Machine Learning Based Analysis of FDG-PET Image Data for the Diagnosis of Neurodegenerative Diseases

van Veen, Rick; Talavera Martinez, Lidia; Kogan, Rosalie Vered; Meles, Sanne; Mudali, Deborah; Roerdink, J. B. T. M.; Massa, Federico; Grazzini, Matteo; Obeso, J.A.; Rodriguez-Oroz, M.C.

Published in:
Applications of Intelligent Systems

DOI:
10.3233/978-1-61499-929-4-280

Machine Learning Based Analysis of FDG-PET Image Data for the Diagnosis of Neurodegenerative Diseases

R. VAN VEEN [a,f,1], L. TALAVERA MARTINEZ [1 b], R.V. KOGAN [d], S.K. MELES [c], D. MUDALI [a], J.B.T.M. ROERDINK [a,e], F. MASSA [g], M. GRAZZINI [g], J.A. OBESO [h], M.C. RODRÍGUEZ-OROZ [h,i], K.L. LEENDERS [d], R.J. RENKEN [e], J.J.G. DE VRIES [f] and M. BIEHL [a,2]

[a] *Bernoulli Institute for Math., CS and AI, Univ. of Groningen, NL*
[b] *Dept. of Maths. and Computer Science, Univ. of Balearic Islands, ES*
[c] *Dept. of Neurology, Univ. Medical Center Groningen, NL*
[d] *Dept. of Nuclear Medicine & Medical Imaging, Univ. Medical Center Groningen, NL*
[e] *Neuroimaging Center, Univ. of Groningen, NL*
[f] *Philips Research - Healthcare, NL*
[g] *Clinical Neurology, Dept. of Neuroscience, Univ. of Genoa and IRCCS AOU San Martino-IST, IT*
[h] *Neurosciences Area, CIMA, Neurology and Neurosurgery, Clínica universidad de Navarra, ES*
[i] *Neurology and Neuroscience, Univ. Hospital Donostia, BioDonostia Research Institute; Ikerbasque, Basque Foundation for Science, ES*

Abstract Alzheimer's disease (AD) and Parkinson's disease (PD) are two common, progressive neurodegenerative brain disorders. Their diagnosis is very challenging at an early disease stage, if based on clinical symptoms only. Brain imaging techniques such as [^{18}F]-fluoro-deoxyglucose positron emission tomography (FDG-PET) can provide important additional information with respect to changes in the cerebral glucose metabolism. In this study, we use machine learning techniques to perform an automated classification of FDG-PET data. The approach is based on the extraction of features by applying the scaled subprofile model with principal component analysis (SSM/PCA) in order to extract characteristics patterns of glucose metabolism. These features are then used for discriminating healthy controls, PD and AD patients by means of two machine learning frameworks: Generalized Matrix Learning Vector Quantization (GMLVQ) with local and global relevance matrices, and Support Vector Machines (SVMs) with a linear kernel. Datasets from different neuroimaging centers are considered. Results obtained for the individual centers, show that reliable classification is possible. We demonstrate, however, that cross-center classification can be problematic due to potential center-specific characteristics of the available FDG-PET data.

Keywords. Machine Learning, Learning Vector Quantization, Neuroimaging, Neurodegenerative Disorders, Alzheimer Disease, Parkinson Disease

[1] These authors contributed equally

[2] Corresponding Author: M. Biehl, Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, Univ. of Groningen, The Netherlands; E-mail: m.biehl@rug.nl

1. Introduction

Alzheimers disease (AD) and Parkinsons disease (PD) are the two most common neurodegenerative brain diseases [1,2]. AD, with an average duration of 2 to 10 years, is a type of dementia that affects over 7 million people in Europe [3], causing problems with memory, thinking and behaviour. On the other hand, PD is manifested clinically by bradykinesia (slowness of movement), muscular rigidity (stiffness), rest tremor and postural instability (balance difficulties). It affects over 4 million people and over 80% of patients with PD will eventually develop dementia (PDD) [4].

Differentiating between neurodegenerative brain disorders and their diagnosis at an early disease stage can be challenging on clinical grounds only, as they have overlapping clinical and pathological features at presentation. A thorough clinical evaluation combined with imaging techniques, able to provide patterns of neuronal dysfunction which are specific to a particular disease, might become an indispensable approach to assist the accurate diagnosis and choice of appropriate treatment [5].

Technological advances have allowed for the acquisition of high-quality, three-dimensional brain image data, which require the development of computer-aided algorithms for efficient analysis. Earlier studies of image-based diagnoses of neurodegenerative diseases include the use of other technical platforms, see for instance [6,7,8]. In this work we focus on the analysis of image data obtained by [^{18}F]-fluoro-deoxyglucose positron emission tomography (FDG-PET). We study the use of machine learning based classifiers for the detection and discrimination of PD, AD and healthy controls (HC). Features are generated using the scaled subprofile model/principal component analysis (SSM/PCA) method as in [9]. SSM/PCA is a multivariate method designed to find spatial covariance patterns in combined samples from patient and control scans [10,11]. Specifically, we look at the performance of Learning Vector Quantization (LVQ) models versus the more popular and widely used Support Vector Machine (SVM) [12]. In contrast to SVMs, the models produced by LVQ are interpretable and can deal with multi-class problems naturally. In particular we look at Generalized Matrix Learning Vector Quantization (GMLVQ) [13], which extends the LVQ model with an adaptive relevance matrix in the distance measure that captures the correlations between different features and their importance for the classification.

The FDG-PET data used in this work is obtained from three different neuroimaging centers located in Europe. We investigate the classifiers' ability to discriminate patient groups when trained from data that was obtained at one specific neuroimaging facility. It is the ultimate goal to build an universal classifier that performs well regardless of the center where the data was obtained. Therefore, we study, besides center specific classification performance, the quality of cross-center classification, i.e., training the systems on data from one facility and testing its performance with data from a different one. In addition, we study whether the classifiers can be trained to assign data samples to their center of origin, to see if any center specific patterns exist in the data. In spite of the fact that similar technical platforms and supposedly identical processing pipelines have been used.

The paper is organised as follows: Section 2 includes a brief discussion of the data acquisition process, feature generation, used classification models and the experiment set ups. The results and discussion are given in Section 3 with the conclusion in Section 4. The paper is completed in Section 5 with a description of future work and a discussion thereof.

2. Data and Methods

In this section we summarize the workflow we followed with respect to data acquisition, feature extraction, and the validation of the SVMs and GMLVQ models.

2.1. Data Acquisition

In this study we use the brain images acquired using [^{18}F]-fluoro-deoxyglucose positron emission tomography (FDG-PET), from three different hospitals. The first part is from the Movement Disorders Unit of the Clinica Universidad de Navarra (CUN), from which the scans of 19 HC subject and 68 patients diagnosed with PD were included, whose clinical and metabolic data were published in [14]. Secondly, 20 PD patients, 21 AD patients and 19 HCs were obtained from the University Medical Center Groningen (UMCG), for whom more details can be found in [15]. Thirdly, we have 58 patients with PD, 55 patients with AD and 44 HCs from the University of Genoa and IRCCS AOU San Martino-IST (UGOSM). All images were spatially normalized onto an [^{18}F]-FDG-PET template in Montreal Neurological Institute (MNI) brain space [16], using SPM12 software (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK), implemented in Matlab (version 2017b).

2.2. Feature Extraction

In this step the FDG-PET brain images from the three different centers were used. Features were extracted based on an implementation described in [9,17,18,19], which uses the SSM/PCA method as introduced by Spetsieris and Eidelberg [10]. SSM/PCA was developed as a way of dealing with the typically high dimensional datasets, that are found in imaging experiments, as well as the variability in brain activity between subjects and regions [11].

The initial step requires the images to be spatially normalized and flattened. The image data is then masked to reduce the number of low values and noise. The mask has been obtained by putting image data from a disjoint group of patients through the SSM/PCA process to obtain the disease related patterns. For instance, the FDG-PET data of HC subjects and disease specific patients specifically PD and AD, results in two disease related patterns, i.e., PDRP and ADRP. These patterns are considered to be true descriptors of the abnormal regional interactions due to the invariability and specificity of the disease-related patterns and can therefore be applied to other centers as well. For this work we have looked at the effect of applying the two different masks separately and their combination, which is done by concatenating the subject scores, and have found no significant difference in performance.

The features, i.e. the subject scores are obtained by applying a log-transform and centering the data, for details see Spetsieris and Eidelberg [10]. The difference between the implementation described in [9,17,18,19] and the method used here is in the computation of the subject scores. The subject scores are computed by projecting (inner product) the processed data onto the disease specific GIS (Group Invariant Subprofile) matrix which is, in contrast to the aforementioned work, obtained not from the training set, but from the same disjoint patient groups that were used to acquire the masks. The final projected subject scores will have a dimensionality equal to the number of subjects in the disjoint group, which contains 36 patients in our case.

2.3. Models and Validation

The subject scores, as described in the previous section, are used to train the classifiers. We consider two specific frameworks for classification: the Support Vector Machine (SVM) and the LVQ variant Generalised Matrix Learning Vector Quantization (GMLVQ).

The SVM is a popular binary classification method introduced in [12,20]. SVMs try to find the maximum margin separation between two hyperplanes, employing the so-called kernel trick for non separable data, see for example [21] for more detail. In this study we have used a SVM with a linear kernel. In particular, we employed the Matlab R2018a (Statistics Toolbox) functions "fitcsvm" and "predict" with default settings, for training and testing, respectively.

Generalized Matrix LVQ was introduced by Schneider et al. [13], and extends the distance measure with an adaptive relevance matrix which defines a generalized quadratic distance measure. In particular, its diagonal elements can be interpreted as a quantification of the relevance of the respective feature dimensions for the classification [13,22,23]. The LVQ models try to find the prototypical representation of a category or class of data. These prototypes together with the relevance matrix, allow GMLVQ to account for correlations between features, resulting in a robust performing and interpretable model [13,24]. In our work we used both the global relevance (GMLVQ) and the local variant (LGMLVQ). The local variant attaches a relevance matrix to every class in contrast to the single matrix of the global variant. Besides a potentially better performance, LGMLVQ also provides insight into which features are important per disease. For applying (L)GMLVQ, we resorted to the publicly available toolbox provided at [25]. The relevance matrix, or matrices in the case of LGMLVQ, were initialized with the identity matrix. The regularization parameter [25] was set to 0 and training was performed over 50 epochs of stochastic gradient descent ("sgd"). All other parameters were left at their default values, see [25].

Unless otherwise specified, for each of the considered classification problems, performance was estimated by averaging 10 randomized realizations of 10 fold-cross validation, see Section 3 for the results. For every fold a z-score transform was applied based on the subject scores selected for training. As performance measures, we considered sensitivity and specificity of the unbiased SVM and Nearest Prototype (L)GMLVQ classifier, as well as the areas under the ROC curve (AUC).

3. Results

In this section we present the results of the classifiers' ability to discriminate between the different patient groups, when trained on data from a specific center. In addition, we looked at the performance of cross-center classification, i.e., the models are trained on the data from one facility and tested on another. Lastly, we have included the results of the experiments where we train the models to distinguish between the center of origin, using only one group, e.g., only HC or only PD patients.

Lastly, we have included the results that show that, from both the disease specific and healthy control data, it is possible to identify nearly perfectly the center the data has been obtained.

	Classifier	Sens. (%)	Spec. (%)	AUC
HC vs. PD	SVM	74.23(18.98)	68.05(25.90)	0.80(0.16)
	GMLVQ	75.13(16.93)	77.50(22.45)	0.84(0.11)
	LGMLVQ	79.23(15.23)	68.15(22.55)	0.83(0.12)
HC vs. AD	SVM	95.40(8.86)	92.00(13.10)	0.99(0.03)
	GMLVQ	88.67(15.02)	92.90(13.37)	0.97(0.06)
	LGMLVQ	91.47(12.26)	91.45(14.40)	0.98(0.04)
PD vs. AD	SVM	82.10(16.18)	83.83(15.99)	0.92(0.07)
	GMLVQ	81.00(17.22)	81.67(15.82)	0.91(0.09)
	LGMLVQ	84.70(15.24)	86.63(14.77)	0.95(0.06)

Table 1. Data provided by UGOSM. Performance of three different models: SVM, GMLVQ, and LGMLVQ on three different classification problems. During pre-processing the PDRP was used to mask the data. Average values are shown together with standard deviations (in brackets).

3.1. Center specific classification

Our experiments confirm that the prime objective to discriminate between condition (HC, PD, AD) can be achieved to good accuracy when classifiers are trained from (and applied to) center specific data. As a representative example of the achievable performance on center specific data, Table 1 displays the results for the three different binary classification problems (HC/PD, HC/AD, AD/PD) with respect to the data from UGOSM. Results can vary slightly when applying the different masks before the feature extraction, but no significant difference was found. All reported results are from applying the PD specific mask. These results are comparable or slightly better than the ones found in [9,17,18,19] although a direct numerical comparison is not possible due to differences in feature extraction, model validation and data sets.

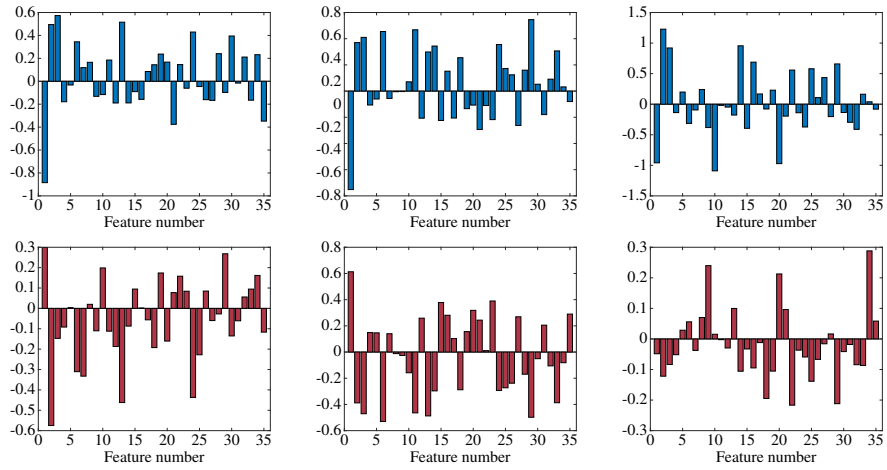
Table 1 shows that there is no significant difference among the performance of the (L)GMLVQ and SVM classifiers. All three classifiers produce good (0.80) to near perfect AUCs (0.99). The observed differences can be explained by potentially suboptimal hyper parameters. A benefit of (L)GMLVQ is that it can deal with multi-class classification naturally and it is interpretable (see [23]), in contrast to the SVM. With (L)GMLVQ we could approach the ternary problem (of distinguishing between HC, AD, and PD) in a more straightforward way, which the SVM could not deal with without the need of binary combination schemes. We did not include other than binary classification results to keep the comparison with the SVM. Due to the interpretable nature of (L)GMLVQ we could gain further insight into the diseases themselves, i.e., which features are more important for one disease and which are for another, and how does that compare to the relevance profile of the healthy subjects.

3.2. Cross-center classification

The results of three cross center models are shown in Table 2. The table includes the results for the HC versus PD classification problem as representative example. We see that the performances relative to the center specific performances have decreased considerably, suggesting different patterns in the data from the centers. These differences could be due to technical equipment or pre-processing pipelines which may differ in details from center to center. Consequently, we do not have the ideal conditions that we in-

	Classifier	Sens. (%)	Spec. (%)	AUC
Train: CUN	SVM	58.62	70.45	0.68
Test: UGOSM	GMLVQ	86.21	31.82	0.72
	LGMLVQ	98.28	4.55	0.57
Train: UGOSM	SVM	100.00	21.05	0.82
Test: UMCG	GMLVQ	70.00	63.16	0.74
	LGMLVQ	95.00	47.37	0.91
Train: UMCG	SVM	54.41	73.68	0.70
Test: CUN	GMLVQ	33.82	89.47	0.70
	LGMLVQ	47.06	84.21	0.72

Table 2. Performance of the three models for the HC versus PD classification problem. The models were trained on the data of one center and tested on a different center. The PDRP was used to mask the data during pre-processing.



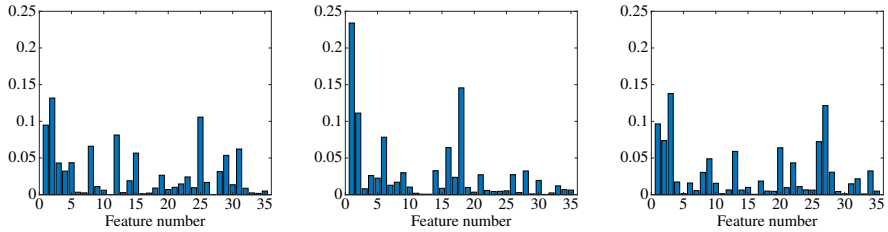
(a) UGMOS (b) UMCG (c) CUN

Figure 1. Comparison among the prototypes obtained when classify HC (top blue) and PD (bottom red) from each center, using GMLVQ.

tended in our objectives to build a universal classifier capable of classifying any subject from any center. In [9] a similar decrease in performance was found between the center specific classification and training and testing using data from different centers.

The prototypes, shown in Fig. 1 confirm this finding. The plots include the prototypes between HC and PD, and we see that the classifiers build different profiles for the same classes depending on the origin of the samples, i.e., CUN, UMCG or UGOSM. This is a sign that the data of the centers are different and could indirectly explain why we can differentiate so easily the center of origin, as each center has different features to characterize the same disease. This is reflected in both the disease (Table 3) as well as the healthy control data (Table 4).

Additionally, in Fig. 2, we can see how the relevance profile, provided by GMLVQ, varies for each center. Thus, for each center different discriminative features are most



(a) UGMOS (b) UMCG (c) CUN

Figure 2. Comparison among the diagonal of relevance matrices obtained when classifying HC (class 1) and PD (class 2) from each center, using GMLVQ.

	Classifier	**Sens. (%)**	**Spec. (%)**	**AUC**
CUN vs. UGOSM	SVM	88.60(12.98)	89.69(10.16)	0.96(0.04)
	GMLVQ	89.77(13.07)	92.81(8.85)	0.98(0.04)
	LGMLVQ	92.60(10.87)	94.12(8.97)	0.99(0.03)
UGOSM vs. UMCG	SVM	97.27(8.24)	87.50(26.94)	0.99(0.04)
	GMLVQ	91.83(11.87)	80.00(26.59)	0.94(0.12)
	LGMLVQ	97.80(5.73)	71.00(32.71)	0.92(0.14)
UMCG vs. CUN	SVM	82.50(26.94)	98.21(4.87)	0.99(0.03)
	GMLVQ	89.00(20.82)	99.83(1.67)	1.00(0.02)
	LGMLVQ	84.00(24.49)	100.00(0.00)	0.96(0.10)

Table 3. Data are from the PD patients from the two centers in the context of the problem. The PDRP was used to mask the data during pre-processing. Average values are shown together with standard deviations (in brackets).

	Classifier	**Sens. (%)**	**Spec. (%)**	**AUC**
CUN vs. UGOSM	SVM	99.75(2.50)	93.00(21.34)	1.00(0.01)
	GMLVQ	97.30(7.40)	91.00(21.77)	0.99(0.04)
	LGMLVQ	100.00(0.00)	89.50(24.92)	0.99(0.06)
UGOSM vs. UMCG	SVM	97.75(6.83)	87.50(25.99)	0.97(0.08)
	GMLVQ	92.75(12.50)	81.00(28.23)	0.95(0.12)
	LGMLVQ	98.85(5.07)	70.00(32.57)	0.95(0.10)
UMCG vs. CUN	SVM	95.50(14.38)	100.00(0.00)	100.00(0.00)
	GMLVQ	100.00(0.00)	100.00(0.00)	100.00(0.00)
	LGMLVQ	100.00(0.00)	99.50(5.00)	100.00(0.00)

Table 4. Data are from the healthy controls from the two centers in the context of the problem. The PDRP was used to mask the data during pre-processing. Average values are shown together with standard deviations (in brackets).

discriminative in the classification, which should be related to the fact that the centers end up being so distinguishable.

The results shown in Table 3 and 4 show the performance achieved classifying the different centers from both the PD data of HC data. Evidently, the difference between images from one center versus another are present, with near perfect AUC scores for almost

all models. The fact that this can be observed in the HC data supports the assumption that the center specific properties are caused by a difference in protocols or equipment.

4. Conclusions

We have demonstrated that it is feasible to achieve reliable diagnosis and discrimination of neurodegenerative diseases from FDG-PET scans based on SSM/PCA features. However, performance deteriorates in cross-center classification. Our results show that the datasets display center specific properties which may hinder the development of a universally applicable classifier. Therefore, a more unified approach should be aimed at when pre-processing and normalizing the data, achieving greater homogeneity with respect to measurements, calibration and processing protocols.

5. Future Work

In future work we will focus on further optimizing the used models as well as investigating the specific differences between the centers. In the best case, all data comes from sources using identical equipment and procedures. This is however not a practical requirement for a universally applicable classifier. One idea is to identify the specific differences in patterns between the centers which can potentially lead to another normalization step to cancel out the center specific artifacts. In forthcoming projects, we will focus on the identification of the specific patterns by, e.g., using the interpretable (L)GMLVQ models. Although these models are interpretable the SSM/PCA features they are trained on are not directly. Therefore, we will propose a visualization of prototypes and relevance profiles in voxel space. This can be done by exploiting the linearity of the transformations as applied during the feature extraction step, on the prototypes and relevance matrices. Example applications of this concept in different data domains can be found in [26,27].

Acknowledgement

This project has received funding from the European Unions Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 676157.

References

[1] European commission, "Public health." `http://ec.europa.eu/health/`. [Online; accessed 11-April-2018].

[2] S. Przedborski, M. Vila, and V. Jackson-Lewis, "Series introduction: Neurodegeneration: What is it and where are we?," *The Journal of Clinical Investigation*, vol. 111, no. 1, pp. 3–10, 2003.

[3] "Website of EU joint programme - neurodegenerative disease research." `http://www.neurodegenerationresearch.eu/`. [Online; accessed 11-April-2018].

[4] M. J. Firbank, A. J. Yarnall, R. A. Lawson, G. W. Duncan, T. K. Khoo, G. S. Petrides, J. T. O'Brien, R. A. Barker, R. J. Maxwell, D. J. Brooks, and D. J. Burn, "Cerebral glucose metabolism and cognition in newly diagnosed Parkinson's disease: ICICLE-PD study," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 88, no. 4, pp. 310–316, 2017.

[5] S. Risacher and A. Saykin, "Neuroimaging biomarkers of neurodegenerative diseases and dementia," *Seminars in Neurology*, vol. 33, no. 4, pp. 386–416, 2013.

[6] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehéricy, and H. Benali, "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI," *Neuroradiology*, vol. 51, no. 2, pp. 73–83, 2009.

[7] S. Haller, S. Badoud, D. Nguyen, V. Garibotto, K. O. Lovblad, and P. R. Burkhard, "Individual detection of patients with parkinson disease using support vector machine analysis of diffusion tensor imaging data: Initial results," *American Journal of Neuroradiology*, vol. 33, no. 11, pp. 2123–2128, 2012.

[8] 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, pp. 1140–1152, 2012.

[9] D. Mudali, M. Biehl, S. Meles, R. Renken, D. Garcia-Garcia, P. Clavero, J. Arbizu, J. A. Obeso, M. Rodriguez-Oroz, and K. Leenders, "Differentiating early and late stage Parkinson's disease patients from healthy controls," *Journal of Biomedical Engineering and Medical Imaging*, vol. 3, no. 6, pp. 33–43, 2016.

[10] P. G. Spetsieris and D. Eidelberg, "Scaled subprofile modeling of resting state imaging data in Parkinson's disease: Methodological issues," *NeuroImage*, vol. 54, no. 4, pp. 2899 – 2914, 2011.

[11] D. Eidelberg, "Metabolic brain networks in neurodegenerative disorders: A functional imaging approach," *Trends in Neurosciences*, vol. 32, no. 10, pp. 548–557, 2009.

[12] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[13] P. Schneider, M. Biehl, and B. Hammer, "Adaptive relevance matrices in learning vector quantization," *Neural Computation*, vol. 21, no. 12, pp. 3532–3561, 2009.

[14] D. Garcia-Garcia, P. Clavero, C. Gasca Salas, I. Lamet, J. Arbizu, R. Gonzalez-Redondo, J. A. Obeso, and M. C. Rodriguez-Oroz, "Posterior parietooccipital hypometabolism may differentiate mild cognitive impairment from dementia in Parkinson's disease," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 39, no. 11, 2012.

[15] L. K. Teune, A. L. Bartels, B. M. de Jong, A. T. M. Willemsen, S. A. Eshuis, J. J. de Vries, J. C. H. van Oostrom, and K. L. Leenders, "Typical cerebral metabolic patterns in neurodegenerative brain diseases," *Movement Disorders*, vol. 25, no. 14, pp. 2395–2404.

[16] P. A. Della Rosa, C. Cerami, F. Gallivanone, A. Prestia, A. Caroli, I. Castiglioni, M. C. Gilardi, G. Frisoni, K. Friston, J. Ashburner, and D. Perani, "A standardized [18F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia," *Neuroinformatics*, vol. 12, no. 4, pp. 575–593, 2014.

[17] D. Mudali, L. K. Teune, R. J. Renken, K. L. Leenders, and J. B. T. M. Roerdink, "Classification of parkinsonian syndromes from FDG-PET brain data using decision trees with SSM/PCA features.," *Computational and mathematical methods in medicine*, vol. 2015. Article ID 136921, 10 pages, 2015.

[18] M. Biehl, D. Mudali, K. L. Leenders, and J. B. T. M. Roerdink, "Classification of FDG-PET brain data by generalized matrix relevance LVQ," in *Brain-Inspired Computing*, (Cham), pp. 131–141, Springer International Publishing, 2016.

[19] D. Mudali, M. Biehl, K. L. Leenders, and J. B. T. M. Roerdink, "LVQ and SVM classification of FDG-PET brain data," in *Advances in Self-Organizing Maps and Learning Vector Quantization* (E. Merényi, M. J. Mendenhall, and P. O'Driscoll, eds.), vol. 428, pp. 205–215, Springer International Publishing, 2016.

[20] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory - COLT '92*, pp. 144–152, ACM Press, 1992.

[21] B. Schölkopf and A. J. Smola, *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT press, 2001.

[22] K. Bunte, P. Schneider, B. Hammer, F.-M. Schleif, T. Villmann, and M. Biehl, "Limited rank matrix learning, discriminative dimension reduction and visualization," *Neural Networks*, vol. 26, pp. 159–173, 2012.

[23] M. Biehl, B. Hammer, and T. Villmann, "Prototype-based models in machine learning," *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 7, no. 2, pp. 92–111, 2016.

[24] M. Biehl, P. Schneider, D. J. Smith, H. Stiekema, and a. E. Taylor, "Matrix relevance LVQ in steroid metabolomics based classification of adrenal tumors," in *European Symposium Artificial Neural Net-*

works (ESANN) (M. Verleysen, ed.), no. April, pp. 423–428, 2012.

[25] M. Biehl, P. Schneider, and K. Bunte, "Relevance and matrix adaptation in learning vector quantization (GRLVQ, GMLVQ and LiRaM LVQ)." `http://matlabserver.cs.rug.nl/gmlvqweb/web/`. [Online; accessed 11-April-2018].

[26] M. Biehl, "Biomedical applications of prototype based classifiers and relevance learning," in *Algorithms for Computational Biology* (D. Figueiredo, C. Martín-Vide, D. Pratas, and M. A. Vega-Rodríguez, eds.), vol. 10252 of *Lecture Notes in Computer Science*, pp. 3–23, Springer International Publishing, 2017.

[27] M. Straat, M. Kaden, M. Gay, T. Villmann, A. Lampe, U. Seiffert, M. Biehl, and F. Melchert, "Prototypes and matrix relevance learning in complex fourier space," in *2017 12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM)*, no. 1, pp. 1–6, IEEE, 2017.