Graph theory for feature extraction and classification: A migraine pathology case study

Fernando Jorge-Hernandez [a,*] Yolanda Garcia Chimeno [a], Begonya Garcia-Zapirain [a], Alberto Cabrera Zubizarreta [b], Maria Angeles Gomez Beldarrain [c] and Begonya Fernandez-Ruanova [d]

[a] *Deustotech Life Department, University of Deusto, Bilbao 48007, Spain*
[b] *Department of Neuroradiology, Osatek, Galdakao 48960, Spain*
[c] *Department of Neurology, Galdakao-Usansolo Hospital, Galdaka 48960, Spain*
[d] *Research Department, Osatek, Bilbao 48011, Spain*

Abstract. Graph theory is also widely used as a representational form and characterization of brain connectivity network, as is machine learning for classifying groups depending on the features extracted from images. Many of these studies use different techniques, such as preprocessing, correlations, features or algorithms. This paper proposes an automatic tool to perform a standard process using images of the Magnetic Resonance Imaging (MRI) machine. The process includes pre-processing, building the graph per subject with different correlations, atlas, relevant feature extraction according to the literature, and finally providing a set of machine learning algorithms which can produce analyzable results for physicians or specialists. In order to verify the process, a set of images from prescription drug abusers and patients with migraine have been used. In this way, the proper functioning of the tool has been proved, providing results of 87% and 92% of success depending on the classifier used.

Keywords: Functional MRI (fMRI), graph theory, migraine, machine learning, synchronization likelihood

1. Introduction

Use of graph theory [1,2] marked a significant step forward in brain connectivity analysis. It has demonstrated that the brain is a complex network, similar to other large networks, and furthermore, that it is divided into interconnected sub-networks. The brain's connectivity network can be shown through a graph G = (V,E) [3] where V is the graph nodes representing each brain region and E denotes the connections or edges between each node or region.

Different characteristics can be extracted from the graph, which make it possible to analyse and conduct studies on pathologies, differences, etc.

At the same time, a great deal of research has been done on resting state fMRI [4–6]. These studies

---
*Corresponding author: Fernando Jorge-Hernandez, Deustotech Life Department, University of Deusto, Bilbao 48007, Spain. Tel.: +34 944 13 90 00; Fax: +34 944 45 68 17; E-mail: fernandojorge@deusto.es.

analyze brain activity while the subject is completely at rest but not asleep. Once the most common images are extracted, Principal Analysis Components (PCA) [7] or Independent Analysis Components (ICA) [8,9] analyses are run to extract the components that describe brain functions. In this case, the images from the resting state are used to apply graph theory.

The importance of a good group classification for a pathology and the use of different methods to create, characterize and classify the graphs are two essential characteristics of these studies [10,11]. This article, therefore describes a tool capable of carrying out these functions for any pathology studied, the images used or the regions of interest analyzed. The graph of the groups studied is obtained and the most commonly used supervised and unsupervised classifiers are compared to determine which one is more suitable.

A group of 4D images from fMRI was used to analyze the proposed tool. The images were from a study on migraine with three groups of subjects: one healthy control group, a group with sporadic migraines and a group suffering from migraine and medication abuse.

Recent studies have shown recurrent errors when conducting fMRI studies [12]. On the one hand, problems arise when homogenizing different subjects' graphs to carry out group analysis, which does not give very reliable results. The statistical analyses conducted [13,14] are also noted as erroneous or are modified to extract better results, known as double dipping. The tool proposed in this study provides a method and a set of independent classifiers which can be used to study any given pathology.

Therefore, the purpose of this study is to provide a tool suitable for any pathological study so that the tool is expected to be one of the most widely used method currently. Furthermore, this tool offers users a set of results depending on the classification methods used.

2. Material and methods

2.1. Subjects

This study was conducted with images and information collected in another study on migraine and prescription drug abuse. The subjects were all right-handed and divided into three groups: fifteen healthy migraine-free subjects (mean age 37, SD: 6.9 years, range 32 to 60, 1 man), twenty subjects with a sporadic migraine (mean age 47, SD: 7.7 years, range 30 to 59, 0 men, 2 ambidextrous) and nineteen subjects suffering from migraine and medication abuse (mean age 45, SD: 7.9 years, range 27 to 55, 3 men).

All of the subjects underwent an MRI session lasting approximately 8 minutes. One subject had to be excluded from the experiment due to claustrophobia.

2.2. Image acquisition

TFE T1 3D, T1 images with high resolution reconstructions of the patients' anatomy, were used as the sequence for morphological evaluation. These images did not contain information on the active regions of the brain but were useful in locating the patients' neural correlation areas once they were obtained. A T1 image was made for each patient. It consisted of 250 slices with a resolution of 256x256 pixels, following the Philips format. Echo-Planar Imaging Blood Oxygen Level Dependent (EPI-BOLD) was the functional sequence used, with the following parameters: TR 2100; TE 29; 30 slices with a thickness of 3 mm; in-plane resolution 3x3 mm; 214 dynamic measurements with a total sequence duration of 7 minutes and 39 seconds.

2.3. Brain atlas

The tool has different digital atlases which can be used to know which regions of the brain should be analyzed. Two of the most well-known are Automated Anatomical Labeling (AAL) [15] and Brodmann [16]. It also includes the possibility of introducing user-created masks. When using AAL, the tool allows the user to decide how to use the 116 available regions of interests (ROIs) or 90 ROIs excluding the 16 cerebellum-related regions [10].

2.4. Preprocessing of fMRI sequences

Previous preprocessing is required for analysis of the images. FMRIB's Software Library (FSL 4.1) [17] and some of its tools such as MCFLIRT (Motion correction software), FLIRT or fslmeants can be used for this purpose. Other tools such as mricron have been used to analyze the atlases proposed.

2.5. Graph analysis characteristics

The method used to elaborate the graph for each subject is similar to that used in other studies [18] although the tool makes it possible to carry out two types of correlation between brain regions. By using a conventional correlation or synchronization likelihood (SL) [19], the correlation can be made by taking other aspects of the signal into account.

Studies on graph theory often use different characteristics such as the number of nodes or node degree. The proposed tool includes the three most widely used, fully described and contrasted characteristics [10]. These characteristics are found in many similar studies, so they are quite relevant to characterizing the connections between different brain areas. When executing the tool, it automatically carries out a set of 50 permutations for each subject's correlations and extracts the corresponding characteristics so that each subject's results are standardized.

– Clustering (C): This measures the grouping of the different nodes that form the graph. A high clustering value indicates that the close connections between brain regions function correctly but it may indicate a connectivity problem with distant regions.

$$C_i = \frac{E_i}{\frac{k_i(k_i-1)}{2}} \tag{1}$$

– Path Length (L): This examines the distant connections of the graph nodes from the greatest distances between the nodes found. A high path length value indicates that connections between distant regions work properly but that a problem may exist between regions where nodes are near each other.

$$L_i = \frac{1}{n-1}\sum_{i\neq j} min\{l_{ij}\} \tag{2}$$
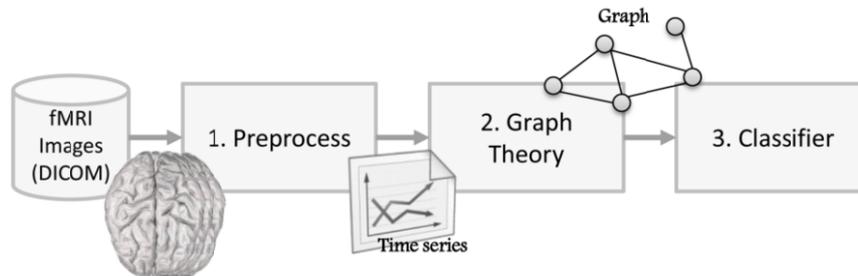
Fig. 1. High-level design of the tool.

– Dispersion (W): This evaluates the connectivity dispersion between different graph nodes [20].

$$\overline{\hspace{4cm}} \qquad (3)$$

2.6. Machine learning

The tool has a set of machine learning algorithms [21] to classify the groups entered in the study: Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Neuronal Networks (NN), K-Means, K-Nearest Neighbor and AdaBoost.

The authors have chosen the selected algorithms due to the incorporation of supervised (LDA, SVM, NN y K Nearest Neighbor), unsupervised (K-means) and semi-supervised (AdaBoost) classifiers because they are the most commonly used.

3. System design

3.1. High-level design

Figure 1 shows a diagram of the high-level design of the tool. It is divided into three large blocks:
1. Preprocess: From the input of the images obtained on the MRI machine, preprocessing is carried out to avoid errors or discard defective images. Following this, the time series obtained from each image are provided.
2. Graph theory: Once the time series have been extracted, different graph theory algorithms are applied, extracting the characteristics needed for later classification.
3. Classifier: When the characteristics have been calculated, this block executes different classification algorithms, obtaining the results for each one.

3.2. Low-level design

As shown in Figure 2, the low-level design is divided into the blocks. The following section gives an in-depth explanation for each block.
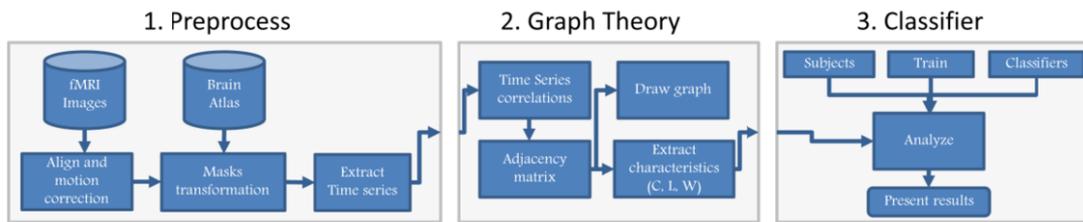
Fig. 2. Low-level design of the tool.

3.2.1. Preprocess

The data from the available fMRI images are entered in the tool. This first step consists of image alignment and motion correction, which are common problems when capturing images. This process avoids erroneous results or artefacts in the signals which make them difficult to analyze. Once it has aligned the images, the system has a 4D image for each subject in the study. The atlases selected are used to form the regions that will serve as masks in each of the subjects. Therefore, each region of the atlas is taken to each subject's individual space, extracting a time series for each region and subject entered.

3.2.2. Graph theory

Correlations between all the regions for each subject are made from the time series obtained in the previous step. The system makes it possible to carry out a simple correlation or SL. A symmetric matrix with the results is obtained from the correlations. The rows and columns are formed by the subjects. The system generates the corresponding graph from the matrix. The nodes represent each region studied and the edges that link the connectivity described in the matrix. Once the matrix has been calculated for each subject, the most relevant characteristics are extracted (C, L, W). In addition to the calculation for each subject, 50 permutations of each matrix are carried out and the characteristics are extracted again in order to homogenize the results.

3.2.3. Classifier

The input for this block consists of the characteristics extracted for each subject. And 10 subjects from one group are taken and 10 healthy subjects are taken as a practice run, leaving the rest as experimental validation. The third group is analyzed after the practice run, evaluating which of the two main groups are most similar. The final result is a table comparing the most commonly used algorithms, indicating the percentage of correct answers for each one.

4. Results

Once the tool was developed, a test based on images from a previous study on migraine was elaborated. To do this, a study is inserted in the tool with 90 and 116 regions of AAL areas, simple correlation and SL with these parameters: $W = 1$, $L = 1$, $M = 6$ y $Pref = 0:05$. All values are standardized from performing 50 random permutations of the data. Table 1 shows the result of the calculations made with graph theory. For each group (CON: controls, MIG: sporadic migraine, ABU: medication abuse), the mean value and the standard deviation of the three characteristics are analyzed (C: Clustering, L: Path Length, W: Dispersion). The results are shown for all the regions (116) and 90 regions excluded the cerebellum regions by using the AAL atlas.

Table 1

Graph theory results. Features extracted with graph theory. The number of areas used, typical correlation or SL, mean and standard deviation for each group for each feature are shown.

Charact.	Areas	CON		MIG		ABU	
		M(N/SL)	SD(N/SL)	M(N/SL)	SD(N/SL)	M(N/SL)	SD(N/SL)
C	90	1.075/1.069	0.024/0.028	1.08/1.069	0.012/0.026	1.076/1.057	0.013/0.02
	116	1.065/1.084	0.02/0.027	1.07/1.09	0.012/0.02	1.066/1.074	0.011/0.021
L	90	1.007/1.044	0.001/0.008	1.007/1.048	0.001/0.007	1.007/1.045	0.001/0.007
	116	1.004/1.03	0.001/0.005	1.004/1.032	0.001/0.005	1.004/1.031	0.001/0.004
W	90	0.995/0.807	0.007/0.018	0.997/0.804	0.001/0.017	0.995/0.8	0.002/0.017
	116	0.996/0.834	0.004/0.016	0.997/0.828	0.001/0.015	0.9962/0.822	0.002/0.016

Table 2

Classifiers. Results of the classifiers examined by the tool. For each classifier, type of correlation (Normal/SL), the number of areas, the percentage of correct answers, percentage of each group, the sensitivity and specificity are indicated.

Classifier	Corr.	%Success	%CON	%AB	Sens.	Spec.
LDA	90	50/64.29	66.67/77.78	20/40	0.45/0.56	0.38/0.64
	116	21.43/64.29	11.11/66.67	40/60	0.16/0.63	0.31/0.64
SVM	90	79.92/87.18	33.33/77.78	40/40	0.36/0.56	0.38/0.64
	116	76.93/87.18	33.33/66.67	40/60	0.36/0.63	0.38/0.64
NN (1 layer)	90	71.43/71.43	100/77.78	20/40	0.56/0.56	1/0.64
	116	71.43/64.29	66.67/44.44	100/100	1/1	0.75/0.64
NN (2 layers)	90	78.57/78.57	77.78/100	80/40	0.8/0.63	0.78/1
	116	92.86/78.57	88.89/100	100/40	1/0.63	0.9/1
NN (3 layers)	90	85.71/71.43	100/100	60/20	0.71/0.56	1/1
	116	85.71/78.57	100/88.89	60/60	0.71/0.69	1/0.84
K -means	90	57.14/57.14	88.89/44.44	0/80	0.47/0.69	0/0.59
	116	57.14/64.29	88.89/44.44	0/100	0.47/1	0/0.64
K – nearest	90	50/64.29	33.33/88.89	80/20	0.63/0.53	0.55/0.64
	116	57.14/64.29	77.78/88.89	20/20	0.49/0.53	0.47/0.64
AdaBoost	90	64.29/57.14	55.56/66.67	80/40	0.74/6.53	0.64/0.55
	116	64.29/64.29	66.67/88.89	60/20.53	0.63/0.53	0.64/0.64

From the data in Table 1, a classification with different classifiers, areas and correlations was carried out. The results are shown in Table 2. Sensitivity and Specificity results are observed. The term Sensitivity relates to the classifier's ability to identify disease in sick subjects, while Specificity indicates the classifier's ability to identify absence of disease.

5. Discussion and conclusion

The tool proposed can perform the entire process, from taking fMRI images to providing final results which can be easily interpreted by doctors or specialists. This is an automated process, where the user only provides fMRI images and chooses the correlations and the best atlas to use.

Studies with subjects suffering from migraine and drug abuse were conducted in order to validate this tool. The tool performs a complete analysis and proposes various classifiers, some of which provide 92.86% accuracy (NN), while others provide 87% (SVM). Other studies in other pathologies with similar machine learning algorithms [21–23] had success rates as between 75% and 87.9%, so the tool has demonstrated satisfactory results.

The existing differences between classifiers' results are due to several reasons such as their type (supervised, unsupervised, semi-supervised) or the difference between classifiers with the same data which might obtain global or local performance. Some classifiers such as NN can show different results because of the train's random procedure.

A more in-depth analysis could be achieved by increasing the number of subjects in each participant group. However, in the case of subjects with migraines, this task is complicated, because people with this pathology suffer from the noise generated by the MRI machine.

Finally, regarding the current limitations of the study, the impossibility of combining automatic classifiers and the complete atlas or a personal correlation use, can be found. In the future, new atlases and new correlations should be added to improve the results by providing a greater number of variables to enable specialists to analyze the pathologies.

Acknowledgement

References

[1] C. Stam, W. De Haan, A. Daffertshofer, B. Jones, I. Manshanden, A.V.C. vanWalsum, T. Montez, J. Verbunt, J. De Munck, B. Van Dijk et al., Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease, Brain **132** (2009), 213–224.

[2] J. Wang, X. Zuo and Y. He, Graph-based network analysis of resting-state functional MRI, Frontiers in Systems Neuroscience **4** (2010), 1–14.

[3] M.P. van den Heuvel and H.E. Hulshoff Pol, Exploring the brain network: a review on resting-state fMRI functional connectivity, European Neuropsychopharmacology **20** (2010), 519–534.

[4] C.F. Beckmann, M. DeLuca, J.T. Devlin and S.M. Smith, Investigations into resting-state connectivity using independent component analysis, Philosophical Transactions of the Royal Society B: Biological Sciences **360** (2005), 1001–1013.

[5] L. Wang, Y. Zang, Y. He, M. Liang, X. Zhang, L. Tian, T. Wu, T. Jiang and K. Li, Changes in hippocampal connectivity in the early stages of Alzheimer's disease: evidence from resting state fMRI, Neuroimage **31** (2006), 496–504.

[6] L. He, D. Hu, M. Wan and Y. Wen, Measuring temporal dynamics of resting-state fMRI data, Bio-Medical Materials and Engineering **24** (2014), 939–945.

[7] R. Baumgartner, L. Ryner, W. Richter, R. Summers, M. Jarmasz and R. Somorjai, Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis, Magnetic Resonance Imaging **18** (2000), 89–94.

[8]   M. Raemaekers, W. Schellekens, R.J. van Wezel, N. Petridou, G. Kristo and N.F. Ramsey, Patterns of resting state con-
      nectivity in human primary visual cortical areas: A 7t fMRI study, NeuroImage **84** (2014), 911–921.

[9]   Y. Xiong, Y. Luo, W. Huang, W. Zhang, Y. Yang and J. Gao, A novel classification method based on ica and elm: A
      case study in lie detection, Bio-Medical Materials and Engineering **24** (2014), 357–363.

[10] E.J. Sanz-Arigita, M.M. Schoonheim, J.S. Damoiseaux, S.A. Rombouts, E. Maris, F. Barkhof, P. Scheltens and C.J.
      Stam, Loss of 'small-world'networks in Alzheimer's disease: graph analysis of fMRI resting-state functional connectiv-
      ity, PloS One **5** (2010), e13788.

[11] B.C. Bernhardt, Z. Chen, Y. He, A. C. Evans and N. Bernasconi, Graph-theoretical analysis reveals disrupted small-
      world organization of cortical thickness correlation networks in temporal lobe epilepsy, Cerebral Cortex **21** (2011),
      2147–2157.

[12] B.C. van Wijk, C.J. Stam and A. Daffertshofer, Comparing brain networks of different size and connectivity density
      using graph theory, PLoS One **5** (2010), e13701.

[13] K. Nakajima, T. Minami, H.C. Tanabe, N. Sadato and S. Nakauchi, Facial color processing in the face-selective re-
      gions: An fmri study, Human Brain Mapping **35** (2014), 4958–4964.

[14] Z.N. Roth and E. Zohary, Fingerprints of learned object recognition seen in the fMRI activation patterns of lateral occi-
      pital complex, Cerebral Cortex, **bhu042** (2014), 27–34.

[15] A.V. Lebedev, E. Westman, A. Simmons, A. Lebedeva, F.J. Siepel, J.B. Pereira and D. Aarsland, Large-scale resting
      state network correlates of cognitive impairment in Parkinson's disease and related dopaminergic deficits, Frontiers in
      Systems Neuroscience **8** (2014), 1–12.

[16] E. Bujnošková, J. Fousek, E. Hladká and M. Mikl, 22 comparison of the methods for brain parcellation, Clinical Neuro-
      physiology **125** (2014), e31–e32.

[17] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich and S.M. Smith, FSL, Neuroimage **62** (2012), 782–790.

[18] J.C. Reijneveld, S.C. Ponten, H.W. Berendse and C.J. Stam, The application of graph theoretical analysis to complex
      networks in the brain, Clinical Neurophysiology **118** (2007), 2317–2331.

[19] C. Stam, B. Jones, G. Nolte, M. Breakspear and P. Scheltens, Small-world networks and functional connectivity in Alz-
      heimer's disease, Cerebral Cortex **17** (2007), 92–99.

[20] M. Boersma, D.J. Smit, H. de Bie, G.C.M. Van Baal, D.I. Boomsma, E.J. de Geus, H.A. Delemarre-van de Waal and
      C.J. Stam, Network analysis of resting state eeg in the developing young brain: structure comes with maturation, Hu-
      man Brain Mapping **32** (2011), 413–425.

[21] F. Pereira, T. Mitchell and M. Botvinick, Machine learning classifiers and fMRI: A tutorial overview, Neuroimage **45**
      (2009), S199–S209.

[22] Z. Wang, A.R. Childress, J. Wang and J.A. Detre, Support vector machine learning-based fMRI data group analysis,
      NeuroImage **36** (2007), 1139–1151.

[23] C. Davatzikos, K. Ruparel, Y. Fan, D. Shen, M. Acharyya, J. Loughead, R. Gur and D.D. Langleben, Classifying spatial
      patterns of brain activity with machine learning methods: application to lie detection, Neuroimage **28** (2005), 663–668.