Computer-aided Diagnostics of Schizophrenia: Comparison of Different Feature Extraction Methods

Radomír Kůs, Daniel Schwarz

Institute of Biostatistics and Analyses, Masaryk University
Kamenice 3, 62500 Brno, Czech Republic

E-mail: radomir.kus@mail.muni.cz, schwarz@iba.muni.cz

Abstract: Receiving an early diagnosis of schizophrenia is a crucial step towards its treatment. However, in current thinking, the diagnosis is based on time-consuming criteria, burdened with subjectivity. Hence, objective and more reliable therapeutic tests are desirable for the clinical practice of Psychiatry. Since schizophrenia is characterized by progressive brain volume changes during the course of the disease, many studies have recently turned attention to machine learning and brain morphometric techniques serving as tools for computer-aided diagnosis of schizophrenia based on neuroimaging data. In our study, the methodology is applied to distinguish between 52 first-episode schizophrenia patients and 52 healthy volunteers on the basis of T1-weighted magnetic resonance images of their brains preprocessed by the means of voxel-based and deformation-based morphometry. The proposed classification schemes vary in the feature extraction and selection steps. Namely, Mann-Whitney testing is implemented as a simple univariate approach playing the role of a comparator to multivariate methods such as inter-subject PCA, the K-SVD algorithm, and pattern-based morphometry. The highest classification accuracy, 70%, is reached with the pattern-based morphometry technique. The study points out the difference between univariate and multivariate approaches towards neuroimaging data. Additionally, the contrast between feature extraction capabilities of voxel-based and deformation-based morphometry is demonstrated.

Keywords: feature extraction; computer-aided diagnosis; schizophrenia; brain morphometry; voxel-based morphometry; deformation-based morphometry; magnetic resonance imaging; classification; machine learning

1 Introduction

As schizophrenia worsens with the progression of the disease [1], its early and accurate diagnosis can be beneficial for patient prognosis and overall treatment strategies [2]. Unfortunately, since psychiatry deals with mental states of patients,

its measurement techniques (such as Schneiderian First-Rank Symptoms) evaluate general symptoms common to a variety of mental disorders rather than the specific ones.

In the case of schizophrenia, the final verdict is partly based on observing patient's actions and noting the constellation of patient's symptoms, partly on psychiatric rating systems, and diagnostic classification and rating scales. Thus, most of the diagnoses are dependent on a subjective perspective and judgment of the psychiatrist assessing a patient [3], leading to the situation when more sophisticated methods, taking into account more aspects than a naked eye does, are desired. For instance, the changes in the brain morphology are only subtle during the first episode, and hence often indistinguishable even by an experienced psychiatrist.

Consequently, many studies have recently turned attention to Magnetic Resonance Imaging (MRI) as it can be utilized to explore the structure of the brain and to understand better the neurobiology of brain disorders. Moreover, neuroimaging data can be to a lesser or greater extent [4] successfully exploited as an input data to machine learning techniques which attempt to reduce or completely eliminate the need for human intuition in the analysis of the neuroimaging data.

2 Schizophrenia through the Optics of MRI

As the evidence of pathological changes in the brain morphology of schizophrenia patients exists [1], the researchers have started exploiting MRI data as a base to the disorder diagnosis. Should the schizophrenia-related manifestation in brain structure be profoundly understood, predictions pertaining to patients diagnoses could be made on the basis of an individual brain MR scan.

However, such an approach faces two confronting requirements where, on the one hand, all the information crucial for the classification must be retained in the data. On the other hand, a successful diagnosis – in terms of an accurate classification – is feasible only when the dimensionality of the problem is properly reduced as the brain image classification algorithms fail to operate on data exhibiting an adverse ratio between its dimensionality and the number of acquired samples [5].

Many studies have attempted to find the connection between schizophrenia neuropathology and brain structure. Although various brain regions have been identified, the results are not entirely consistent [6-9]. Nevertheless, even though the general consensus upon what brain structures are affected in schizophrenia is yet to be achieved, it has been revealed that structural changes happen in both gray [9] and white matter [10] and that these changes are not bound to a specific region but rather they are distributed throughout the entire brain following spatially complex and unknown patterns.

Thus, reduction of brain image data dimensionality using regions of interest (ROIs) methods may be misleading as they are prone to human error due to manual brain segmentation [11]. In contrast, automated and whole-brain morphometry methods, such as voxel-based morphometry (VBM) and deformation-based morphometry (DBM), are two main concepts used by the neuroimaging community for assessing MRI brain scans without the need to limit the analysis to arbitrarily predefined anatomical hypotheses or ROIs.

Although both of the methods are designed to assess the brain structure, they differ in workflow and interpretation. Whereas VBM segments the images in order to generate gray matter (GM) maps and uses low-dimensional registration, DBM utilizes full brain scans and the employed registration algorithm is high-dimensional [12-13]. The images resulting from those techniques are interpreted as local GM volume and local brain volume changes respectively. Their application is therefore advised to be chosen accordingly, with the knowledge of the disease process [14].

A strong critique of those techniques stems from the fact that, as they deal with brain images on a voxel-to-voxel basis, they neglect multivariate group differences [15]. Advantageously, multivariate approaches known from machine learning can be employed to conjointly account for voxels interactions [16] hence extracting complex patterns suitable for schizophrenia classification, leaving the brain morphometry methods a place among data preprocessing tools rather than feature extraction techniques.

3 Research Problem and Proposed Methodology

Typically, studies on computer-aided diagnostics of schizophrenia employ several machine learning algorithms in order to achieve the highest classification accuracy. However, as classification performance considerably depends on a preceding feature extraction step, an equal effort should be made in finding what algorithms are the most suitable for each application domain.

Thus, the purpose of this study is to elicit conditions under which one feature extraction method outperforms the other and vice versa. Namely, multivariate machine learning methods are put in contrast with univariate statistics. The comparison takes place on a dataset of first-episode schizophrenia patients preprocessed by the means of VBM and DBM separately, allowing us also to comment on whether and how the brain morphometry techniques influence the ensuing feature extraction step.

The whole study is organized accordingly to a well-known scheme in schizophrenia classification, starting by presenting a dataset and its preprocessing (Section 3.1), describing employed feature extraction algorithms (Section 3.2) and

a classification pipeline utilized to evaluate the performance of all the algorithms (Section 3.3).

As we aim at analyzing and comparing feature extraction algorithms, we include a short experiment about the anticipated behavior of the algorithms performed on a synthetic dataset (Section 4.1). Next, an elaboration on parameters of the algorithms and the process of their tuning is stated (Section 4.2). Last, the classification results are introduced (Section 5), followed by a commentary on capabilities of the feature extraction algorithms (Section 6).

3.1 Datasets

3.1.1 Subjects

The datasets consisted of 104 individual T1-weighted MRI whole-head scans, where exactly one-half of the scans belonged to 52 first-episode schizophrenia patients (FES) who were recruited at the Department of Psychiatry, Masaryk University in Brno. The patients were all male with the mean age of 24 years (±5.1). The diagnosis was based on diagnostic interviews regarding patient's history, substance abuse, etc., and evaluated using the Positive and Negative Syndrome Scale (PANSS [17]). A senior psychiatrist reviewed the tests and, in compliance with International Statistical Classification of Disease and Related Health Problems (ICD-10), established the diagnosis. Additionally, the patients were physically examined and, given specific criteria such as suffering from another neurological disease, substance dependence, etc. were met, excluded from the study.

The other 52 scans were acquired from volunteering healthy controls (HC) whose mean age (24 ± 3.1 years) and handedness matched with the patients.

3.1.2 Acquisition & Preprocessing

The images were obtained using a 1.5 T MR scanner with a resolution of 160×512×512 voxels per scan and, subsequently, using the VBM8 toolbox available in the SPM8 Matlab software package, they were corrected for bias-field inhomogeneity and spatially normalized by affine co-registration to the standard SPM T1 template.

Acquired and co-registered images were preprocessed correspondingly to VBM and DBM approaches resulting in two datasets in here referred to as GM Densities and Volume Changes.

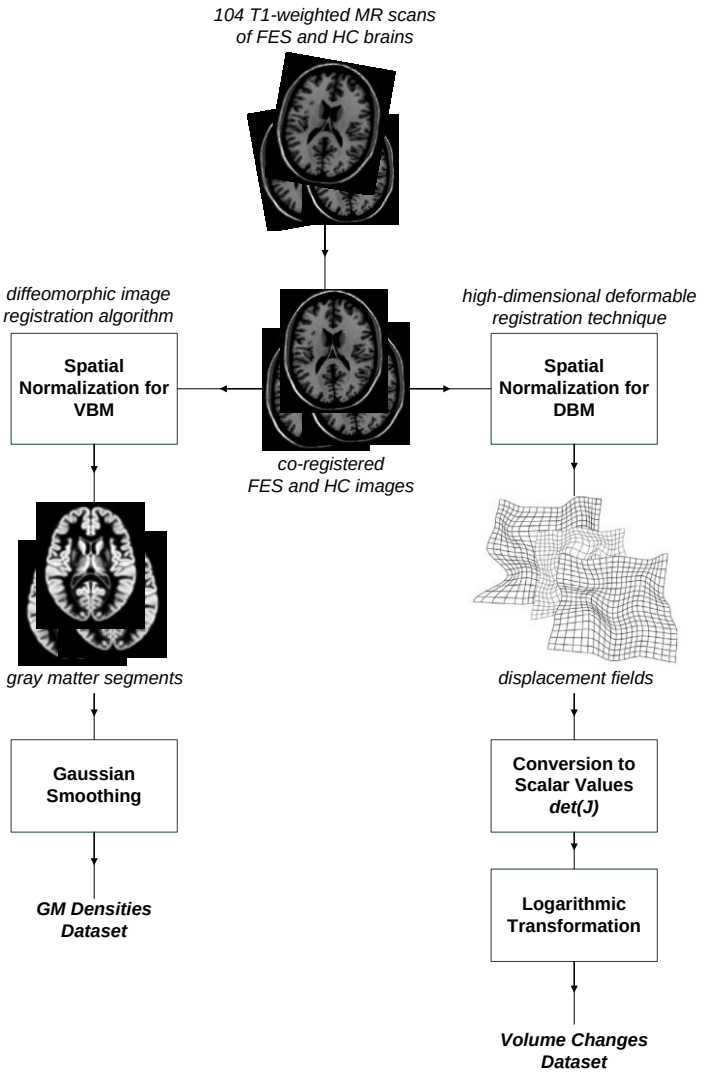

Figure 1
Scheme of the Datasets Preprocessing

In order to create the GM densities dataset, additional steps needed to be performed following the VBM pipeline. After the affine registration of the T1-weighted images, the images were non-linearly registered using fast diffeomorphic image registration algorithm (DARTEL [18]). Resulting GM tissue segments were modulated with the determinant of Jacobian matrices of the deformations to account for registration related changes in local volumes.

Subsequently, the modulated GM segment images were smoothed with the 8 mm FWHM Gaussian kernel to enable inter-subject comparisons.

As the Volume Changes dataset resulted from the DBM method, it was based on an additional spatial normalization which output displacement fields referring to volume adjustments needed for each image to match the template. Thus, after the images were normalized to the same stereotactic space, a high-dimensional deformable registration [19] was performed. The obtained 3-D displacement fields were converted to scalars by computing the Jacobian determinants at each voxel. Additionally, the scalar values were logarithmically transformed in order to distribute the values symmetrically around zero instead of an asymmetric distribution of solely positive values which the determinant of Jacobian matrix normally yields.

For better illustration, the datasets preprocessing is schematically depicted in Figure 1. The same datasets have been successfully used in a previous study [11].

3.2 Feature Extraction Methods

3.2.1 Univariate statistics (Mann-Whitney testing)

In order to reveal structural differences between schizophrenic and healthy brains, both VBM and DBM utilize a voxel-wise comparison between the groups, i.e. they employ univariate statistical analysis [12-13]. Therefore, Mann-Whitney testing (MW) was implemented as a univariate approach playing the role of a comparator to multivariate feature extraction methods.

MW indicates whether the tested variables come from the same distribution. Applying MW on each voxel, we selected those voxels which statistically belonged to different populations, i.e. they were important for distinguishing between FES and HC.

In general, when testing multiple hypotheses, one should correct for the number of false discoveries either with the familywise error rate (FWER [20]) or the false discovery rate (FDR [21]) corrections. However, those techniques are often too stringent [13]. Moreover, statistical significance does not necessarily imply discriminative power. Therefore, we regarded the resulting p-values as a selection criterion rather than a level of significance. In other words, we manually set the threshold for p-values dividing the voxels to those which were to be incorporated into classification and which were to be disregarded.

3.2.2 Intersubject PCA (isPCA)

Principal component analysis (PCA [22]) is a classic multivariate procedure seeking a transformation converting data to a set of orthogonal principal

components ordered according to the amount of variance they explain in the original data. However, PCA requires a covariance matrix of descriptors to be computed which, in the case of our data, was not feasible since the number of voxels in each image was over a half of a million.

Fortunately, it has been proven [23-24] that the eigenvectors v_j, corresponding to new components, can be computed from the eigenvectors w_j of a covariance matrix of subjects as:

$$v_j = \frac{X^T w_j}{\sqrt{q_j (N-1)}}$$

greatly reducing the demands on computation. The matrix X^T represents a transposed data matrix containing N subjects and q_j are the eigenvalues of the intersubject covariance matrix. Such a method, later named intersubject PCA (isPCA [25]), allowed us to preserve all the dataset variability using solely $N–1$ eigenvectors.

The feature space dimensionality can be progressively reduced by disregarding some of the new components. Since the eigenvectors are sorted in an ascending order of explained data variance, at first sight it may be tempting to get rid of the last ones. However, the amount of explained variance does not necessarily imply schizophrenia-related differences between FES and HC and therefore just as the first component can be, for instance, related to differences in liquor, the last component might be crucial for recognizing the proper affiliation of the subject.

Thus, before removing components from the ensuing analysis, we sorted the components according to their discriminative power measured by the level of their significance once tested with the subjects projected into the new feature space spanned by the components.

3.2.3 K-SVD

The aim of K-SVD [26] is to find the best sparse representation of the images x_i captured in X by solving

$$\min_{\Phi,C} \| X - \Phi C \|_2^F \ \textit{subject to} \ \| c_i \|_0 \leq s$$

where $i \in \{1,..,N\}$ and $\| . \|^F$ is the Frobenius norm.

Firstly, the dictionary Φ is initialized with l_2-normalized columns. The subsequent optimization process iteratively alternates between the sparse coding phase, when the optimization of each sparse coefficient vector c_i takes place, and the dictionary update phase. Here, for every atom in the dictionary, an error matrix representing the error of discarding the atom from the dictionary is computed, restricted to the

columns that correspond to non-zero sparse coefficients and finally it is decomposed using singular value decomposition (SVD). The update of both the dictionary and the loading matrix C is dependable on the matrices resulting from the SVD factorization.

When applied to brain imaging data, resulting atoms in the dictionary represent complex morphological patterns revealed by the algorithm in the brain scans. As the sparsity constraint s controls for the maximum number of atoms utilized to compose each image, its adjustment allows for extraction of small regions as well as global patterns [27].

Again, in order to gain the best set of atoms for schizophrenia diagnostic inference, the atoms can be sorted and the least discriminative ones can be discarded. However, due to the optimization process, we decided not to discard any atoms once they were learned.

3.2.4 Pattern-based Morphometry (PBM)

Although the K-SVD algorithm has emerged relatively recently, it has already been incorporated into a new methodology, pattern-based morphometry (PBM [27]). Despite its name, it is not a morphometry preprocessing technique as VBM and DBM described above. Instead, it provides a new perspective to multivariate pattern extraction using K-SVD, which is why we categorize it as a feature extraction method.

Unlike in the above-mentioned case where the dictionary is built upon data matrix of images, PBM introduces the idea of generating atoms from the so-called difference images. The generation of a difference images matrix is diagrammatically depicted in Figure 2.

For each image, using the Euclidean distance, a set of its k-nearest neighbors with a different affiliation is found. In other words, for an image a belonging to the group A (e.g. FES), its k most similar images belonging to the group B (e.g. HC) are searched for and vice-versa. Subsequently, the images are subtracted from their neighbors N. In the end, the resulting difference images matrices D_A and D_B are put together into a single matrix X. Assuming the images are in columns, the new matrix will have k-times more columns than the original matrix. At this point, the extracted atoms straightforwardly represent structural changes between FES and HC.

We created the new dictionary accordingly and used it in the same manner as with the K-SVD algorithm.

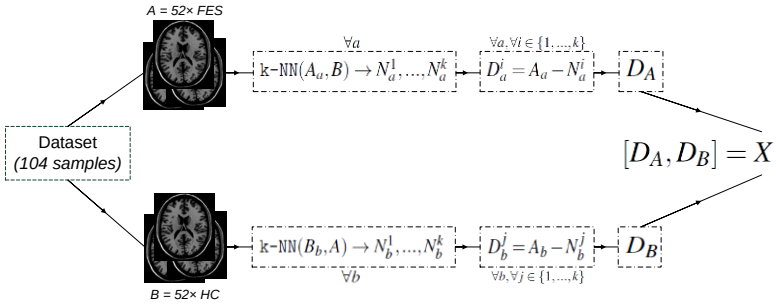

Figure 2
Generation of the Difference Images Matrix

3.3 Classification Pipeline

In order to evaluate the algorithms in real situations, they were incorporated into a classification pipeline (Figure 3).

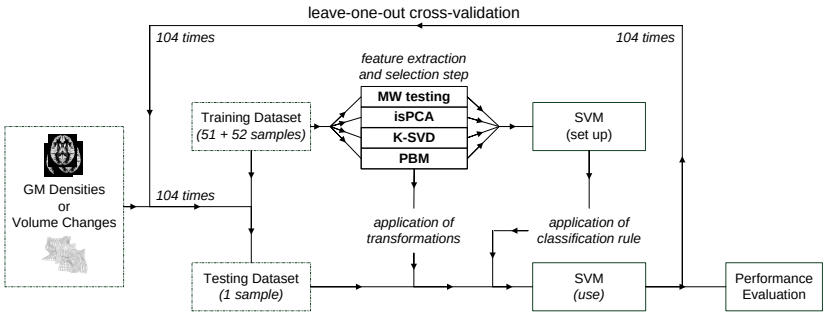

Figure 3
Classification Pipeline

Leave-one-out cross-validation scheme was used to assess the performance of the linear SVM classifier based on features extracted from the GM densities or the Volume changes datasets. Note that the feature extraction and selection steps were performed for each iteration of the cross-validation.

The rationale of the design was that alternating only the datasets and feature extraction methods in otherwise rigidly fixed pipeline settings facilitated their later comparison.

In terms of the performance evaluation, classification accuracy, sensitivity and specificity were used as its metrics.

4 Preliminary Experiments

4.1 Anticipated Behavior – a Toy Example

Before proceeding to classification on real datasets, we created a synthetic dataset consisting of 2-D images of 10 hand-drawn circles and 10 hand-drawn triangles in order to illustrate the difference between the behavior of univariate and multivariate feature extraction approaches.

Each image in the synthetic dataset consisted of 50,184 pixels with values ranging from 0 to 255. Figure 4 shows the pixels selected by MW and the most discriminative patterns revealed by isPCA, K-SVD, and PBM respectively from left to right. The gray scale patterns are displayed in colors, where yellow represents the most and dark blue the least significant pixels.

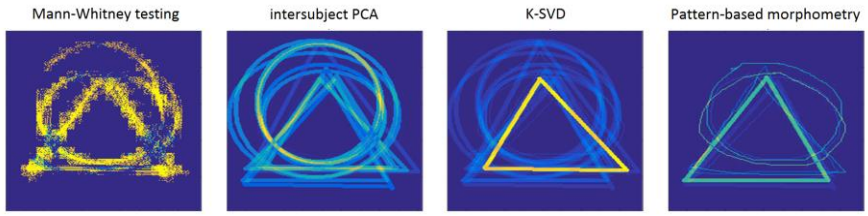

Figure 4
Features Extracted from the Synthetic Dataset

The toy example underlines what is known from the theory. Whereas univariate statistics dismantled geometric shapes into pixels, multivariate methods were capable of recognizing complex patterns[1] while dealing with the same data. Moreover, in the case of K-SVD and PBM, the most discriminative feature resembled a representative from the group of triangles.

4.2 Parameters Tuning

The last step preceding final classification was tuning the parameters of the employed algorithms as their proper settings enhance the classification performance. Table 1, summarizes the parameters included for tuning.

As their influence on the classification was unknown, we evaluated the classification cross-validated accuracies for various parameters settings equidistantly distributed over the parameter space in a way to capture a behavior of each of the parameters separately for the GM Densities and the Volume

[1] whole circles and triangles

Changes datasets, hence mapping the parameter spaces. In order to reduce computational costs, random projection (RP), with a random matrix suggested in [28], was utilized to reduce the dimensionality of the problem.

Table 1
List of Parameters of the Feature Extraction Algorithms

Algorithm	Parameter	Token
MW	p-values threshold	t
isPCA	number of retained components	c
K-SVD	number of atoms	$a_{k\text{-}svd}$
	sparsity constraint	$s_{k\text{-}svd}$
PBM	number of atoms	a_{pbm}
	sparsity constraint	s_{pbm}
	number of nearest neighbors	k

4.3 Final Parameters Settings

The parameters settings reaching the highest classification accuracies are displayed in Table 2. As sparsity constraint and the number of nearest neighbors did not exhibit any trend, we set them, in accordance with [27], to 5 and 3 correspondingly.

Table 2
List of Final Parameters Settings for Both of the Datasets

Algorithm	Token	Value	
		GM Densities	*Volume Changes*
MW	t	0.01	0.05
isPCA	c	11	84
K-SVD	$a_{k\text{-}svd}$	1	103
	$s_{k\text{-}svd}$	5	5
PBM	a_{pbm}	1	309
	s_{pbm}	5	5
	k	3	3

5 Classification Results

All tested feature extraction algorithms with their final parameters settings being put through the classification pipeline. Cross-validated classification accuracies along with sensitivities and specificities for each of the methods and both the datasets are shown in Table 3.

On average, the classification methods with the use of Volume Changes features resulting from DBM outperformed the classification methods using the GM Densities features resulting from VBM. Comparing the classification algorithms, the highest accuracy, slightly over 70%, was attained by PBM.

Whereas the across-datasets performance increased for all multivariate methods when switched from GM Densities to Volume Changes, it diminished for the univariate statistics.

Table 3
Classification Performance on Both Datasets

	GM Densities			***Volume Changes***		
Algorithm	Accuracy [%]	Sensitivity [%]	Specificity [%]	Accuracy [%]	Sensitivity [%]	Specificity [%]
MW	67.31	63.46	71.15	66.35	65.38	67.30
isPCA	68.27	63.46	73.08	69.23	71.15	67.31
K-SVD	65.38	63.46	67.31	69.23	69.23	69.23
PBM	64.42	63.46	65.38	**70.19**	**69.23**	**71.15**

6 Discussion

The main distinction we would like to stress here is the difference for the results in the different types of features: GM Densities and Volume Changes. Considering the datasets are two modalities of the same data, we were able to evaluate the differences between the VBM and DBM approaches to the preprocessing of the MRI data.

The most noteworthy piece of information stems from the parameter settings, indicating the number of features (components, atoms) that are optimal for the classification. In the case of the GM Densities dataset, the best classification results were achieved with the minimum of features retained. On the contrary, the Volume Changes dataset yielded the best results when the number of features was set at its highest values.

Also, the most discriminative isPCA component of GM Densities captured 12.5 times more variance of the original data than the one calculated from the covariance matrix corresponding to deformations. When comparing components with the most variance explaining the ratio was approximately 2.5. Such findings are in correspondence with [25], where isPCA components are evaluated in more detail.

Furthermore, for the Volume Changes dataset, multivariate approaches slightly outperformed univariate MW, serving as a mere feature selection. However, the differences are not statistically significant.

All the aforementioned behavior indicates that Volume Changes concealed more sophisticated patterns, than can be discovered, disregarding voxel-to-voxel interactions. Consequently, our results confirm that whereas VBM serves mainly for extracting information about changes on a local scale, DBM preserves information from a wider region.

At this point, it should be stressed that accuracies around 70% are insufficient for clinical practice. Nevertheless, our findings can serve as guidelines to those dealing with unknown parameter spaces. With VBM, the best parameter settings in terms of the number of retained features will most likely lay among low values. In the case of DBM, the opposite statement is the most probable.

We also suggest that studies utilizing DBM as a preprocessing tool should reach for multivariate feature extraction approaches as they appear to be superior on such data. Interestingly, a novel PBM technique provided superb results in comparison to others and thus it should be considered as a valid candidate when deciding on a method of extracting brain differences patterns. Moreover, PBM improves the ratio of the number of subjects over the number of features, as it generates a dataset consisting of more images.

Conclusions

This work presented an analysis of two brain morphometry techniques and various feature extraction methods often utilized in the computer-aided diagnostics of schizophrenia. The methodology was incorporated into a classification pipeline and applied to distinguish between first-episode patients and healthy controls on the basis of magnetic resonance images of their brains. First, each method was thoroughly examined in order to explore its parameters and their influence on the classification. Then, the methods were evaluated in terms of classification performance. Our findings confirmed the distinction between VBM and DBM and resulted in recommendations on the numbers of retained features. We also showed that by applying multivariate machine learning techniques, such as, PBM on data preprocessed with the DBM approach have beneficial effects on classification results.

References

[1] N. E. van Haren, W. Cahn, H. E. H. Pol, and R. S. Kahn: The Course of Brain Abnormalities in Schizophrenia: Can We Slow the Progression?, Journal of Psychopharmacology, Vol. 26, No. 5 suppl, 2012, pp. 8-14

[2] D. O. Perkins, H. Gu, K. Boteva, and J. A. Lieberman: Relationship Between Duration of Untreated Psychosis and Outcome in First-Episode Schizophrenia: A Critical Review and Meta-Analysis, American Journal of Psychiatry, Vol. 162, No. 10, 2005, pp. 1785-1804

[3] S. M. Lawrie, B. Olabi, J. Hall, and A. M. McIntosh: Do We Have Any Solid Evidence of Clinical Utility about the Pathophysiology of Schizophrenia?, World Psychiatry, Vol. 10, No. 1, 2011, pp. 19-31

[4] G. Orrù, W. Pettersson-Yeo, A. F. Marquand, G. Sartori, and A. Mechelli: Using Support Vector Machine to Identify Imaging Biomarkers of Neurological and Psychiatric Disease: A Critical Review, Neuroscience & Biobehavioral Reviews, Vol. 36, No. 4, 2012, pp. 1140-52

[5] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller: Introduction to Machine Learning for Brain Imaging, NeuroImage, Vol. 56, No. 2, 2011, pp. 387-99

[6] I. C. Wright, S. Rabe-Hesketh, P. W. Woodruff, A. S. David, R. M. Murray, and E. T. Bullmore: Meta-Analysis of Regional Brain Volumes in Schizophrenia, The American Journal of Psychiatry, Vol. 157, No. 1, 2000, pp. 16-25

[7] E. Antonova, T. Sharma, R. Morris, and V. Kumari: The Relationship between Brain Structure and Neurocognition in Schizophrenia: A Selective Review, Schizophrenia Research, Vol. 70, No. 2-3, 2004, pp. 117-45

[8] R. Honea, T. J. Crow, D. Passingham, and C. . Mackay: Regional Deficits in Brain Volume in Schizophrenia: A Meta-Analysis of Voxel-Based Morphometry Studies, American Journal of Psychiatry, Vol. 162, No. 12, 2005, pp. 2233-45

[9] A. M. Shepherd, K. R. Laurens, S. L. Matheson, V. J. Carr, and M. J. Green: Systematic Meta-Review and Quality Assessment of the Structural Brain Alterations in Schizophrenia, Neuroscience & Biobehavioral Reviews, Vol. 36, No. 4, 2012, pp. 1342-56

[10] D. Antonius, V. Prudent, Y. Rebani, D. D'Angelo, B. A. Ardekani, D. Malaspina, and M. J. Hoptman: White Matter Integrity and Lack of Insight in Schizophrenia and Schizoaffective Disorder, Schizophrenia Research, Vol. 128, No. 1-3, 2011, pp. 76-82

[11] D. Schwarz, and T. Kašpárek: Brain Morphometry of MR Images for Automated Classification of First-Episode Schizophrenia, Information Fusion, Vol. 19, 2014, pp. 97-102

[12] C. Gaser, I. Nenadic, B. R. Buchsbaum, E. A. Hazlett, and M. S. Buchsbaum: Deformation-based Morphometry and Its Relation to Conventional Volumetry of Brain Lateral Ventricles in MRI, NeuroImage, Vol. 13, No. 6, 2001, pp. 1140-45

[13] A. Mechelli, C. J. Price, K. J. Friston, and J. Ashburner: Voxel-based Morphometry of the Human Brain: Methods and Applications, Current Medical Imaging Reviews, Vol. 1, No. 2, 2005, pp. 105-13

[14] C. Scanlon, S. G. Mueller, D. Tosun, I. Cheong, P. Garcia, J. Barakos, M. W. Weiner, and K. D. Laxer: Impact of Methodologic Choice for Automatic Detection of Different Aspects of Brain Atrophy by Using Temporal Lobe Epilepsy as a Model, American Journal of Neuroradiology, Vol. 32, No. 9, 2011, pp. 1669-76

[15] C. Davatzikos: Why Voxel-Based Morphometric Analysis Should Be Used with Great Caution When Characterizing Group Differences, NeuroImage, Vol. 23, No. 1, 2004, pp. 17-20

[16] E. Zarogianni, T. W. J. Moorhead, and S. M. Lawrie: Towards the Identification of Imaging Biomarkers in Schizophrenia, Using Multivariate Pattern Classification at a Single-Subject Level, NeuroImage: Clinical, Vol. 3, 2013, pp. 279-89

[17] S. R. Kay, A. Fiszbein, and L. A. Opler: The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia, Schizophrenia Bulletin, Vol. 13, No. 2, 1987, pp. 261-76

[18] J. Ashburner: A Fast Diffeomorphic Image Registration Algorithm, NeuroImage, Vol. 38, No. 1, 2007, pp. 95-113

[19] D. Schwarz, T. Kašpárek, I. Provazník, and J. Jarkovský: A Deformable Registration Method for Automated Morphometry of MRI Brain Images in Neuropsychiatric Research, IEEE Transactions on Medical Imaging, Vol. 26, No. 4, 2007, pp. 452-61

[20] H.-Y. Kim, Statistical notes for clinical researchers: post-hoc multiple comparisons, Restorative Dentistry & Endodontics, Vol. 40, No. 2, 2015, pp. 172-76

[21] Y. Benjamini, and Y. Hochberg: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society, Vol. 57, No. 1, 1995, pp. 289-300

[22] I. T. Jolliffe: Principal Component Analysis, Springer, New York, 2012

[23] C. E. Thomaz, J. P. Boardman, S. Counsell, D. L. G. Hill, J. V. Hajnal, A. D. Edwards, M. A. Rutherford, D. F. Gillies, and D. Rueckert: A Multivariate Statistical Analysis of the Developing Human Brain in Preterm Infants, Image and Vision Computing, Vol. 25, No. 6, 2007, pp. 981-94

[24] O. Demirci, V. P. Clark, V. A. Magnotta, N. C. Andreasen, J. Lauriello, K. A. Kiehl, G. D. Pearlson, and V. D. Calhoun: A Review of Challenges in the Use of fMRI for Disease Classification / Characterization and A Projection Pursuit Application from A Multi-Site fMRI Schizophrenia Study, Brain Imaging and Behavior, Vol. 2, No. 3, 2008, pp. 207-26

[25] E. Janoušová, D. Schwarz, and T. Kašpárek: Combining Various Types of Classifiers and Features Extracted from Magnetic Resonance Imaging Data

in Schizophrenia Recognition, Psychiatry Research: Neuroimaging, Vol. 232, No. 3, 2015, pp. 237-49

[26] M. Aharon, M. Elad, and A. Bruckstein: K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE Transactions on Signal Processing, Vol. 54, No. 11, 2006, pp. 4311-22

[27] B. Gaonkar, K. Pohl, and C. Davatzikos: Pattern Based Morphometry, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011, Lecture Notes in Computer Science, Vol. 6892, No. Pt 2, 2011, pp. 459-66

[28] D. Achlioptas: Database-Friendly Random Projections: Johnson-Lindenstrauss with Binary Coins, Journal of Computer and System Sciences, Vol. 66, No. 4, 2003, pp. 671-87