Wavelet Features for Recognition of First Episode of Schizophrenia from MRI Brain Images

Petr DLUHOŠ[1], Daniel SCHWARZ[1], Tomáš KAŠPÁREK[2,3]

[1] Inst. of Biostatistics and Analyses, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic
[2] Dept. of Psychiatry, University Hospital Brno and Masaryk University, Jihlavská 20, 625 00 Brno, Czech Republic
[3] Behavioral and Social Neuroscience Group, CEITEC – Central European Inst. of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic

schwarz@iba.muni.cz

Abstract. *Machine learning methods are increasingly used in various fields of medicine, contributing to early diagnosis and better quality of care. These outputs are particularly desirable in case of neuropsychiatric disorders, such as schizophrenia, due to the inherent potential for creating a new gold standard in the diagnosis and differentiation of particular disorders. This paper presents a scheme for automated classification from magnetic resonance images based on multiresolution representation in the wavelet domain. Implementation of the proposed algorithm, utilizing support vector machines classifier, is introduced and tested on a dataset containing 104 patients with first episode schizophrenia and healthy volunteers. Optimal parameters of different phases of the algorithm are sought and the quality of classification is estimated by robust cross validation techniques. Values of accuracy, sensitivity and specificity over 71% are achieved.*

Keywords

Schizophrenia, machine learning, neuroimaging, classification, wavelet transform, MRI.

1. Introduction

Schizophrenia is a disabling psychiatric disorder affecting many people worldwide. It manifests in a variety of symptoms ranging from misinterpretation of reality and delusions to disorganization of thinking and behavior. It is associated with progressive altered brain functions during the course of the illness [1]. Findings in different areas of the brain were published in numerous reviews and meta-analyses [2–5]. However, many of these findings are inconsistent or even contradictory, which could indicate the heterogeneity of this severe disorder [6].

During the last two decades, methods for quantitative evaluation of the morphological changes in the human brain have been rapidly developed thanks to the invention of magnetic resonance imaging (MRI), which provides a good contrast among tissues as well as the necessary spatial resolution. The historically first quantitative approach to in-vivo brain morphology is called MRI volumetry. The region of interest (ROI) is manually marked in the MRI images in voxel-by-voxel manner and its total volume is computed by multiplying the number of voxels by a voxel volume. The most commonly used morphometric method so far has been the voxel-based morphometry (VBM) [7]. It seeks statistical difference in a volume of brain tissues at the level of individual voxels. The method employs spatial normalization based on image registration, segmentation to binary tissue images (white matter, gray matter, cerebrospinal fluid) and Gaussian smoothing. These pre-processing steps are then followed by univariate statistical analysis aimed at identifying the voxels differing in the amount of a selected tissue between groups of patients and normal healthy controls. Different approach to brain morphometry is represented by the method called deformation-based morphometry (DBM) [8]. It evaluates changes in the position, shape and volume of the brain areas using deformation fields obtained during nonlinear registration to a common template. The vector values in the deformation field describe the translations of each voxel of the image. These values can be either evaluated directly by means of multivariate statistics [9] or used to compute a scalar value expressing the local change in each voxel (e.g. Jacobian determinants representing local volume changes [10]). In this case, similarly to the VBM approach, univariate analysis is applied to the scalar field. Tensor-based morphometry (TBM) [9] is a method also based on the information derived during initial registration. Complete information about local deformations is contained in a tensor field and multivariate statistics are applied to the tensor manifold instead of the derived indices [11]. A relatively new approach, source-based morphometry (SBM) [12], utilizes independent component analysis (ICA) [13] to identify maximally independent sources responsible for the variability of the images. After the same pre-processing steps as in VBM, several "source patterns" are extracted and statistical tests are applied on all components. Besides the most frequent methods described above, other approaches to brain morphometry exist, e.g., surface-based morpho-

metry [14] utilizing spherical representation of the gray matter surface, and pattern-based morphometry [15], that is based on sparse image representation.

The methods of automated brain morphometry are often used for extracting interesting areas and features that are subsequently used for classification and computer-aided diagnostics of neuro-psychiatric disorders. The most common classifier used for recognition of schizophrenia patients based on their MRI data is support vector machines (SVM) [16]. The values of accuracy achieved by using this classifier vary between 66% and 90% in recent studies [17–22]. It is worth to note that a substantially lower accuracy of only 70% was reported on a significantly greater dataset (277 subjects) [23]. SVM was also used recently for classification of patients with first episode of schizophrenia (FES) with accuracy reported from 54% to 73% [24], [25]. Other popular methods for classification are based on various discriminant analyses, such as the Fisher linear discriminant analysis [26] or the maximum-uncertainty linear discriminant analysis [27]. The accuracy achieved using these methods ranges between 80% to 98% for chronic schizophrenia patients [28–30] and from 61% to 81% for FES only patients [27], [31], [32]. Other studies use also artificial neural networks (accuracy 70% to 76%) [33], [34], logistic regression (accuracy 86%) [35], k-nearest neighbor (accuracy 80%) [36] or the projection pursuit algorithm (accuracy 80% to 90%) [37], [38].

The aim of the presented study is to investigate the possibility of using multiresolution representation of medical images in the wavelet domain for automated recognition of neuro-psychiatric disorders, particularly schizophrenia. Detailed analysis of the topic is provided in the thesis [39], which introduced the preliminary version of this paper.

2. Methods

The proposed algorithm for recognizing schizophrenia patients from healthy subjects based on their structural MRI brain images consists of three main steps. Firstly, the images are transformed into a domain providing sparse representation. Secondly, the best discriminating features in the new domain are selected. And lastly, a supervised classifier is applied to the selected features. In this study, we implemented several variants for each step of the proposed classification scheme and then performed systematic experiments, in order to find a setting showing the best classification results.

2.1 Study Design and Subjects

Fifty-two patients (mean age 24, SD 5.1 years) admitted to the all-male unit of the Department of Psychiatry, Masaryk University in Brno, for first episode of schizophrenia were recruited. Their symptoms fulfilled the criteria for schizophrenia for the first time when admitted to the department, including the time criterion – duration of symptoms longer than 1 month. Diagnosis was established during clinical interviews held in compliance with the International Statistical Classification of Disease and Related Health Problems (ICD-10) research criteria. Exclusion criteria pertained to substance dependence detected by clinical evaluation and urine toxicology tests, neurological or systemic disease with known relationship to brain alteration detected by clinical evaluation, physical and neurological examinations, serum and urine chemistry and blood count, serological examination for neurotropic agents, clinical evaluation of MRI scans, and contraindications for MRI. Fifty-two healthy subjects (matched for age – mean age 24, SD 3.7 years – gender and handedness) were recruited from the community, the local staff, and medical students. The exclusion criteria (as assessed during clinical interviews performed by a trained psychiatrist) were substance dependence, family history of axis I psychiatric conditions, personal history of axis I psychiatric condition, neurological or somatic conditions affecting the structure or function of the brain, and the contraindications for MRI examination. The study was approved by the local ethics committee and all subjects signed the informed consent.

The dataset contained 104 T1-weighted images of the entire head obtained with the 1.5 T MR device (sagittal tomographic plane thickness was 1.17 mm, the in-plane resolution was 0.48 mm x 0.48 mm, 3-D field of view contained 160 x 512 x 512 voxels). Gray matter (GM) tissue segments were obtained from all images after correction for bias-field inhomogeneity, spatial normalization and segmentation [40] with the use of VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm/) implemented in SPM8 framework (http://www.fil.ion.ucl.ac.uk/spm/). Spatial normalization steps involved affine registration to standard SPM T1 template followed by fast diffeomorphic registration algorithm DARTEL [41]. GM tissue segments were modulated with the determinant of Jacobian matrices of the deformations to account for registration related changes in local volumes. The modulated GM segment images were finally smoothed with 8 mm FWHM Gaussian kernel to enable intersubject comparisons and to render data distribution more normal.

2.2 Sparsity and Wavelet Transform

Generally, we call a signal sparse if most of its samples are equal to zero. Natural signals such as images are usually not sparse in the space domain. However, they can be often transformed into a suitable representation, in which they are sparse or at least weakly sparse in the sense that most of the coefficients in the new domain are almost zero [42]. For natural images, one of the transforms producing such behavior is based on wavelets.

Wavelet transform decomposes a signal into a weighted sum of wavelets - functions of certain form [43]. This new representation captures not only the time

course of the signal, but also its properties in the frequency domain. A small fraction of the representation coefficients with the highest magnitudes retains the major part of information contained in the signal. Moreover, it is usually the substantial part of the information, because noise tends to be contained mainly in the small coefficients [43]. For practical applications on discrete signals, the discrete wavelet transform (DWT) was developed. It originated in the Mallat's multiresolution decomposition scheme [44] and led to a huge number of applications in various fields. The signal is iteratively decomposed into detail and approximation coefficients by combination of two operations: 1) convolution with special finite response filters and 2) subsampling. The approximation coefficients are then taken as input for a new level of decomposition. The output of this procedure is several sequences of coefficients describing details of the signal at different levels and one sequence of coefficients composing its rough approximation. DWT can be easily generalized into more dimensions. For 2-D images, one dimensional DWT is applied on the rows, columns and diagonals leading to three sets of detail coefficients for each level. Similarly, seven sets of detail coefficients are generated for 3-D images.

The number of wavelet coefficients approximately corresponds to the number of voxels in the transformed image, which was around 2 million in our case. In order to reduce the noise contained in the data and to lower its dimensionality, the coefficients from all levels of DWT decomposition were sorted according to their maximum magnitude among all subjects and those below a certain threshold were removed. The optimal value of this threshold was one of the parameters which had to be determined experimentally, since it represents the trade-off between lower dimensionality and better noise reduction on the one hand and lesser losses of potentially useful information on the other hand. This operation led to a reduction in the number of coefficients by the factor of 5-100, depending on the selected threshold. The remaining coefficients continued to the next steps as features describing the subjects.

Systematic optimization of the wavelet family and the level of decomposition used for DWT could not be performed due to high computational demands of the experimental procedure. Based on the results of our preliminary experiments, sym5 wavelet from the Symlet family and four levels of decomposition were chosen in our computations. This wavelet family was shown to provide good results in natural image compression [45].

2.3 Feature Selection

After the feature extraction using DWT, a limited number of features with the best discriminative power are selected. Further reduction of the feature space dimensionality helps to match better the number of subjects in the dataset as well as to avoid the features carrying only low information about the differences between the studied groups. We examined several criteria for determining the discriminative power of individual features while testing the effect of the number of the best features selected for subsequent classification. The studied criteria taken from the literature were:

Fisher's discriminant ratio [26]:

$$FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^{\,2} + \sigma_2^{\,2}} \tag{1}$$

Bhattacharyya distance [46]:

$$Bha = \frac{1}{4}\frac{(\mu_1 - \mu_2)^2}{\sigma_1^{\,2} + \sigma_2^{\,2}} + \frac{1}{2}\ln\left(\frac{\sigma_1^{\,2} + \sigma_2^{\,2}}{2\sigma_1\sigma_2}\right) \tag{2}$$

and divergence [47]:

$$Div = \frac{1}{2}\left(\frac{\sigma_1^{\,2}}{\sigma_2^{\,2}} + \frac{\sigma_2^{\,2}}{\sigma_1^{\,2}} - 2\right) + (\mu_1 - \mu_2)^2\left(\frac{1}{\sigma_1^{\,2}} + \frac{1}{\sigma_2^{\,2}}\right) \tag{3}$$

For each feature, μ_1 and μ_2 represent the mean values of this feature in the first and second group and σ_1 and σ_2 represent the variances of the feature values in each group. Apart from the criteria (1)-(3), we proposed and tested three others. Two of them were modifications of FDR designed for better robustness in case of non-normally distributed data:

$$medFDR = \frac{(med_1 - med_2)^2}{\sigma_1^{\,2} + \sigma_2^{\,2}}; \tag{4}$$

$$quantileFDR = \frac{(med_1 - med_2)^2}{(\sigma_1^{\,*})^2 + (\sigma_2^{\,*})^2} \tag{5}$$

where med_1 and med_2 are medians and $\sigma_1^{\,*}$ and $\sigma_2^{\,*}$ are estimates of standard deviations by interquartile range $\sigma^* = Q_{84} - Q_{16}$. The last criterion designed is:

$$variances = \frac{\Sigma^2}{\sigma_1^{\,2} + \sigma_2^{\,2}}$$

where $\sum^2$ stands for the variance of the tested feature among all subjects. High values of *variances* are expected for the features, which show a high variance in the whole dataset and are homogenous inside the studied groups at the same time.

2.4 Classification

Features extracted and selected in the previous steps were used for training a classifier – the SVM classifier was selected based on the results of our preliminary experiments. Three implementations of the SVM classifier from the PRTools (http://prtools.org) toolbox for MATLAB were tested. They differed in the kernel functions (linear: SVC, NUSVC and radial basis functions: RBSVC were used) and in the regularization method.

The entire procedure of predicting a class for a new subject works as follows: MRI image is pre-processed as

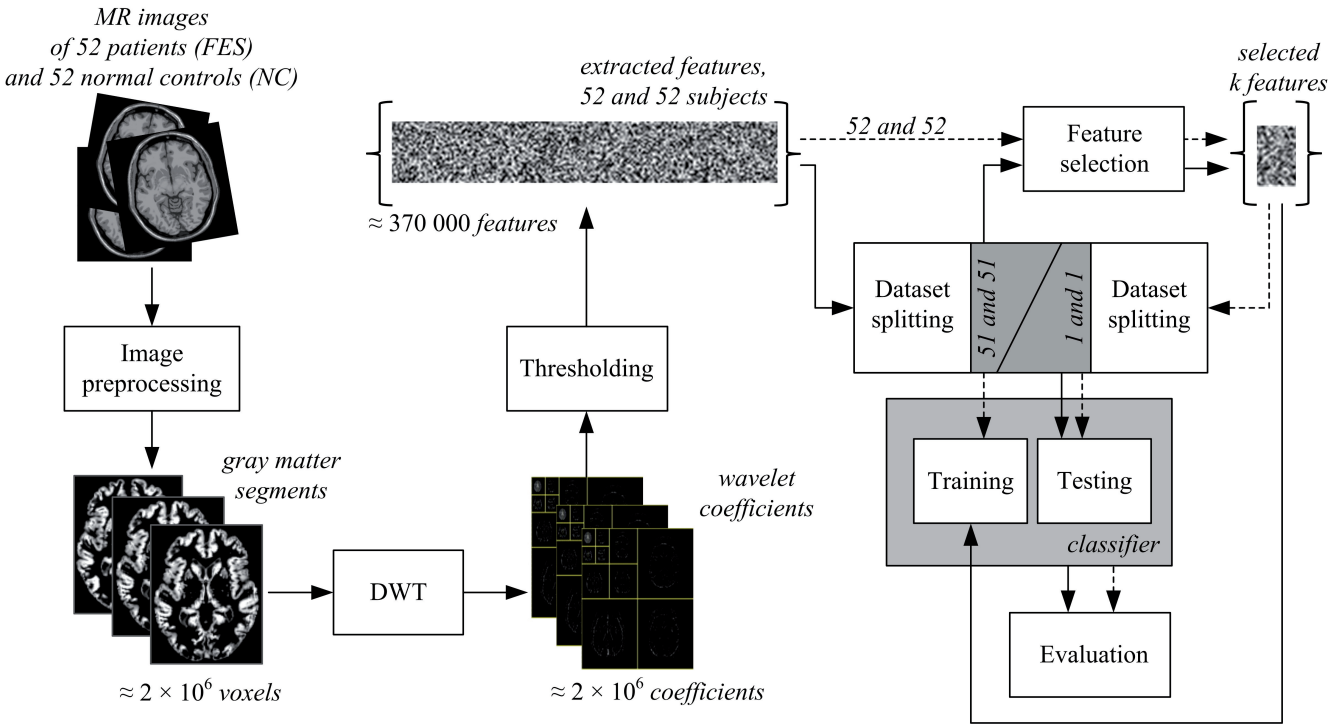

Fig. 1. The scheme of the proposed classification algorithm with correct (solid lines) and incorrect (dotted lines) cross validation. MRI images are pre-processed and the gray matter is segmented. The resulting images are transformed into DWT coefficients. Only the coefficients with magnitude greater than a chosen threshold are extracted as potential features. In case of correct cross validation, the dataset is divided into testing and training subsets and further steps are performed repeatedly with the subjects in the training subset only. A limited number of the most discriminant features is selected and used for training a classifier. The performance of the classification is tested on subjects in the testing subset. Incorrectness of the dotted variant lies in the reversed order of dataset splitting and feature selection, as the feature selection step relies on the information about subjects in the testing subset.

described above, the pre-processed image is transformed into the wavelet domain and the coefficients with magnitude under a threshold, computed over the whole data set, are removed. Then the most discriminative coefficients are selected as features for the classification. The discrimination criteria computations and training of the classifier are performed only over the training subjects. The class for the unknown subject is finally predicted using its values of the selected features.

3. Experiments and Results

We performed several series of experiments in order to find the best parameters for each step of the classification algorithm. The parameters were:

- usage of approximation coefficients (YES, NO),
- the threshold for removing small magnitude coefficients (0, 0.001, 0.01, 0.1, 1),
- the criterion for evaluating the discriminative power of the features,
- an algorithm for choosing the number of selected features (nested cross validation – fold sizes of 3, 17 and 51; bootstrap selection – testing subset sizes 2, 6, 10 and 34; voting of independent classifiers trained progressively on the first 1, 2, …, k features – k = 20, 100, 1000), and
- the SVM classifier implementation.

The quality of classification for each parameter setting was evaluated using stratified 52-fold cross validation on the whole dataset. All runs were repeated one hundred times to improve robustness of the estimates. As testing all 1800 possible combinations of the setting parameters would not be computationally feasible, we chose a default setting and tested the effects of the parameters one at a time. The default setting was: using approximation coefficients, removing coefficients < 0.01, FDR criterion, voting of classifiers with 1-1000 best features and implementation of SVC. The results of the experiments are summarized in Tab. 1.

4. Discussion

The results in Tab. 1 show that almost all parameters in the proposed algorithm had a significant impact on the quality of classification. In the first step, the best results were achieved with using only the detail coefficients greater than 0.01. Lower thresholds left too much noise in the data while higher ones probably removed information necessary for discriminating the studied groups. In the second step, the criteria FDR, Bhattacharyya and variances performed similarly well and significantly better than the three remaining criteria. Quite surprisingly, voting methods were superior to the methods based on selection via nested validation cycle. The reason may be the high heterogeneity

of schizophrenia manifestations in the images [6], in combination with the relatively small dataset. This might have caused for the optimal number of features selected on a smaller group in a nested cycle not to be generalized for all subjects in the testing set. Ensemble methods, on the other hand, are better in such situations [48]. This is probably also the reason for the rather inferior results of RBSVC classifier implementation because the radial basis kernel functions were optimized with 5-fold cross validation.

Parameter	Value	Sensitivity [%]	Specificity [%]	Accuracy [%]
approx coeffs.	**YES**	**71.06**	**74.75**	**72.90**
	NO	71.63	72.71	71.87
removing coeffs.	0	64.21	67.81	66.01
	0.001	64.69	69.90	67.30
	0.01	**71.06**	**74.75**	**72.90**
	0.1	59.35	66.71	63.03
	1	59.46	64.85	62.15
discrimination crit.	**FDR**	**71.06**	**74.75**	**72.90**
	Bhattacharyya	**71.31**	**74.79**	**73.05**
	Divergence	65.23	72.08	68.65
	medFDR	57.83	58.31	58.07
	quantileFDR	53.77	58.19	55.98
	Variances	**71.73**	**74.67**	**73.20**
number of features	51-fold	56.15	52.50	54.33
	17-fold	52.98	54.13	53.56
	3-fold	53.04	56.94	54.99
	bootstrap-2	58.00	48.23	53.12
	bootstrap-6	54.73	52.08	53.40
	bootstrap-10	52.85	54.08	53.46
	bootstrap-34	51.50	56.38	53.84
	voting 1-20	57.52	66.33	61.92
	voting 1-100	60.23	64.27	62.25
	voting 1-1000	**71.06**	**74.75**	**72.90**
Classifier	**SVC**	**71.06**	**74.75**	**72.90**
	NUSVC	62.90	73.98	68.44
	RBSVC	64.23	75.00	69.62

Tab. 1. Effects of different settings on the accuracy, sensitivity and specificity of the proposed classification algorithm. Best values for each parameter are highlighted in bold. Each value was estimated by averaging the results of 100 independent runs of stratified 52-fold cross validation on the whole dataset. The implementations of SVC and NUSVC both train the SVM classifier with linear kernel and differ only in the regularization method. The function RBSVC searches optimal kernel in the form of radial basis functions by nested cross validation.

The overall best quality of classification was achieved for default configuration of all parameters, except for the discriminating criterion for which the criterion variances showed the best results – accuracy 73.20% (SD 2.07), sensitivity 71.73% (SD 3.00) and specificity 74.67% (SD 2.58). When comparing these results with other studies, only those working with patients with FES should be taken into account. Chronic schizophrenia patients are expected to have considerable morphological changes caused by long-term progression of the disease and medication [49], [50]. Another factor that can cause overestimation of the classification quality is an incorrectly performed cross validation [38]. Information about correct classification of the subjects in the testing set must not be used in any step of the learning process. If the most discriminative features were computed on the whole data set (see the dotted path in Fig. 1), the results would be biased towards the correct classification. Due to the way in which the proposed classifier works, separation into training and testing sets had to be done just prior to selecting the most discriminative features. This way of cross validation corresponds to real application of the algorithm for prediction. The stratified 52-fold cross validation was chosen over the more frequent leave-one-out approach, in order to avoid possible bias caused by uneven proportions of subjects from different classes in the training and testing subsets [51].

The computations that led to the optimal setting parameters of the classification algorithm took several days due to thousands of repetitions of the whole classification procedure (100 repetitions × 52 validation runs × best feature selection × classifier optimization). For this reason, it was not feasible to test the other two key parameters of DWT – the wavelet mother function and the level of decomposition. The fourth level of decomposition was preset based on our preliminary results and the wavelet sym5 was selected based on the results in the study [45] as well as from the previous studies of the same authors, who reported a good performance of the wavelets from Symlet family in coding of natural images. Inclusion of these two parameters into properly cross-validated experiments focused on pattern recognition provides space for further potential improvement in the classification quality [52].

The achieved results are comparable with the recent studies aimed at automated classification of patients with FES (accuracy 54% to 81%) [24], [25], [27], [31], [32].

5. Conclusion

We have proposed a scheme for automated discrimination between MRI images of patients with first episode of schizophrenia and healthy controls using the multiresolution representation in the wavelet domain. For the best found settings of our classification algorithm we achieved 73.20% accuracy, 71.73% sensitivity and 74.67% specificity robustly estimated using 100 repeats of 52-fold stratified cross validation. These values are comparable to state-of-the-art MRI-based methods for automated classification of schizophrenia. Robust estimates of sensitivity and specificity in combination with correctly performed cross validation and relatively large dataset show that wavelet transform provides a useful tool for extracting important information from medical images. On the other hand, accuracy values around 70% are not sufficient enough for applying in automated diagnosis of schizophrenia based on MRI data in clinical practice.

Acknowledgements

This work was supported by the research grant of the Ministry of Health, Czech Republic No. 13359-4. The

access to the CERIT-SC computing and storage facilities provided under the program Center CERIT Scientific Cloud, part of the Operational Program Research and Development for Innovations, reg. no. CZ. 1.05/3.2.00/08.0144 is appreciated.

References

[1] HAREN, N. V., CAHN, W., POL, H. H., KAHN, R. S. The course of brain abnormalities in schizophrenia: can we slow the progression? *Journal of Psychopharmacology / British Association for Psychopharmacology,* 2012, vol. 26, no. 5, p. 8 – 14.

[2] SHENTON, M. E., DICKEY, C. C., FRUMIN, M., MCCARLEY, R. W. A review of MRI findings in schizophrenia. *Schizophrenia Research,* 2001, vol. 49, no. 1 – 2, p. 1 – 52.

[3] HONEA, R., CROW, T. J., PASSINGHAM, D., MACKAY, C. E. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *The American Journal of Psychiatry,* 2005, vol. 162, no. 12, p. 2233 – 2245.

[4] ELLISON-WRIGHT, I., GLAHN, D. C., LAIRD, A. R., THELEN, S. M., BULLMORE, E. The anatomy of first-episode and chronic schizophrenia: An anatomical likelihood estimation meta-analysis. *The American Journal of Psychiatry,* 2008, vol. 165, no. 8, p. 1015 to 1023.

[5] SUN, J., MALLER, J. J., GUO, L., FITZGERALD, P. B. Superior temporal gyrus volume change in schizophrenia: A review on Region of Interest volumetric studies. *Brain Research Reviews,* 2009, vol. 61, no. 1, p. 14 – 32.

[6] NENADIC, I., GASER, C., SAUER, H. Heterogeneity of brain structural variation and the structural imaging endophenotypes in schizophrenia. *Neuropsychobiology,* 2012, vol. 66, no. 1, p. 44 to 49.

[7] MECHELLI, A., PRICE, C. J., FRISTON, K. J., ASHBURNER, J. Voxel-Based Morphometry of the human brain: Methods and applications. *Current Medical Imaging Reviews,* 2005, vol. 1, no. 2, p. 105 – 113.

[8] ASHBURNER, J., HUTTON, C., FRACKOWIAK, R., JOHNSRUDE, I., PRICE, C., FRISTON, K. Identifying global anatomical differences: deformation-based morphometry. *Human Brain Mapping,* 1998, vol. 6, no. 5 – 6, p. 348 – 357.

[9] FRACKOWIAK, R. S. J., ASHBURNER, J. T., PENNY, W. D., ZEKI, S. *Human Brain Function.* Second edition. Academic Press, 2004.

[10] GASER, C., NENADIC, I., BUCHSBAUM, B. R., HAZLETT, E. A., BUCHSBAUM, M. S. Deformation-based morphometry and its relation to conventional volumetry of brain lateral ventricles in MRI. *NeuroImage,* 2001, vol. 13, no. 6, p. 1140 – 1145.

[11] LEPORE, N., BRUN, C., CHOU, Y. -Y., CHIANG, M. -C., DUTTON, R. A., HAYASHI, K. M., LUDERS, E., LOPEZ, O. L., AIZENSTEIN, H. J., TOGA, A. W., BECKER, J. T., THOMPSON, P. M. Generalized Tensor-Based Morphometry of HIV/AIDS using multivariate statistics on deformation tensors. *IEEE Transactions on Medical Imaging,* 2008, vol. 27, no. 1, p. 129 – 141.

[12] XU, L., LIU, J., ADALI, T., CALHOUN, V. D. Source based morphometry using structural MRI phase images to identify sources of gray matter and white matter relative differences in schizophrenia versus controls. In *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008.* Las Vegas (NV, USA), 2008, p. 533 – 536.

[13] CALHOUN, V. D., ADALI, T., PEARLSON, G. D., PEKAR, J. J. A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping,* 2001, vol. 14, no. 3, p. 140 – 151.

[14] PANTAZIS, D., LEAHY, R. M., NICHOLS, T. E., STYNE, M. Statistical surface-based morphometry using a non-parametric approach. In *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano*. Arlington (VA, USA), vol. 2, 2004, p. 1283 – 1286.

[15] GAONKAR, B., POHL, K., DAVATZIKOS, C. Pattern based morphometry. *Medical Image Computing and Computer-Assisted Intervention: MICCAI.,* 2011, vol. 14, no. 2, p. 459 – 466.

[16] VAPNIK, V. N. An overview of statistical learning theory. *IEEE Transactions on Neural Networks,* 1999, vol. 10, no. 5, p. 988 to 999.

[17] CASTELLANI, U., ROSSATO, E., MURINO, V., BELLANI, M., RAMBALDELLI, G., PERLINI, C., TOMELLERI, L., TANSELLA, M., BRAMBILLA, P. Classification of schizophrenia using feature-based morphometry. *Journal of Neural Transmission,* 2012, vol. 119, no. 3, p. 395 – 404.

[18] LIU, Y., TEVEROVSKIY, L., CARMICHAEL, O., KIKINIS, R., SHENTON, M., CARTER, C. S., STENGER, V. A., DAVIS, S., AIZENSTEIN, H., BECKER, J. T., LOPEZ, O. L., MELTZER, C. C. Discriminative MR image feature analysis for automatic schizophrenia and Alzheimer's disease classification. In *MICCAI 2004, Lecture Notes in Computer Science.* Saint-Malo (France), 2004, vol. 3216, p. 393 – 401.

[19] FAN, Y., SHEN, D., GUR, R. C., GUR, R. E., DAVATZIKOS, C. COMPARE: Classification of Morphological Patterns using Adaptive Regional Elements. *IEEE Transactions on Medical Imaging,* 2007, vol. 26, no. 1, p. 93 – 105.

[20] SHEN, H., WANG, L., LIU, Y., HU, D. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. *NeuroImage,* 2010, vol. 49, no. 4, p. 3110 – 3121.

[21] LIU, M., WANG, L., SHEN, H., LIU, Z., HU, D. A study of schizophrenia inheritance through pattern classification. In *Proceedings of the 2nd International Conference on Intelligent Control and Information Processing, ICICIP 2011*. Harbin (China), 2011, p. 152 – 156.

[22] INGALHALIKAR, M., KANTERAKIS, S., GUR, R., ROBERTS, T. P. L., VERMA, R. DTI based diagnostic prediction of a disease via pattern classification. *Medical Image Computing and Computer-Assisted Intervention,* 2010, vol. 13, no. 1, p. 558 – 565.

[23] NIEUWENHUIS, M., VAN HAREN, N. E. M., HULSHOFF POL, H. E., CAHN, W., KAHN, R. S., SCHNACK, H. G. Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *NeuroImage,* 2012, vol. 61, no. 3, p. 606 – 612.

[24] MOURAO-MIRANDA, J., REINDERS, A. A. T. S., ROCHA-REGO, V., LAPPIN, J., RONDINA, J., MORGAN, C., MORGAN, K. D., FEARON, P., JONES, P. B., DOODY, G. A., MURRAY, R. M., KAPUR, S., DAZZAN, P. Individualized prediction of illness course at the first psychotic episode: a support vector machine MRI study. *Psychological Medicine,* 2012, vol. 42, no. 5, p. 1037 to 1047.

[25] ZANETTI, M. V., SCHAUFELBERGER, M. S., DOSHI, J., OU, Y., FERREIRA, L. K., MENEZES, P. R., SCAZUFCA, M., DAVATZIKOS, C., BUSATTO, G. F. Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* 2013, vol. 43, p. 116 – 125.

[26] BISHOP, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York: Springer-Verlag, 2006.

[27] KAŠPÁREK, T., THOMAZ, C. E., SATO, J. R., SCHWARZ, D.,

JANOUŠOVÁ, E., MAREČEK, R., PŘIKRYL, R., VANÍČEK, J., FUJITA, A., ČEŠKOVÁ, E. Maximum-uncertainty linear discrimination analysis of first-episode schizophrenia subjects. *Psychiatry Research,* 2011, vol. 191, no. 3, p. 174 – 181.

[28] SHI, F., LIU, Y., JIANG, T., ZHOU, Y., ZHU, W., JIANG, J., LIU, H., LIU, Z. Regional homogeneity and anatomical parcellation for fMRI image classification: Application to schizophrenia and normal controls. *Medical Image Computing and Computer-Assisted Intervention: MICCAI*. Springer-Verlag Berlin Heidelberg, vol. 4792, 2007.

[29] FORD, J., SHEN, L., MAKEDON, F., FLASHMAN, L. A., SAYKIN, A. J. A combined structural-functional classification of schizophrenia using hippocampal volume plus fMRI activation. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. Houston (TX, USA), 2002, vol. 1, p. 48 – 49.

[30] ARDEKANI, B. A., TABESH, A., SEVY, S., ROBINSON, D. G., BILDER, R. M., SZESZKO, P. R. Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Human Brain Mapping,* 2011, vol. 32, no. 1, p. 1 – 9.

[31] YOON, J. H., NGUYEN, D. V., MCVAY, L. M., DERAMO, P., MINZENBERG, M. J., RAGLAND, J. D., NIENDHAM, T., SOLOMON, M., CARTER, C. S. Automated classification of fMRI during cognitive control identifies more severely disorganized subjects with schizophrenia. *Schizophrenia Research,* 2012, vol. 135, no. 1-3, p. 28 – 33.

[32] TAKAYANAGI, Y., TAKAHASHI, T., ORIKABE, L., MOZUE, Y., KAEASAKI, Y., NAKAMURA, K., SATO, Y., ITOKAWA, M., YMASUE, H., KASAI, K., KURACHI, M., OKAZAKI, Y., SUZUKI, M. Classification of first-episode schizophrenia patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. *PLoS ONE,* 2011, vol. 6, no. 6, e21047.

[33] JAFRI, M. J., CALHOUN, V. D. Functional classification of schizophrenia using feed forward neural networks. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. New York (NY, USA), 2006, p. 6631 to 6634.

[34] ARRIBAS, J. I., CALHOUN, V. D., ADALI, T. Automatic Bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from fMRI data. *IEEE Transactions on Biomedical Engineering,* 2010, vol. 57, no. 12, p. 2850 – 2860.

[35] SUN, D., VAN ERP, T. G. M., THOMPSON, P. M., BEARDEN, C. E., DALEY, M., KUSHAN, L., HARDT, M. E., NUECHTERLEIN, K. H., TOGA, A. W., CANNON, T. D. Elucidating a Magnetic Resonance Imaging-based neuroanatomic biomarker for psychosis: Classification analysis using probabilistic brain atlas and machine learning algorithms. *Biological Psychiatry,* 2009, vol. 66, no. 11, p. 1055 – 1060.

[36] WANG, P., VERMA, R., In: METAXAS, D., AXEL, L., FICHTINGER, G., SZÉKELY, G. (Eds.) On classifying disease-induced patterns in the brain using diffusion tensor images. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*. Berlin (Germany), 2008, p. 908 – 916.

[37] DEMIRCI, O., CLARK, V. P., CALHOUN, V. D. A projection pursuit algorithm to classify individuals using fMRI data: Application to schizophrenia. *NeuroImage,* 2008, vol. 39, no. 4, p. 1774 – 1782.

[38] DEMIRCI, O., CLARK, V. P., MAGNOTTA, V. A., ANDREASEN, N. C., LAURIELLO, J., KIEHL, K. A., PEARLSON, G. D., CALHOUN, V. D. 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,* 2008, vol. 2, no. 3, p. 207 to 226.

[39] DLUHOŠ, P. Multiresolution feature selection for recognition in magnetic resonance brain images (in Czech). *Diploma Thesis*, Masaryk University, Faculty of Science, 2013.

[40] ASHBURNER, J., FRISTON, K. J. Unified segmentation. *NeuroImage,* 2005, vol. 26, no. 3, p. 839 – 851.

[41] ASHBURNER, J. A fast diffeomorphic image registration algorithm. *NeuroImage*, 2007, vol. 38, no. 1, p. 95 – 113.

[42] STARCK, J. -L., MURTAGH, F., FADILI, J. M. *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*. Cambridge University Press, 2010.

[43] MISITI, M., MISITI, Y., OPPENHEIM, G., POGGI, J. -M. (Eds.). *Wavelets and Their Applications*. Wiley-ISTE, 2007.

[44] MALLAT, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 1989, vol. 11, no. 7, p. 674 – 693.

[45] KUMARI, S. Effect of symlet filter order on denoising of still images. *Advanced Computing: An International Journal,* 2012, vol. 3, no. 1, p. 137 – 143.

[46] KAILATH, T. The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology,* 1967, vol. 15, no. 1, p. 52 – 60.

[47] WANG, S., LI, D., SONG, X., WEI, Y., LI, H. A feature selection method based on improved Fisher's discriminant ratio for text sentiment classification. *Expert Systems with Applications,* 2011, vol. 38, no. 7, p. 8696 – 8702.

[48] DIETTERICH, T. G. Ensemble methods in machine learning. In *Proceedings of the First International Workshop on Multiple Classifier Systems*. London (UK), 2000, p. 1 – 15.

[49] VITA, A., DE PERI, L., DESTE, G., SACCHETTI, E. Progressive loss of cortical gray matter in schizophrenia: A meta-analysis and meta-regression of longitudinal MRI studies. *Translational Psychiatry,* 2012, vol. 2, e190.

[50] PURI, B. K. Brain tissue changes and antipsychotic medication. *Expert Review of Neurotherapeutics*, 2011, vol. 11, no. 7, p. 943 to 946.

[51] KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*. San Francisco (CA, USA), 1995, p. 1137 – 1143.

[52] ZHAO, Y., OGDEN, R. T., REISS, P. T. Wavelet-based LASSO in functional linear regression. *Journal of Computational and Graphical Statistics*, 2012, vol. 21, no. 3, p. 600 – 617.

About Authors ...

Petr DLUHOŠ graduated from the Masaryk University, Brno in 2012 with MSc. degree in Artificial Intelligence. Currently, he is studying the Ph.D. program in Neuroscience and the MSc. program in Mathematical Biology at the same university. He focuses his research on the methods of artificial intelligence and pattern recognition and their applications in neuroscience.

Daniel SCHWARZ graduated from the Brno University of Technology in 2000 with MSc. degree in Electronics and Communication Technology. Then he studied for the Ph.D. degree at the same university at the Department of Biomedical Engineering in Biomedical Electronics and Biocybernetics. He finished his studies with the thesis entitled

Automated Morphometry of MRI Brain Images with Use of Deformable Registration in 2005. Since then he has been working for the Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University. He focuses his research on medical image processing and pattern recognition.

Tomáš KAŠPÁREK graduated from the Charles University in Prague in 2000 with M.D. degree. Then he started postgraduate studies in the field of Psychiatry at the Masaryk University, Brno. At the same time, he started his clinical career as a psychiatrist at the Department of Psychiatry of the University Hospital Brno-Bohunice. In 2004 he successfully defended his thesis on volumetric analysis of hippocampus in patients with first-episode schizophrenia and obtained the Ph.D. degree; in 2007 he finished his clinical training. In 2010, he successfully defended his habilitation thesis "Imaging brain morphometry in schizophrenia" and obtained the academic post – of Assoc. Professor. Since then he works as chief doctor and associate professor at the Department of Psychiatry. His research interests are focused on the neurobiology of schizophrenia and its analysis using neuroimaging methods.