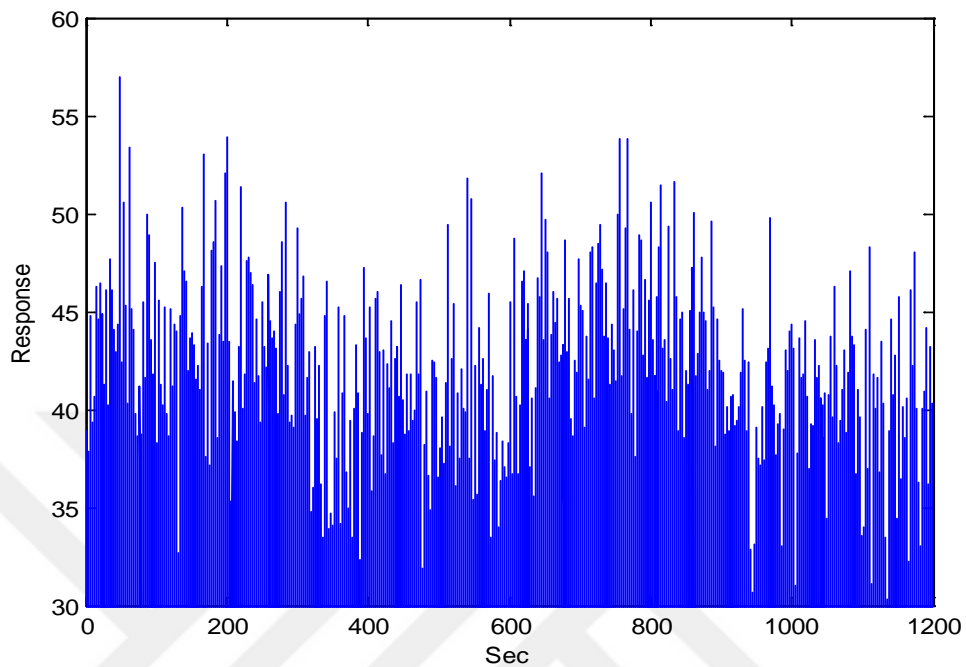


Figure 4.17 A hypothesis BOLD data.

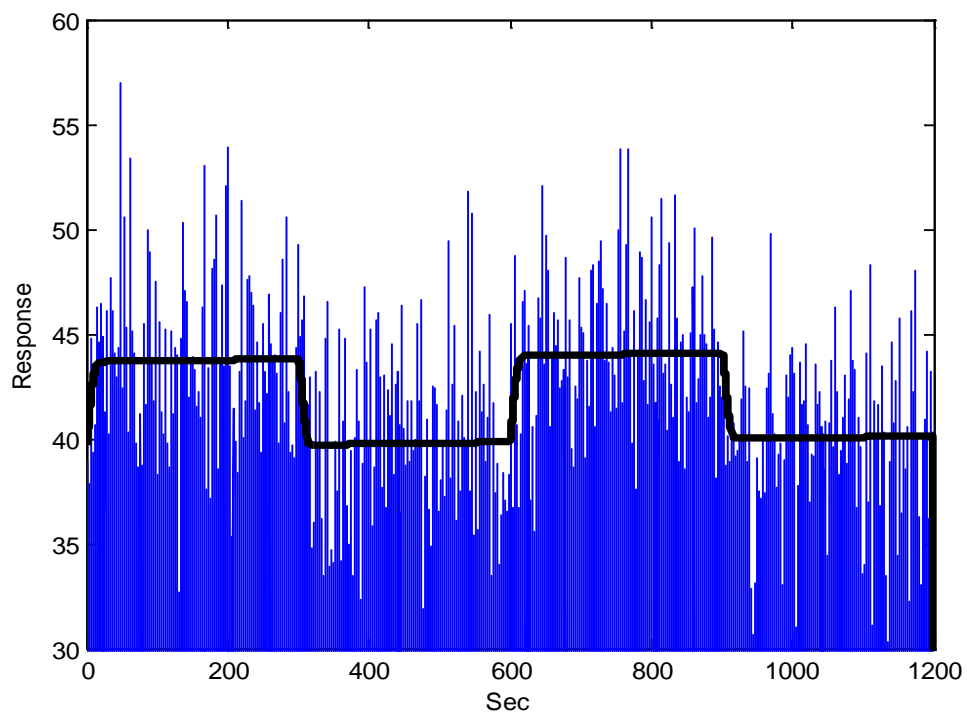


Figure 4.18 Predicted BOLD response TC.

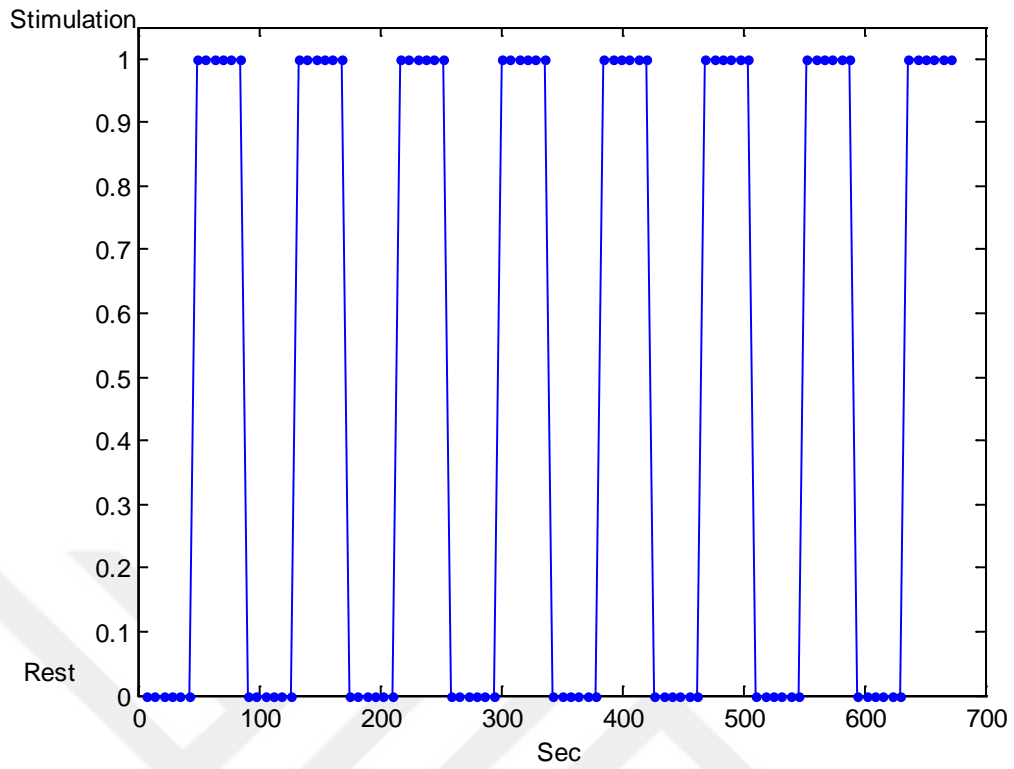


Figure 4.19 Experimental paradigm "silence" and "talk".

In clustering algorithms, they try to classify the TC signals of the voxels into different groups according to the similarity among them. The temporal information is ordered in clusters and is independent of their spatial neighborhood. These clusters are described by an average TC or a cluster center obtained by averaging all the TCs of the cluster. The fMRI data are transformed into a TC of voxel intensity variations proportional to its average as follows:

$$I_{av}^a = \frac{1}{n} \sum I_i^a \quad (4.2)$$

$$X_a = \{w_1, w_2, \dots, w_n\} \quad (4.3)$$

$$w_i = I_{av}^a - I_i^a \quad (4.4)$$

where:

I_{av}^a : average intensity of voxel s of a series of n images;

X_a : fMRI signal

The distances between two fMRI signals X_a and X_b may be computed as Euclidian distance:

$$d_E = \sqrt{(X_{ai} - X_{bi})^2} \quad (4.5)$$

4.4 Discussion

SPM is one of a free and open source package used for processing and analysis of fMRI data. SPM is a software package that runs within the MATLAB environment which is based on the GLM and is one of the commonly used approaches for fMRI data analysis. GLM is a hypothesis or model-based approach not model free or exploratory analysis like clustering techniques.

This chapter introduced the preprocessing steps, which included: Realignment, Coregistration, Segmentation, Normalization and Smoothing within SPM. FMRI data preprocessing is an important step before starting with an fMRI data analysis as well as removing the unwanted data or noises from the dataset to make data satisfy the later processing assumptions.

The real auditory fMRI data used in this work was preprocessed with above steps, and the output results were shown using SPM. Then the conventional GLM method that used by SPM was introduced and presented the output results related to the model used and paradigm.

CHAPTER 5

PROPOSED ALGORITHMS WITH FMRI DATA ANALYSIS

In this chapter the proposed learning algorithm to be used with fMRI dataset is explained in details using a novel RGNG approach. Moreover, complete explanation of NG and GNG algorithms is presented for the next comparison with the RGNG. Flowcharts are designed for each of the three algorithms and presented as well as a comparison among the three artificial neural network approaches are presented in a table. These algorithms are prototype-based unsupervised clustering used for fMRI data Analysis in this work, which are data-driven approaches.

5.1 FMRI Data-Driven Techniques

There are a number of techniques applied to fMRI for data analysis. These techniques or algorithms can be divided mainly according to the based approaches into three groups: transformation based techniques, SPM and clustering based techniques.

PCA and ICA are two main transformations based methods applied to fMRI data analysis. Both methods based on transforming the original data as a matrix into a high-dimensional vector space in order to separate complex patterns of correlation (as different functional responses and types of noise) between the element vectors of the original data.

ICA regarded as the reference method to extract underlying networks and identify spatial nodes from rest fMRI that are independent and sparse [42, 63]; also it has the ability to divide data spatially into non-overlapping and specific sets, but this technique has some limitations. First, it is based on a linear mixing approach that involves other probabilistic models and changes the nature of the problem [49], so this approach attempt to find maximally independent maps and split the wide activation areas into a number of maps which have a strong correlation between TCs of different components. Also the ICs from ICA decomposition are not ordered i.e., it

is associated with the model order selection for linear model-based region extraction is still an open problem. So, it is difficult to identify any ICs are non-linear activation correlated or not.

In contrast, PCA decomposes the correlated variable data into sets of linearly uncorrelated variables with orthogonal directions space through the identification of principal components in the original data. One weakness of this approach is related to some task of experiences in fMRI data analysis when it is difficult to inform small changes with fMRI signal variance. For clarification, some important signals may lose when the signals of interest and the distorted signals from hardware scanner or physiological artifact are non-orthogonal. In sum, ICA and PCA algorithms are able to separate the fMRI signals into a group of the defined components, but each must deal with the limitations of their independence and orthogonality respectively [50].

SPM is based on the conventional GLM which is one of the commonly used approaches for fMRI data analysis. This approach is well done with the fMRI data analysis, but it is hypothesis or model-based approach not model free or exploratory analysis like clustering techniques. So, it needs a precise estimate of the fMRI signals corresponding to the performance of the paradigm. But in many cases it is difficult to provide accurate models for different reasons. For example the volunteers may have been doing the task incorrectly in the experiment. Even if the subjects perform perfectly during the scan, different volunteers may still give different BOLD signals with same paradigm. The same subject may also give different response signals at different time.

There are many clustering based algorithms that have already been applied to fMRI for data mining instead the previous classical methods that face difficulty to predict what occurs during acquisition such as: K-means, fuzzy classification, hierarchical classifications, LBG, CURE, the neural models SOM, NG and Fritzke's GNG algorithms. But one of the main problems in fMRI clustering algorithms is to decide the clusters' number as an input [53, 54]. Also good results in a higher level of interpretation were obtained using the clustering approaches, but here another problem associated with the cost for computing time and memory space.

From these different algorithms we evaluated the GNG algorithm as the best clustering performance and give an acceptable robustness, but also have some limitations associated with the sensitivity for initialization, the order of input vectors and existence of many outliers. Therefore, in this work a novel approach is proposed that relies on using the RGNG algorithm into fMRI data to detect the active areas in the brain with a comparison to the GNG algorithm which have not been used for that purpose yet. GNG was applied with fMRI previously [56- 58], but for the better robustness properties associated with RGNG network, it is suggested to be used in this work which was proposed by Qin and Suganthan (2004) within the GNG structure. In this thesis, this algorithm proofed its ability of finding activated regions in the brain and has different and important features in comparison with others clustering approaches. These features are insensitive to initialization, input sequence ordering, the presence of outliers and determining the optimal number of underlying clusters during different growth stages as well as it is able to deal well with multimodal datasets as the fMRI and detect the active zones in the specified area of the brain.

The suggested model in this thesis produced a robust unsupervised learning system (RGNG with fMRI) for detecting the active zones in the brain as well as its ability to work on the optimal number of underlying clusters in synthetic and real data. The following sections include the details of the three prototypes-based unsupervised clustering algorithms introduced in this work, NG, GNG and RGNG algorithms.

5.2 Clustering Approaches with fMRI

Cluster analysis [64] is a robust tool for exploring the underlining structures in data and grouping them with similar objects that is called clusters. It attempt to group voxels according to their TC signals into a similar HDR over time.

Cluster analysis found applications in different fields ranging from a main task of data mining applications [65] such as: scientific data exploration, spatial database applications, Web analysis, marketing, medical diagnostics, computational biology, and etc., and a common technique for statistical data analysis that used in many fields

including machine learning, pattern recognition [66], image analysis [67], information retrieval [68], and Bioinformatics [69].

In fMRI, clustering algorithms depending not only on a 3D distance between voxels, but based on the activity of the voxels without any influence on their localization. So, clustering is different from segmentation techniques, but it groups data in a few similar segments. In fMRI images, the active area is identified according to a comparison between voxels and their neighbours.

There are many clustering algorithms were applied to fMRI data such as: PCA [70, 71], ICA [42, 72, 73], and K-means [39, 49, 74]; fuzzy classification [75, 76], Hierarchical classifications [39, 77, 78], LBG [50, 57], CURE [79], neural models SOM [50, 57, 80, 81], NG [53, 82] and GNG [50, 57, 83].

As illustrated previously, ICA has some limitations as its dependency on a linear mixing approach that involves other probabilistic models and changes the nature of the problem, it is not ordered and difficult to inform any ICs are activation correlated or not. However, ICA and PCA algorithms have to deal with limitations of their independence and orthogonality. As for many of clustering algorithms, that have already been applied to fMRI for data mining instead the previous classical methods that face the difficulty to predict what occurs during acquisition, have a main problem in choosing the true number of clusters as an input; as well as the cost for computing time and memory space.

From these different algorithms GNG algorithm is evaluated as the best clustering performance and give an acceptable robustness, but also has some limitations associated with the sensitivity for initialization, the order of input vectors and existence of many outliers. For that in this thesis a novel application of the RGNG algorithm into fMRI data is proposed to detect the active areas in the brain with a comparison to the GNG algorithm. The GNG is originating from the NG algorithm [83, 84]. Before starting with introducing the proposed RGNG algorithm for feeding with fMRI data, a review of the NG and GNG algorithms are presented. Because of the length and complexity of the NG, GNG and RGNG algorithms, flowcharts are

designed for the three algorithms in this work as well as the mathematical model in order to be understandable and make easier in writing the related codes.

5.3 Neural Gas (NG) Algorithm

The NG network algorithm is a simple artificial neural network algorithm for finding optimal data representations based on reference vectors (prototype vectors), it was first introduced in 1991 [82] and based on Kohonen's SOM [85]. Because of the dynamics of the reference vectors during the adaptation process, this algorithm was coined "neural gas" that spread themselves as a gas during the data space. NG is unlike other methods that consider distance as a rank like Euclidean distance, but it proposes a new way of calculating the influence. Nearer prototypes in NG algorithm are more affected, but it does not depend directly on the influence of the distance.

NG has been successfully applied to clustering [86], speech recognition [87], image processing [88], vector quantization, pattern recognition and topology representation, etc. [89, 90] especially where there is a problem getting along to vector quantization or data compression. An interested result was obtained by applying this algorithm to fMRI data [51- 55].

It adapts the reference vectors (prototype vectors) ' w_i ' without any fixed topological arrangement within the network. NG is not just adapts the winner vector for a specific input vector as a single-layered soft competitive learning neural network, but also updates the residual reference vectors according to the input vector nearness using a soft-max updating rule [91]. The main advantages of NG network [92] are: (1) lower distortion error resulting than other clustering algorithms (k-means, maximum-entropy and SOM), (2) faster assemblage to low distortion errors, (3) submission a stochastic gradient descent on a specific energy surface.

The NG algorithm is represented by the dependence of updating strengths for c reference vectors $w_{ci}(i_0, i_1, \dots, i_{N-1})$ on their positions in the 'neighborhood ranking' list. When an input vector is presented by x , the determination of the neighborhood ranking $(w_{i_0}, w_{i_1}, \dots, w_{i_k})$ of the reference vectors w_{ci} as follows:

w_{i_0} , being the closest to x

w_{i1} , being second closest to x

w_{ik} (for $k = 1, 2, \dots, N - 1$), being the reference vector for which there is a k vectors w_j with $\|x - w_j\| < \|x - w_{ik}\|$

$k_i(x, w)$: is the ranking index associated with each weight w_i .

The updating step of adjusting w_i according to a Hebb-like learning rule is given by:

$$\Delta w_i = \varepsilon(t) \cdot h_\lambda(k_i(x, w)) \cdot (x - w_i), \quad i = 1, 2, \dots, c \quad (5.1)$$

where:

$h(\dots)$: deterministic function with some regularity condition imposed on it.

$\varepsilon(t) \in [0, 1]$: the learning rate (step size) that describes the total extent of the modification. This extent usually takes an exponential decreasing form $\{\varepsilon(t) = \varepsilon_i \cdot (\varepsilon_f / \varepsilon_i)^{t / (Maxiteration)}\}$, so t and $Maxiteration$ denote the iteration step t and the maximum number of iterations.

$h_\lambda(k_i(v, w)) \in [0, 1]$: accounts for the topological arrangement of the w_i within the input space.

for: $h_\lambda(k) \in [0, 1]$, the exponential form $exp(-k/\lambda)$ was proposed [92] to obtain the best extensive result with comparison to other options like the Gaussian function.

λ : finds the number of reference vectors that significantly changing their positions in the updating steps and usually individually decreases with the iteration step t as:

$$\lambda(t) = \lambda_i \cdot (\lambda_f / \lambda_i)^{t / (Max_iter)}.$$

NG Algorithm is widely related to the structure of fuzzy clustering methods [93]. So NG used the uncertainty of relationship value $(h_\lambda(k_i(x, w))) / (C(\lambda))$ to set each input vector ' x ' to all the reference vectors w_i ($i = 1, 2, \dots, c$) instead of using u_{ij} ($2 \leq i \leq c, 1 \leq j \leq N$) in FCM algorithm. This algorithm is based on solving a

cost function using iterative methods plus the familiarity with linear optimization methods essentially the gradient descent method and Newton's method. Therefore, the NG cost function to optimize [92] is:

$$E_{ng} = \frac{1}{2C(\lambda)} \sum_{i=1}^c \int P(x) h_{\lambda}(k_i(x, w)) \|x - w_i\|^2 \quad (5.2)$$

with

$$C(\lambda) = \sum_{i=1}^c h_{\lambda}(k_i) = \sum_{k=0}^{c-1} h_{\lambda}(k) \quad (5.3)$$

Martinetz et al. (1993) introduced this cost function and proved that the updating in the Hebb-like learning rule can be derived by the stochastic gradient descent on this function. By starting with a large value of λ and reduces it slowly, a good results reference vector will be obtained.

Due to the sequential learning scheme in NG algorithm and use of neighborhood dealing rule, NG became less sensitive to various initializations due to the sequential learning scheme and use of neighborhood cooperation rule with comparison to other clustering algorithms like k-means and FCM.

Before feeding the NG algorithm, there are some parameters have to be defined:

N : maximal number of neurons

$\varepsilon_i, \varepsilon_f$: step size

λ_i, λ_f : decay constant

T_i, T_f : life-time

$t_{max} = Max_iter$ (maximal number of iterations)

The complete NG algorithm [84] is illustrated as follows:

1. Initialize the set 'w' to contain N units c_i

$$w = \{c_1, c_2, \dots, c_N\} \quad (5.4)$$

with reference vectors ($w_{ci} \in R^n$) chosen randomly according to $P(x)$. Initialize the time parameter t , with $t = 0$. Initialize the connection C :

$$C = 0 \quad (5.5)$$

2. Generate at random an input signal 'x' according to $P(x)$.

3. Order all elements of w according to their distance to 'x', i.e., find the sequence of indices $(i_0, i_1, \dots, i_{N-1})$ such that w_{i_0} is the reference vector closest to 'x', ' w_{i_1} ' is the reference vector second-closest to x and w_{i_k} (for $k = 1, 2, \dots, N - 1$) is the reference vector such that k vectors w_j exist with $\|x - w_j\| < \|x - w_{i_k}\|$. Following Martinetz et al. (1993) we denote with $k_i(x, w)$ the number k associated with each weight w_i .

4. Adapt the reference vectors according to:

$$\Delta w_i = \varepsilon(t) \cdot h_\lambda(k_i(x, w)) \cdot (x - w_i), \quad i = 1, 2, \dots, c. \quad (5.6)$$

with the following time-dependencies:

$$\lambda(t) = \lambda_i \cdot (\lambda_f / \lambda_i)^{t/t_{max}} \quad (5.7)$$

$$\varepsilon(t) = \varepsilon_i \cdot (\varepsilon_f / \varepsilon_i)^{t/t_{max}} \quad (5.8)$$

$$h_\lambda(k) = \exp(-k/\lambda(t)) \quad (5.9)$$

5. Create a new connection between two neurons if there is no connection:

$$C_{i_0 i_1} = 1 \quad (5.10)$$

set the age of the connection between two neurons to zero (reset the age):

$$\text{age}_{i_0 i_1} = 0 \quad (5.11)$$

6. Increase the age for each edge emanating from the winner neuron by one, and remove edges larger than maximum $T(t)$:

$$T(t) = T_i \cdot (T_f / T_i)^{t/t_{max}} \quad (5.12)$$

increase the time parameter t :

$$t = t + 1 \quad (5.13)$$

7. If $t < t_{max}$ return with step 2.

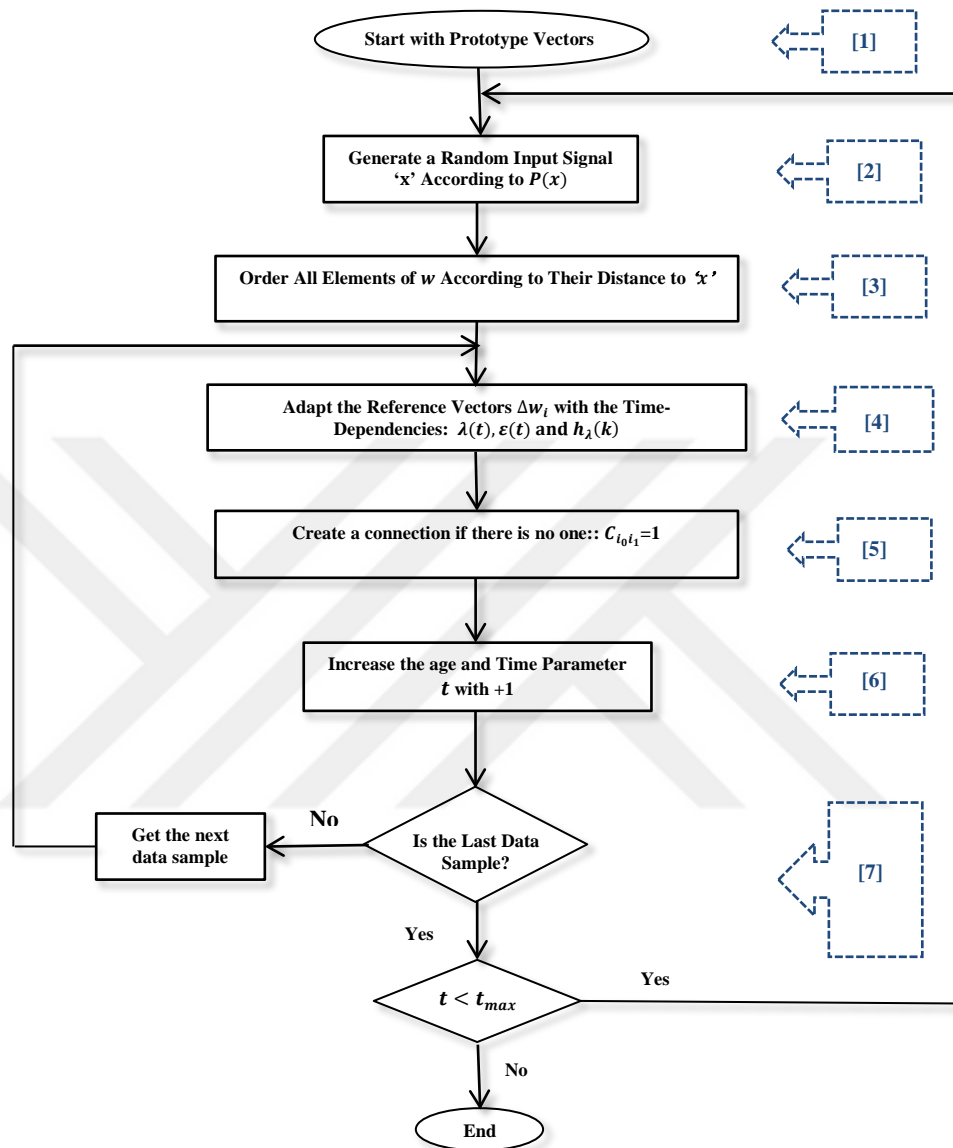


Figure 5.1 The flowchart of NG Algorithm.

Figure 5.1 introduces the flowchart of the NG Algorithm. As well as the advantages of NG model that has introduced, but it also has some limitations like it depends on decaying parameters which change over time and it is incapable to find a network size and structure automatically and continue learning. So, and originating from the NG algorithm, the GNG algorithm introduced by Fritzke (1995, 1997) which has an advantage against NG algorithm through its ability to modify the network topology

by removing edges with its age variable and during the growth process associated with the neighborhood updating rule leads to the no need for the neighborhood sorting step [83, 84]. It has the ability to find a network size and structure automatically, continue learning, adding units and connections until a performance criterion fulfilled.

5.4 Growing Neural Gas (GNG) Algorithm

Originating from the NG algorithm, GNG algorithm introduced by Fritzke (1995, 1997), he proposed changing the unit numbers (mostly increased) during SOM network with a variable topological structure [83, 84]. This growth mechanism is combined with topology formation rules by the Competitive Hebbian Learning (CHL) [94] with the earlier proposed growing mechanism inherited from the Growing Cell Structures [95] to a new model. Interesting results were obtained by applying the algorithm to fMRI data with Lachiche et al., (2005) and Korczak, (2007).

The GNG algorithm needs only constant parameters, but deciding the amount of prototypes is not required. The main idea of the GNG starts with minimal network size and inserting few numbers of new neurons and connections respectively in a growing structure by using a vector quantization until the desired quality of the model is fulfilled (e.g., net size, time limit, predefined numbers of neurons inserted, quality or some performance measure). To determine where to insert new units, local error measures are gathered during the adaptation process. Each new unit is inserted near the unit that has accumulated the highest error, and a connection between the winner and the second nearest neuron is formed using the competitive Hebbian learning algorithm.

Before feeding the GNG algorithm, there are some parameters have to be defined:

N : maximal number of neurons

$\varepsilon_b, \varepsilon_n$: constant learning rate for the winner and its topological neighbors, respectively

λ : iteration number

α : reduction of error counter by inserting a new neuron

β : it will reduce the overall value of the error counter every iteration step

Max_iter : maximal number of iterations

For each reference vector w_i , $i = 1, 2, \dots, c$, a set of edges emanating from it. It is defined to connect with its direct topological neighbors; the GNG algorithm starts in step 1 with initializing a few prototype vectors (usually 2) $W = \{w_1, w_2\}$ with reference vectors that are chosen randomly and new prototype vectors are successively inserted. The learning rates $\varepsilon_b, \varepsilon_n$ used in the training procedure and a connection C , $C \subset w \times w$, to the empty set: $C = \emptyset$.

Set the pre-specified maximum number of prototypes or neurons to grow as $pre_numnode$ and the maximum predefined training epoch Max_iter during each growth stage with the largest local accumulated error measure, the data set used for training is $X = \{x_1, x_2, \dots, x_N\}$. Then set the initial training epoch number: $m = 0$ and the iteration step in training epoch m : $t = 0$.

The complete implementation of the GNG algorithm [84] is as follows:

1. Start with two nodes selected from input data set 'w' to contain two neural units w_1 and w_2 :

$$W = \{w_1, w_2, \dots, w_N\} \quad (5.14)$$

with reference vectors chosen randomly according to $P(x)$. Initialize the connection set C , $C \subset w \times w$, to the empty set:

$$C = \emptyset \quad (5.15)$$

2. Generate at random an input signal 'x' according to $P(x)$; and for $m = 0$ to $Max_iter - 1$

- Set $t = 1$, thus $iter = m \cdot N + t$.
- Set trainingset = X , i.e. include all input vectors into trainingset.
- Draw randomly an input vector x_t^m ; at the iteration step t in training epoch m , from the trainingset.

3. Determine first two closest neurons (winner neurons) s_1 and the second-nearest unit s_2 , ($s_1, s_2 \in w$), to the new data sample (i.e. calculate Euclidean distances to the input neurons) by:

$$s_1 = \arg \min_{i \in w} \|x_t^m - w_i^{iter}\| \quad (5.16)$$

$$s_2 = \arg \min_{i \in w \setminus \{s_1\}} \|x_t^m - w_i^{iter}\| \quad (5.17)$$

4. Compare the Distance:

- i. Modify Age of Edges: Increase the age of all edges emanating from s_1 :

$$\text{age}_{(s_1,i)} = \text{age}_{(s_1,i)} + 1 \quad (\forall_i \in N_{s_1}) \quad (5.18)$$

- ii. Update the Local Error of Winner Neuron: Add the squared distance between the input signal vector and the winner neuron to a local counter error variable for s_1 :

$$\Delta E_{s_1} = \|x_t^m - w_{s_1}\|^2 \quad (5.19)$$

- iii. Modify Weights: The updating rule of GNG algorithm is expressed by moving the reference vectors of the winner neuron s_1 and its direct topological neighbors (neurons connected to s_1) towards x by fractions (learning step) ε_b and ε_n respectively. The total distance to the input signal:

$$\Delta w_{s_1} = \varepsilon_b(x - w_{s_1}), \Delta w_i = \varepsilon_n(x - w_i) \quad (\forall_i \in N_{s_1}) \quad (5.20)$$

where

N_{s_1} : is the set of direct topological neighbors of s_1 that are connected by an edge with w_{s_1} .

The prototype's updating is only operated on the winning prototype w_{s_1} and its direct topological neighbors w_i , $\forall_i \in N_{s_1}$ at the presence of an input vector x . The updating strengths stay constant over time but are irregular for the winning prototype and its topological neighbors.

iv. Create Edges: If s_1 and s_2 are connected by an edge, then set the age of this edge connection between s_1 and s_2 to zero (“refresh” the edge):

$$\text{age}_{(s_1, s_2)} = 0 \quad (5.21)$$

If the edge connection between s_1 and s_2 does not exist already, create it

$$C = C \cup \{(s_1, s_2)\} \quad (5.22)$$

5. Dead Node Removal Procedure:

- Deleting all connections with $\text{age} > \alpha_{\max}$
- If the results in units have no more emanating connections, remove them as well.
- Increment the age of all emanating connection of s_1 by 1 and remove the used vector x_t^m from the set trainingset.

6. If the current number of prototypes $\leq \text{pre_numnode}$ (the number of input signals generated so far is an integer multiple of a parameter λ) and some predefined performance measure, then return to (2).

7. Node Insertion Procedure (Keeping Topology by inserting new neurons), if the number of input signals generated so far is an integer multiple of a parameter λ , insert a new unit as follows:

- Determine the prototype neuron q with the maximum accumulated error:

$$q = \arg \max_{i \in w} E_i \quad (5.23)$$

- Determine among the topological neighbors of q , the prototype unit f with the maximum accumulated error:

$$f = \arg \max_{i \in q} E_i \quad (5.24)$$

- create a new prototype neuron r to the network and interpolate its reference vector between q and f :

$$w = w \cup \{w_r\}, \quad w_r = (w_q + w_f)/2 \quad (5.25)$$

- create edges connecting the new prototype unit r with units q and f , and remove the original edge between q and f :

$$C = C \cup \{(r, q), (r, f)\}, \quad C = C \setminus \{(q, f)\} \quad (5.26)$$

8. Modify Error Counters

- Decrease the error variables of q and f by a fraction α

$$\Delta E_q = -\alpha E_q, \quad \Delta E_f = -\alpha E_f \quad (5.27)$$

- Initialize the error variable counter of the new neuron r with the values of q and f

$$E_r = (E_q + E_f) / 2 \quad (5.28)$$

9. Decrease the error variables of all prototypes neurons by multiplication with a fraction β

$$\Delta E_i = -\beta E_i \quad (\forall i \in w) \quad (5.29)$$

10. If a stopping criterion (e.g., net size or some performance measure) is not yet fulfilled; i.e. number of epoch $m \leq Max_iter$; continue with step 2., else End.

Figure 5.2 presents the flowchart of the GNG algorithm. This figure shows that the nonfunctional prototypes do not win during a long time interval may be detected by tracing the changes of an age variable associated with each edge. So, the GNG algorithm has an advantage against NG algorithm through its ability to modify the network topology by removing: edges with its age variable (not being refreshed for a time interval α_{max}) and the resultant nonfunctional prototypes. At GNG the growth process associated with the neighborhood updating rule used is somewhat amounting to the neighborhood, decreasing procedure in NG, therefore and unlike the NG algorithm, there is no need for the neighborhood sorting step.

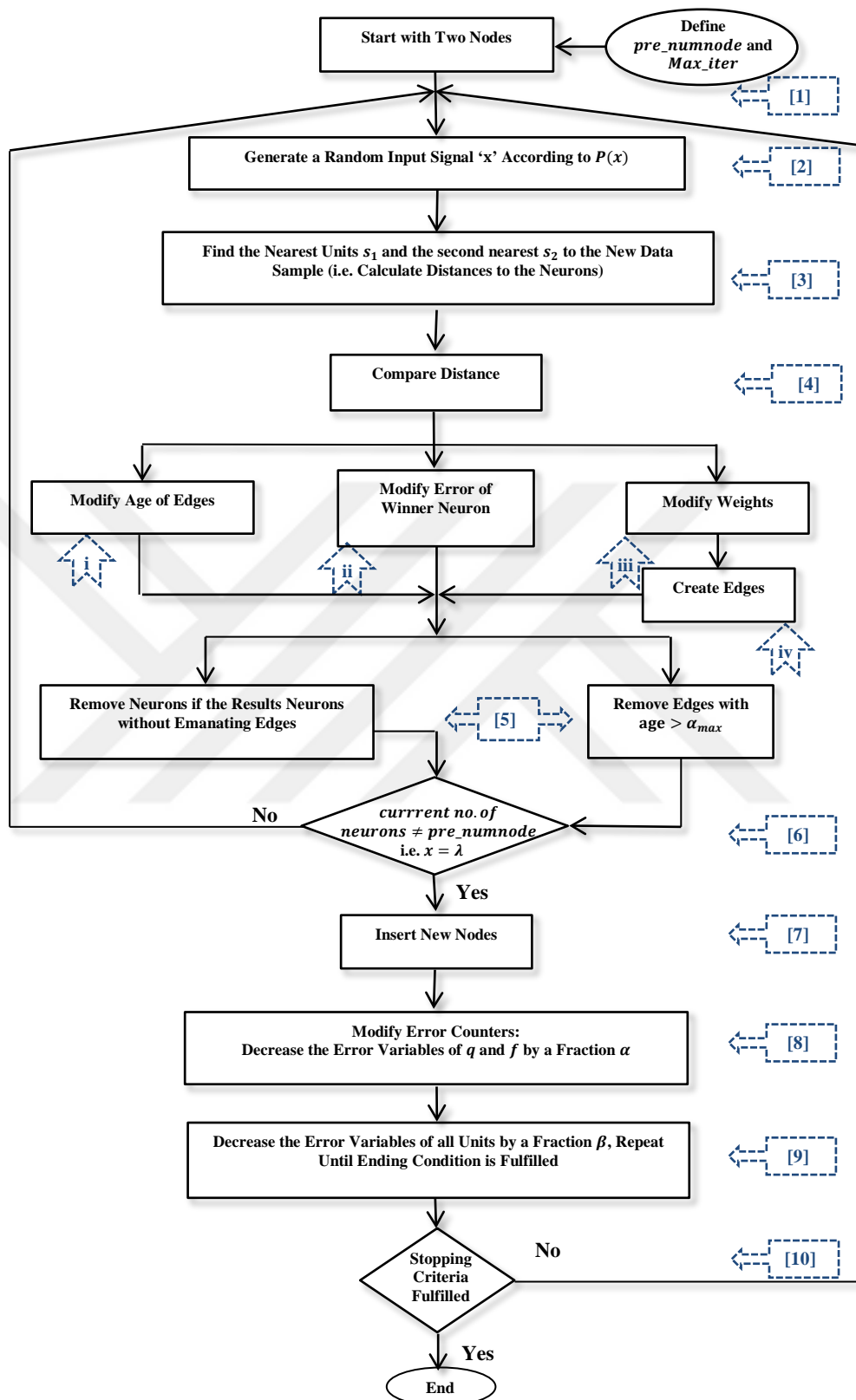


Figure 5.2 The flowchart of GNG Algorithm.

5.5 Robust Growing Neural Gas (RGNG) Algorithm

In the proposed work, RGNG technique applied for the first time to fMRI for detecting active zones in the brain. This application is novel and no such study is done according to the researcher's knowledge, so it will be the first time in the literature. RGNG has the ability to detect the active zones in the brain, analyze brain function as well as its ability to work on the optimal number of underlying clusters with respect to the MDL value in fMRI dataset and define the active zones in a specified region of the brain.

Before the introduction of a RGNG technique algorithm in detail, here a brief illustration is introduced about why robustness is used. Robustness is a significant characteristic of related clustering algorithms. Any robust algorithm should have these features [96]:

- (1) it should achieve a good precision at the given model
- (2) The performance of the given model may have little deviations from its assumptions, but these deviations should not weaken the performance with a small degree only.
- (3) The presence of large deviations from the model assumption should not cause disaster.

According to classical clustering methods as a prototype based clustering algorithms, the major robustness problems are the sensitivity for initialization, the order of input vectors and existence of many outliers, but each well executed regarding condition 1. Due to the growth scheme associated with the GNG algorithm; the 'dead nodes' problem appears with the GNG algorithm. Dead nodes problems occur due to inappropriate initializations that led to some prototypes may never win through the training process. Even with the initialization insensitive clustering methods, good clustering results may not be obtained if the order of the input sequence is not chosen properly.

Even with the initialization insensitive clustering methods, good clustering results may not be obtained if the order of the input sequence is not chosen properly. As well as the introduced problem gets along with the sensitivity for initialization and the

order of input vectors, there also another problem attributable to the existence of many outliers. This implies the GNG network may fail to differentiate the outliers from the inliers through the original prototype updating rule when many of outliers exist in a data set. These outliers can be regarded as input vectors that different from data points belonging to the ordinary clusters (inliers).

For these limitations of GNG algorithm, a novel robust clustering algorithm was proposed [59] within the GNG structure, named Robust Growing Neural Gas (RGNG) network. RGNG possesses better robustness properties by succession the properties of the original GNG algorithm that is overcoming the robustness issues associated with it and incorporating it with several robust strategies, such as outlier resistant scheme, adaptive modulation of learning rates and cluster repulsion method.

Therefore, with comparison to GNG network, the RGNG network is insensitive to initialization, input sequence ordering, the presence of outliers and determining the optimal number of clusters. The Minimum Description Length (MDL) value was used with RGNG as one of the famous clustering validity index [97, 98]. MDL value is used to find the optimal number of clusters and their center positions corresponding to the smallest MDL. Hence, they determined automatically the optimal number of clusters by searching the extreme value of the MDL measure through the network growing process.

In this thesis a novel application for the RGNG algorithm with fMRI dataset is proposed to detect the active areas; GNG was applied with fMRI previously [56- 58] but for the better robustness properties associated with RGNG network, RGNG is suggested to detect the active zones in the brain.

Before feeding the RGNG algorithm, there are some parameters have to be defined:

- N : maximal number of neurons
- ε_b^l : learning rate of the winner
- ε_n^l : learning rate of its topological neighbors
- $\varepsilon_{bf}^l, \varepsilon_{bi}^l, \varepsilon_{nf}^l, \varepsilon_{ni}^l$: initial and final values of ε_b^l and ε_n^l

- α_{max} : maximal age of a connection.
 β : mobility of the winners neighborhood towards the input vector.
 k, η : used to determine the MDL value.
 Max_iter : maximal number of iterations

Set the maximum number of prototypes to grow as $pre_numnode$ and the maximum predefined training epoch Max_iter during each growth stage with a certain prototype number. Also, set the initial or current training epoch number: $m = 0$ and the iteration step in training epoch m : $t = 0$. Hence, the total iteration step $iter$ during each growth stage is: $iter = m.N + t$, where N is the actual number of neuron. The data set used for training is $X = \{x_1, x_2, \dots, x_N\}$.

Figure 5.3 presents the flowchart of the RGNG algorithm. The complete RGNG algorithm is as follows:

1. Start with a few of prototype vectors (usually 2) $W = \{w_1, w_2\}$ randomly with reference vectors chosen randomly according to $P(x)$. Initialize the connection set C , $C \subset w \times w$, to the empty set:

$$C = \emptyset \quad (5.30)$$

2. Generate at random an input signal 'x' according to $P(x)$, and for:

$$m = 0 \text{ to } Max_iter - 1$$

- Calculate the harmonic average distance value $d_m^i(0)$ for each prototype $w_i, \forall_i \in w$, w.r.t. its current position by:

$$d_k^m(0) = \left[\frac{1}{N} \sum_{j=1}^N \frac{1}{\|x_j - w_k^m\|} \right]^{-1} \quad (5.31)$$

- Calculate the learning rates for the current prototype l , when it serves as a winner or its topological neighbors, respectively, by:

$$\varepsilon_b^l = \varepsilon_{bi}^l \left(\frac{\varepsilon_{bf}^l}{\varepsilon_{bi}^l} \right)^{\frac{pre_numnode^l}{pre_numnode}} \quad (5.32)$$

$$\varepsilon_n^l = \varepsilon_{ni}^l (\varepsilon_{nf}^l / \varepsilon_{ni}^l)^{pre_numnode^l / pre_numnode} \quad (5.33)$$

where:

$$l = 1, \dots, C$$

c : current number of neurons

$pre_numnode^l$: ranking index

$pre_numnode^l$ is zero for the new inserted neurons and one if a newer neuron is inserted.

The learning rates remain same in the following training for the prototype l , before the new prototype is inserted. Afterwards:

- Increase the control variable $t = 1$ and $iter = m.N + t$.
 - Set trainingset = X , i.e. include all input vectors into trainingset.
 - Draw randomly an input vector x_t^m ; at the iteration step t in training epoch m , from the trainingset.
3. Determine first two closest neurons (winner neurons) s_1 and the second-nearest unit s_2 , ($s_1, s_2 \in W$), to the new data sample (i.e. calculate distances to the neurons) by:

$$s_1 = \arg \min_{i \in W} \|x_t^m - w_i^{iter}\| \quad (5.34)$$

$$s_2 = \arg \min_{i \in W \setminus \{s_1\}} \|x_t^m - w_i^{iter}\| \quad (5.35)$$

4. Compare the Distance:

i. Modify Age of Edges: Increase the age of all edges connections emanating from s_1 :

$$age_{(s_1,i)} = age_{(s_1,i)} + 1 \quad (\forall i \in N_{s_1}) \quad (5.36)$$

ii. Modify Weights: Adapt the new reference weight vector for the winner neuron with:

$$\Delta w_{s_1} = \varepsilon_b^{s_1} \sigma_{s_1}(iter) \frac{(x - w_{s_1})}{\|x - w_{s_1}\|} \quad (5.37)$$

and for its topological neighbors:

$$\Delta w_i = \varepsilon_n^i \sigma_i(iter) \frac{(x-w_i)}{\|x-w_i\|} + \exp\left(-\frac{d_{s_1 i}}{\zeta}\right) \times \beta \frac{\sum_i d_{s_1 i} (w_i - w_{s_1})}{|N_{s_1}| \|w_i - w_{s_1}\|}, \quad \forall_i \in N_{s_1} \quad (5.38)$$

where:

$$\sigma_k(iter) = \sigma_k^m(t) = \begin{cases} d_k^m(t) & , \text{if } \|x_t^m - w_k^{iter}\| \geq d_k^m(t-1) \\ \|x_t^m - w_k^{iter}\| & , \text{if } \|x_t^m - w_k^{iter}\| < d_k^m(t-1) \end{cases} \quad (5.39)$$

$$d_k^m(t) = \begin{cases} \left\{ \frac{1}{2} \left[\frac{1}{d_k^m(t-1)} + \frac{1}{\|x_t^m - w_k^{iter}\|} \right] \right\}^{-1} & , \text{if } \|x_t^m - w_k^{iter}\| \geq d_k^m(t-1) \\ \frac{1}{2} [d_k^m(t-1) + \|x_t^m - w_k^{iter}\|] & , \text{if } \|x_t^m - w_k^{iter}\| < d_k^m(t-1) \end{cases} \quad (5.40)$$

$$d_k^m(0) = \left[\frac{1}{N} \sum_{j=1}^N \frac{1}{\|x_j - w_k^{mN}\|} \right]^{-1}; \quad i = 1, 2, \dots, N \quad (5.41)$$

ζ : controls the weakening effect in terms of the distance between winner s_1 and its neighbors

$d_k^m(t)$: restricting the distance for prototype w_k

iii. Create Edges: If s_1 and s_2 are connected by an edge, then set the age of this edge connection between s_1 and s_2 to zero (“refresh” the edge):

$$\text{age}_{(s_1, s_2)} = 0 \quad (5.42)$$

Otherwise, i.e. if the edge connection between s_1 and s_2 does not exist already, create it:

$$C = C \cup \{(s_1, s_2)\} \quad (5.43)$$

5. Dead Node Removal Procedure

- Removing all edges or connections with age $> \alpha_{max}$.
- If the results in units having no more emanating edges, remove those units as well.
- Increment the age of all emanating connection of s_1 by 1 and remove the used vector x_t^m from the set trainingset.

6. If the current number of prototypes $\leq pre_numnode$ and some predefined performance measure, then return to (2).

7. As training proceeds, the MDL value is calculated by the following equation at the completion of each growth stage until the predefined maximum number of prototypes is reached:

$$MDL(V, W) = cK + N \log_2 c + k \sum_{i=1}^c \sum_{x \in S_i} \sum_{k=1}^d \max \left(\log_2 \left(\frac{\|x_k - w_{ik}\|}{\eta} \right), 1 \right) + |O|K \quad (5.44)$$

i.e., the MDL value can be calculated according to the current prototypes' positions by the two equations (one above and below):

$$\begin{aligned} \Delta L = & \{L((I - \{x\})(W)) + K\} - \{L(I(W)) + L(x - w_i)\} - \varphi_{S_i}K = \\ & \{(N - 1) \log_2(c - \varphi_{S_i}) + K\} - \{N \log_2 c + \sum_{k=1}^d \max \left(\log_2 \left(\frac{\|x_k - w_{ik}\|}{\eta} \right), 1 \right)\} - \\ & \varphi_{S_i}K \end{aligned} \quad (5.45)$$

where:

c : current number of prototypes

d : dimension of input vectors

η : resolution of the data source

S_i : receptive field of neuron w_i

I : inlier set

O : outlier set

φ_{S_i} : equals 1 if x is the only vector in S_i before movement, otherwise it equals 0

K : number of bits needed for encoding a single data vector

The MDL value is calculated by the above equations at the completion of each growth stage until the predefined maximum number of prototypes is reached.

8. If the calculated MDL value is smaller than the previous one, then save the current positions of all prototypes.

The cluster center positions that correspond to the optimal cluster number can be located according to the smallest MDL value, because it is a better network representation of the given problem. Therefore, if the calculated MDL value is smaller than the previous one, the current positions of all prototypes will be saved.

9. Determine the neuron with greatest local accumulated error:

$$\sum_{x \in S_i} \exp\left(-\frac{\|x - w_i\|}{\text{harmdist}^i}\right) \|x - w_i\| \quad (5.46)$$

where:

$$\text{harmdist}^i = \left[\frac{1}{|S_i|}\right]^{-1} \sum_{x \in S_i} \frac{1}{\|x - w_i\|}$$

S_i : is the receptive field of prototype w_i

harmdist^i : is defined as the harmonic average distance from all data points in S_i to the prototype vector w_i .

10. Node Insertion Procedure:

If the number of input signals generated so far is an integer multiple of parameter λ iteration, insert a new unit near the existing prototype with the largest local accumulated error as follows:

- Determine the prototype neuron q with the maximum accumulated error:

$$q = \arg \max_{i \in W} E_i \quad (5.47)$$

- Determine among the topological neighbors of q , the prototype unit f with the maximum accumulated error:

$$f = \arg \max_{i \in N_q} E_i \quad (5.48)$$

- Create a new prototype neuron r , add it to the network and interpolate its reference vector between q and f :

$$w = w \cup \{w_r\}, \quad w_r = 2 \cdot w_q + \frac{w_f}{3} \quad (5.49)$$

- Create edges connecting the new prototype unit r with units q and f , and remove the original edge between q and f :

$$C = C \cup \{(r, q), (r, f)\}, \quad C = C \setminus \{(q, f)\} \quad (5.50)$$

11. Set the ranking counter $prenode^r$ of the newly inserted prototype r for 0 and for other existing prototypes, increment their corresponding counter variables by 1, i.e.

$$prenode^r = 0, \quad prenode^l = prenode^l + 1, \quad \forall l \in W\{r\}$$

12. If a stopping criterion (e.g., net size or some performance measure) is not yet fulfilled, continue with step 2, else End.

5.6 Comparison among Three Clustering Approaches

There are three prototype-based unsupervised clustering algorithms where presented in the previous sections. NG approach has a little sensitivity to different initializations and lower distortion error than other clustering algorithms as k-means or SOM. In this approach, the number of prototypes must be fixed in advance and it depends on decaying parameters which change over time. In contrary, the number of related prototypes is not needed to be predefined in GNG method. The number of prototypes and the network topology can be increased or decreased dynamically during the classification which gives the ability to cluster a large volume of data. GNG model has the property of aging connections and cancelling when they reach the maximum defined age; the neurons can be eliminated if they not connected anymore. Inserting a new prototype near the prototype with the maximum accumulated error, gives the ability to create new neurons at regular time intervals. The local error of a prototype is estimated by using the sum of the distances from the prototype to the voxels that started with this prototype as a winner. Thus the network is fixed near the prototypes that measure the longest distances.

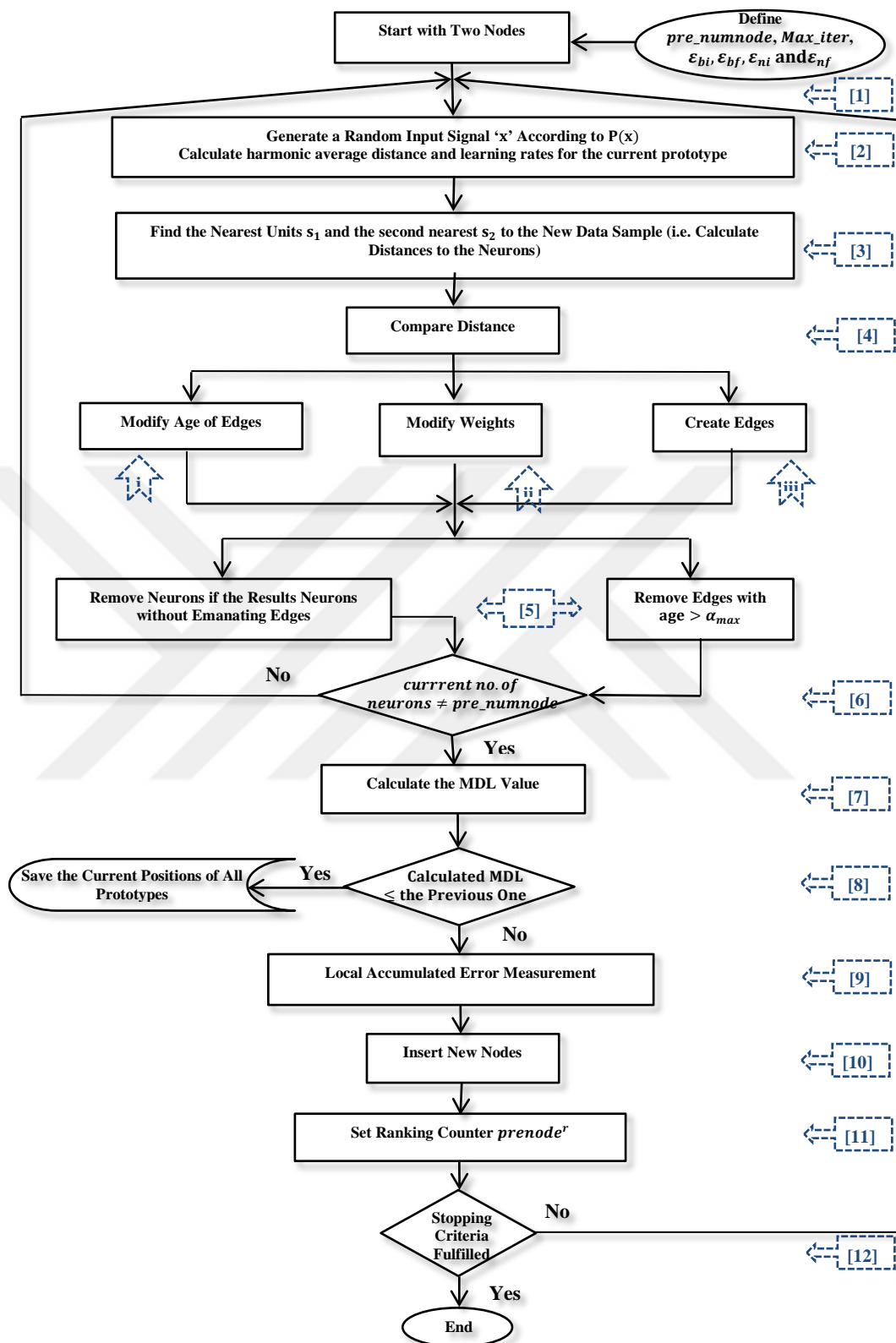


Figure 5.3 The flowchart of RGNG Algorithm.

Table 5.1 Comparison among Three Artificial Neural Network Approaches Based on Unsupervised Clustering for fMRI Analysis.

Methods	NG	GNG	RGNG
Who First Introduced	Martinetz and Schulten (1991)	Fritzke (1995)	Qin and Suganthan (2004)
Used with fMRI	<ul style="list-style-type: none"> • Evgenia et al. (2003) • Meyer-Baese et al. (2004) • Wismuller et al. (2004) • Dimitriadou et al. (2004) • Ana et al. (2006) 	<ul style="list-style-type: none"> • Lachiche et al., (2005) • Korczak J., (2007) • Heydar et al., (2009) 	No One till now
Advantages	<ul style="list-style-type: none"> • Little sensitive to different initializations due to the sequential learning scheme and use of neighborhood cooperation rule. • Lower distortion error resulting than other clustering algorithms (k-means, maximum-entropy and SOM) • Faster assemblage to low distortion errors • Submission a stochastic gradient descent on a specific energy surface. 	<ul style="list-style-type: none"> • Its ability to modify the network topology by removing edges with its age variable • There is no need for the neighborhood sorting step • It has the ability to find a network size and structure automatically and continue learning. • The number of classes is not fixed in advance as in most clustering algorithms 	<ul style="list-style-type: none"> • Insensitive to initialization, input sequence ordering and the presence of outliers during different growth stages. • Can automatically determine the optimal number of clusters • Deal well with multimodal data sets
Limitations	<ul style="list-style-type: none"> • NG model depends on decaying parameters which change over time • It is incapable to find a network size and structure automatically. • Require to predefine the number of related prototypes • Existence of many outliers 	Its sensitivity for: <ul style="list-style-type: none"> • Initialization • The order of input vectors • Existence of many outliers 	Not detected limitations yet

Despite GNG is still sensitive for initialization, the order of input vectors and existence of many outliers. In contrast, RGNG face the sensitivity problem during different growth stages and it can automatically determine the optimal number of clusters by detecting the minimum MDL value.

Comparison among three artificial neural network approaches that are based on unsupervised clustering for fMRI Analysis proposed in Table 5.1. Table 5.1, shows who introduced each technique firstly, who used them with fMRI research, advantages and limitations of each one.

5.7 Summary

In summary, this chapter has presented a general introduction to the clustering approaches that was used to fMRI, review of the NG algorithm, the complete GNG algorithm and the proposed model which has different and important features in comparison with other clustering approaches by using a novel RGNG approach. The algorithm is applied into fMRI dataset to detect the active areas in the brain with a comparison to the NG and GNG algorithm, which have not been used for such purpose yet. Flowcharts are designed for each algorithm and presented, also a comparison among the three artificial neural network approaches that are based on unsupervised clustering for fMRI Analysis are proposed and presented in a table.

All the results related to the proposed algorithm and others will be introduced in details in the next chapter, when applying for synthetic and real fMRI dataset. The output results are introduced and implemented as a toolbox software package based on brain fMRI clustering by using new model design for neuroscience data analysis.

CHAPTER 6

ALGORITHMS IMPLEMENTATION AND RESULTS

Clustering techniques are used to group data depending on a distance. In fMRI, a 3D distance between voxels has to be defined according to their activities. So, the clustering techniques used to identify active areas in fMRI data rely on a comparison of neighboring voxels. In this chapter, the performance of the proposed RGNG algorithm on the synthetic and real auditory fMRI data is calculated. The cases of study are carried out to compare the performance of the proposed approach to the NG and the GNG methods as well as the hypothesis-driven GLM analysis. In this thesis, two interactive tools are proposed, designed and implemented successfully. One interacts with the prototype-based unsupervised clustering techniques (Prototype-Based Clustering software package), while the other interacts with visualization of the fMRI dataset and highlights the active zones using the introduced techniques (FMRI Clustering and Analysis software package).

In the following section, the validity of using the proposed clustering technique with fMRI dataset according to some distinguished parameters is pointed out. The performance of the proposed method is compared to NG and GNG methods in terms of both robustness and sensitivity to initializations as well as the presence of outliers. Section 6.2 devotes to the output of the experimental results on the synthetic data in terms of tables and figures. These results are introduced in a simple software package in Section 6.3. Section 6.4 introduces the experimental output results on the real auditory fMRI dataset running NG, GNG and RGNG. Moreover, the proposed model-free clustering technique is examined over fMRI analysis and compared to the statistical model-based GLM analysis using SPM12, in Section 6.5. The fMRI dataset is visualized in the designed package in Section 6.6. Finally, in Section 6.7, the results of these analyses are discussed and some concluding remarks are given.

6.1 Case of Studies

The robust performance of the proposed RGNG algorithm is tested on synthetic and real auditory fMRI datasets with a comparison to the NG and GNG algorithms. There are four parameters where are used in this work to evaluate the performance of the proposed clustering technique. These performance measures are: Classification Rate (CR), Average Partition Quality (PQ), Minimum Cluster Number (MCN) and Mean Square Error (MSE). The robust clustering technique should be less sensitive to the parameter configurations and give better performance under the same parameter settings in all experiments.

In the following experiments, the parameters are fixed for each technique with typical values suggested in literatures. RGNG technique was set as the typical values in [59]: $\varepsilon_{bi} = 0.1$, $\varepsilon_{bf} = 0.01$, $\varepsilon_{ni} = 0.005$, $\varepsilon_{nf} = 0.0005$, $\alpha_{max} = 100$, $k = 1.3$, $\eta = 1 \times 10^{-4}$. GNG and NG techniques were set as the typical values in [84]: $\varepsilon_b = 0.05$, $\varepsilon_n = 0.006$, $\alpha_{max} = 100$, $\beta = 0.0005$, $\lambda = 300$ for GNG; while $\varepsilon_i = 0.5$, $\varepsilon_f = 0.005$, $\lambda_i = 10$, $\lambda_f = 0.01$, $t_{max} = 40000$ for NG network.

The performance of the RGNG technique on fMRI data shows superior performance with respect to the other two methods as well as the model-based GLM approach. Each index of the performance measures is explained in the following sections.

6.1.1 Classification Rate

This index refers to the Classification Rate (CR) for the whole dataset so each data point is classified regarding to its nearest prototype. CR is based on using a majority voting classifier [99], by labeling all prototypes by using a simple voting mechanism. According to the proposed technique, numbers of prototypes are small, so the resulting CR will not be high.

6.1.2 Partition Quality

This index refers to average Partition Quality (PQ) measurement, which is averaged over all the independent runs in the experiments. PQ was defined by [100], as:

$$PQ = \frac{\sum_{i=1}^{n_{cs}} \sum_{j=1}^{n_{ct}} p(i,j)^2}{\sum_{i=1}^{n_{cs}} p(i)^2} \quad (6.1)$$

where:

n_{cs} : True number of classes

m_{ct} : Minimum number of clusters founded by the technique

$p(i, j)$: Probability of a point vector in cluster j belonging to the class i

$p(i)$: class probability

Number of classes n_{cs} should equal the actual number of clusters if each natural cluster assumed to stand for an individual class. The minimum cluster number m_{ct} can be obtained by running the techniques.

The $p(i, j)$ term represents the frequency based on probability that a data point is labeled by clusters i and j . The $p(i, j)$ quality is normalized by the sum of true probabilities then squared. This statistic is related to the rand statistic for comparing partitions [101]. PQ index is maximized when the number of clusters m_{ct} is correctly detected and induces the same partition of n_{cs} , i.e. $m_{ct} = n_{cs}$, so all points in each cluster are the same as those in one of the natural clusters.

6.1.3 Minimum Cluster Number

Minimum Cluster Number (MCN) is the average number of the detected clusters by the techniques. MCN index the ability of the techniques to find the underlying natural clusters. During the training of the techniques and according to the MCN value, only the proposed RGNG approach can find the actual number of clusters successfully.

During the growing process, this value is defined as the number of natural clusters in which the algorithm placed at least one prototype when the number of prototypes in the network reaches the actual number of clusters. Cluster numbers are detected by NG and GNG during the growing process deviate from the actual value of the clusters when the number of prototypes is the same as the actual number of clusters.

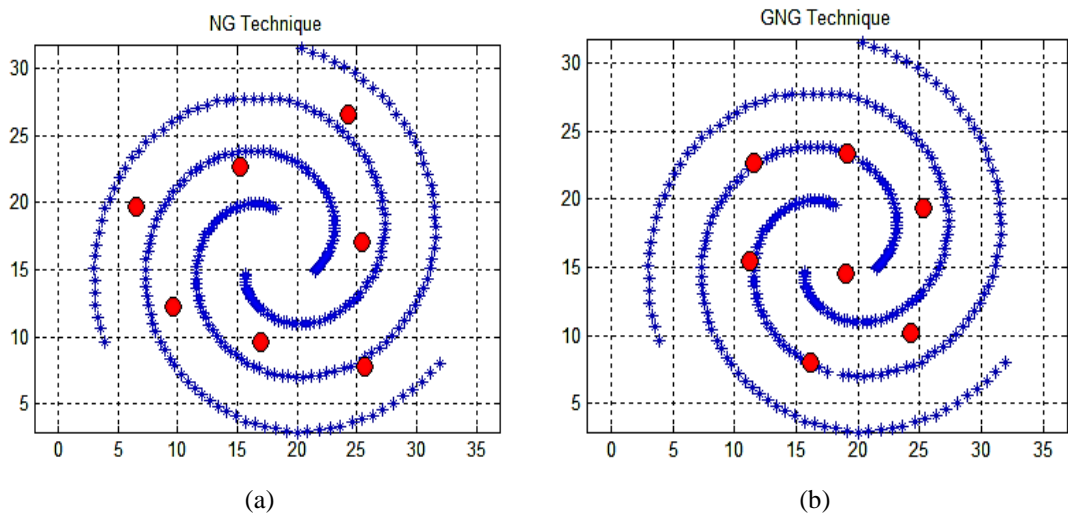
6.1.4 Mean Square Error

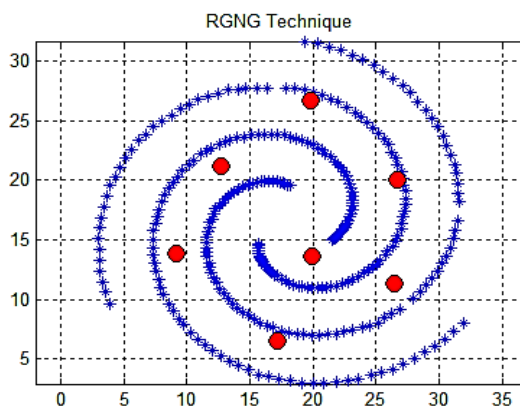
Mean Square Error (MSE) is another criterion used for evaluating the performance of the proposed clustering technique. The MSE value represents the mean distance between the current nearest prototypes' positions resulting from the application of the techniques and the actual cluster centers.

The average MSE value in this experiment is higher for NG and GNG techniques with respect to RGNG. This indicates that the RGNG approach achieves the best accuracy with the strongest stability among the other two approaches.

6.2 Experimental Results with Synthetic Data

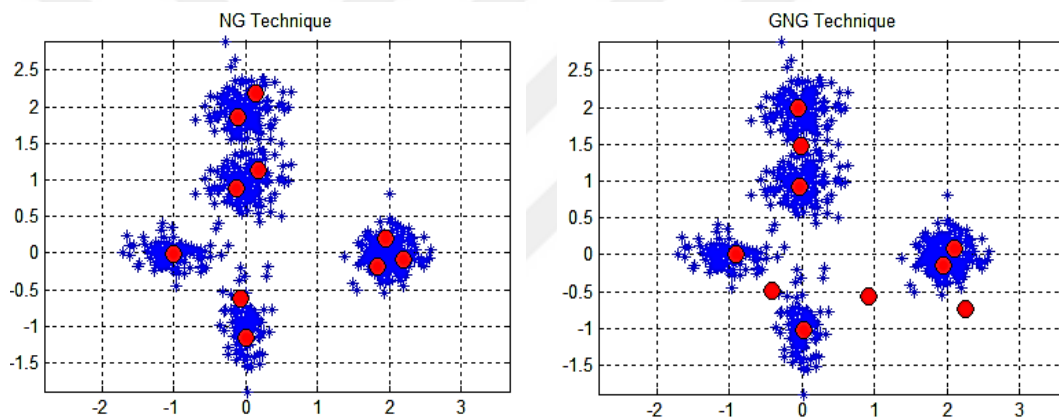
There are six different types of 2D synthetic dataset [59, 102] which are used in this work they are: snail, screw, ring, set3, set5 and set25 dataset. Figures 6.1-3 show the plots of NG, GNG and RGNG clustering with three types of 2D synthetic dataset (screw, set5 and snail) as an example. The number of neurons are selected randomly, $N=7, 10$ and 12 .





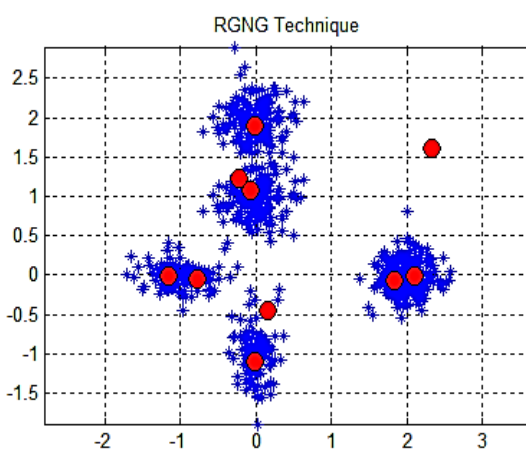
(c)

Figure 6.1 Clustering with screw synthetic dataset for $N=7$, by running NG, GNG and RGNG techniques.



(a)

(b)



(c)

Figure 6.2 Clustering with set5 synthetic dataset for $N=10$, by running NG, GNG and RGNG techniques.

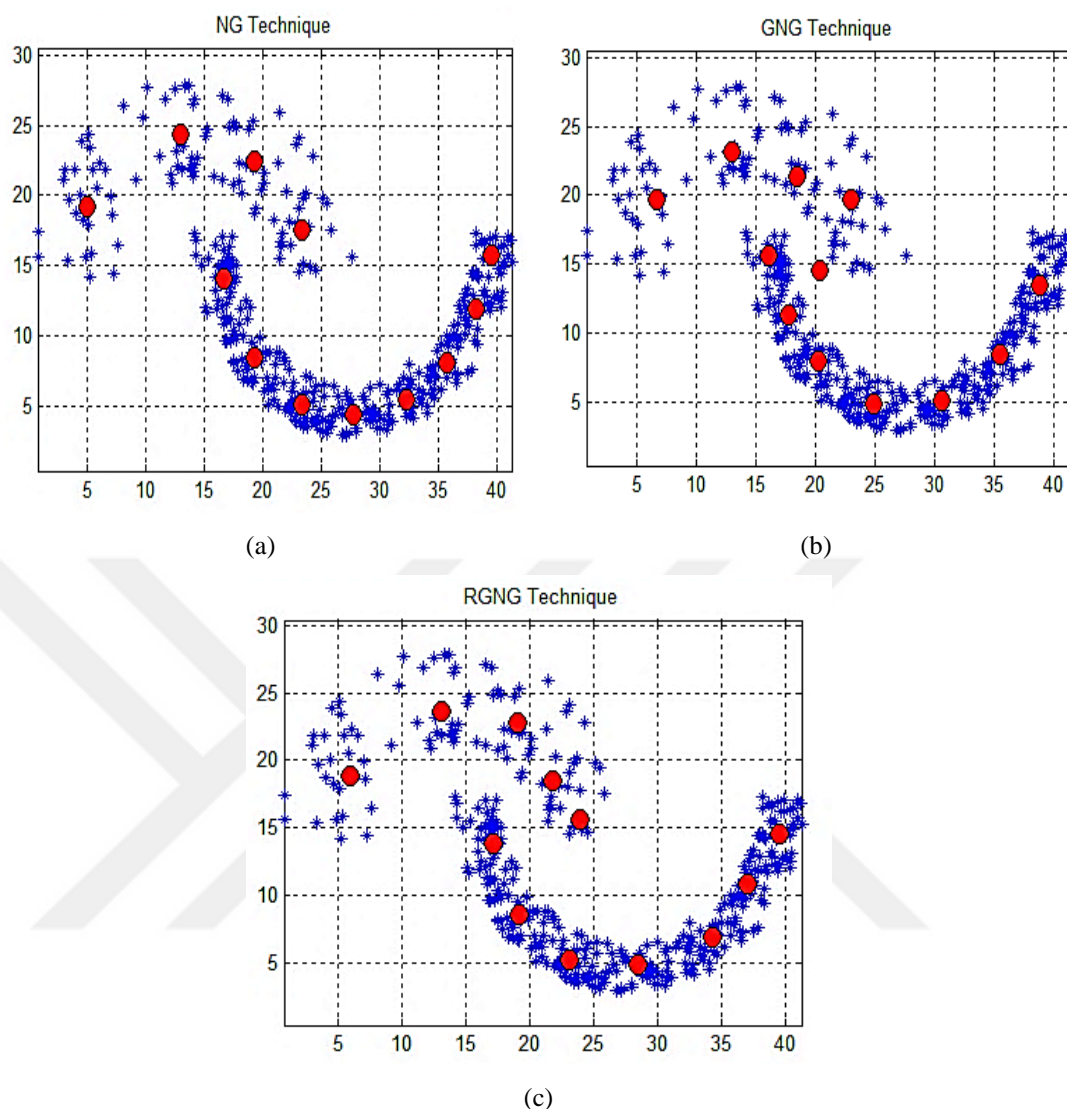


Figure 6.3 Clustering with snail synthetic dataset for $N=12$, by running NG, GNG and RGNG techniques.

These figures cannot give clearly the differences between each method. So, there are four parameters used in this work to evaluate the performance of the proposed clustering technique; CR, PQ, MCN and MSE introduced in the previous section. For the best comparison with RGNG, MDL criterion is added to NG and GNG techniques. The training results of these techniques with the synthetic data are shown in Table 6.1, while the number of neurons chosen randomly as $N= 7, 10$ and 12 .

Regarding to the literatures [59, 103], the clustering output results introduced in the Table 6.1 clarified that RGNG approach is insensitive to different initializations and

the presence of outliers. In these techniques, the number of neurons used is small, so the CR values registered in the table are not high. All the three clustering techniques are applied as the number of neurons equal to the actual cluster number, so RGNG can effectively locate the actual number of clusters with respect to other two methods, NG and GNG, which fails with higher cluster numbers in the synthetic case.

Table 6.1 Clustering results of the synthetic data.

Parameters	Number of Neurons	NG	GNG	RGNG
CR	N=7	0.8718	0.9686	0.9929
	N=10	0.8514	0.9786	0.9843
	N=12	0.8010	0.9647	0.9759
MCN	N=7	9	8	7
	N=10	12	11	10
	N=12	15	14	12
PQ	N=7	0.8990	0.9465	0.9869
	N=10	0.8531	0.9288	0.9841
	N=12	0.8279	0.9043	0.9807
MSE	N=7	2.8032e+004	2.7608e+004	2.6493e+004
	N=10	2.7913e+004	2.7378e+004	2.6351e+004
	N=12	2.7703e+004	2.6940e+004	2.6188e+004

The registered values of the MCN show that the number of the detected prototypes or clusters in RGNG technique is less than the others; which means that its ability to group data in actual number of clusters. For example, when N set to 10, the MCN

value for RGNG is 10 which is less than that for NG and GNG values. The MCN value of running RGNG is equal to the number of neurons 10 and has the same rate when compared with other N values; while MCN value of running NG and GNG deviated from the actual cluster number.

Regarding to PQ value, it is noticed that RGNG approach possesses higher PQ values than the NG and GNG. For example, when N set to 12, the PQ value for RGNG is 0.9807 which is higher than that of NG and GNG values. These high values of PQ indicate that the RGNG technique has a better partitioning quality with respect to others, and finds more representative clusters.

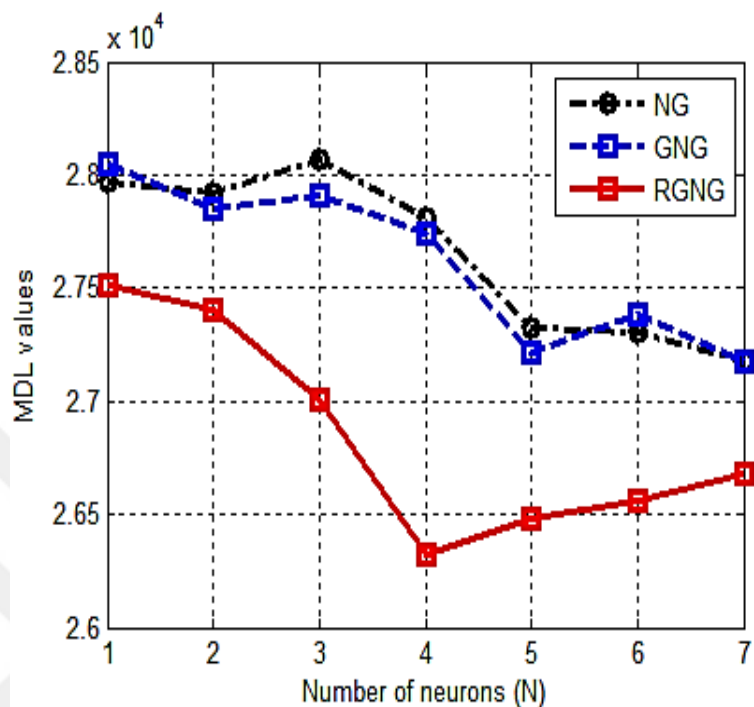
Moreover, RGNG method can find all the natural clusters during the growing stage with the correct number of prototypes. So, MSE values are lower which indicates obtaining better robustness of the RGNG technique. For example, when N set to 7, the MSE value for RGNG is $2.6493e+004$ which are lower than that for NG and GNG values. NG and GNG may not detect all the actual clusters, so they yield higher MSE value.

In RGNG technique, MDL value is one of the famous information theory evaluation measures which was used as the clustering validity index [104]. MDL criterion gives the ability of finding the optimal number of clusters and their center positions, corresponding to the smallest MDL value.

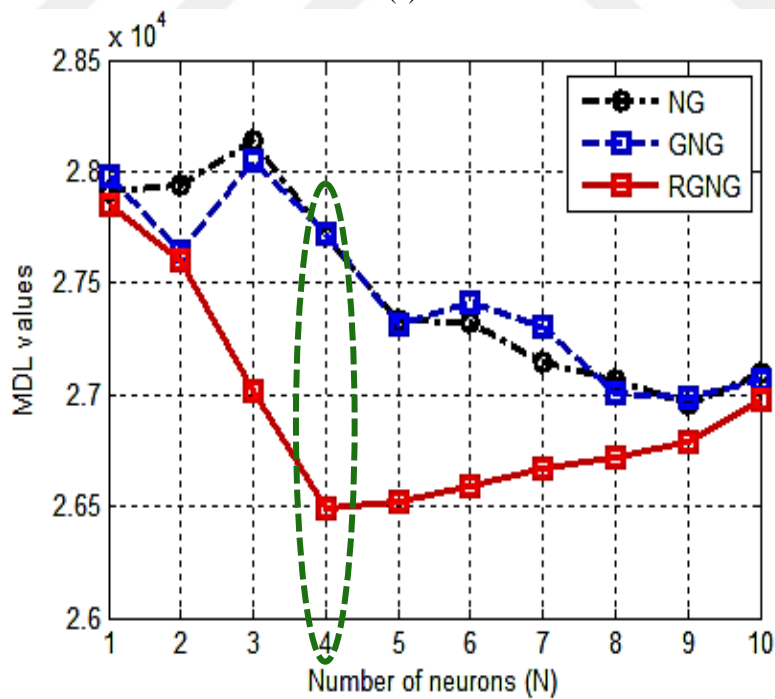
The average MDL values during the growth stages are plotted versus the number of clusters or prototypes. Figure 6.4 shows the curves for NG and GNG combined with MDL criterion, as well as RGNG approaches on the synthetic dataset for different number of neurons which are selected randomly as $N=7, 10$ and 12 . Each detected is the cluster number corresponding to the MDL value.

In RGNG the smallest MDL value was recorded on average with respect to NG and GNG combined with MDL principle. For example, in Figure 6.4 (b) the smallest MDL value is 2.65 that are obtained from running RGNG when N is equal to 4. While in the same $N=4$, higher MDL value is recorded as 2.77 from running NG and GNG. From the presented figures, it is concluded that the proposed RGNG approach

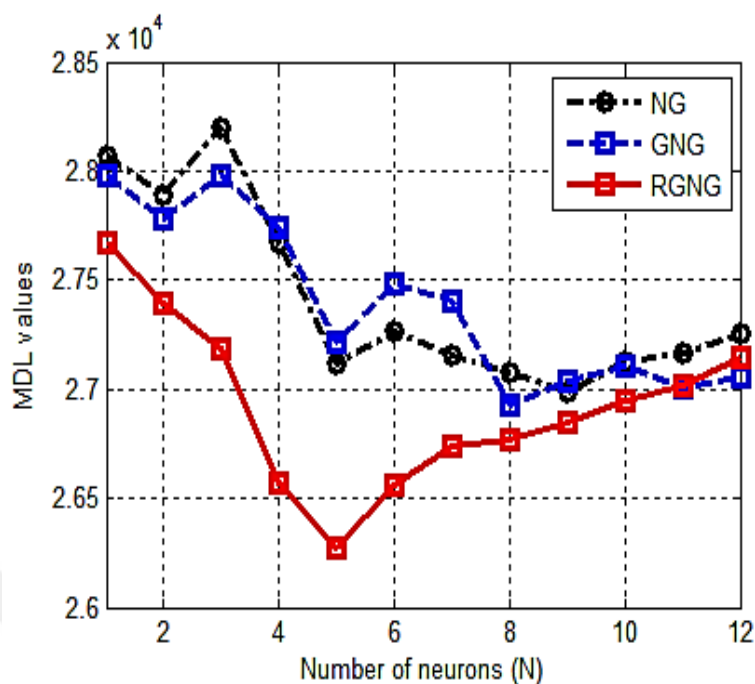
is insensitive to different initializations and the presence of outliers and can successfully find the actual number of clusters.



(a)



(b)



(c)

Figure 6.4 MDL values versus the number of clusters running NG, GNG and RGNG techniques on synthetic data, for: (a) N=7; (b) N=10; (c) N=12.

6.3 Prototype-Based Clustering Package

The techniques introduced in this work are designed and implemented in a simple software package tool which allows users to interact with the clustering techniques and output data easily [105]. Figure 6.5 shows the main window with the most important features of the designed Prototype-Based Clustering software package.

1. *Selection Data*: The user can select one type of data from the different synthetic 2D datasets inside the Pop-up Menu. Ring data is a 2D synthetic data selected as an example in Figure 6.5.
2. *Load Data*: The selected data are loaded and all information related to the selected data ('Dimension', 'Name' and 'Type of Data') are appeared in the 'info' window. The dimension of the selected 'Ring' data is 400x2 double. The selected data is plotted on the sketch1 inside the main clustering window of Figure 6.5.

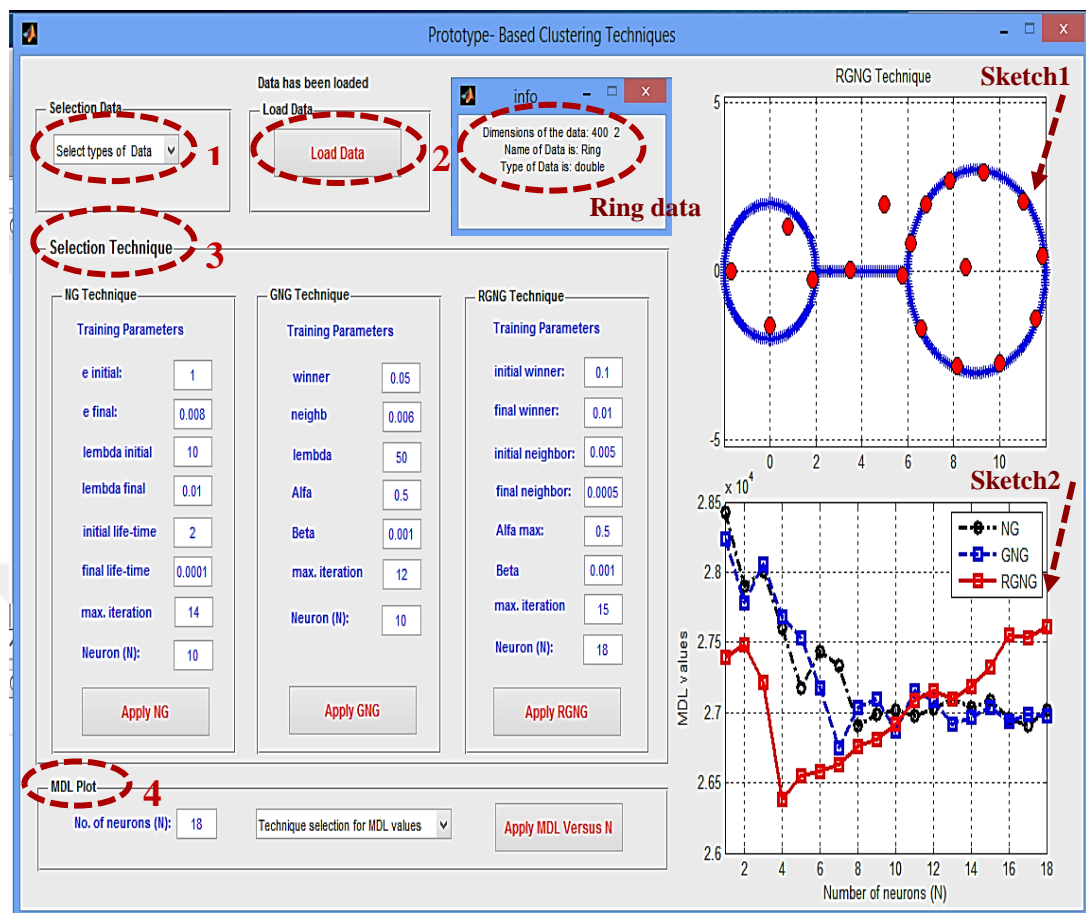


Figure 6.5 Main window of the Prototype-Based Clustering software package.

Figure 6.6 shows some of selected 2D synthetic dataset from the different dataset that used in this work. Beside each plot, the information related to it is shown in the 'info' window, in the left side of each plot.

3. *Selection Technique*: The user can select one of the clustering techniques NG, GNG or RGNG. RGNG technique is selected as an example for the training in Figure 6.5 in the Ring data with $N=18$, which is selected roughly.

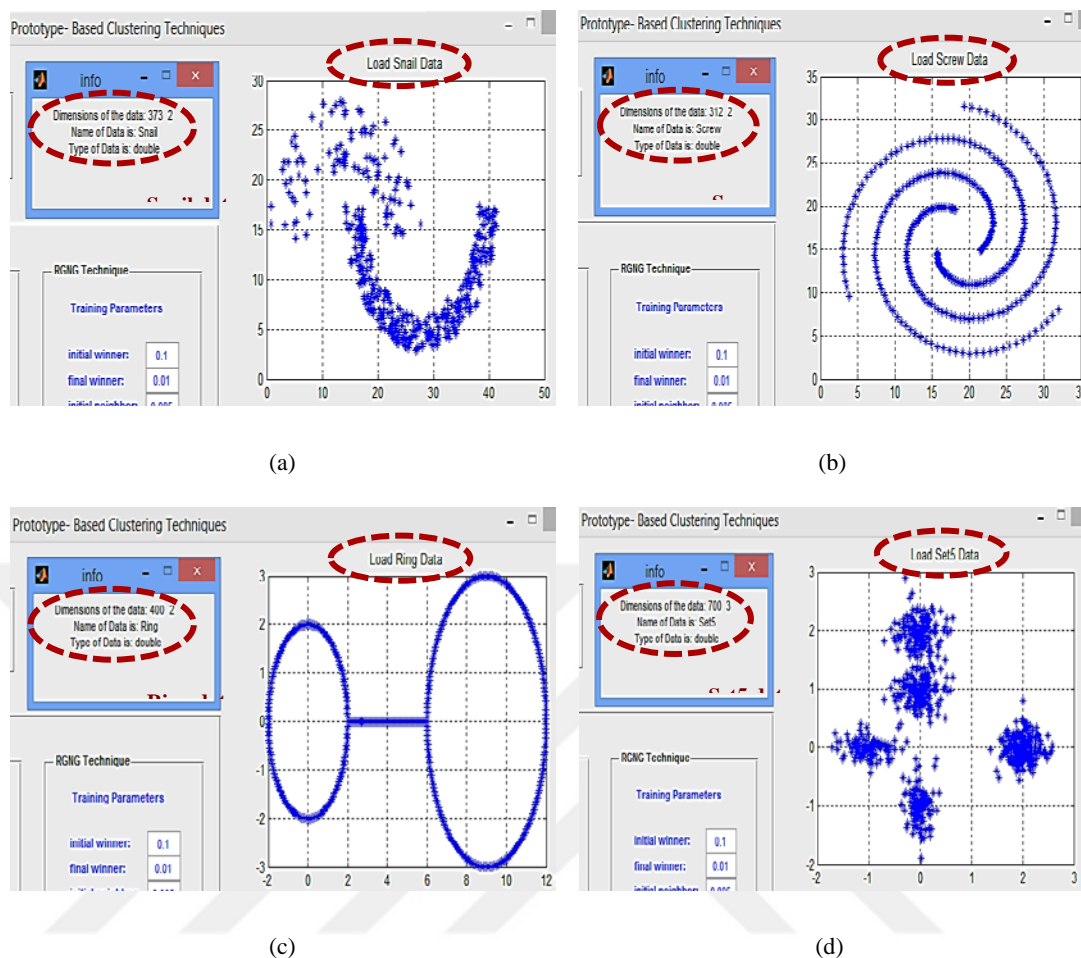


Figure 6.6 Different datasets with their information, (a) Snail data; (b) Screw data; (c) Ring data; (d) Set5 data.

Before clicking on “Apply NG”, “Apply GNG” or “Apply RGNG” button, the training parameters related to each technique must be defined. As it is explained in Section 6.1, the training parameters must be set carefully within the limited range. Number of neurons (N) must be defined also as the other parameters related to the selected technique. Another example of using RGNG technique with Set3 dataset is shown in Figure 6.7. RGNG training parameters are set as the typical values in the literature to: $\varepsilon_{bi} = 0.1$, $\varepsilon_{bf} = 0.01$, $\varepsilon_{ni} = 0.005$, $\varepsilon_{nf} = 0.0005$, $\alpha_{max} = 100$, $k = 1.3$, $\eta = 1 \times 10^{-4}$; the number of neurons (N) is chosen randomly to 14. When the algorithm’s training is started, the program sketches the output running of the implemented technique on the Sketch1. In Sketch1, a Set3 data is shown with firm red circles which represent the actual cluster centers.

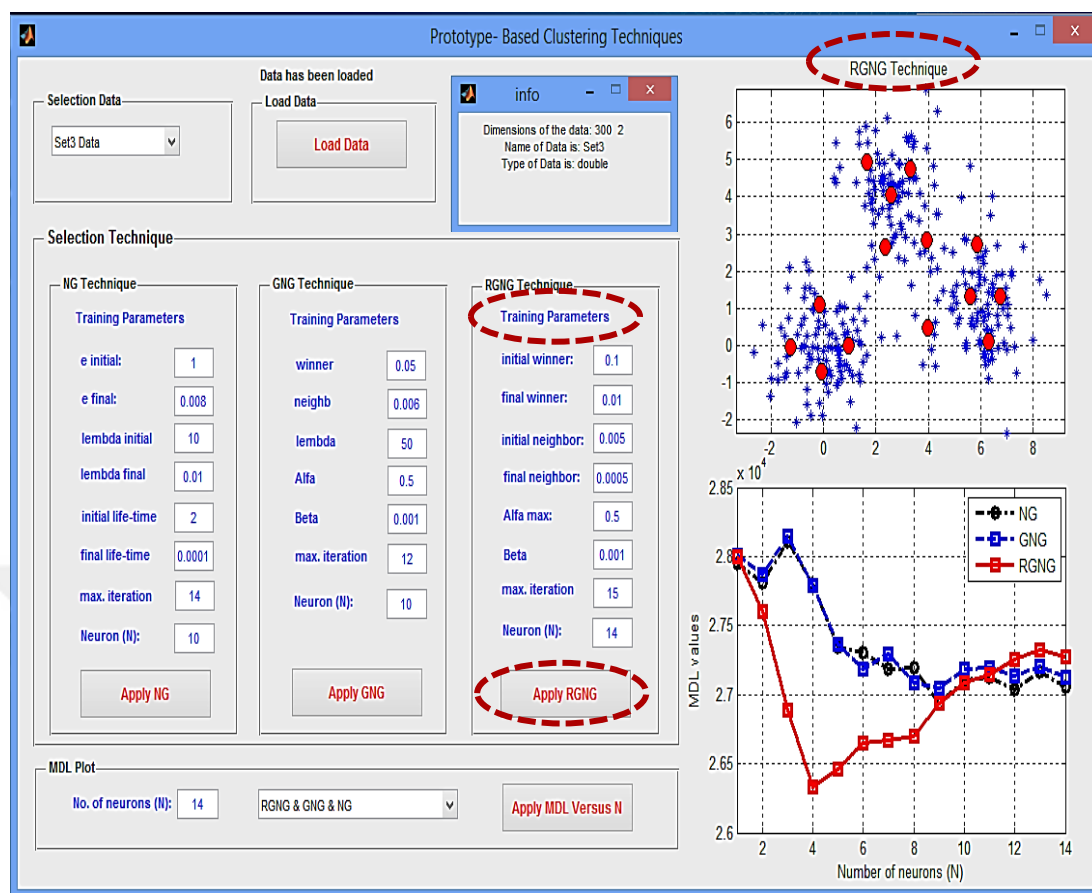


Figure 6.7 RGNG clustering with Set3 data (N=14).

4. *MDL Plot*: This panel related to plotting MDL values versus the number of neurons (N) running the RGNG, GNG and NG combined with MDL criterion. This panel includes three main buttons include “No. of neurons (N)”, “Technique selection for MDL value” and “Apply MDL versus N” buttons, as shown in Figure 6.8.

After defining the number of neurons (N); one, two or three of the training techniques have to be selected for comparing the MDL results. Inside the “Technique selection for MDL value” Pop-up Menu, there are seven selections; either show the result of each technique alone, two of them or three of them for the comparison easily. After clicking on the “Apply MDL versus N” button, the output results of MDL values are plotted with respect to the number of neurons (N) in Sketch2.

Figure 6.8 shows an example of the MDL plot, defining N=16 and choose “RGNG & GNG & NG” for comparison the results of the three techniques in Sketch2. For easy

and best comparison between the MDL values of the three techniques, the output results sketch in the same figure.

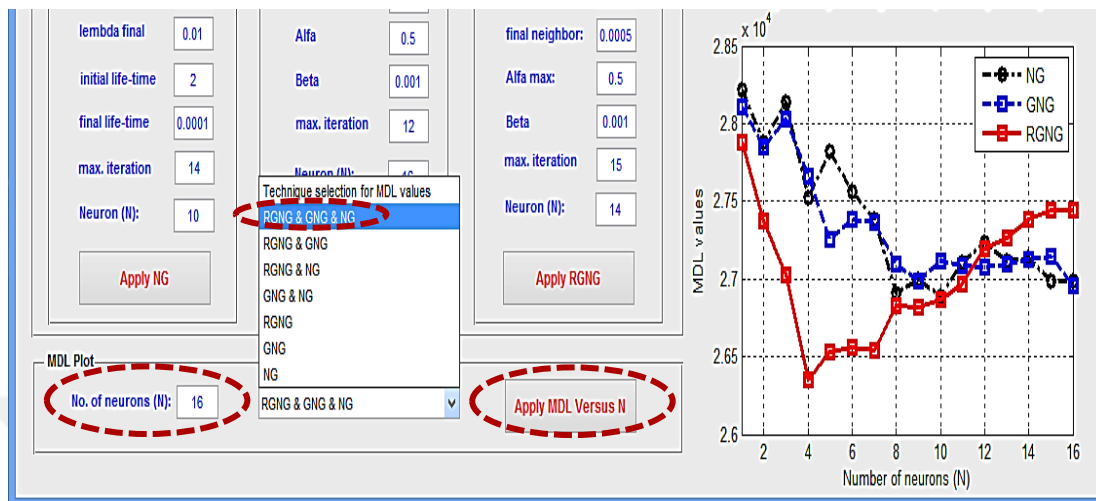


Figure 6.8 Comparison of MDL values for N=16.

6.4 Experimental Results with the Real fMRI Data

The principles behind the prototype-based clustering techniques are designed and introduced. So in this section, the validity of RGNG performance with fMRI experiments is proved. Since fMRI analyzing involves known areas and functions of the brain, the well-known expected results have to be used in the experiments.

One of the known areas is the auditory cortex. So, a real auditory fMRI dataset has selected to be used in this work. It is freely available for education and evaluation purposes. These dataset comprises whole brain BOLD/EPI images acquired on a modified 2T Siemens MAGNETOM Vision system. The auditory fMRI dataset was conducted by Geraint Rees under the direction of Karl Friston and the FIL methods group [60]. It is the first ever collected and analyzed in the FIL and is known locally as the mother of all experiments (MoAE). The data were used by some works as Lachiche et al., (2005), Korczak J., (2007) and Heydar et al., (2009); as shown in the literatures previously. Lachiche et al. demonstrated a new interactive data mining approach on a typical auditory fMRI dataset, and showed that the GNG has successfully recognized the active areas in brain fMRI images. Korczak introduced a

new interactive data mining technique to fMRI images by applying five clustering data-driven techniques to synthetic and real auditory fMRI data. The experimental results showed that the GNG technique performance was the best among the presented clustering technique. Heydar et al. developed the algorithm of the GNG network, which can run the optimal number of clusters automatically. The experimental results used artificial and real auditory fMRI dataset with the proposed technique.

One of the decisive fMRI advantages is that fMRI studies do not require the analysis of a group of volunteers, but can produce valuable results at the level of single individuals. The analysis of single volunteers is crucial for studying small structures which show strong inter-individual variation [106, 107], as it is true for the auditory cortex shown in Figure 6.9.

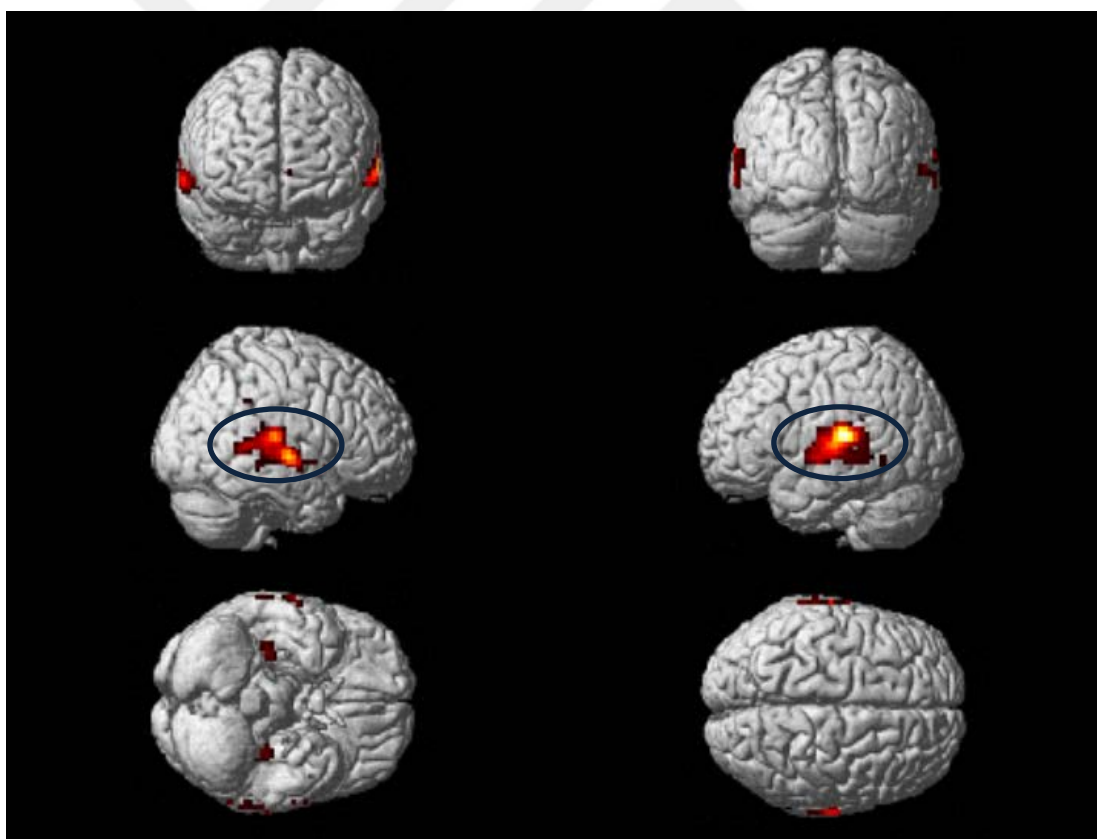


Figure 6.9 Active areas in the brain auditory cortex area using SPM package.

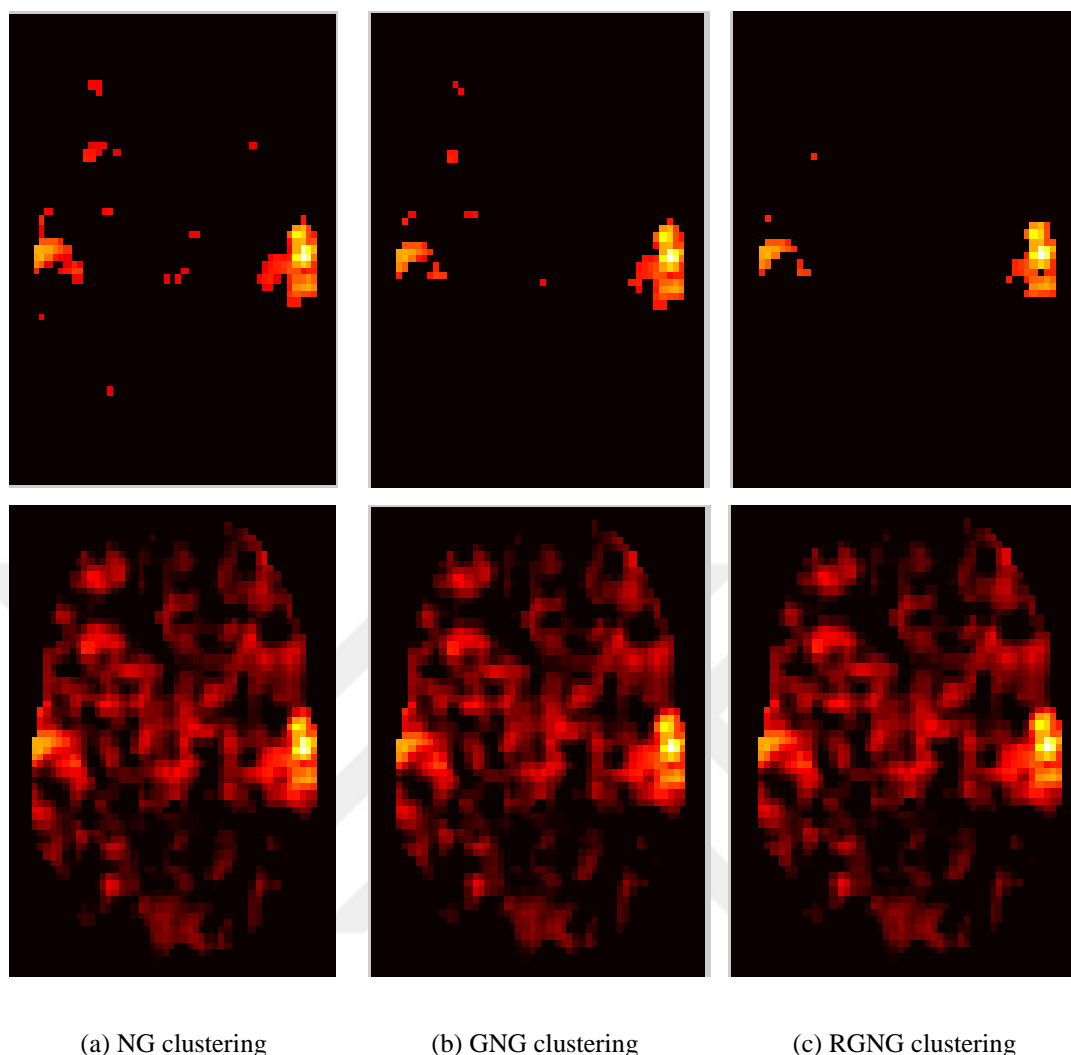


Figure 6.10 Results of the clustering techniques for a transparent brain image.

A block-design experiment was performed by using auditory stimulus. Figure 6.10 and Figure 6.11 show the active areas in the auditory cortex of whole brain running NG, GNG and RGNG techniques. Figure 6.10 shows the clusters in a transparent or glass brain image while Figure 6.11 shows the alignment of the obtained clusters into a structural space of the brain running NG, GNG and RGNG in Figure 6.11 a, b and c respectively.

Regarding to the output results obtained by running the three unsupervised clustering techniques; spatial information is visualized fine clusters in the auditory cortex area. In comparison to the previous literatures, NG did not have any application with the auditory data in the literatures; while auditory data had used with a GNG technique

previously [56- 58]. ROI has been obtained within the auditory cortex running the GNG technique as shown in Figure 6.11 (b) is similar to those delivered by the same approach in those literatures. In general, cluster is corresponds to a group of voxels with a similar Hemodynamic Response (HDR) over a Time Course (TC).

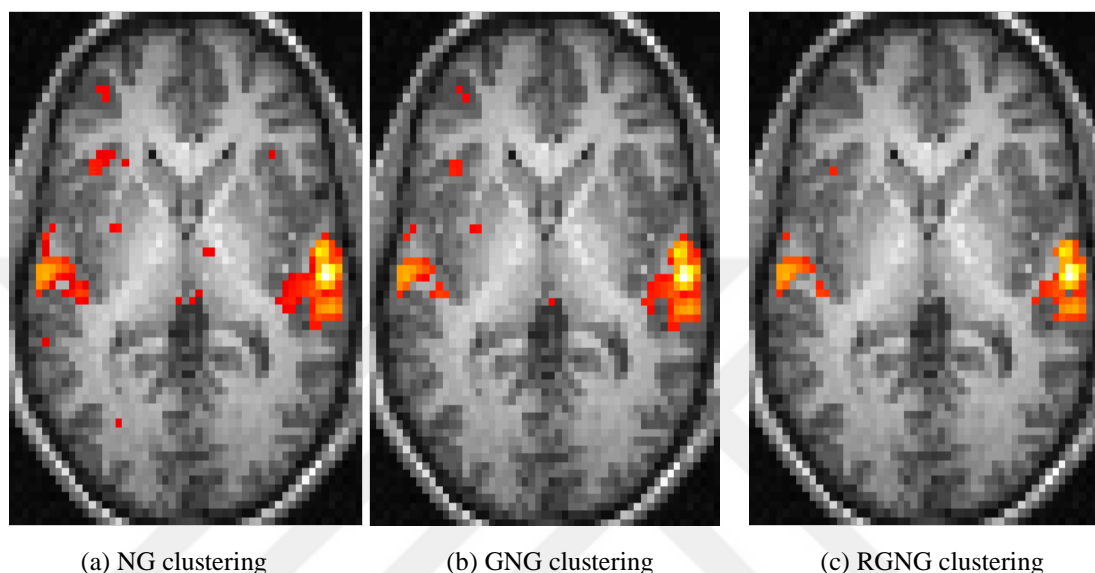


Figure 6.11 Clusters overlaid into the anatomical image.

The block-design experiment was performed by running the proposed RGNG approach using auditory data. The activation shown in Figure 6.11 (c) is located in the temporal lobe. The spatial information shows the areas of activation obtained are similar to those which expected from the auditory cortex experiments. The areas of activation are detected as variations of voxels intensity over time. In this case, the separating of the data delivered by the RGNG approach is used according to the TC signals of voxel intensity variations relative to its average. As all clustering techniques, RGNG attempted to portion homogeneous areas of activation in the brain. The homogeneous areas are comparable to those located in other approaches and that are found in the recognized cortexes related to the experiment. These areas or clusters are described by an average TC or a cluster center was obtained by averaging all the TCs of the cluster.

The novel application of RGNG output clustering results can be recognized as the best with respect to NG and GNG, due to clusters results which is defined in the specific auditory cortex area. Moreover the cluster's definition in the specific area, RGNG output clustering results as same as the results obtained from SPM using the same dataset with the same paradigm as shown in the next section.

6.5 Comparison of RGNG and SPM using Auditory Data

The paradigm of the block-design experiment alternates two conditions: without stimulus and auditory stimuli which is consisting of repetitions of two-syllable words as: “mother”, “house”, “weather”, “movie” and etc. Figure 6.12 (a) shows the ability of RGNG clustering technique to identify winner nodes. RGNG technique works on the optimal number of underlying clusters, and output a TC for activation detection in auditory dataset. Figure 6.12 (b) shows the stimulation brain activation of SPM with f -contrast test results with family-wise error (FWE) threshold, with no masking, the FWE-corrected p value = 0.05.

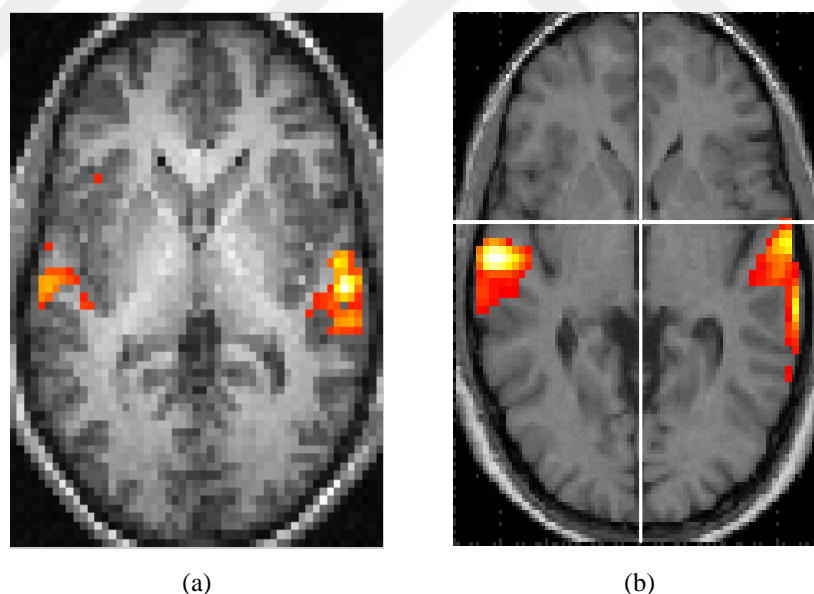


Figure 6.12 Activation zones of whole brain running: (a) RGNG; (b) SPM package.

The results of SPM based on GLM using the paradigm as a reference signal introduced bias in this experiment. In contrast, RGNG approach did not use the paradigm as the reference signal since it works as a model-free method. A model-

based method using GLM as the statistical analysis of fMRI image and hypothesis tests would be formed and consider the timing variation in the HDR. In sum, RGNG results were within the expectations and have similar results to those found by the hypothesis method in detecting active areas within the expected auditory cortexes.

The Receiver Operating Characteristic (ROC) analysis is another revealed index of the performance of RGNG in comparison with the SPM [108]. ROC is very popular in medical imaging and machine learning applications; ROC space consists of False Positive Ratio (FPR) on the x-axis and True Positive Ratio (TPR) on the y-axis [109]. Figure 6.13 shows the ROC space with good and bad classifiers in (a) and (b) respectively. The good classifier space is shown in a high TPR and low FPR; while the bad classifier space is shown in a low TPR and high FPR.

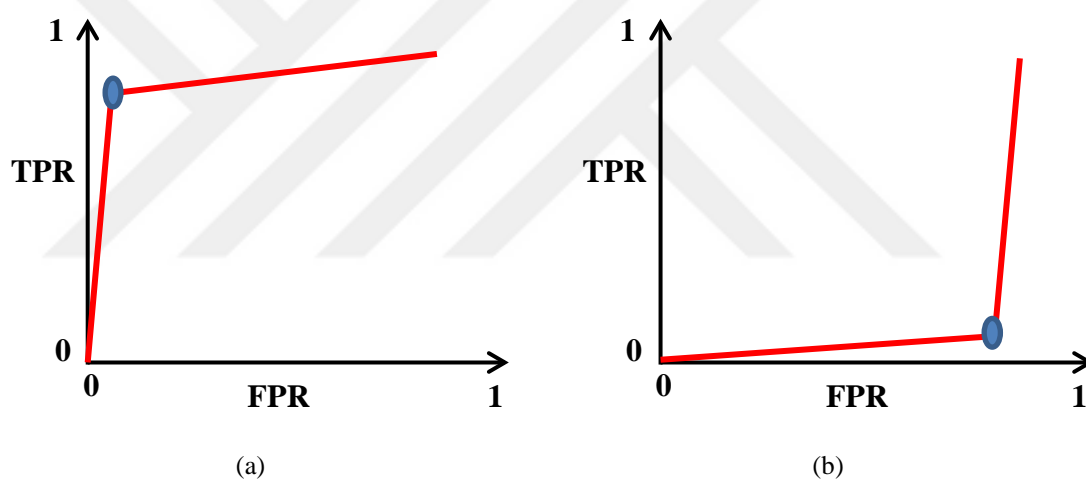


Figure 6.13 ROC analyses, (a) Good classifier; (b) Bad classifier.

In fMRI, the FPR is calculated by dividing the number of misclassified inactivated voxels by the total number of voxels considered, while the TPR is calculated by dividing the number of correct classifications of activated voxels by the total number of voxels considered [110]. In the same situation, the ROC curves for RGNG and SPM methods are compared in Figure 6.14.

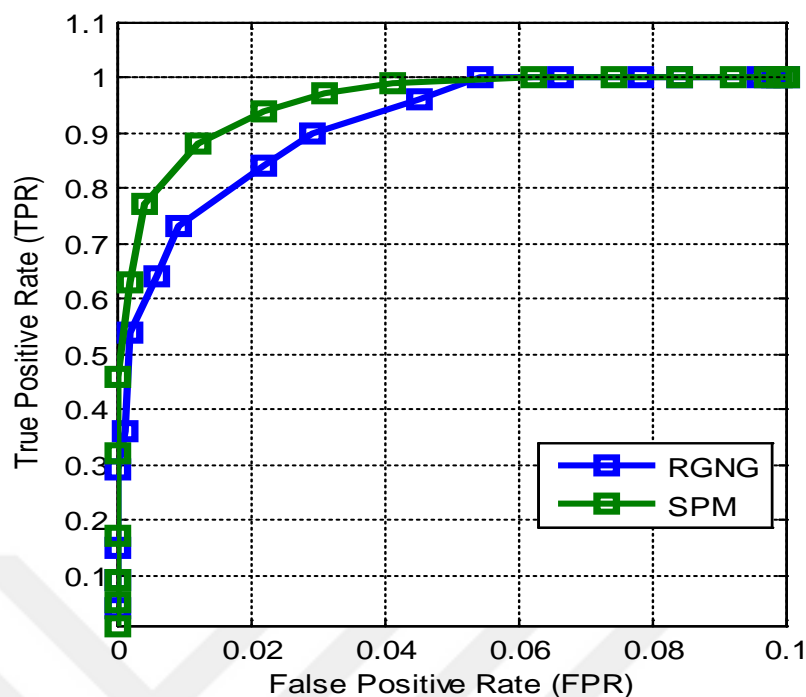


Figure 6.14 ROC curves of the auditory fMRI dataset.

These curves generally indicated the two methods work as a good classifier, with a high TPR and low FPR. RGNG method could detect more real activations under the same FPR ratio.

6.6 Visual fMRI Data Mining Package

The RGNG with a comparison to other unsupervised artificial neural network clustering (NG and GNG) techniques could be aligned in a certain software package to be able to detect the active zones in the brain fMRI images as well as viewing the fMRI images as a whole brain or slices. Figure 6.15 shows the main window with the most important features of the designed fMRI Clustering and Analysis package.

1. *Load fMRI Data*: The raw fMRI data are loaded with .nii or .img and .hdr format
2. *fMRI preprocessing*: By calling the SPM tool within the designed package. The spatial preprocessing functions are:
 - Realignment of the fMRI images
 - Coregistration between structural and functional data
 - Anatomical segmentation for CSF, grey and white matter (optional)

- Normalization of the data onto a standard anatomical MNI template
- Smoothing of the data using a Gaussian kernel of 8

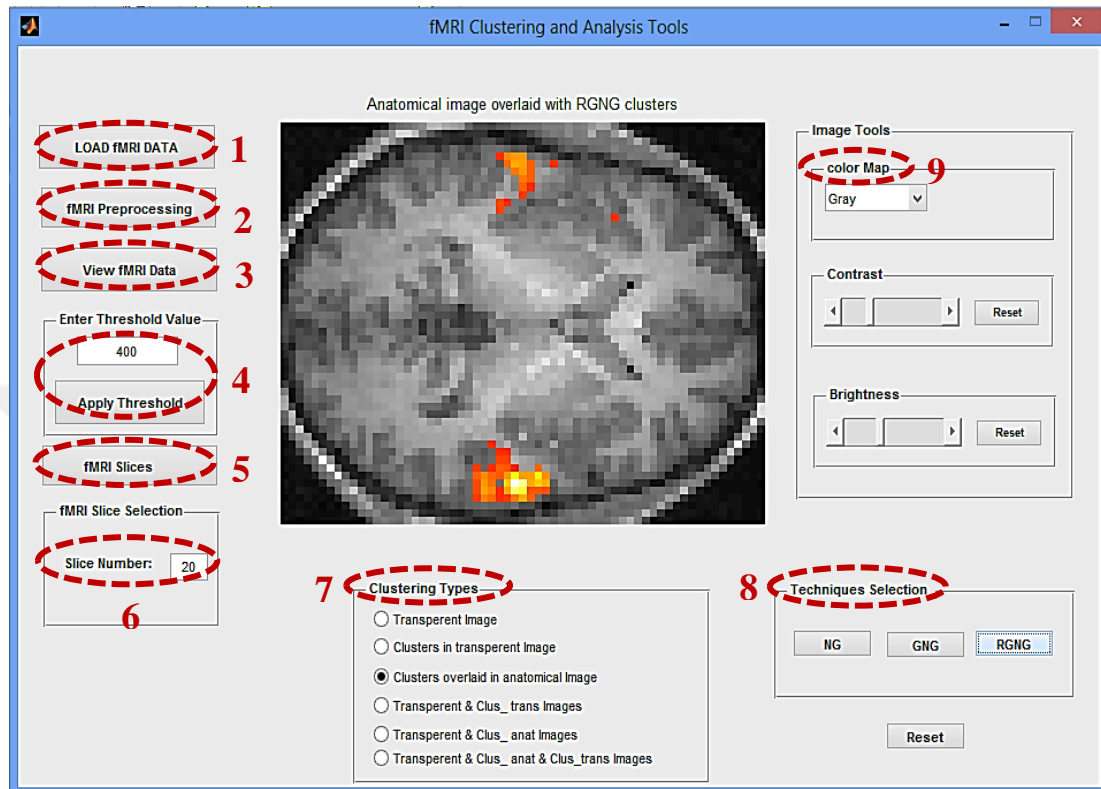


Figure 6.15 FMRI Clustering and Analysis software package.

Figure 6.16 and Figure 6.17 show the loaded fMRI data before and after preprocessing respectively. There is a pretty difference between fMRI images before and after preprocessing. In the analysis stage, the results that are used fMRI dataset without preprocessing will lead to the presence of a lot of unreal active areas in the shape of the image. These areas are looked as an activated, but in fact they are just an artifact and scattered all around the image and even outside the image.

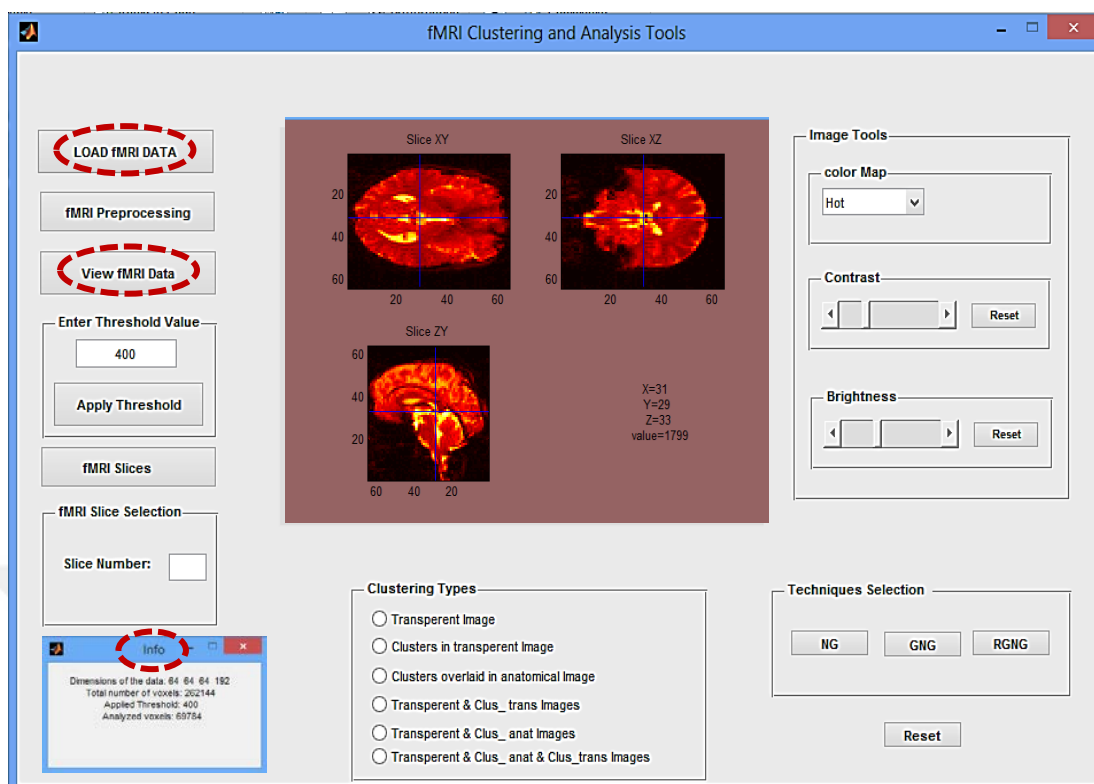


Figure 6.16 FMRI dataset for whole brain before the preprocessing

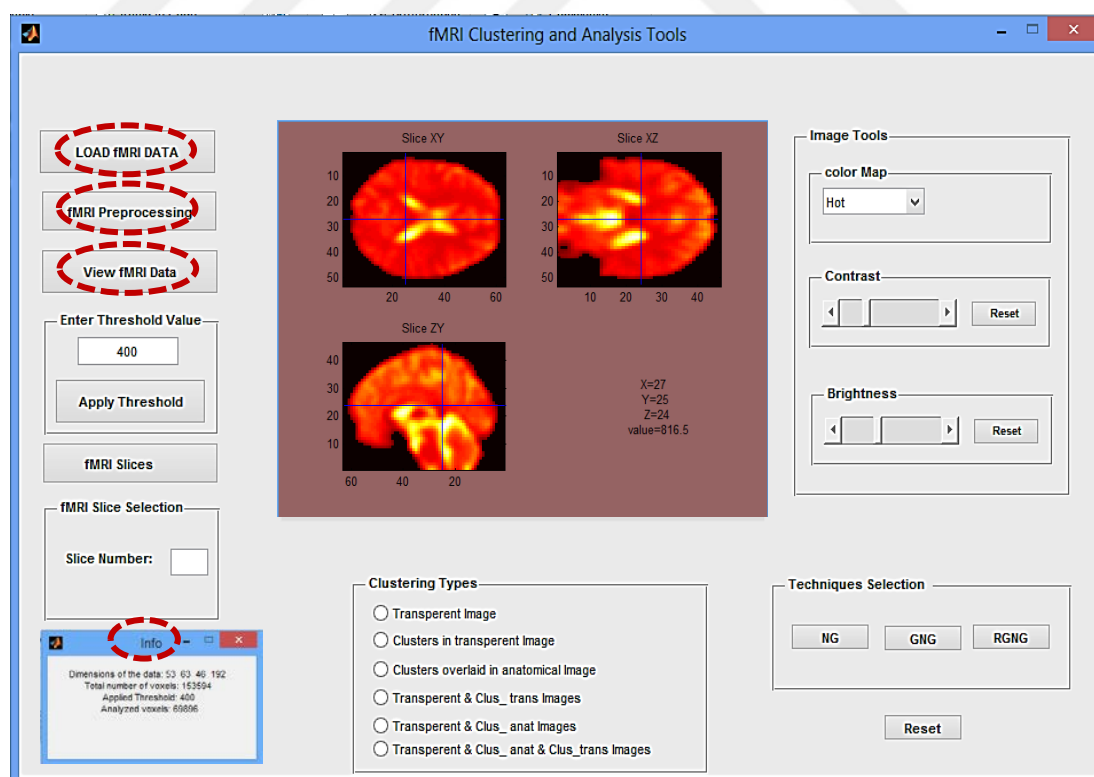


Figure 6.17 FMRI dataset for whole brain after the preprocessing.

Figure 6.16 shows the fMRI images for whole brain as it is obtained from the scanner, i.e. before preprocessing the raw data. In this figure, the auditory data is loaded with its dimension of 64X64X64 with 96 number of slices (192 slice in total included .img and .hdr for each slice). Figure 6.17 shows the fMRI images for whole brain after the preprocessing the raw fMRI data, i.e. after the smoothing process which is the final stage of the preprocessing steps. By clicking on the 'fMRI processing' button, the SPM tool is called within the designed package. In Figure 6.17, the same auditory data are loaded (before the preprocessing) with the same number of slices 96, but the dimension of data is changed to 53X63X46.

3. *View fMRI data*: View the raw fMRI data loaded by button 1 as an image.

Figure 6.16 and Figure 6.17 show the fMRI images when raw fMRI dataset is loaded using 'Load fMRI data' button and then view it using 'View fMRI data' button. When the raw data are loaded, 'info' window will be appeared, including the details of fMRI data as:

- Dimension of Data: represents the dimensionality of the data in X, Y and Z dimension as well as the total number of fMRI slices which selected as the whole brain
- Total number of voxels: the total number of voxels of the selected fMRI dataset
- Applied threshold: the defined threshold value
- Analyzed voxels: represent the number of analyzing voxels after applying the threshold value. For example, if the threshold value is zero, then the analyzed voxels value is the same as the total number of voxels. This value is considered in the clustering or training process.

4. *Apply Threshold*: After the definition of an appropriate threshold value for the best active area detection, press this button. FMRI data is huge, so the dimensionality reduction must be done before feeding the learning technique. A threshold definition is important in the data mining and decrease the size of dataset in order to exclude the voxels outside of the brain structure before analysis. In Figure 6.15, the threshold

number is chosen to 400, which is the best in this case to show the image clear without distortion.

5. *fMRI Slices*: The loaded fMRI data can be viewed as a demo of image slices. Figure 6.18 shows the fMRI images as brain slices before detecting active areas, slices 20 and 46 is taken as an example from the auditory fMRI dataset.

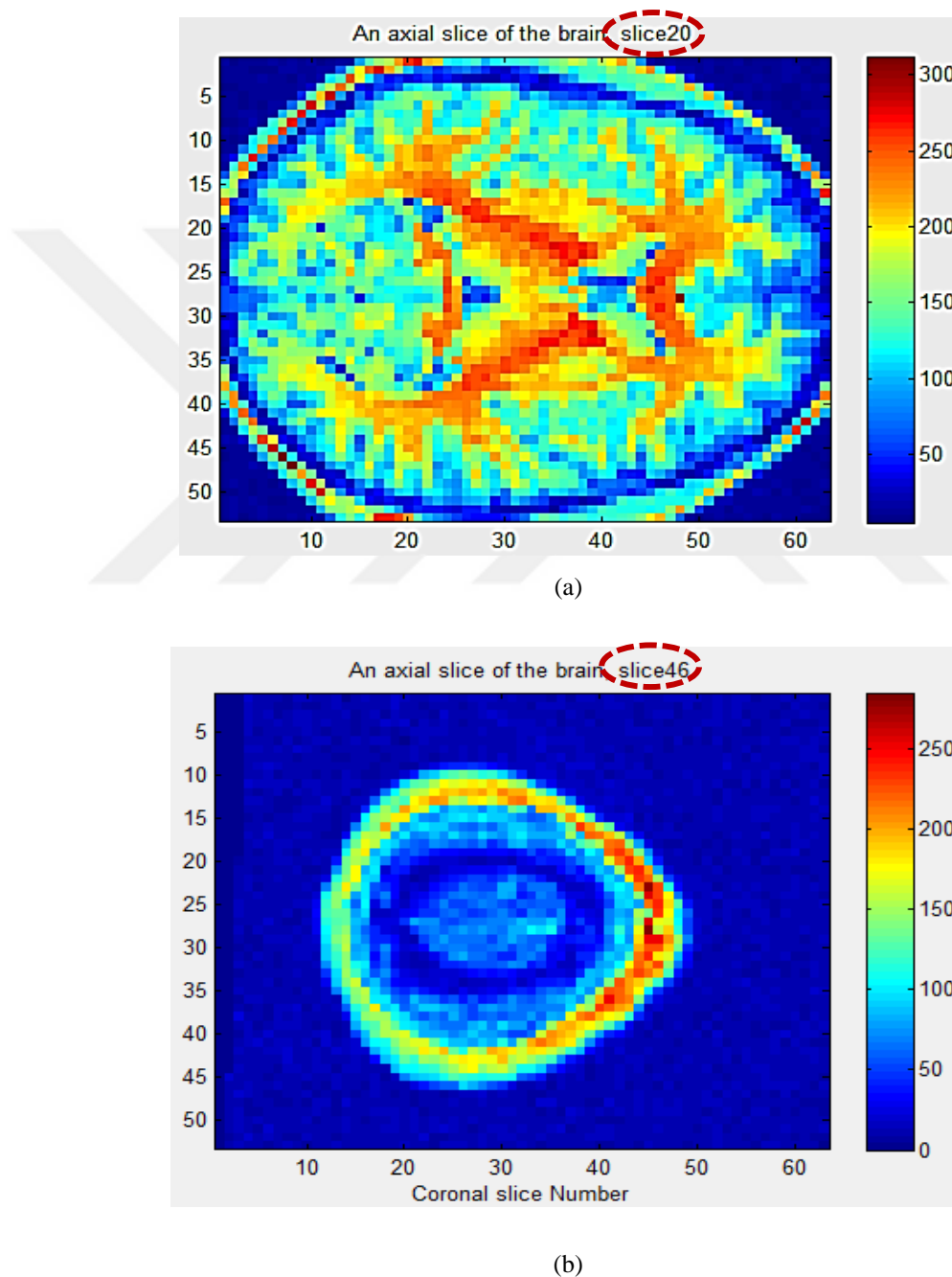


Figure 6.18 fMRI brain slices (a) slice no: 20; (b) slice no: 46.

6. *Slice Number*: The number of slices is choosing before passing to the next step and feed the clustering technique. In Figure 6.15, slice number 20 was chosen as an example which can show the responses or the areas of activation clearly.

7. *Clustering Types*: This panel can give a facility in appearing the clusters in the image. The clusters can be shown as: a transparent image, clusters in the transparent image or overlaid with the structural image. Other choices are viewing two or three of them together. Figure 6.19 shows the RGNG clusters on a transparent or clear image as an example.

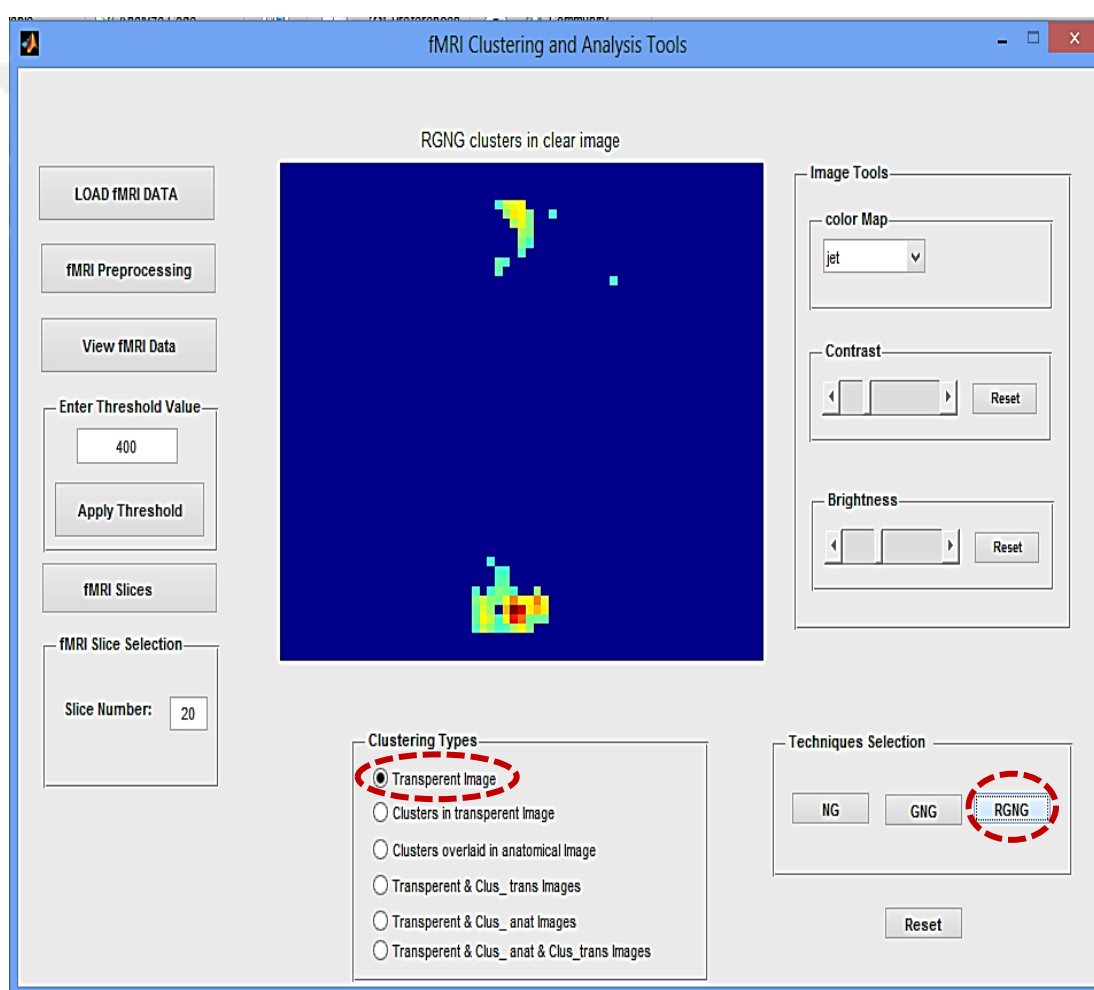


Figure 6.19 RGNG clusters on a transparent image.

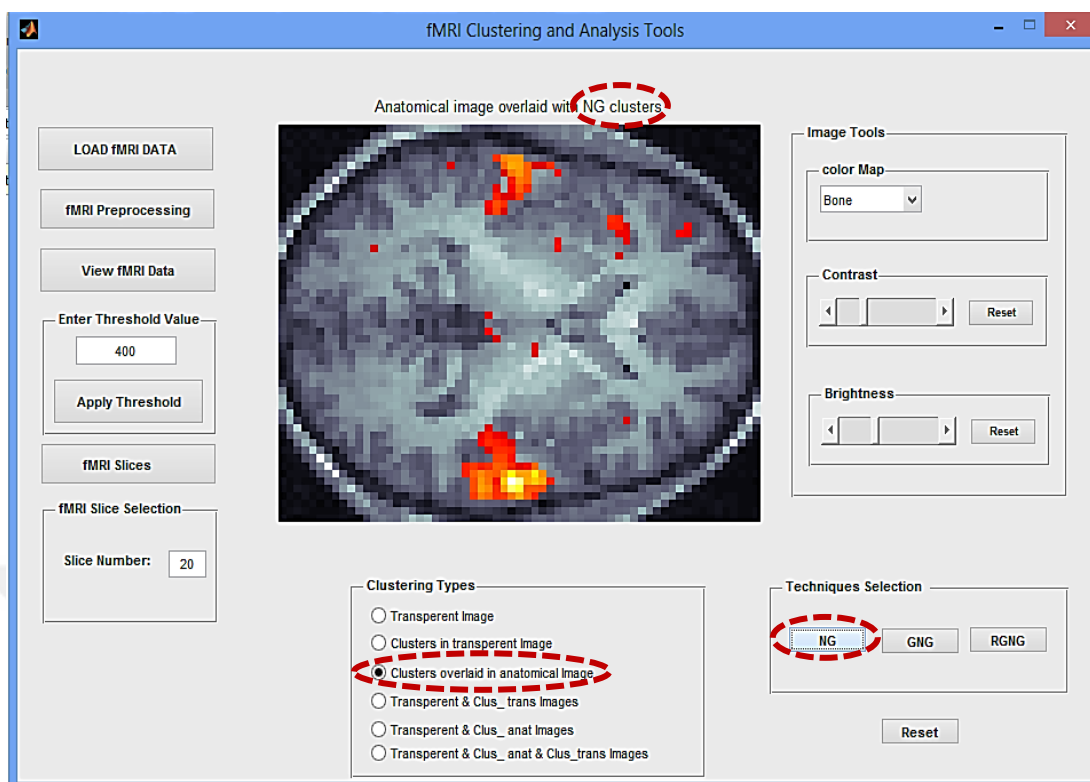


Figure 6.20 NG clusters overlaid in the anatomical image.



Figure 6.21 GNG clusters overlaid in the anatomical image.



Figure 6.22 RGNG clusters overlaid in the anatomical image.

8. *Techniques Selection*: After defining the type of clusters to be shown as an image in the main window, one of the three unsupervised techniques NG, GNG or RGNG is chosen. The output results of running NG, GNG and RGNG as clusters overlaid onto the anatomical brain image as shown in Figure 6.20, Figure 6.21 and Figure 6.22 respectively.

9. *Color Map*: By selecting one of the color map defined in this Pop-Menu; the image color can be changed. The image color in Figure 6.15 is 'gray', Figure 6.16 and Figure 6.17 are 'hot', Figure 6.19 is 'jet', while Figure 6.20, Figure 6.21 and Figure 6.22 are 'bone'. Brightness and contrast adjustments are used in this GUI for easy and better viewing of fMRI image.

In this system, a number of clustering methods are implemented within a user-friendly data mining software package based on GUI using MATLAB [111, 112]. This tool gives the facility of selecting which number of slices can show the active zones in the specified area. With the help of designing, the physician and

neuroscience can process any fMRI raw data directly within this package as well as dealing with the introduced techniques easily.

6.7 Discussion

The proposed research is highly interdisciplinary and combines different aspects together, including artificial neural network technique, software package design, data mining and dealing with fMRI dataset. A novel and extensive simulation studies on real fMRI datasets are carried out using the RGNG unsupervised clustering technique in this work. Results show that this method could complement the model based method to cope with the difficulties and challenges in fMRI data analysis. This may participate in better recognition of the nature of the fMRI data and the underlying mechanisms. The software package are designed and implemented in this work which can provide:

- an easy way to import the raw fMRI dataset;
- view the imported data as image;
- select desired processing and analysis methods as NG, GNG or RGNG;
- save all the output results by clicking a few buttons.

The main contribution of this research is running a new and robust data-driven RGNG technique on fMRI data mining, which is not used previously in any medical applications. This technique has different features in comparison to other techniques used with fMRI as its insensitivity to different initializations and the presence of outliers, as well as its ability to determine the actual number of clusters successfully. For example, the statistical GLM model is the most popular method which is used with analyzing of fMRI dataset. A potential problem associated with GLM model is the requirement of an accurate estimate of the fMRI paradigm design. In different cases, it is difficult to provide precise model designs; either the problem from the subjects who doing the task incorrectly (also the same subject may give a different response for same paradigm at different time) or different subjects may still give different BOLD signals during the same paradigm.

RGNG clustering results were the best according to different proofs using:

- MDL: RGNG recorded the smallest MDL value (as an indication of finding the optimal number of clusters and their center positions) with respect to NG and GNG combined with MDL principle.
- ROC: RGNG detects more real activations under the same FPR ratio, in spite of RGNG and SPM work as a good classifier.
- Clustering brain images: RGNG was defined in the specific area with no unknown areas of activation clusters in comparing to NG and GNG. So, RGNG results were within the expectations and as same as SPM output.

The findings from this research work can deal with various difficulties that the neurologist and psychologist facing while analyzing the measured fMRI data. Also, the different software packages introduced in this work can help healthcare specialists and researchers deal with this subject easily.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

The research work presented in this work provides an illustrative overview of the fMRI technique, including: the physical principles, fMRI brain clustering and parcellation technique principle, preprocessing techniques and data analysis approaches. The source of fMRI dataset acquires from the scanner machine in a form of raw data as sequences of 3D images due to the variations of voxels intensity over time. There are different noises factors interfere with the fMRI signals of interest and typically the subject is never completely motionless. So the pre-processing steps must be adapted to each identified artifact before the clustering phase.

The major objective of this work is the detection and classification of the activated areas of the brain using a robust and efficient RGNG technique. There is no such study has been done according to the researcher knowledge, and this thesis related to using RGNG with fMRI dataset would be the first time in the literature.

RGNG has the ability to detect the active zones in the brain, analyze brain function as well as its ability to work on the optimal number of underlying clusters with respect to the MDL value in fMRI dataset. This technique can define the positions of the output cluster's center corresponding to the minimal MDL value.

The validity of the performance of RGNG technique is tested by a real auditory fMRI data which is based on the stimulation of the auditory cortex. As performed in this thesis, the process of the brain function data analysis is composed of five stages:

- preprocessing of the raw data
- clustering voxels together based on the similarity of their intensity profile in TCs of the image
- overlay with the structural image
- visual fMRI image

- validation

There are some difficulties addressed by using the conventional clustering techniques, such as the number of clusters must be defined earlier and the cluster detection problem with different dimensions within the same data set. RGNG merges the GNG structure with robust properties as well as using MDL for defining the problems of optimal network representation and further parameter. These features made RGNG insensitive to the initializations, input sequence ordering, and the presence of outliers; and more robust towards noisy input data. During the network growing process, RGNG can find effectively the optimal number of clusters and its corresponding positions which are closer to the actual cluster centers (with the smallest MDL value) with little influence by the outliers.

Experimental output results showed the superior performance of RGNG over model-based and several existing prototypes-based clustering techniques on both 2D and real fMRI datasets in static data clustering tasks as revealed by their performance measured by MDL and ROC analysis. This research work proposed a novel and powerful methods for fMRI data analysis which integrates the advantages of both hypothesis and exploratory analysis methods. The findings from this work could help and cope with various difficulties that the neurologist and psychologist are facing while analyzing for better interpreting of the fMRI data.

In the presented work there are two packages which are implemented and designed to demonstrate the output results. First one included the dealing with the prototype-based unsupervised clustering approaches proposed in the thesis (NG, GNG and RGNG), while the second was used as a visual tool for fMRI data mining. First package deals with the synthetic dataset, while the second with the real fMRI dataset to define the active zones. The interactive tools are integrating the researcher into the process of the analysis of the used clustering techniques as well as the detection of functional areas in the brain and their system.

The proposed fMRI package is designed for the Brain fMRI Parcellation using new model design for neuroscience data analysis, and the proposed RGNG, GNG and NG techniques are performed in the same package using a MATLAB-based GUI tool.

There are several tools for fMRI processing and analyses exist, but they are usually sophisticated, need a significant amount of time to learn, and there is still the need for programming to use those tools. So, this visual tool, with a comparison to the model-based approaches, is easy to learn and can be used by most of doctors, physicians and healthcare specialists who do not have a good programming background as the engineers and programmers.

7.2 Future Directions

Two types of fMRI analysis methods were presented in the thesis as a comparison: GLM and data-driven analyses using machine learning classifiers. GLM is the most common method for fMRI data analysis, but it is based heavily on priori BOLD model design. In some cases, GLM couldn't be used for brain activation detection when the previous information about the data is unavailable. For example, the subject with mental cases or during daydreaming and mind-wandering researching (default mode of brain function). Also, the same action could not be performed in similar with different subject during the same paradigm, and the same subject could not repeat the same stimuli in different times.

In this work, an effective alternative approach using data-driven analysis was introduced for detecting brain activity based on the data structure itself. The proposed application of the RGNG on real fMRI dataset is reviewed on a single subject auditory fMRI data. It would be good to see the extension of this method to multi-subject data driven analysis (multiple subject data) of fMRI dataset as well. Paradigm design of auditory dataset is block type data design, the output results from this work allows to extend this work towards event-related design data experiments where the signal to-noise ratio is weak and noisy data mask the relevant information.

Since RGNG can deal well with fMRI which is multimodal data sets, so it is suggested to be applied to other real multimodal data sets such as MRI image segmentation not only the brain but other regions of the body. So, the clusters of different organs shapes of the body can be detected using other distance metric measures since the Euclidean distance metric used with RGNG is able to detect the

clusters of the brain, which is an approximately spherical or ellipsoidal region in shape with small differences in the variance in each dimension.

As a future work, the researcher can suggest other cluster validity measures rather than the common MDL criterion used with RGNG in this work. For example, Minimum Message Length (MML), Bayesian Information Criterion (BIC) and Akaike's Information Criterion (AIC) could be proposed to tackle this issue as.

The fMRI package implemented in this work using MATLAB GUI can be advanced in the future to be used with many fMRI data analysis approaches as well as other medical imaging modalities. Furthermore, there is an ability of applying this work on ASP .NET in the future; so the packages could be shared to website for telemedicine experiences between centers and researchers using just a Web browser. With the presented appropriate application, the proposed work could provide a good chance for more advanced neuroscience studies.

REFERENCES

- [1] Huettel, S. A., Allen W. S., & Gregory M. Functional Magnetic Resonance Imaging, (2nd ed.). Sinauer Associates, Inc. 2009
- [2] McKie, S., & Brittenden J. (ii) Basic Science: Magnetic Resonance Imaging. Current Orthopaedics, 19 (1), 13-19. 2005
- [3] Peter, J., Paul, M. M., & Stephen, M. S. Functional MRI: An Introduction to Methods. Oxford University Press. 2001
- [4] Anderson, IM., Haddad, PM., & Scott, J. Bipolar Disorder. BMJ, Clinical research ed. 2012
- [5] Anthony, G. New Drugs May Help with Parkinson's disease. eMissourian, New Drugs Ma. 2014
- [6] Clery, H., Andersson, F., Bonnet-Brilhault, F., Philippe, A., Wicker, B., & Gomot, M. fMRI Investigation of Visual Change Detection in Adults with Autism. NeuroImage: Clinical 2, Elsevier Inc., 303–312. 2013
- [7] Donohoe, G., Duignan, A., Hargreaves, A., Morris, D.W., Rose, E., Robertson, D., Cummings, E., Moore, S., Gill, M., & Corvin, A. Social Cognition in Bipolar Disorder Versus Schizophrenia: Comparability in Mental State Decoding Deficits. Bipolar Disorder, John Wiley and Sons A/S., 14 (7), 743–748. 2012
- [8] Anthony, D. W. Early Detection of Alzheimer's Disease: An fMRI Marker for People at Risk?. Nature America Inc., Nature Neuroscience, 3 (10), 973-4. 2000
- [9] Brian, M. D., Mark, A. B., & Richard, C. S. MRI: Basic Principles and Applications. (5th ed.), Wiley-Blackwell. 2015
- [10] Luigi, L., Vincenzo, P., & Maria, F. S. Advanced Image Processing in Magnetic Resonance Imaging. Boca Raton: Taylor & Francis Group, LLC. 2005

- [11] Russell, A. P., Jeanette, A. M., & Thomas, E. N. Handbook of Functional MRI Data Analysis. Cambridge University Press. 2011
- [12] Richard, B. B. Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques. (2nd ed.), Cambridge University Press. 2009
- [13] Amaro, E.J., & Barker, G.J. Study Design in fMRI: Basic Principles. *Brain and Cognition*, 60, 220-232. 2006
- [14] Aljobouri, H. K., Çankaya, I., & Karal, O. From Biomedical Signal Processing Techniques to fMRI Parcellation. *Biosciences Biotechnology Research Asia*, 12 (2), 1115–1138. 2015
- [15] Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Roger, W., Tomas, P., Gregory, S., Bruce, P., Colin, H., Louis, C., Paul, T., David, M.D., Marco, I., Thorsten, S., Katrin, A., Nicola, P.-G., Stefan, G., Larry, P., Katherine, N., Noor, K., Georges, L. G., Dorret, B., Tyrone, C., Ryuta, K., & Bernard, M. A Probabilistic Atlas and Reference System for the Human Brain: International Consortium for Brain Mapping (ICBM) *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 356, 1293–1322. 2001
- [16] Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *Neuroimage*, 15 (1), 273–89. 2002
- [17] Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., Poldrack, R.A., Bilder, R.M., & Toga, A.W. Construction of a 3D Probabilistic Atlas of Human Cortical Structures. *Neuroimage*, 39, 1064–1080. 2008
- [18] Bohland, J. W., Bokil, H., Allen, C. B., & Mitra, P. P. The Brain Atlas Concordance Problem: Quantitative Comparison of Anatomical Parcellations. *PLoS One*, 4 (9), 1-18. 2009
- [19] Retrieved 1991, from <http://www.fil.ion.ucl.ac.uk/spm/>. 1991

- [20] Retrieved 2012, from <http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>. 2012
- [21] Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. Functional Connectivity: the Principal Component Analysis of Large PET Data Sets. *Journal of Cerebral Blood Flow Metabolism*, 13 (1), 5–14. 1993
- [22] Friston, K. J., Poline J. B., Strother, S., Holmes, A. P., Frith, C. D., & Frackowiak, R. S. A Multivariate Analysis of PET Activation Studies. *Human Brain Mapping*, 4, 140–151. 1996
- [23] Hyvarinen, A., Karhunen, J., & Oja, E. *Independent Component Analysis*. John Wiley & Sons. 2001
- [24] McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., & Sejnowski, T. Analysis of fMRI Data by Blind Separation into Independent Spatial Components. *Human Brain Mapping*, 6, 160–188. 1998
- [25] Calhoun, V., Adali, T., Pearlson, G., & Pekar, J. Spatial and Temporal Independent Component Analysis of Functional MRI Data Containing a Pair of Task-Related Waveforms. *Human Brain Mapping*, 13, 43-53. 2001
- [26] Mckeown, M. J. Detection of Consistently Task-Related Activations in fMRI Data with Hybrid Independent Component Analysis. *Neuroimage*, 11, 24–35. 2000
- [27] Friman, O., Borga, M., Lundberg, P., & Knutsson, H. Exploratory fMRI Analysis by Autocorrelation Maximization. *Neuroimage*, 16, 454–464. 2002
- [28] Chen, H., Yuan, H., Yao, D., Chen, L., & Chen, W. An Integrated Neighborhood Correlation and Hierarchical Clustering Approach of Functional MRI. *IEEE Transactions on Biomedical Engineering*, 53 (3), 452–458. 2006
- [29] Seghier, M. L., Friston, K. J., & Price, C. J. Detecting Subject-Specific Activations Using Fuzzy Clustering. *NeuroImage*, 36, 594–605. 2007
- [30] Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., & Turner R.

Dynamic Magnetic Resonance Imaging of Human Brain Activity During Primary Sensory Stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, 89 (12), 5675–5679. 1992

[31] Ogawa, S., Tank, D.W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. Intrinsic Signal Changes Accompanying Sensory Stimulation: Functional Brain Mapping with Magnetic Resonance Imaging. *Proceedings of the National Academy of Sciences USA*, 89 (13), 5951–5955. 1992

[32] Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. Time Course EPI of Human Brain Function During Task Activation. *Magnetic Resonance in Medicine: Official Journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 25 (20), 390–397. 1992

[33] Friston, K. J., Jezzard, P., & Turner, R. Analysis of Functional MRI Time-Series. *Human Brain Mapping 1*, Wiley-Liss Inc., 153–171. 1994

[34] Aguirre, G.K., Zarahn, E., & D'Esposito, M. The Variability of Human, Bold Hemodynamic Responses. *Neuroimage*, 8 (4), 360–369. 1998

[35] Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. Variation of BOLD Hemodynamic Responses across Subjects and Brain Regions and Their Effects on Statistical Analyses. *NeuroImage*, Elsevier Inc., 21 (4), 1639–1651. 2004

[36] Cao, J., & Worsley, K.J. The Geometry of Correlation Fields, with an Application to Functional Connectivity of the Brain. *Ann Appl Probab*, 9, 1021–1057. 1999

[37] Woolrich, M. W., Behrens, T. E. J., & Smith, S. M. Constrained Linear Basis Sets for HRF Modelling Using Variational Bayes. *NeuroImage*, Elsevier Inc., 21 (4), 1748–1761. 2004

[38] Backfrieder, W., Baumgartner, R., Samal, M., Moser, E., & Bergmann, H. Quantification of Intensity Variations in Functional MR Images Using Rotated Principal Components. *Physics in Medicine and Biology*, 41 (8), 1425–1438. 1996

- [39] Goutte, C., Toft, P., Rostrup, E., Nielsen, E.F., & Hansen, L. K. On clustering fMRI Time Series. *NeuroImage*, by Academic Press, 9 (3), 298–310. 1999
- [40] Gao, J., Yee, S., & Jan. Iterative Temporal Clustering Analysis for the Detection of Multiple Response Peaks in fMRI. *Magnetic Resonance Imaging*, Elsevier Science Inc., 21 (1), 51–53. 2003
- [41] McIntosh, A. R., Chau, W. K., & Protzner, A. B. Spatiotemporal Analysis of Event-Related fMRI Data Using Partial Least Squares. *NeuroImage*, Elsevier Inc., 23 (2), 764–775. 2004
- [42] Beckmann, C., & Smith, S.M. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. *Medical Imaging, IEEE Transactions on Medical Imaging*, 23 (2), 137–152. 2004
- [43] Yi-Ou, L., Adali, T., & Calhoun, V. D. A Feature-Selective Independent Component Analysis Method for Functional MRI. *Journal of Biomedical Imaging*, Hindawi Publishing Corporation, 2007 (2), 1–12. 2007
- [44] Wang, Z., & Peterson, B. S. Partner-Matching for the Automated Identification of Reproducible ICA Components from fMRI Datasets: Algorithm and Validation. *Human Brain Mapping*, PMID: 18058813, Wiley-Liss, Inc., 29 (8), 875–93. 2008
- [45] Flandin, G., Kherif, F., Pennec, X., Rivière, D., Ayache, N., & Poline, J.-B. Parcellation of Brain Images with Anatomical and Functional Constraints for fMRI Data Analysis. In *Proceedings 1st International Symposium on Biomedical Imaging*, Washington, DC, 907–910. 2002
- [46] Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., & Poline, J.-B. Dealing with the Shortcomings of Spatial Normalization: Multi-Subject Parcellation of fMRI Datasets. *Human Brain Mapping*, PMID: 16281292, Wiley-Liss, Inc., 27 (8), 678–693. 2006

- [47] Yongnan, Ji., Pierre-Yves, H., Uwe, A., & Alain P. Parcellation of fMRI Datasets with ICA and PLS- a Data Driven Approach. MICCAI 2009, part I, LNCS 5761, by Springer Berlin Heidelberg, 984-991. 2009
- [48] Thomas, B., Saad, J., Matthew, F.G., David, C.V.E., Kamil, U., Timothy, E. J. B., & Stephen, M.S. Spatially Constrained Hierarchical Parcellation of the Brain with Resting-State fMRI. *Neuroimage*, 76, 313-24. 2013
- [49] Thirion, B., Gaël, V., Elvis, D., & Jean-Baptiste, P. Which fMRI Clustering Gives Good Brain Parcellations?. *Frontiers in Neuroscience*, 8 (167), 1-13. 2014
- [50] Korczak, J. Visual Exploration of Functional MRI Data. Chapter 11 of *Data Mining Applications in Engineering and Medicine*, Edited by Adem Karahoca, INTECH, 249-264. 2012
- [51] Evgenia, D., Kurt, H., Ewald, M., & Markus B. Application of the Neural Gas Algorithm in fMRI Analysis. 20th Annual Meeting of the European Society for Magnetic Resonance in Medicine and Biology (ESMRMB). 2003
- [52] Meyer-Baese, A., Wismueller, A., & Oliver L. Comparison of Two Exploratory Data Analysis Methods for fMRI: Unsupervised Clustering Versus Independent Component Analysis. *IEEE Transactions on Information Technology in Biomedicine*, 8 (3), 378-398. 2004
- [53] Wismuller, A., Meyer-Base, A., Lange, O., Auer, D., Reiser, M. F., & Summers, D. Model-Free Functional MRI Analysis Based on Unsupervised Clustering. *Journal of Biomedical Informatics*, 37 (1), 10-18. 2004
- [54] Dimitriadou, E., Barth, M., Windischberger, C., Hornika, K., & Moser, E. A Quantitative Comparison of Functional Cluster Analysis. *Artificial Intelligence in Medicine*, 31 (1), 57-71. 2004
- [55] Ana, I. G., Alicia, D., Maite, G., & Manuel, G. SOM and Neural Gas as Graduated Nonconvexity Algorithms. *International Conference on Computational*

Science and its Applications (ICCSA), Springer Verlag, Berlin Heidelberg, part III, 1143-1152. 2006

[56] Lachiche, N., Hommet, J., Korczak, J., & Braud, A. Neuronal Clustering of Brain fMRI Images. Proceeding of Pattern Recognition and Machine Inference, 300-305. 2005

[57] Korczak, J. Interactive Mining of Functional MRI Data. Signal-Image Technologies and Internet-Based System, (SITIS '07), IEEE Computer Society Washington, DC, USA, 912-917. 2007

[58] Heydar, D., Ali, T., & Emad F. Extracting Activated Regions of fMRI Data Using Unsupervised Learning. Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA, 641 – 645. 2009

[59] Qin, A.K., & Suganthan, P.N. Robust Growing Neural Gas Algorithm with Application in Cluster Analysis. Neural Networks, Elsevier Ltd., 17, 1135–1148. 2004

[60] Retrieved May 13, 1991, from <http://www.fil.ion.ucl.ac.uk/spm/data/auditory/>. 1991

[61] John, A., Gareth, B., Chun-Chuan, C., Jean, D., Guillaume, F., Friston, K., Stefan, K., James, K., Vladimir, L., Rosalyn, M., Will, P., Maria, R., Klaas, S., Darren, G., Rik, H., Chloe, H., Volkmar, G., Jeremie, M., & Christophe, P. SPM8 Manual. Functional Imaging Laboratory, Trust Centre for Neuroimaging, Institute of Neurology, London, UK. 2013

[62] Everitt, B. S., & Skrondal, A. The Cambridge Dictionary of Statistics. 4th ed., Cambridge, UK New York: Cambridge University Press. 2010

[63] Friston, K. J. Modes or Models: a Critique on Independent Component Analysis for fMRI. Trends in Cognitive Sciences, Elsevier Science Ltd., 2, 373-375. 1998

- [64] Jain, A.K., Murty, M.N., & Flynn, P.J. Data Clustering: a Review. *ACM Computing Surveys (CSUR)*, 31 (3), 264–323. 1999
- [65] Berkhin, P. Survey of Clustering Data Mining Techniques. Technical Report, Accrue Software Inc. 2002
- [66] Kato, N., & Nemoto, Y. Large Scale Handwritten Character Recognition System Using Subspace Methods. *Proceeding of IEEE International Conference on Systems, Man and Cybernetics, Beijing, China*, 432–437. 1996
- [67] Ray, S., & Turi, R. H. Determination of Number of Clusters in k-means Clustering and Application in Color Image Segmentation. *Proceeding of the Fourth International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99), Calcutta, India*, 137–143. 1999
- [68] Bhatia, S. K., & Deogun, J. S. Conceptual Clustering in Information Retrieval. *IEEE Transactions on System, Man, Cybernetics, Part B*, 28 (3), 427–436. 1998
- [69] Resson, H., Wang, D., & Natarajan, P. Adaptive Double Self-Organizing Maps for Clustering Gene Expression Profiles. *Neural Networks*, 16 (5-6), 633–640. 2003
- [70] Viviani, R., Gron, G., & Spitzer, M. Functional Principal Component Analysis of fMRI Data. *Human Brain Mapping*, 24 (109-129). 2005
- [71] Jiang, Z., Xianguo, T., Zhen, Y., Wei, L., & Huafu, Ch. Analysis of fMRI Data Using an Integrated Principal Component Analysis and Supervised Affinity Propagation Clustering Approach. *IEEE Transactions on Biomedical Engineering*, 58 (11), 3184 – 3196. 2011
- [72] Daubechies, I., Roussos, E., Takerkart, S., Benharrosh, M., Golden, C., D'Ardenne, K., Richter, W., Cohen, J.D., & Haxby, J. Independent Component Analysis for Brain fMRI Does Not Select for Independence. In *Proceeding of the Natural Academy of Sciences of the USA*. 2009

- [73] Anderson, A., Bramen, J., Douglas, P.K., Lenartowicz, A., Cho, A., Culbertson, C., Brody, A.L., Yuille, A.L., & Cohen, M.S. Large Sample Group Independent Component Analysis of Functional Magnetic Resonance Imaging using Anatomical Atlas-based Reduction and Bootstrapped Clustering. *International Journal of Imaging Systems and Technology*, 21 (2), 223–231. 2011
- [74] Lindquist, M.A. The Statistical Analysis of fMRI Data. *Statistical Science*, 23 (4), 439-464. 2008
- [75] Baumgartner, R., Windischberger, C., & Moser, E. Quantification in Functional Magnetic Resonance Imaging: Fuzzy Clustering vs Correlation Analysis. *Magnetic Resonance Imaging*, 16, 115-125. 1998
- [76] Hidetomo, I., Katsuhiko, H., Akira, N., & Takafumi, K. Exploratory Approach to fMRI Study with Fuzzy Clustering and General Linear Model. *IEEE International Conference on Fuzzy Systems, Canada*, 1167–1174. 2006
- [77] Danial, L., Ramesh, S., Edward, V., Po-Jang, H., Nancy, K., & Polina, G. Search for Patterns of Functional Specificity in the Brain: A Nonparametric Hierarchical Bayesian Model for Group fMRI Data. *Neuroimage*, 59 (2), 1368-1368. 2011
- [78] Liao, W., Chen, H., Yang, Q., & Lei, X. Analysis of fMRI Data Using Improved Self- Organizing Mapping and Spatio-temporal Metric Hierarchical Clustering. *IEEE Transaction on Medical Imaging*, 27 (10), 1472-1483. 2008
- [79] Guha, S., Rastogi, R., & Shim, K. CURE: An Efficient Clustering Algorithm for Large Databases. *Proc. SIGMOD'98, Seattle*, 73-84. 1998
- [80] Kohonen, T. Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*, 43, 59-69. 1982
- [81] Katwal, S.B., Gore, J. C., & Rogers, B. P. Analyzing fMRI Data with Graph-Based Visualizations of Self-Organizing Maps. *IEEE International Symposium on Biomedical Imaging, Chicago*, 1577-1580. 2011

- [82] Martinetz, T., & Schulten, K. A "Neural Gas" Network Learns Topologies. *Artificial Neural Networks*, Elsevier, 397–402. 1991
- [83] Fritzke, B. A Growing Neural Gas Network Learns Topologies. *Advances in Neural Information Processing Systems 7*, MIT Press, Cambridge, 625- 632. 1995
- [84] Fritzke, B. Some Competitive Learning Methods (draft). Technique report, Institute for Neural Computation, Ruhr-University, Bochum. 1997
- [85] Kohonen, T. *Self-Organizing Maps*. (3rd ed.), Berlin: Springer. 2001
- [86] Fernando, C., & Max C. Modification of the Growing Neural Gas Algorithm for Cluster Analysis. *Progress in Pattern Recognition, Image Analysis and Applications*, Lecture Notes in Computer Science, Springer, 4756, 684-693. 2007
- [87] Curatelli, F., & Mayora-Iberra, O. Competitive Learning Methods for Efficient Vector Quantizations in a Speech Recognition Environment. *MICAI 2000: Advances in Artificial Intelligence*, Lecture Notes in Computer Science, Springer, 1793, 108-114. 2000
- [88] Anastassia, A., Alexandra, P., José, G. R., & Kenneth, R. Automatic Landmarking of 2D Medical Shapes Using the Growing Neural Gas Network. *Computer Vision for Biomedical Image Applications*, Lecture Notes in Computer Science, Springer, 3765, 210-219. 2005
- [89] Atukorale, A. S., Downs, T., & Suganthan P. N. Boosting the HONG Network”, *Neurocomputing*, 51, 75–86. 2003
- [90] Winter, M., Metta, G., & Sandini, G. Neural-Gas for Function Approximation: a Heuristic for Minimizing the Local Estimation Error. *Proceeding of International Joint Conference on Neural Network (IJCNN)*, Italy, 535–538. 2000
- [91] Haykin, S. *Neural Networks: a Comprehensive Foundation*. (2nd ed.), Prentice-Hall, Englewood Cliffs, NJ. 1998

- [92] Martinetz, T.M., Berkovich, S.G., & Schulten, K.J. "Neural Gas" Network for Vector Quantization and Its Application to Time Series Prediction. *IEEE Transactions on Neural Networks*, 4 (4), 558–569. 1993
- [93] Bezdek, J. C., Keller, J. M., Krishnapuram, R., & Pal, N. R. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Norwell, MA: Kluwer. 1999
- [94] Martinetz, T. M. Competitive Hebbian Learning Rule Forms Perfectly Topology Preserving Maps. *Proceedings of International Conference on Artificial Neural Networks (ICANN93)*, Amsterdam, the Netherlands, 427–434. 1993
- [95] Fritzke, B. Growing Cells Structures—a Self-Organizing Network for Unsupervised and Supervised Learning. *Neural Networks*, 7 (9), 1441–1460. 1994
- [96] Huber, P. J. *Robust Statistics*. New York: Wiley. 1981
- [97] Tenmoto, H., Kudo, M., & Shimbo, M. MDL-Based Selection of the Number of Components in Mixture Models for Pattern Classification. *Advance in Pattern Recognition, Lecture Notes in Computer Science*, 1451, 831–836. 1998
- [98] Zemel, R. S. *A Minimum Description Length Framework for Unsupervised Learning*. Ph.D. Thesis, University of Toronto. 1994
- [99] Gareth, J. *Majority Vote Classifiers: Theory and Applications*. Ph.D. dissertation submitted to Stanford University. 1998
- [100] Hamerly, G., & Elkan C. Learning the k in k-means. *Proceeding of 17th Annual Conference on Neural Information Processing Systems (NIPS2003)*, Canada. 2003
- [101] Hubert, L., & Arabie, P. Comparing Partitions. *Journal of Classification*, 2, 193–218. 1985
- [102] Retrieved 2015, from www.yarpiz.com. 2015

- [103] Qin, A.K., & Suganthan, P.N. Enhanced Neural Gas Network for Prototype-Based Clustering. *Pattern Recognition*, Elsevier Ltd., 38, 1275–1288. 2005
- [104] Rissanen, J. A Universal Prior for Integers and Estimation by Minimum Description Length. *Annals of Statistics*, 11, 416-431. 1983
- [105] AlJobouri, H. K., Jaber, H. A., & Çankaya, I. Performance Evaluation of Prototype-Based Clustering Algorithms Combined MDL Index. *Computer Applications in Engineering Education*, Wiley Inc., DOI: 10.1002/cae.21824. 2017
- [106] Campain, R., & Minckler, J. A note on the Gross Configurations of the Human Auditory Cortex. *Brain Lang*, 3, 318–23. 1976
- [107] Francesco, D.S., Fabrizio, E., Tommaso, S., Elia, F., Elio, M., Claudio, S., Sossio, C., Raffaele, E., Klaus, S., & Erich, S. fMRI of the Auditory System: Understanding the Neural Basis of Auditory Gestalt. *Magnetic Resonance Imaging*, 21, 1213–1224. 2003
- [108] Skudlarski, P., Constable, R. T., & Gore J. C. ROC Analysis of Statistical Methods Used in Functional MRI: Individual Subjects. *NeuroImage*, 9, 311–329. 1999
- [109] Sun, X., & Xu, W. Fast Implementation of DeLong's Algorithm for Comparing the Areas under Correlated Receiver Operating Characteristic Curves. *IEEE Signal Processing Letters*, 21 (11), 1389-1393. 2014
- [110] Lange, N., Strother, S. C., Anderson, J. R., Nielsen, F. A., Holmes, A. P., Kolenda, T., Savoy, R., & Hansen, L. K. Plurality and Resemblance in fMRI Data Analysis. *NeuroImage*, 10, 282–303. 1999
- [111] Alziarjawey, H. A., & Çankaya, I. Heart Rate Monitoring and PQRST Detection Based on Graphical User Interface with Matlab. *International Journal of Information and Electronics Engineering*, 5 (4), 311-316. 2015

[112] AlJobouri, H. K., Alziarjawey, H. A., & Çankaya, I. Biosignal Processing, Medical Imaging and fMRI (BSPMI) Software Package Based on MATLAB GUI for Education and Research. International Journal of Scientific Research in Information Systems and Engineering, 1 (2), 2380-8128. 2015



PAPERS ORIGINATED FROM THIS WORK

Published/Accepted

[1] Aljobouri, H. K., Çankaya, I., & Karal O. From Biomedical Signal Processing Techniques to fMRI Parcellation. *Biosciences Biotechnology Research Asia*, 12 (2), 1115–1138. 2015

[2] AlJobouri, H. K., Alziarjawey, H. A., & Çankaya, I. Biosignal Processing and Medical Imaging (BSPMI) Software Package Based on Matlab GUI for Education and Research. *Abstract Proceedings on 3rd International Symposium on Engineering, Artificial Intelligence and Applications (ISEAIA 2015)*, Girne American University, p.39. 2015

[3] AlJobouri, H. K., Alziarjawey, H. A., & Çankaya I. Biosignal Processing, Medical Imaging and fMRI (BSPMI) Software Package Based on MATLAB GUI for Education and Research. *International Journal of Scientific Research in Information Systems and Engineering*, 1 (2), 2380-8128. 2015

[4] AlJobouri, H. K., Jaber, H. A., & Çankaya I. Performance Evaluation Of Prototype-Based Clustering Algorithms Combined MDL Index. *Computer Applications in Engineering Education*, Wiley Inc., DOI: 10.1002/cae.21824. 2017

RESUME (CV)

PERSONAL INFORMATION

Name Surname : Hadeel Kassim ALJOBOURI
Date of Birth : 18/11/1977
Phone : 5545938217
E-mail : hadeel_bme77@yahoo.com



EDUCATION

Bachelor : University of Baghdad, Baghdad, Iraq, in 2000.
Master Degree : Al-Nahrain University, Baghdad, Iraq, in 2004.
Doctor Degree : Electrical and Electronics Engineering Department-Biomedical Engineering, Graduate School of Natural Science / Ankara Yıldırım Beyazıt University, Turkey.

WORK EXPERIENCE

Instructor : Biomedical Engineering Department / Al-Nahrain University, Baghdad, Iraq. (2004-continued)

TOPICS OF INTERESTS

- Biomedical Signal Processing
- Medical Imaging
- Data Mining, clustering Techniques
- Machine Learning