Squizofrenia: Classification and Correlation from MRI

Ana Filipa Rodrigues, Mariana Barros, Pedro Furtado , *FCT, U. Coimbra, Portugal*

Abstract— **In order to better understand what structural and functional brain components changes are associated with schizophrenia, various investigations have been conducted. Functional Network Connectivity (FNC) generally interpreted as an indirect measure of brain activity, measures the functional component, and Structural Based Morphometry (SBM), an indirect measure of concentration of Gray Matter (GM), assesses the structural component. This work investigates the possibility of performing an automatic diagnosis of schizophrenia using FNC or SBM, considering each component individually and also the correlation between them. The best classification obtained for the diagnosis of schizophrenia was based on a Naïve Bayes classifier with an accuracy of 83.7%. In general, the accuracy of classifiers varied between 62% to 84% in both FNC and SBM attributes. The components that have experienced higher correlation were those related to the basal ganglia, the posterior components, motor and media components and frontal components. But we also show that the best results are obtained by considering both FNC and SBM at the same time.**

I. Introduction

Schizophrenia is a complex mental disorder that affects about 1% of the population [1]. This condition results in a moderate or severe disability in approximately 60% of cases and is characterized by delusions, hallucinations and loss of initiative and cognitive impairment [2]. Their origin is unknown, however, it is believed to be associated with a combination of genetic and environmental factors. Changes in brain dopamine and neurotransmitters, such as glutamate, also appear to be involved in the cause of the pathology [3]. Recent studies showed that schizophrenic patients have differences in the structure of the brain and central nervous system. It has also been shown that this condition is characterized by strong connectivity disruption, both at the anatomical level (temporal lobe: top spin, lower spin and medial turn, frontal and parietal lobe [4]) and functional (frontal lobe [5], hippocampus, thalamus among others [6] [7]). In order to obtain structural and functional data to study organization and brain function, Magnetic Resonance Imaging (MRI) is used. MRI is an imaging technique that can capture many physiological and anatomical markers by various acquisition protocols [8][9]. It can be further classified as functional imaging (fMRI) and structural (SMRI). There is evidence that schizophrenia is associated with changes (increase and decrease) in gray matter density in certain brain regions, and the calculation of density is fundamental for the study of the pathology [10][11]. This study is conducted through Source-Based Morphology (SBM). Functional data has been used to carry out the study of several brain areas activated over time in healthy and schizophrenic individuals. The relationship between the activated brain areas has been performed by Functional Network Connectivity (FNC), which exploits the properties of neuronal interactions between brain networks based on temporal information contained in fMRI scans. Usually, during a MRI exam, several tests are conducted in order to activate specific brain regions through the application of certain stimuli. In the case of schizophrenia, there is evidence of spontaneous fluctuations in the brain by changing the pattern of FNC without the application of any stimuli [33]. Given this evidence, in this work the FNC values are studied without the application of any stimulus when performing the MRI scan, allowing detection of functional changes associated with schizophrenic volunteers. According to [12], correlations between functional and structural information can be used for better assessment of brain organization and to understand associated pathologies and mental illnesses. This work has as main objective to evaluate the combination of functional and structural information in order to classify schizophrenia, allowing an automatic diagnosis of the pathology in question.

II. Related Work

The diagnosis of schizophrenia has traditionally been based on clinical manifestations. However, there is increasing interest in automatic diagnosis of the pathology. Schizophrenia is hypothesized to involve disordered connectivity between brain regions. Currently, there are no direct measures of brain connectivity; functional and structural connectivity used separately provide only limited insight. In [24] the authors tested the hypothesis that schizophrenia is a disorder related to the connectivity between components of large-scale brain networks. They concluded that people with schizophrenia tend to have a less strongly integrated, more diverse profile of brain functional connectivity, and a less hub-dominated configuration of complex brain functional networks [24]. In [25] global, regional, and voxel measures and K-means network analysis were employed to identify group differences and correlation with clinical symptoms. Statistical analysis allows us to understand general aspects about the data. This analysis can be performed through various tests. The t-test with a sample allows an analysis of characteristics inside each group, while the two-sample t-test allows an analysis of characteristics between groups. This test is based on assumptions, and usually tests the null hypothesis [14][15]. When the data is characterized by many features, Principal Components Analysis (PCA) can be used to reduce the number of features [16]. PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components [27]. In what concerns classification, three approaches recently employed in the context of squizofrenia [12] included Rating Gaussian Process (GP) [17], Support Vector Machines

978-1-5090-4179-4/17/$31.00 ©2017 IEEE

(SVM) [11] and Distance Weighted Discrimination (DWD). Classification by Support Vector Machine (SVM) uses a hyperplane responsible for maximizing the margin between the two classes in question [11]. Classification by DWD is similar in approach to SVM, however, this method gives higher relevance to points closer to the hyperplane [18].

III. Acquisition and Pre-Processing

The dataset used for this analysis was obtained in [32]. The structural and functional data there were obtained from conducting MRI (using a scanner 3 Tesla Siemens Trio with a 12 channel radio frequency coil) on 86 volunteers. The volunteers were instructed to relax and stay awake, resulting in 149 volumes of functional weighted images. After scanning, the data was pre-processed with SPM5 software [12], and analyzed using independent component analysis groups (ICA) [12]. Depending on the formation of the data matrix, temporal or spatial ICA (SICA) can be performed on the fRMI data. SICA is the most used approach and breaks down the data set into another set of maximally independent spatial maps, and their corresponding time courses. Each SICA map consists of several remote brain regions that form a functional brain network. SICA generates consistent spatial maps (SMs) while modeling complex fRMI data acquired during the task or in standby mode. The signal within a dynamic component is described by the course of time (CT). The regions that contribute significantly within a given component are functionally linked to each other [14].

The results of the ICA are two matrices (X=AxS): matrix A whose rows are points in time and columns are selected independent components, representing how each independent component varies over time; and S matrix, where each row corresponds to a separate component and each column to a voxel [27]. The results of this matrix S are the spatial maps used to calculate the FNC. An example is shown in Figure 1, where the left image corresponds to line 7 of the array S according to three directions. Of the 75 components obtained by ICA, only 28 were identified as components associated with brain network at rest state [27].

FNC features (378) correspond to the correlation between the time courses of each independent component (matrix A). In a way, the FNC indicates the general level of 'synchronicity' between brain areas associated with each volunteer [26]. The FNC was estimated using the Pearson correlation coefficient between pairs of brain maps. The correlations resulted in values presented through a symmetric C1xC1 correlation matrix for each volunteer. For all FNC analysis, correlations were transformed to 'z-score' using the Fisher transformation, z = atanh(k), where k is the correlation between two brain maps by ICA [27]. The dataset of functional data has 378 features and 86 samples. From the 86 samples, 46 belonged to class non-squizofrenia, and 40 to class squizofrenia. The structural images were obtained the same way as the functional images, however these were acquired with a high spatial resolution. The density (concentration) of Gray Matter (GMC) is especially superior in the outermost region of the brain as well as in subcortical regions, their study being fundamental in the context of this type of brain diseases [12].

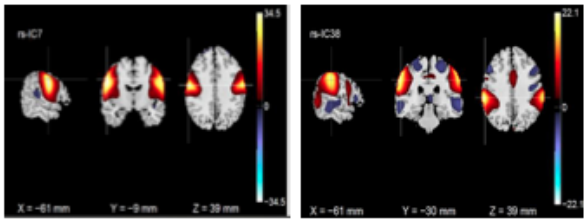

Figure 1. Two Maps output by ICA

Through Source Based Morphometry (SBM), the data of the Gray Matter density were decomposed by Independent Component Analysis (ICA) in a set of spatially independent components and corresponding weight coefficients (normalized). These weights indicate the contribution of each component on each individual; however, the interpretation of the difference between the weight factors depends on the component analysis [13]. The SBM approach was first performed on a set of data of 603 healthy people, yielding 75 maps of independent components. In this study, the data corresponding to the volunteers was projected into the space of 75 components obtained earlier. This projection was performed by ada Toolbox GIFT. Of the 75 components, only 32 were identified as components without artifacts.

In order to allow an understanding of which functional and anatomical structures are being used as attributes, and thus to allow new data selection strategies, a set of files was provided containing the spatial maps for the FNC and SBM characteristics. This additional information indicates the spatial extent of the components in the analysis and is useful in combination with the functional and structural characteristics [12].

IV. Results

Table 1 shows the results obtained with statistical t-test on two samples (a) and with attribute selection (b). Attribute selection ranks attributes according to their relevance to classification, according to information gain metric in this case [27]. Attribute selection with PCA creates 62 linear combinations of FNC attributes and 23 linear combinations for SBM attributes. Table 2 shows the classification accuracy results of the functional (FNC) and Structural (SBM) data, using various classifiers, after attribute selection and using cross validation with four folds as test option.

The previous analysis evaluated accuracy of classification based only on FNC or SBM data. The next analysis concerns using FNC and SBM together to improve classification. Table 3 shows the results. In (a) we show the ranking of attributes according to information gain metric, in (b) we show classification outcome with all attributes, versus with only the attributes selected in (a) (IG).

The last analysis results, shown in figure 2 concern functional-structural correlation analysis. The IG highest ranking 28 functional and 32 structural components were

analyzed, with 16 functional-structural pairs found above the correlation threshold of 0.2 (the red dots).

Statistically meaninful attrs	
FNC	SBM
13,30,33,35,38,40,41,42,43,48, 61,62,63,64,71,78,88,102,151,1 65,169,170,171,182,183,185,18 9,193,194,200,208,210,211,21 5,220,226,243,244,253,279,285 ,290,295,297,301,302,304,328, 333,337,347,350,353,368.	3,7,11,16,22,2 4,25,26,32

(a) T-test results

Feature	Info Gain
FNC	
FNC295	0.166
FNC302	0.16
FNC226	0.157
FNC243	0.146
FNC244	0.146
FNC183	0.141
FNC33	0.128
FNC194	0.126
FNC220	0.125
FNC289	0.125
FNC292	0.111
SBM	
SBM68	0.139
SBM61	0.137
SBM75	0.115

(b) Attribute selection

Table 1. T-test and attribute selection

Weka		
IG, Cross Val	FNC	SBM
J48	71%	65.70%
NN	73.80%	74.30%
Naive Bayes	82.60%	67.90%
k-NN (5)	77.90%	65%
SVM	79.40%	75.10%
Logistic Regr	69.70%	65%

Table 2. Classification Accuracy (% correct)

Feature	Info Gain
FNC+SBM	
FNC295	0.166
FNC302	0.16
FNC226	0.157
FNC243	0.146
FNC244	0.146
FNC183	0.141
SBM_map67	0.139
SBM_map61	0.137
FNC33	0.128
FNC194	0.126
FNC289	0.125
FNC220	0.125
SBM_map75	0.115
FNC292	0.111

(a) Attribute Selection

	all attributes	IG
NN	76.74%	79.40%
Naive Bayes	70.90%	83.70%
k-NN (5)	69.76%	77.90%
SVM	79.50%	79.30%
Logistic Regr	62.80%	68.60%

(b) Classification accuracy (%)

Table 3. Attr Selection and Rank FNC+SBM - Information Gain

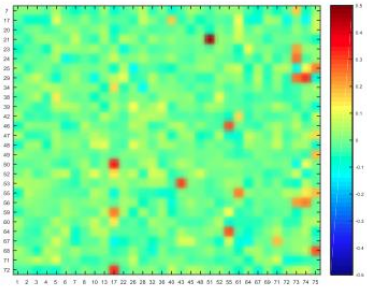

Figure 2. Functional-Structural Correlation Map

V. ANALYSIS OF RESULTS

1) Statistical analysis - the t-test on two samples (Table 1) indicates which FNC and SBM attributes have statistical significance for classification of schizophrenia. 55 of the 378 FNC attributes are statistically significant. The corresponding main areas associated with schizophrenia are: the motor area, visual, olfactory, attention, actions enforcement, and emotion. These results agree with common knowledge about the illness, especially in what concerns emotional and motor areas [19], [20]. In what concerns SBM, results agree with those reported in [21][22], especially with regard to the Temporal region (SBM3 and SBM24), Front, Limbic and Hippocampal Gyrus (SBM16, SBM32).

2) Features Selection - according to the information gain metric, the most important attribute for prediction of schizophrenia is FNC295. This attribute correlates brain maps 50 and 71, belonging to Brodmann areas 7 and 22, respectively. Brodmann area 7 is mainly related to motor areas (motor imagery: mental simulation of a motor act, saccadic movements: rapid eye movement), somatosensory (sensations such as touch, temperature and vibration), linguistics (language comprehension), memory (working memory), attention (visual-motor attention) and processing of pain, while Brodmann area 22 is mainly linked to auditory (processing complex sounds), language (verbal comprehension) and attention (attention associated with speech). The second selected attribute was FNC302, corresponding to correlation between brain maps 53 and 68 belonging to Brodmann areas 30 and 8 respectively. Brodmann area 30 is related to the motor area (motor learning), language (verbal comprehension), memory (working memory, movement memory), attention (speech attention) and emotions (experience and processing of emotions) while Brodmann area 8 is mainly related to motor (secondary motor functions, motor learning and saccades (rapid eye movement)), language (verbal expression), memory (working memory, memory retrieval) , attention (visual-motor attention), enforcement actions (planning actions and inhibitory behavior) and mental calculation. The remaining attributes distinguished by information gain also correspond to brain areas with similar functions. In summary, the main functions selected were: motor, somatosensory, language, memory, attention and emotion. These results are consistent with the pathology, since schizophrenia is known to be associated with social cognition anomalies and is associated with various brain disorders, particularly on the amygdala, cortex ventromedial prefrontal, insula and somatosensory cortex (activation decrease) [21][22].

3) Classification and Class (Functional + Structural) – The most relevant classification results are summarized in Table 2 and especially Table 3. The best possible classification accuracy (obtained with Naïve Bayes) was high (83.7%), and most accuracy results were also high (usually between 70% to 84%), meaning that squizofrenia is correctly classified in a high fraction of the cases using this automated method (note: per class precision and recall on Naybe Bayes were 83%, 85%, 84%, 78%). There are some variations in classifiers accuracy, and feature selection based on information gain (IG) had best results.

4) Functional-Structural Correlation – Basal Ganglia components were those with greatest structural-functional spatial correlation, in particular the Putamen and Superior Temporal Gyrus. There are studies that correlate the size of the Putamen with schizophrenia, since the volume of this structure appears increased in schizophrenic individuals [28]. The Superior Temporal Gyrus, corresponding to Brodmann area 38, has as one of its functions experience

and processing of emotions, as well as the understanding of language. It is also related to schizophrenia, according to [29] [31]. Three other correlated pairs match the Precuneus zone, which is connected to the Network Default Mode (DMN) - network resting state [30]. Another correlation zone is the Posterior Cingulate Cortex (PCC), also related with DMN. The PCC has been associated with schizophrenia in several studies [31]. In terms of motor and medial components, there are 6 pairs of structural-functional correlations. These correlations refer to areas in the Pre-Central Gyrus, Middle Frontal Gyrus and Superior Frontal Gyrus. All of these areas correspond to Primary Motor and Premotor Cortex, with many functions, the main ones being motor and memory functions [31]. The last two correlations involve Frontal components, where there is evidence of relationship with schizophrenia. The front area, particularly in the Middle Frontal Gyrus, is related with affection, social judgment, execution memory, abstract and intentional thinking [31]. The most correlated pairs are consistent with the literature, e.g. [30].

VI. Conclusions

This paper studied functional and structural MRI data pertaining to schizophrenia and non-schizophrenia (control group) to answer how strongly one can predict schizophrenia automatically from that data. Statistical analysis (t-test, information gain rankings) selected highly correlated attributes, and we showed that machine learning classification techniques are undoubtedly a fundamental tool that can diagnose schizophrenia with a high precision. Using correlation analysis between functional and structural components, we were also able to analyze correlation on 16-functional structural pairs. It is possible that schizophrenia may cause a number of changes, without any of them being singled out as the main evidence of the disease, therefore the study of evidence associated with the disease can be critical for the diagnosis. In the future we would like to hold similar studies on larger populations and to do follow-up analysis on schizophrenic patients as the disease progresses.

References

[1] WHO site [URL, Access in 5/5/2016]: http://www.who.int/en/, OMS.

[2] Mental Health Statistics: Schizophrenia. [URL, Accessed in 5/5/2016]: http://www.mentalhealth.org.uk/help-information/mental-health-statistics/schizophrenia-statistics/.

[3] MinhaVida Brasil. [URL, Accessed in 5/5/2016]: http://www.minhavida.com.br/saude/temas/esquizofrenia.

[4] Bassitt, Débora Pastore, et al. "Insight and regional brain volumes in schizophrenia." European archives of psychiatry and clinical neuroscience 257.1 (2007): 58-62.

[5] Liu, Hu, et al. "Differentiating patterns of amygdala-frontal functional connectivity in schizophrenia and bipolar disorder." Schizophrenia bulletin 40.2 (2014): 469-477.

[6] Kraguljac, Nina Vanessa, et al. "Hippocampal-parietal dysconnectivity and glutamate abnormalities in unmedicated patients with schizophrenia." Hippocampus 24.12 (2014): 1524-1532.

[7] Zhuo, Chuanjun, et al. "Functional connectivity density alterations in schizophrenia." Frontiers in behavioral neuroscience 8 (2014).

[8] Bois, C., et al. "Structural magnetic resonance imaging markers of susceptibility and transition to schizophrenia: a review of familial and clinical high risk population studies." Journal of Psychopharmacology 29.2 (2015): 144-154.

[9] Demirci, Oguz, and Vince D. Calhoun. "Functional magnetic resonance imaging–implications for detection of schizophrenia." European neurological review 4.2 (2009): 103.

[10] Segall, Judith M., et al. "Voxel-based morphometric multisite collaborative study on schizophrenia." Schizophrenia bulletin 35.1 (2009): 82-95.

[11] Turner, Jessica A., et al. "Heritability of multivariate gray matter measures in schizophrenia." Twin Research and Human Genetics 15.03 (2012): 324-335.

[12] Silva, Rogers F., et al. "The tenth annual MLSP competition: schizophrenia classification challenge." 2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2014.

[13] Hilgetag, Claus, Thomas R. Knosche, and Max Planck. "Correspondence between structure and function in the human brain at rest." Mapping the connectome: Multi-level analysis of brain connectivity (2012).

[14] Arbabshirani, Mohammad Reza, et al. "Classification of schizophrenia patients based on resting-state functional network connectivity." Frontiers in neuroscience 7 (2013): 133.

[15] Stone, J. (2004). Independent component anal. Cambridge: MIT Press.

[16] Thais Mayumi Oshiro, 2013, Uma abordagem para a construção de uma única árvore a partir de uma Random Forest para classificação de bases de expressão génica, PhD Thesis, USP, Brasil.

[17] Solin, Arno, and Simo Särkkä. "MLSP 2014 Schizophrenia Classification Challenge: Winning Model Documentation."

[18] Karolis Koncevicius, 2014, MLSP 2014 Schizophrenia Classification Challenge: 3rd position (solution).

[19] Judith Butman, Ricardo F. Allegri, 2011, A Cognição Social e o Córtex Cerebral Universidade de Buenos Aires, Argentina.

[20] Schröder, J., et al. "Sensorimotor cortex and supplementary motor area changes in schizophrenia. A study with functional magnetic resonance imaging." The British journal of psychiatry: the journal of mental science 167.2 (1995): 197-201.

[21] Yang, Chunlan, et al. "A Review: Detecting Alterations of Brain Connectivity in Schizophrenia based on Structural MRI." Journal of Psychiatry: Open Access 2014 (2014).

[22] Nieuwenhuis, Mireille, et al. "Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples." Neuroimage 61.3 (2012): 606-612.

[23] Definition of PCA in Wikipedia [URL, Accessed in 5/5/2016]: https://en.wikipedia.org/wiki/Principal_component_analysis

[24] Lynall, Mary-Ellen, et al. "Functional connectivity and brain networks in schizophrenia." The Journal of Neuroscience 30.28 (2010): 9477-9487.

[25] Skudlarski, Pawel, et al. "Brain connectivity is not only lower but different in schizophrenia: a combined anatomical and functional approach." Biological psychiatry 68.1 (2010): 61-69.

[26] Software to obtain the regions corresponding to the respective coordinates, Version 2.4.3 client-Talairach available in http://www.talairach.org/client.html. [URL, Accessed in 20/3/2016].

[27] Weka Software: http://www.cs.waikato.ac.nz/ml/weka/. [URL, Access in 20/3/2016].

[28] Mamah, Daniel, et al. "Structural analysis of the basal ganglia in schizophrenia." Schizophrenia research 89.1 (2007): 59-71.

[29] Matsumoto, Hideo, et al. "Superior temporal gyrus abnormalities in early-onset schizophrenia: similarities and differences with adult-onset schizophrenia." American Journal of Psychiatry 158.8 (2001): 1299-1304.

[30] Buckner, Randy L., Jessica R. Andrews-Hanna, and Daniel L. Schacter. "The brain's default network." Annals of the New York Academy of Sciences 1124.1 (2008): 1-38.

[31] http://www.internationaljournalssrg.org/IJECE/2015/Special-Issues/ICFTCC/IJECE-ICFTCC-P115.pdf. [Access: 20/1/2016].

[32] Kaggle fMRI Challenge (www.kaggle.com/c/mlsp-2014-mri).

[33] Allen, Elena A., et al. "A baseline for the multivariate comparison of resting-state networks." Frontiers in systems neuroscience 5 (2011): 2.