SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM

JBEMi JOURNAL OF BIOMEDICAL ENGINEERING AND MEDICAL IMAGING

ISSN: 2055-1266
VOLUME 3 ISSUE 6

Differentiating Early and Late Stage Parkinson's Disease Patients from Healthy Controls

[1]D. Mudali, [1]M.Biehl, [3]S.K. Melesc, [2] R.J. Renken, [4,7]D. Garc´ıa-Garc´ıa, [4,7]P. Clavero, [4]J. Arbizu, [4,5,9]J.A. Obeso, [4,5,6]M.C. Rodriguez-Oroz, [3]K.L. Leenders and [1,2]J.B.T.M. Roerdink

[1]Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, The Netherlands;

[2]University of Groningen, University Medical Center Groningen, Neuroimaging Center, The Netherlands;

[3]Department of Neurology, University Medical Center Groningen, University of Groningen, The Netherlands;

[4]Neurosciences Area, CIMA, Department of Neurology and Neurosurgery, Cl´ınica Universidad de Navarra, University of Navarra, Pamplona, Spain;

[5]Centro de Investigaci´on Biom´edica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain;

[6]Department of Neurology and Neuroscience, University Hospital Donostia, BioDonostia Research Institute, SanSebastian; Ikerbasque, Basque Foundation for Science, Bilbao, Spain;

[7]Departamento de Bioingenier´ıa e Ingenier´ıa Aeroespacial. Universidad Carlos III de Madrid. Instituto de Investigaci´on Sanitaria Gregorio Mara˜non. Madrid. Spain;

[8]Department of Neurology, Complejo Hospitalario de Navarra, Pamplona, Spain;

[9]CINAC, HM Puerta del Sur, Hospitales de Madrid, and Medical School, CEU-San Pablo University, Madrid, Spain;

debmudal@gmail.com; dagarciagarcia@yahoo.es; dgarcia@hggm.es; p.clavero@hotmail.com; jarbizu@unav.es; jobeso52@gmail.com; maria.rodriguezoroz@biodonostia.org; j.b.t.m.roerdink@rug.nl

ABSTRACT

Parkinson's disease (PD) is a neurodegenerative disease which is difficult to diagnose at an early stage. Brain imaging techniques like [18F]-fluorodeoxyglucose positron emission tomography (FDG-PET) may aid to identify disease-related changes in cerebral glucose metabolism. The scaled subprofile model with principal component analysis (SSM/PCA) is applied to FDG-PET data to extract features and corresponding patterns of glucose metabolism which can be used to distinguish PD subjects from healthy controls. From a previous study, the decision tree (DT) classifier's performance to separate the PD group from healthy controls was below chance level. This could be attributed to the small number of subjects in the dataset, combined with the early disease progression. In this study, we make use of an additional PD dataset, consisting of subject brain images obtained at a later disease stage. The features extracted by the SSM/PCA method are used for distinguishing PD subjects from healthy controls using three classification methods, that is, decision trees, Generalized Matrix Learning Vector Quantization (GMLVQ), and Support Vector Machine (SVM) with linear kernel. The classifiers are validated to determine their capability of classification given new subject data. We compare the classifiers'

DOI: 10.14738/jbemi.36.2280
Publication Date: 08th December 2016
URL: http://dx.doi.org/10.14738/jbemi. 36.2280

performances on the distinct early-stage and late-stage datasets, as well on the combined datasets. We also use the early and late-stage datasets interchangeably for training and testing the classifiers. We find that the DT classification performance on the late-stage dataset is considerably better than in the previous study, where we used early-stage data. For early-stage patients, the application of the GMLVQ and SVM classifiers gives a significant improvement as compared to the DT classifier.

Keywords: Parkinson's disease; SSM/PCA; decision tree classification.

1 Introduction

Parkinson's disease (PD) and other parkinsonian disorders such as progressive supranuclear palsy (PSP) and multiple system atrophy (MSA) often show overlap in symptoms at an early disease stage. An accurate diagnosis can only be achieved after long-term serial assessment by a movement disorder specialist [1,2]. This is problematic because early diagnosis is important for selecting appropriate treatments. We use typical patterns of glucose metabolism delineated by [18F]-Fluoro-deoxyglucose (FDG) PET with the purpose of differentiating among parkinsonian syndromes. Such patterns are extracted by applying the scaled subprofile model and principal component analysis (SSM/PCA, [3]) to FDG PET data of healthy controls and patients [4]. The expression of previously identified patterns can be computed from the scans of new individuals. These pattern expression values are useful markers for disease [5].

The decision tree method [6] was used in the previous study [7] to classify parkinsonian syndromes based on SSM/PCA features. However, it was quite a challenge to separate the PD subjects from the healthy controls. This could be because the number of subjects in the dataset was not sufficient enough to train a robust decision tree classifier.

In this study, in addition to the dataset of early stage PD and healthy controls used in [7], a larger dataset consisting of brain images of healthy controls and patients with PD obtained at a later disease stage is also used. It is desirable to generate a large dataset consisting of brain data obtained at all stages of disease progression to extract features which can be used to train a robust classifier. Therefore, we will investigate whether features that are more suitable to separate the PD and healthy groups can be extracted from the advanced disease stage dataset, showing evident disease patterns in the data; in other words, to extract patterns which are evidently associated with PD [8].

In our earlier study [7] the number of subjects was too small to separate the dataset in a training and test set to assess classifier accuracy. Therefore to estimate classification performance the Leave-One-Out Cross Validation (LOOCV) method was used. However, as is well known, the LOOCV performance results are only an indication of what can be achieved when training and test sets are defined by different input data. Since we now have independent PD data from different sources we can use one as training and the other as test set to determine classifier accuracy. For comparison with earlier results we also compute LOOCV performance for the case of single datasets.

The scaled subprofile model with principal component analysis (SSM/PCA) method [3, 9] is used to extract the discriminative features from the brain data. Based on these features, the C4.5 decision tree classification algorithm [6] is used for building classifiers to separate the PD patients from healthy controls. Decision trees have the advantage of being easily constructed and understood, hence they provide an insight into the most important features for classification [10].

In previous brain imaging studies several other classifiers have been used with promising results. An example is the Support Vector Machine (SVM) which has been used to detect various neurological and psychiatric diseases [11, 12, 13]. Another example is Generalized Matrix Learning Vector Quantization (GMLVQ), which has been used in many disciplines including image analysis and bioinformatics, see [14]. A strong point of the prototype-based GMLVQ classifier is that it is intuitive and easy to interpret. In addition, it provides insight into the relevance of individual features for the classification [15]. For this reason, in addition to the decision tree method, we applied the SVM and GMLVQ classifiers to the subject scores extracted from the FDG-PET brain image data, with the aim to study classification accuracy of the methods given larger and different datasets.

2 Method

2.1 Subjects

Subject brain images were acquired from two hospitals. First, data of forty nine patients diagnosed with PD according to the UK Parkinson's Disease Society Brain Bank criteria were obtained from the Movement Disorders Unit of the Clinica Universidad de Navarra (CUN), Spain. Clinical and metabolic data of these patients was previously published in [16]. In addition, 19 age-and gender-matched control subjects without a history of neurologic, psychiatric illness and no abnormalities on MRI were included. From the 49 PD subjects, we randomly selected 20 PD subjects for training the classifier (PD subjects of dataset D1 CUN, see table 2), and 29 for testing the classifier (PD subjects of dataset D2 CUN/UMCG see table 2). Age, gender, disease duration, Unified Parkinson's Disease Rating Scale (UPDRS) motor ratings and Hoehn & Yahr (H&Y) scores did not differ significantly between PD patients in the two cohorts. Ethical permission for the procedures was obtained from the Ethics Committee for Medical Research of the University of Navarra. Written consent was obtained at each institution from all subjects following detailed explanation of the testing procedures. All the 19 healthy controls were added to the training set to make a total of 39 subjects (dataset D1 CUN).

Second, 20 PD subjects and 18 healthy controls were obtained from the University Medical Center Groningen (UMCG), more details are found in [17]. The 18 healthy controls (from UMCG) were added to the test set of 29 PD (dataset D2 CUN/UMCG, see table 2) from CUN to make 47 subjects. These 18 HC subjects from UMCG were considered for the test set because the 19 HC from CUN were too few to divide into the training and test sets. Also, the 20 PD and the earlier mentioned 18 healthy controls both from [17] (dataset D3 UMCG, see table 2) were considered for training and testing the classifiers. This particular dataset D3 UMCG was obtained at an early disease stage.

The original datasets from the University Medical Center Groningen (UMCG) and the Clinica Universidad de Navarra (CUN) are shown in Table 1;

Table 1. The original datasets as provided from their respective sources.

Subjects	Source
49 PD and 19 HC	CUN
20 PD and 18 HC	UMCG [17]

The following Table 2 shows the arrangement of the derived datasets from the original datasets for experiments, i.e., for training and testing classifiers.

Table 2. The arrangement of the datasets as used for both training and testing of classifiers.

Dataset	Description
D1_CUN	20 PD & 19 HC both groups from CUN
D2_CUN/UMCG	29 PD from CUN 18 HC from UMCG
D3_UMCG	20 PD & 18 HC both groups from UMCG

2.2 Image acquisition and preprocessing

The CUN subjects were scanned with [18F]fluorodeoxyglucose Positron Emission Tomography (FDG-PET) under resting conditions. Patients were studied in the 'on' pharmacological condition (under the effect of anti-parkinsonian medication). Central nervous system depressant drugs were withdrawn, and subjects fasted overnight before FDG-PET scanning. FDGPET imaging was performed in 3D mode using a Siemens ECAT EXAT HR+ scanner (Siemens, Knoxville, TN). Image acquisition was performed in a resting state with the subject's eyes closed in a dimly lighted room with minimal auditory stimulation. Images were reconstructed by means of a filtered back-projection method using ECAT software (version 7.2; Siemens). Preprocessing of imaging data was performed by SPM8 software (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) implemented in Matlab 8.0 (Mathworks Inc, Sherborn, MA). All images were spatially normalized onto a PET template in Montreal Neurological Institute (MNI) brain space and then smoothed by a Gaussian filter of 10 mm FWHM. The UMCG FDG-PET brain data was scanned as described previously by [17]) and preprocessed in the same way as the CUN data.

2.3 Feature extraction, classification and classifier validation

The same steps as those of [7] were followed to extract features from the brain image data in the form of subject scores on principal components using the SSM/PCA method [3, 9]. These subject scores were the features provided to the decision tree inducer, GMLVQ and SVM to train and test the classifiers for the different cohorts. All the extracted features were considered for building the classifiers.

For SVM binary classification, we use the Matlab R2014a functions "fitcsvm" and "predict" for training and testing, respectively, with default parameters and a linear kernel, representing a large margin linear separation in the original feature space. Also, all features are centered at their mean in the dataset and scaled to have unit standard deviation. The "fitcsvm" returns an SVM classifier which can be used for classification of new data samples. In addition, the LIBSVM library [18] with the one-against-one method is used for the SVM multi-class classification. The one-against-one method has a shorter training time than the one-against-all, as reported in [19]. As for GMLVQ, we employ it in its simplest setting with one prototype $\boldsymbol{w_k}$ per class. A global quadratic distance measure of the form $d(\boldsymbol{w_k},\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{w_k})^T \Lambda (\boldsymbol{x} - \boldsymbol{w_k})$ is used to quantify the dissimilarity of an input vector $\boldsymbol{x}$ and the prototypes $\boldsymbol{w_k}$. The measure is parameterized in terms of the positive semi-definite relevance matrix Λ. Both prototypes and relevance matrix are optimized in the training process which is guided by a suitable cost function [15]. We employed the gmlvq-toolbox, which performs a batch gradient descent minimization with automated step size control, see [20] for details.

The classifiers' performance was determined using leave-one-out cross validation (LOOCV). In each LOOCV iteration, a subject was removed from the training set before the SSM/PCA process to obtain features for training the classifiers. The left-out-subject was then used for testing the trained classifier.

In anticipation of better classification performance, we used the dataset D1 CUN which was obtained at a later disease stage to train the classifier. Then we tested the classifier using the subject scores extracted from both dataset D2 CUN/UMCG (PD group obtained at a later disease stage) and D3 UMCG (the PD subjects obtained at an earlier disease stage and healthy controls), see Table 2.

The decision tree classifiers are built using the C4.5 decision tree algorithm designed by [6]. This algorithm takes subject scores as inputs and outputs corresponding decision trees as classifiers [7]. Additionally, we use the gmlvq-toolbox by [20] to train and test GMLVQ classifiers with default parameters. Further, for the SVM we use the linear kernel since the dataset is still small. The Matlab R2014a functions "fitcsvm" and "predict" are used to train and test the classifiers respectively, see [21].

3 Classification Results

The data was used to train and test three different types of classifiers i.e., decision trees (DT), GMLVQ and SVM. The LOOCV results and the performances of training the classifiers on one cohort and testing on another are included.

3.1 Classifier Leave-one-out cross validation (LOOCV) on dataset D1_CUN

In this section we present the results obtained after the LOOCV of the DT, GMLVQ and SVM classifiers on dataset D1 CUN (39 subjects).

Table 3. GMLVQ, SVM, and DT LOOCV performance: Perf. = total accuracy, Sens. = Sensitivity and Spec. = Specificity with respect to detecting the disease.

Classifiers	GMLVQ	SVM	DT
Sens. %	100	100	90
Spec. %	89.5	94.7	84.2
Perf. %	94.9	97.4	87.2

Although both GMLVQ and SVM outperform DT in the LOOCV of dataset D1 CUN (the training set), the DT classifier is competitive since the difference in the performances is relatively small as can be seen in Table 3. We observe that with the CUN dataset of PD subjects obtained at a later disease stage, the DT is capable of separating the groups to a satisfactory extent.

3.2 GMLVQ, SVM and DT performance with dataset D1 CUN as the training set and D2 CUN/UMCG as the test set

Here we used the later disease stage dataset D1_CUN for training and dataset D2_CUN/UMCG for testing, which contains advanced PD subjects from CUN and a HC group from UMCG.

Table 4. GMLVQ, SVM, and DT performance. D1_CUN as the training set and D2_CUN/UMCG as the test set: The table shows the confusion matrix for the classification of dataset D2_CUN/UMCG with the overall performance (perf.) in percentage.

	GMLVQ		SVM		DT	
	HC	**PD**	**HC**	**PD**	**HC**	**PD**
HC(18 subjects)	**14**	4	**14**	4	**12**	6
PD (29 subjects)	3	**26**	2	**27**	3	**26**
Class accuracy (%)	77.8	89.7	77.8	93.1	66.7	89.7
Overall perf. (%)	85.1		87.2		80.9	

As can be seen in table 4, for DT only 3 out of 29 PD subjects from dataset D2 CUN/UMCG are misclassified as healthy controls with an overall performance of 80.9%. With respect toGMLVQ and SVM, the only difference is in the PD group, where SVM correctly classifies just one more PD subject than GMLVQ. However, both GMLVQ and SVM perform better than DT due to a higher accuracy on the HC group.

3.3 Classifier performance with dataset D1 CUN as the training set and D3 UMCG as the test set

In the setting discussed in this subsection, the training and test sets are from two different sites, that is, dataset D1_CUN is used for training and D3_UMCG for testing. The classifier results are shown in Table 5;

Table 5. GMLVQ, SVM, and DT performance. D1_CUN as the training set and D3_UMCG as the test set: The table shows the confusion matrix for the classification of dataset D3_UMCG with the overall performance in percentage.

	GMLVQ		SVM		DT	
	HC	**PD**	**HC**	**PD**	**HC**	**PD**
HC(18 subjects)	**14**	4	**14**	4	**12**	6
PD (20 subjects)	6	**14**	6	**14**	7	**13**
Class accuracy (%)	77.8	70	77.8	70	66.7	65
Overall perf. (%)	73.7		73.7		65.8	

The DT performance as seen in Table 5 is lower than that of 80.9% in Table 4 when testing with dataset D2_CUN/UMCG. However, it is higher than the PD group performance of 63.2% in [7]. Again this means that the decision tree classifier's ability to separate the two groups has improved. On the other hand, both GMLVQ and SVM register the same performance of 73.7% which is better than that of DT.

3.4 Classifier performance with dataset D3_UMCG as the training set and D1_CUN as the test set

The setting in this subsection is the reverse of that in subsection 3.3. At first sight, it may seem surprising to use the early-stage data set for training. Our motivation for this experiment is to see whether the early-stage data perhaps already do contain some features that help to differentiate PD subjects from healthy controls.

As can be seen in Table 6, using D3_UMCG to train the classifiers and testing with D1_CUN yielded the same performance of 92.3% for both GMLVQ and SVM which is better than 66.7% for DT. It is interesting that the performance is better than in the setting of section 3.3 (training with the CUN dataset and testing with respect to the UMCG dataset).

Table 6. GMLVQ, SVM, and DT performance. Dataset D3 UMCG as the training set and D1 CUN as the test set: The confusion matrix and the overall performance in percentage.

	GMLVQ		SVM		DT	
	HC	PD	HC	PD	HC	PD
HC(19 subjects)	**18**	1	**18**	1	**16**	3
PD (20 subjects)	2	**18**	2	**18**	10	**10**
Class accuracy (%)	94.7	90	94.7	90	84.2	50
Overall perf. (%)	92.3		92.3		66.7	

3.5 LOOCV of the combined datasets D1_CUN and D3_UMCG

Datasets D1 CUN and D3 UMCG were combined to make a dataset of 77 subjects arranged into two classes, i.e., 37 HC and 40 PD subjects. The GMLVQ classifier performance on the combined dataset was validated using the leave-one-out method, so as to determine the capability of the classifier to distinguish between the PD and HC subjects. Table 7 shows the confusion matrix for the two-class problem.

Table 6. GMLVQ, SVM, and DT LOOCV performance of the combined datasets D1 CUN and D3 UMCG in two classes: The confusion matrix and the overall performance.

	GMLVQ		SVM		DT	
	HC	PD	HC	PD	HC	PD
HC(37 subjects)	**35**	2	**32**	5	**24**	13
PD (40 subjects)	4	**36**	4	**36**	10	**30**
Class accuracy (%)	94.6	90	86.5	90	64.9	75
Overall perf. (%)	92.2		88.3		70.1	

The GMLVQ classifier can separate the two groups with a 92.2% accuracy as seen in Table 7, with sensitivity of 90% and specificity of 94.6%. SVM is fairly competitive, with a clearly lower performance of DT.

Having obtained good GMLVQ accuracy in Table 7, we next applied the GMLVQ classifier where we arranged the data from the D1 CUN and D3 UMCG datasets into four distinct classes, i.e., 18 HC from UMCG, 20 PD from UMCG, 19 HC from CUN, and 20 PD from CUN. This was done in anticipation of the GMLVQ classification accuracy in separating the CUN subjects from the UMCG subjects. The results for the four-class problem are shown in Table 8.

In Table 8, the GMLVQ classifier is able to separate all the CUN PD subjects from the rest of the subjects. However, 8 out of 20 UMCG PD subjects are misclassified as UMCG HC (5 subjects) and CUN PD (3 subjects).

Table 8. GMLVQ LOOCV performance on the combined datasets D1 CUN and D3 UMCG in four classes: The table shows the number of test subject images correctly classified for each class (in bold) with the overall performance in percentage.

	CUN HC	UMCG HC	CUN PD	UMCG PD
CUN HC (19 subjects)	**17**	1	1	0
UMCG HC (18 subjects)	1	**17**	0	0
CUN PD (20 subjects)	0	0	**20**	0
UMCG PD (20 subjects)	0	5	3	**12**
Class accuracy (%)	89.5	94.4	100	60
Overall performance. (%)	85.7			

4 Discussion and Conclusion

This study has focused on the differentiation between Parkinson's disease and healthy control brain patterns. In the previous study by [7], the decision tree (DT) classifier displayed relatively poor classification performance as assessed by leave-one-out cross validation (LOOCV). This poor performance was attributed to the small number of subjects in the dataset used and/or the brain data being obtained at an early disease stage.

The present study shows that one can obtain high LOOCV performances for patients at a more advanced disease stage using different classifiers; see Table 3 for the D1 CUN data. Although GMLVQ and SVM reach the highest performance, the decision tree classifier also performs very well. It reaches a performance around 87%, which is a significant improvement with respect to the results in [7], which were obtained for the D3 UMCG data. The difference between these data sets is not the number of subjects, but the fact that the D1 CUN data set corresponds to a later disease stage, with more metabolic changes than the early disease stage dataset. Hence, the disease pattern is more pronounced and the extracted features apparently are more informative with respect to separating the late-stage PD subjects from healthy controls.

The availability of a data set from CUN with a larger number of subjects, as well as data sets from different sites (i.e., CUN and UMCG), allowed us to perform a number of additional tests. When D1 CUN was used as the training set and D2 CUN/UMCG as the test set (in both data sets the PD subjects are from CUN), the performances of GMLVQ and SVM were still very good (85% and 87%, resp), while with 81% the DT was still competitive; see Table 4.

When D1 CUN was used as the training set and D3 UMCG as the test set (now the PD subjects are from CUN and UMCG, respectively), the performances are significantly lower for all classifiers; see Table 5. Comparing the results with Table 4 we see that the main reason is the higher percentage of PD subjects in the test set that are misclassified as healthy controls. As before, the explanation is that in this experiment the PD subjects in the test set are early stage patients from UMCG, which are hard to distinguish from healthy controls.

Somewhat surprisingly, training with the early-stage UMCG data and testing with respect to the late-stage CUN data yields much better performance than vice versa for the GMLVQ and SVM classifiers, as can be observed by comparing Tables 5 and 6. Training on early stage data with the GMLVQ and SVM classifiers seems to infer the subtle differences between early-stage PD and HC subjects, which then can be successfully used for the distinguishing late-stage PD from HC. Although in late stage data the differences between PD and HC will be more pronounced, training on such data is apparently less effective for classification when the test set contains early-stage patients which are quite similar to healthy controls. For the DT classifier, no significant improvement is seen when comparing Tables 5 and 6. The decision tree needs to take separate decisions on many features, while GMLVQ and SVM can handle linear combinations of feature values. For the DT this leads to the problem of overfitting and limited generalizability, especially when the number of subjects is relatively small. With pruning (feature selection) the overfitting problem can be reduced, but at the cost of lower performance; see [7] for a more extensive analysis.

The fact that early-stage PD subjects from the UMCG dataset are closer to healthy controls than to the late-stage PD samples in the CUN dataset can also clearly be inferred from Table 8. In this 4-class problem, the CUN PD subjects can be perfectly identified (no misclassification errors). Most errors occur for UMCG PD subjects that are misclassified as UMCG HC. Also, three of the PD UMCG subjects are misclassified as CUN PD subjects, suggesting that these three subjects are closer to late-stage than early-stage patients. In fact, two of the three subjects misclassified as CUN had a disease duration of 8 and 9 years, respectively, which is closer to the average 10 years of the CUN patients as compared to the UMCG patients disease duration, which is in the range of 1-4 years. So it makes sense that the two patients were classified as CUN (late disease stage). The disease duration of the third PD subject misclassified as CUN is not clear. Apparently the disease duration was 3 years but at the time of diagnosis the patient could not tell exactly when the symptoms started. Therefore, there is a possibility that the disease duration is longer than 3 years according to the results in Table 8 which show that this subject belongs to the late disease stage class.

On the combined datasets D1 CUN and D3 UMCG (2-class problem, all PD combined, all HC combined), the classifiers are also able to differentiate between Parkinson's disease and healthy controls with good performances as seen in Table 7, especially for the GMLVQ and SVM classifiers.

In conclusion, this study has shown that by applying state-of-the-art classifiers to FDG-PET brain data, Parkinson's disease subjects can be separated from the healthy controls with high accuracy. We have shown that the application of the GMLVQ and SVM classifiers can give a significant improvement as compared to the DT classifier, especially for classifying early-stage patients.

With respect to understanding the behavior of the classification methods, the GMLVQ and DT methods have proven to be more intuitive than SVM. Moreover, they can handle computationally very large feature sets. When both high accuracy and intuitive understanding of the classifier is desired, the GMLVQ method can be recommended.

We expect that the classifier performance will show further improvement, even for early stage brain data, when the number of subjects in the training dataset further increases. This is the ultimate goal of the GLIMPS project [22], which aims at establishing a large database of FDG-PET scans from the Netherlands and abroad.

ACKNOWLEDGEMENT

D. Mudali was supported by the Netherlands Fellowship Program of NUFFIC, under grant no. CF6695/2010.

REFERENCES

[1]. Hughes, A. J., Daniel, S. E., Ben-Shlomo, Y., Lees, A. J., 2002. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. Brain 125 (4), 861–870.

[2]. Osaki, Y., Ben-Shlomo, Y., Lees, A. J., Daniel, S. E., Colosimo, C., Wenning, G., Quinn, N., 2004. Accuracy of clinical diagnosis of progressive supranuclear palsy. Movement disorders 19 (2), 181–189.

[3]. Moeller, J. R., Strother, S. C., Sidtis, J. J., Rottenberg, D. A., 1987. Scaled subprofile model: a statistical approach to the analysis of functional patterns in positron emission tomographic data. J Cereb Blood Flow Metab 7 (5), 649–58.

[4]. Eidelberg, D., 2009. Metabolic brain networks in neurodegenerative disorders: a functional imaging approach. Trends in Neurosciences 32 (10), 548–557.

[5]. Niethammer, M., Eidelberg, D., 2012. Metabolic brain networks in translational neurology: concepts and applications. Annals of neurology 72 (5), 635–647.

[6]. Quinlan, J. R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo, USA.

[7]. Mudali, D., Teune, L. K., Renken, R. J., Leenders, K. L., Roerdink, J. B. T. M., 2015. Classification of Parkinsonian syndromes from FDG-PET brain data using decision trees with SSM/PCA features. Computational and Mathematical Methods in Medicine Article ID 136921, 1–10.

[8]. Eckert, T., Tang, C., Eidelberg, D., 2007. Assessment of the progression of Parkinson's disease: a metabolic network approach. The Lancet Neurology 6 (10), 926–932.

[9]. Moeller, J. R., Strother, S. C., 1991. A regional covariance approach to the analysis of functional patterns in positron emission tomographic data. J Cereb Blood Flow Metab 11 (2), A121–135.

[10]. Al Snousy, M. B., El-Deeb, H. M., Badran, K., Al Khlil, I. A., 2011. Suite of decision tree-based classification algorithms on cancer gene expression data. Egyptian Informatics Journal 12 (2), 73–82.

[11]. Magnin, B., Mesrob, L., Kinkingn´ehun, S., P´el´egrini-Issac, M., Colliot, O., Sarazin, M., Dubois, B., Leh´ericy, S., Benali, H., 2009. Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology 51 (2), 73–83.

[12]. Haller, S., Badoud, S., Nguyen, D., Garibotto, V., Lovblad, K., Burkhard, P., 2012. Individual detection of patients with Parkinson disease using support vector machine analysis of diffusion tensor imaging data: initial results. American Journal of Neuroradiology 33 (11), 2123–2128.

[13]. Orr`u, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., Mechelli, A., 2012. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. Neuroscience & Biobehavioral Reviews 36 (4), 1140–1152.

[14]. Yeo, L., Adlard, N., Biehl, M., Juarez, M., Smallie, T., Snow, M., Buckley, C., Raza, K., Filer, A., Scheel-Toellner, D., 2015. Expression of chemokines CXCL4 and CXCL7 by synovial macrophages defines an early stage of rheumatoid arthritis. Annals of the rheumatic diseases, annrheumdis–2014.

[15]. Schneider, P., Biehl, M., Hammer, B., 2009. Adaptive relevance matrices in learning vector quantization. Neural Computation 21 (12), 3532–3561.

[16]. Garc´ıa-Garc´ıa, D., Clavero, P., Salas, C. G., Lamet, I., Arbizu, J., Gonzalez-Redondo, R., Obeso, J. A., Rodriguez-Oroz, M. C., 2012. Posterior parietooccipital hypometabolism may differentiate mild cognitive impairment from dementia in Parkinson's disease. European Journal of Nuclear Medicine and Molecular Imaging 39 (11), 1767–1777.

[17]. Teune, L. K., Bartels, A. L., de Jong, B. M.,Willemsen, A. T., Eshuis, S. A., de Vries, J. J., van Oostrom, J. C., LeendersK. L., 2010. Typical cerebral metabolic patterns in neurodegenerative brain diseases. Movement Disorders 25 (14), 2395–2404.

[18]. Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1–27:27, software available at http://www.csie.ntu.edu.tw/˜cjlin/libsvm.

[19]. Hsu, C.-W., Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. Neural Networks, IEEE Transactions on 13 (2), 415–425.

[20]. Biehl, M., 2015. A no-nonsense Matlab (TM) toolbox for GMLVQ. Software available at http://www.cs.rug.nl/biehl/gmlvq.html.

[21]. Mudali, D., Biehl, M., Leenders, K. L., Roerdink, J. B. T. M., 2016. LVQ and SVM classification of FDG-PET brain data. In: et al., E. M. (Ed.), Advances in Self-Organizing Maps and Learning Vector Quantization. Proc. WSOM 2016, 11thWorkshop on Self-Organizing Maps. No. 428 in Advances in Intelligent Systems and Computing. Springer International Publishing Switzerland.

[22]. Teune, L. K., Mudali, D., Renken, R. J., Jong, B. M. D., Segbers, M., Roerdink, J. B. T. M., Dierckx, R. A., Leenders, K. L., 2012. Glucose IMaging in ParkinsonismS. In: 16th International Congress of Parkinson's Disease and Movement Disorders, Dublin, Ireland June 17-21. Abstract # 783.