8.

Improved image-based classification of Alzheimer's disease using multi-modality brain imaging data

Sofie M Adriaanse, Bart NM van Berckel, Jan C de Munck, Betty Tijms, Philip Scheltens, Frederik Barkhof, Alle Meije Wink

Submitted.

Abstract

Purpose Amyloid-PET is an accurate marker of Alzheimer's disease. However, this technique is invasive and not widely available. Here we investigated whether other neuroimaging techniques can yield similar diagnostic accuracy as PET imaging. This proof of concept study presents two approaches of analyzing multi-modality combinations of metabolic, functional and anatomical brain imaging data using kernel-based pattern recognition techniques.

Methods Input images were taken from a systematically collected database of previous studies. Data included [^{18}F]FDG-PET, rs-fMRI and structural MR data of 15 controls and 25 AD patients. [^{11}C]PIB was used as reference test for classification of AD using imaging data. From rs-fMRI data, both eigenvector centrality (EC) images and default mode network (DMN) regression maps were calculated. Structural MR data was used to create gray matter (GM) maps. Intermediate results were combined and analyzed with two types of existing pattern classification software; 1) using spatial concatenation of voxel data and, 2) a regional multi-kernel approach.

Results Both types of multimodal pattern classification software gave very similar results. Classification of only [^{11}C]PIB data as reference test, provided an accuracy of 92% for the regional approach and 94% for voxelwise concatenation. The regional approach yielded strongest results for combinations of imaging techniques; [^{18}F]FDG-GM (91%), EC-GM-[^{18}F]FDG (88%), DMN-GM-[^{18}F]FDG (88%), DMN-[^{18}F]FDG (86%), and EC-DMN-[^{18}F]FDG (86%). One combination that does not require PET data was included, using DMN, EC and GM, showed good accuracy (81%) in the regional analysis.

Conclusion This study highlights the possibilities of multi-modality image-based classifiers in AD diagnosis and shows that combining information from structural and functional MRI can yield similar accuracies as previously achieved with PET.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease, characterized by brain atrophy and cognitive decline. The benefit of early detection and intervention for treatment effect has prompted a search for new biomarkers that enable diagnosis before atrophy and severe cognitive symptoms appear. The amyloid cascade hypothesis [1], the leading theory about the cause of AD, defines the accumulation of amyloid-β in the brain as the central event in AD progression. Amyloid-β can be measured in vivo using carbon-11 labeled Pittsburgh compound B ([11C]PIB) positron emission tomography (PET) [2]. Elevated levels of [^{11}C]PIB binding in the posterior cingulate cortex/precuneus can be seen in prodromal AD, and are therefore sensitive markers in a very early stage of the disease. Amyloid-β deposition is followed by the formation of intracellular neurofibrillary tangles of hyperphosphorylated tau protein, and finally neuronal damage due to both Aβ42 toxicity and tau-related cell damage.

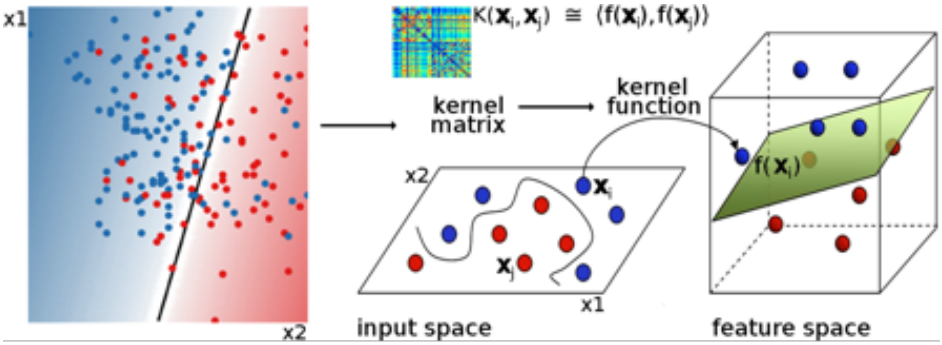
Traditional image-based classification of AD focuses on localized structural brain damage of the medial temporal lobes [3-6], at the end of this cascade, and after when many biomarkers of pathological processes have reached a plateau. Functional imaging methods, such as [^{18}F]FDG-PET, and functional magnetic resonance imaging (fMRI), can quantify physiological changes in the brain in early stages of AD [7-10]. [11C]PIB can detect the disease at very early stages, but PET it is expensive, invasive and not widely available. Using fMRI, connectivity differences in the default mode network have been found in healthy young carriers of the APOE ε4 allele, a genetic risk factor for developing AD in later life [11]. These changes occur decades before the age measurable Aβ42 deposits of have been reported, suggesting that functional brain changes can be measured very early in people at risk of developing AD. Including fMRI measures could therefore be beneficial in early diagnosis of AD. Another advantage of (f)MRI measures is that MR imaging is already part of standard clinical screening in AD.

Due to variability between individuals, single brain regions do not reach sufficient accuracy in classification of AD in common univariate analysis [12]. To improve classification, most current image-based classification methods use support vector machines [13] and similar implementations. An SVM is a two-class classifier that sets a separating hyperplane (in two dimensions this is a straight line, see Bennett & Campbell [14]) in feature space, such that individuals of either class (patient or control) are located as much as possible at one side of this plane. In practice, it turns out that the hyperplane is determined by a small number of feature vectors -the support vectors- closest to the hyperplane. The SVM algorithm can deal with misclassifications by using a penalty term, and the *kernel trick* provides a way to represent high-dimensional feature spaces (Figure 1). With SVM, subtle changes across the brain can be detected, not limited by localization in space [15]. Current image-based pattern classification methods focus on single-modality input data [16] and have reported good accuracies, but these may further be improved by integrating multiple inputs. In the case of multiple input modalities, outcomes of single modality classifications are currently analyzed separately [17,18]. Few studies have examined the integrated analysis of multi-modality classifiers in AD [5,19-21]. One region-

based multimodal approach [22] integrates imaging data from different modalities into the classifier. Multimodality data is pre-processed by sampling the same regions of interest in different modalities, and feed these regional data to the classifier. This region-wise multi kernel learning (MKL) approach has the advantage that different weights can be assigned to the different modalities [21]. Alternatively, data from multiple modalities can be combined at a more detailed voxel level, *e.g.* using spatial concatenation [23]. For both approaches, existing software packages were used, which has the benefit of reproducibility.

In the present study we investigated the performance of classifiers based on single and multimodal imaging data and compared the performance to amyloid-pet classification as a reference. In addition we explored whether other neuroimaging techniques can yield similar diagnostic accuracy as PET imaging. We specifically examined the effect of multi-modality input data with respect to (i) the accuracy of the classifier: using the region-wise MKL or voxelwise concatenation yielded the most accurate classifier; (ii) identification of the most accurate classification without relying on PET scans; (iii) and examine the spatial properties of the discrimination maps, i.e., the feature-space classifier weights projected back into image space.

Figure 1: The 'kernel trick' transforms a non-separable problem in input space into a separable problem in feature space. The kernel function, computed as an inner product between input vectors for every coefficient in the matrix K, represents an inner product in feature space (where the points are separable) without actually using feature vectors.



Materials and Methods

Participants

All AD patients visited the Alzheimer Center of the VU University Medical center and are part of the Amsterdam Dementia Cohort [24]. Healthy controls (HC) were included through advertisement. All patients and controls gave written informed

consent after a complete written and verbal description of the study. The local medical ethics committee approved the study. Ethics review criteria conformed to the Helsinki declaration. Resting-state fMRI and [^{11}C]PIB PET images of 43 subjects (18 HC and 25 AD) were acquired; 29 subjects (15 HC and 14 AD) underwent additional [^{18}F]FDG PET scans [10,25]. The voxelwise analysis required equal group sizes, therefore 18 of the 25 AD patients were selected (Table 1). Data from these AD patients were also used for the regional analysis. For all analyses with [^{18}F]FDG PET data, 14 of the 15 controls were selected. All subjects received a structural MR scan and screening which included medical history, physical and neurological exams and screening laboratory tests. Clinical diagnosis was established by a multidisciplinary team. AD patients met the NINCDS-ADRDA criteria [26]. Exclusion criteria for AD patients were (i) a history of major psychiatric or neurological illness other than AD, (ii) use of non-steroidal anti-inflammatory drugs, and (iii) clinically important abnormalities other than AD on the MRI scan. Additional exclusion criteria for HC were (iv) subjective complaints, (v) a history of major neurological or psychiatric illness, and (vi) any abnormalities on the MRI scan. Global cognitive functioning was assessed using the Mini-Mental State Exam (MMSE, [27]).

PET acquisition and analysis

PET imaging used an ECAT EXACT HR+ scanner (Siemens/CTI, see [28]). A 15.5 cm axial field of view allowed whole-brain coverage in one bed position with volumes containing 63 axial slices of 2.4 mm, matrix size 256×256 of 1.2×1.2 mm2. Visual assessments of parametric PET images by experts were passed on to the neurologists for aiding diagnosis. For the [^{11}C]PIB scan, a 10-minute transmission scan to correct for photon attenuation was followed by a dynamic emission scan of 90 minutes. For the [^{18}F]FDG scan, subjects rested for 10 minutes with eyes closed in adimly lit room with minimal background noise when [^{18}F]FDG was injected. Thirty-five minutes later, patients underwent a 10-minute transmission scan followed by a 15-minute (static) emission scan. Further details of [^{11}C]PIB and [^{18}F]FDG scans can be found elsewhere [25]. Standard preprocessing of PET data was performed. Structural T1 images were aligned to corresponding PET images using a mutual information algorithm. Data were analyzed using PVE-lab, a software program that makes use of a probability map based on 35 defined ROIs [29]. No partial volume correction (PVC) was performed. Parametric images of non-displaceable binding potential (BP_{ND}), a quantitative measure of specific binding, of [^{11}C]PIB were generated using a 2-step basis function implementation of the simplified reference tissue model with cerebellar gray matter as reference tissue (RPM2) [30,31]. If global cortical [^{11}C]PIB BP_{ND} was higher than 0.54, subjects were considered to be PIB-positive [10]. For [^{18}F]FDG, parametric images of standardized uptake value ratio (SUVr), using cerebellar gray matter as reference, were extracted from the interval between 45 and 60 minutes after injection. Cerebellar gray matter lacks Congo red and thioflavin-S positive plaques and was therefore chosen as reference tissue for analysis of both [^{11}C]PIB and [^{18}F] FDG data [32]. All PET images were co-registered to the MNI standard space [33] using affine transforms, and re-sampled at a resolution of 2mm isotropic, in the FLIRT program [34].

Structural MRI acquisition and analysis

MRI scans were acquired on a 1.5T Siemens Sonata scanner (see [8]) and included a high-resolution anatomical scan using an MPRAGE sequence (TR 2700 ms, TE 3.97 ms, TI 950 ms, FA 8o, 160 slices of 1.5 mm, matrix size 256×192 of 1×1 mm2 pixels). Structural MRI analysis was done with FMRIB [35]. To obtain individual GM maps, several steps were performed. First, non-brain tissue was removed from all T1-images. Second, the spatial scaling factor between the individual T1-images and standard space (MNI 152) was estimated. Then, tissue segmentation (GM, white matter (WM) and cerebral spinal fluid (CSF)), yielded the total GM tissue volume, which was multiplied by the scaling factor. Individual GM probability maps were registered to standard space using non-linear registration [36]. These images were averaged and flipped along the x-axis to create a symmetric, study specific GM-template. All native GM images were non-linearly registered to the template and subsequently corrected for local expansion or contraction. These GM probability maps were used as input for the classification calculations.

Functional MRI acquisition and analysis

Resting-state fMRI scans were acquired immediately after the structural T1 images on the same MR-scanner, using an EPI sequence (TR 2850 ms, TE 60 ms, FA 90o, 200 volumes, each 36 axial slices of 3.3 mm, matrix size 64×64 of 3.3×3.3 mm2 pixels). Data was processed with FSL. Standard preprocessing steps were performed, including removal of non-brain tissue, motion correction, high-pass temporal filtering (0.01 Hz), and spatial smoothing using a 5 mm full-width-at-half-maximum Gaussian kernel. Resting-state fMRI images were co-registered to the high-resolution anatomical scans using the same program. The anatomical scans were co-registered to the MNI standard space using affine and non-linear transforms, and resampled at 2mm isotropic, in the FNIRT program [37]; the combined registration was used to resample the resting-state fMRI data in MNI space at 4mm isotropic.

To produce single-subject maps of resting-state network consistency, a dual regression technique was used (see [8]); using standard RSNs masks [38]. This analysis regressed each single-subject fMRI data set against the standard RSNs masks, producing subject-specific representative time signals for the masks, after which a second regression of the data against these time signals produced subject-specific maps of regression slopes for all time signals. Each subject's regression map of the default-mode resting network (DMN), which is known as a group of regions undergoing significant change in AD [10,39-42], was resampled at a resolution of 2 mm isotropic.

In addition, a second fMRI-measure was examined; eigenvector centrality (EC). EC was calculated using in-house fECM software [43]. Since comparing the properties of different network topologies is a non-trivial problem, a mask of in-brain voxels across all subjects' pre-processed data sets (i.e. in the intersection of all single-subject masks) was applied before the EC maps were computed. EC mapping does

not require thresholding of connections and therefore ambiguities resulting from binarizing the connectivity matrix are avoided.

Regional classifiers

To remove the effect of input range on multi-kernel weight factors, all images were converted to z-scores, i.e., after subtracting its mean intensity; each image was divided by its intensity standard deviation. This regional approach can be used for multi-kernel learning (MKL) as previously described [21] which starts by computing regional averages inside regions of an anatomical atlas, in this study the atlas described by Gong [44], a subset of 78 cortical AAL regions [45] excluding deep brain structures and the cerebellum. This resulted in 78 regional averages per subject, per modality. For each modality in the classification, a linear kernel was computed from the Z-scores of every subject's regional averages. Then a coarse-grid search of relative modality-specific kernel contributions trained the SVM with different kernels, whose coefficients were weighted sums of single-modality kernel coefficients [46]. The LibSVM library was used for training and prediction [46], embedded together in a 10-fold cross-validation strategy to determine classification accuracy.

Voxel-based image classifiers

As mentioned above, all images were mapped to standard space and converted to z-scores. As each image already had the same dimensions Nx × Ny × Nz in voxels, multimodal combination by concatenation was trivial. SVM was used to construct image-based classifiers [47], as implemented in the PROBID software package (http://www.brainmap.co.uk/probid.htm). For each combination, the following analysis pipeline was used; Images were organized in a unified format (vectors of Nb in-brain voxels). A classification-specific kernel of size *S*x*S* (S = no of subjects) was computed from the input vectors. PROBID uses a linear kernel SVM in order to allow direct extraction of the weight vector as an image (i.e., the SVM discrimination map). Patients were defined as +1 and controls as -1. Accuracy is the mean of the true positives (TP), the proportions of AD classifications that were correct and true negatives (TN), the proportion of healthy classifications that were correct. A non-parametric test was performed to assess significance of the classifier performance for each combination of modalities [48]. For this purpose, 5000 randomized re-labelings were generated. Classifier assessment metrics were computed and p-values correspond to where the randomized statistic was greater than the true statistic (for either TP or TN).

Non-imaging analysis

Performance of the two approaches and the different combinations were ordered by accuracy. All subsequent statistical analyses were performed using SPSS (version 20; SPSS, Chicago, III). A Chi-squared test was used to examine frequency distributions of sex and PIB-status (PIB-/PIB+) within diagnostic groups. Differences between diagnostic groups on age and MMSE were explored with an independent samples

t-test. A p-value below 0.05 was considered significant.

Results

Subject characteristics

For subject demographics see Table 1. There was no difference in gender distribution (p=0.85) between AD patients and controls. On average, controls were older than patients (p=0.03). AD patients had lower MMSE scores (p=0.01), and a higher proportion of PIB-positive outcomes (p=0.01).

Table 1: Subject demographics of Alzheimer's disease (AD) patients and controls.

	Controls	AD patients	p-value
N	18	18	
Sex (%F)	22%	28%	0.70
Age	67 ± 6.2	62 ± 5.6	0.01
MMSE score	29 ± 0.6	23 ± 2.2	<0.01
PIB-/PIB+	14/4	2/16	<0.01

Data are presented as means ± standard deviation. MMSE score = Mini Mental State Examination (Folstein et al., 1975); PIB+ = positive rated [11C]PIB scan after visual inspection.

Classifier accuracies – regional approach

Seven possible single modalities and combinations showed good accuracies (≥80%) using the regional approach (Table 2). As a reference test, [^{11}C]PIB distinguished AD patients from controls with an accuracy of 92%. This was closely followed by combinations of modalities; [^{18}F]FDG together with GM performed almost as good as [^{11}C]PIB (accuracy 91%). Two combinations showed an accuracy of 88%: the combination of EC, GM and [^{18}F]FDG and the combination of DMN, GM and [^{18}F]FDG. The combination of DMN with [^{18}F]FDG had 86% accuracy. This did not change when combining EC, DMN and [^{18}F]FDG maps (86%). One combination that did not use PET imaging reached an accuracy of 81%: the three-modality combination of EC, DMN and GM maps. The regional analysis outperformed the voxelwise approach for combinations of images in classifying AD patients. The classification accuracies achieved with the voxelwise multimodality approach are listed in Table 3. With this approach, 3 single modalities and 1 combination yielded accuracies above 80%. As with the regional approach, highest accuracy was acquired with [^{11}C]PIB (94%). Single modalities performed better in the voxelwise approach when compared to the regional analysis; both GM and [^{18}F]FDG separately performed good (88% and 86%). Finally, the combination of GM with [^{18}F]FDG maps showed an accuracy of 86%.

Table 2; Regional classification accuracy

	[^{11}C] PIB	[^{18}F] FDG	GM	DMN	EC
Accuracy %					
92	v				
91		v	v		
88		v	v		v
88		v	v	v	
86		v		v	
86		v		v	v
81			v	v	v
79			v	v	
78			v		
78		v			
78		v			v
78			v		v
67				v	v
59					v
53				v	

GM=gray matter. DMN=Default mode network. EC=Eigenvector centrality.

Table 3; Voxelwise classification accuracy

		[^{11}C] PIB	[^{18}F] FDG	GM	DMN	EC
Accuracy %	Corr-p					
94	<0.01	v				
88	<0.01			v		
86	<0.01		v	v		
86	<0.01		v			
74	0.01			v		v
68	<0.01		v	v		v
65	0.08					v
64	0.02		v			v
56	0.02			v	v	v
54	0.11		v		v	v
50	0.48			v	v	
50	0.28		v		v	
50	0.87				v	
47	0.03		v	v	v	
47	0.21				v	v

Corr-p=corrected p-value. GM=gray matter. DMN=Default mode network. EC=Eigenvector centrality

Discrimination weight-maps

Examining the spatially distributed data in the discrimination weight maps helps to obtain a better understanding of changes in the AD brain. Weight maps of classification with accuracy > 80% the voxelwise analysis are shown in Figure 2 (weights are arbitrary units); for [^{11}C]PIB (2A), for GM (2B), [^{18}F]FDG weight maps (2C) and the combination [^{18}F]FDG with GM maps (2D). Figure 2 demonstrates that similar brain regions show strong weights for [^{11}C]PIB (red: positive weights higher in AD), GM and [^{18}F]FDG (blue: negative weights higher in controls) in bilateral parietal cortex and the posterior cingulate, all indicative of AD-pathology. For GM-density, the hippocampus, hallmark region of AD, was also identified, but not visible in the axial slice. Combining GM and [^{18}F]FDG improved accuracy compared to their single-modality performances. In the weight maps for this combination (Figure 2D) the overall distribution of weights was similar to the corresponding individual modalities, although somewhat lower weights for GM. Our main interest was to study to what extend PET data could be eliminated without compromising

classification accuracies, since MR acquisition is part of standard clinical screening and thus reduces the added inconvenience for patients. For this purpose, best accuracy was obtained with DMN, EC and GM in the regional analysis. The regional

analysis does not produce discrimination maps; therefore weight maps from the voxelwise analysis were examined. It is important to note that corresponding accuracy of the voxelwise analysis for this combination was much lower (56%) than acquired in the regional approach. Figure 3 shows the weight maps of the single modalities; for DMN (3A), EC-maps (3B), GM weight maps (3C), and the combination of the three modalities (3D). For the single DMN weight map, regions with high negative weights are seen in bilateral parietal and posterior cingulate cortex, similar to parietal areas of the [^{18}F]FDG weight map (Figure 2C). EC weight maps (Figure 3B) exists largely of positive weights, where strongest weights in medial frontal cortex can be observed, opposite to what was observed in the DMN maps. Dominant features of the DMN and EC weight maps remained the same when adding GM information (Figure 3D). Contrasts of DMN, EC and GM maps changed in a similar manner; for all three modalities lower weights were observed in the combination rather than altered directions or new areas. When examining importance of the three modalities in the regional approach, contribution of all three modalities was necessary for best classification. For corresponding weights of kernels for DMN, EC and GM see Figure 4.

Figure 2: Discrimination maps produced by the voxelwise analysis with good accuracies (>80%); A) [^{11}C]PIB maps, B) GM maps, C) [^{18}F]FDG maps, D) combination of [^{18}F]FDG and GM maps. All results are displayed in radiological orientation on MNI2mm space (z=28). Red-yellow colors represent positive weights (+1=AD patients), where blue-light blue colors represent negative weights (-1=Controls). AD = Alzheimer's disease. GM = Gray matter.

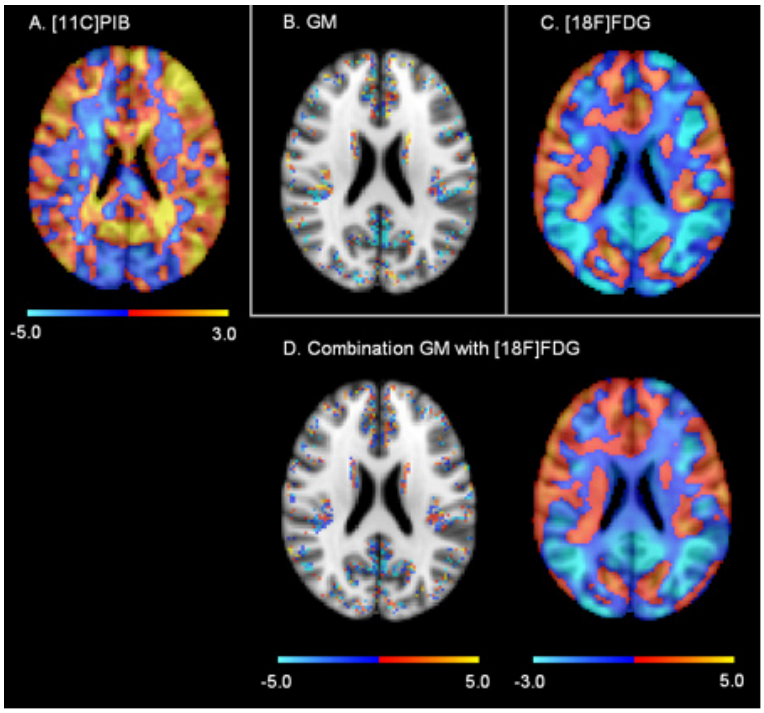


Figure 3; Discrimination maps produced by the voxelwise analysis; A) DMN map, B) EC map, C) GM map, D) combination of DMN, EC, and GM maps. All results are displayed in radiological orientation on MNI2mm space (x=-2). Red-yellow colors represent positive weights (+1=AD patients), where blue-light blue colors represent negative weights (-1=Controls). AD = Alzheimer's disease. DMN = Default mode network. EC = Eigenvector centrality. GM = Gray matter.

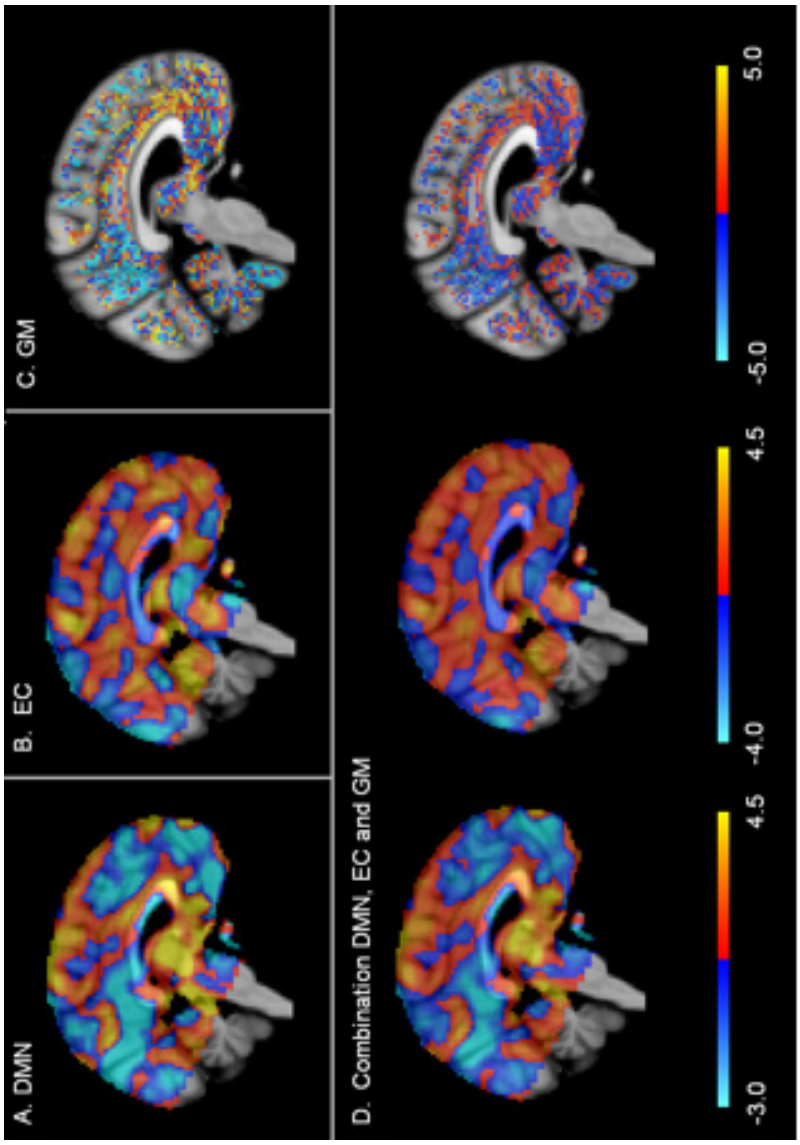
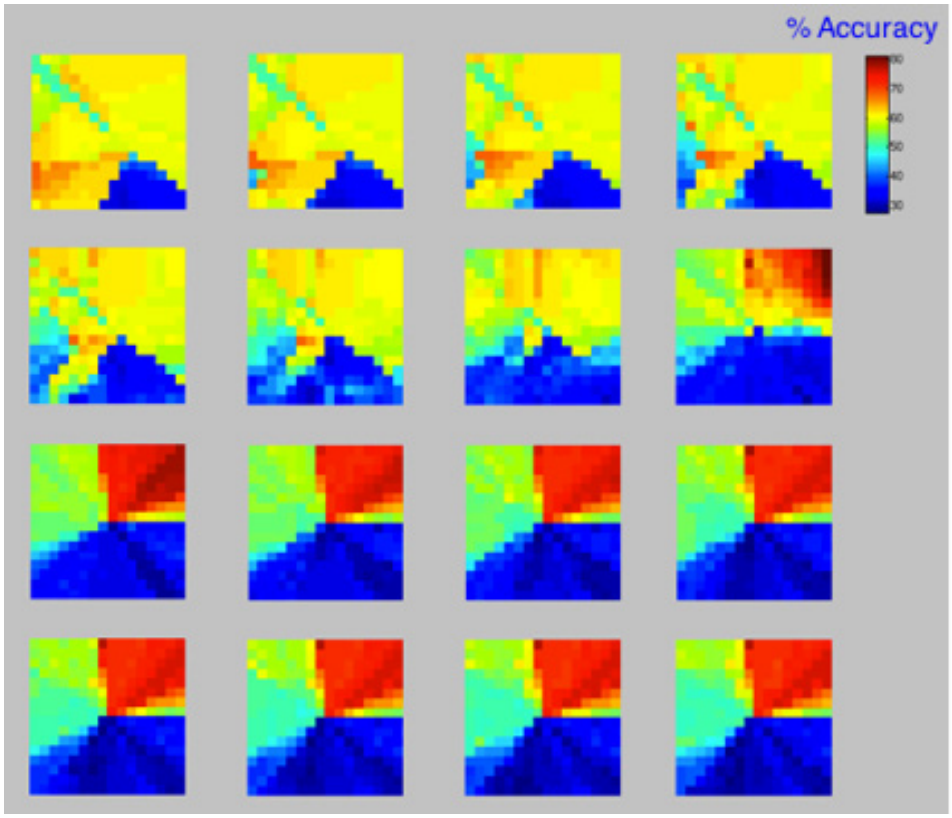


Figure 4; Vector weights and the corresponding accuracies of the regional analysis for the combination of DMN, EC and GM. With this combination an accuracy of 81% was acquired with corresponding weights of vectors: 16,14,8. DMN vector weights are displayed on the x-axis (range from -1 to 1). EC vector weights are displayed on the y-axis (range from -1 to 1). GM vector weights are represented by the different images (upper left to lower right corner represent weights from -1 to 1). The color scale indicates accuracy acquired with the corresponding vector weights (red=high accuracy, blue=low accuracy). DMN = Default mode network. EC = Eigenvector centrality. GM = Gray matter.



Discussion

Classification accuracy in the regional approach was better for multi-modality combinations of MR data than for the single modalities. Combining fMRI with GM is promising for detecting early changes in AD. For single modality processing, the voxelwise analysis appears more suitable. Discrimination maps provide important information on subtle, inter-regional, changes in AD.

Diagnostic accuracy – regional vs. voxelwise classification

The aim of this study was to examine the clinical applicability of multimodal inputs

for automatic classification. The integrated use of multiple inputs in AD has been examined in few other studies, mostly integrating 2 imaging modalities [12,23,49]. Diagnostic accuracy of both [^{18}F]FDG [50,51] and GM information based on structural MRI [52,53] has been described extensively. This study confirms that these measures successfully distinguish AD patients from controls. Similar accuracies have been reported for structural MR [5,12,23,49] and [^{18}F]FDG [21] as single modalities. For classification based on single modalities, the voxelwise approach performed best, similar to a recent study [23]. In line with previous research, combining modalities, often improved classification accuracy [12,23,49], and the regional approach was most suitable for this purpose [21]. Within single-modalities the voxelwise analyses were more accurate than regional approach. However, when combining multiple modalities, the regional approach yielded more accurate classifiers. An important advantage of the regional approach is that not just a single kernel is constructed, but different weights for the modalities (and different kernels) [21]. In contrast, the voxelwise approach includes only a single feature vector for the whole brain, by spatially concatenating the images from multiple modalities. Our results suggest that higher accuracies can be obtained when assigning significance to the different imaging modalities. Assigning differential significance to the different modalities in the voxelwise approach may further enhance accuracy.

For most modalities, adding either GM or [^{18}F]FDG improved classifier accuracy. This was no surprise since both modalities also perform well as single modalities and visual inspection of structural MRI is used in clinical diagnosis, so GM density can be expected to align well with diagnosis. The advantage of the voxelwise approach is that it allows for more detailed localization of features, which may be especially beneficial in the early stages of AD (i.e. MCI stage), when subtle changes occur. High dimensional information, as in the voxelwise analysis, was expected vital for correct classification. However, the voxelwise analysis outperformed regional analysis only for single modalities. This suggests that relative weighting in the kernel-space is more effective than combination of feature space. In addition, dimensionality reduction is important for classification with large feature vectors compared to the sample size [12,54]. Automatic selection of anatomical regions in a voxelwise analysis proved very useful in classification of MCI patients [12]. Enforcing similarity of selected features across subjects while specific selections are done on a subject-level basis outperformed other feature selection methods using task-fMRI data [54].

Discrimination maps

Examining the spatially distributed data in discrimination maps may help our understanding of subtle changes in AD, since complex regularities in data are identified. Since inter-regional correlations can drive classification, discrimination maps should be interpreted as a whole, and not focus on single regions [55]. Overall, adding information from other modalities reduced weights for that single modality. This implies that the voxelwise analysis incorporates information from all modalities for classification. This is similar to regional MKL, as was seen from weights applied

to the vectors. When we examine GM weight maps in more detail, strong (negative) weights were seen in hippocampus, (medial) temporal cortex, posterior cingulate cortex (PCC) and precuneus. These regions are reported vulnerable in AD [52,53,56]. Also, for [^{11}C]PIB [10,57] and [^{18}F]FDG [58-60] regions known to be affected in AD show strong weights (Figure 2).

At present, single rs-fMRI measures do not seem suitable for clinical applications. Separately, DMN and EC maps achieved accuracies of 53% and 59% in the voxelwise analysis, perhaps due to the high inter-individual variability of fMRI [54]. In contrast, Dai and colleagues [49] reached accuracy of 74% based on regional functional connectivity. Including GM seems indispensable for achieving good classification with fMRI measures only. The combination of DMN, EC and GM maps yielded good accuracy of 81% in our regional analysis. Somewhat higher discrimination power was found by Dai et al., [49] when combining fMRI measures with GM (89%), where information on amplitude of low-frequency fluctuations, intra-regional and inter-regional synchronization was included. Unfortunately, no spatial information can be obtained from the regional MKL approach. Therefore, discrimination maps resulting from the voxelwise analysis were examined. Interestingly, DMN weight-maps show spatial similarities to [^{18}F]FDG and [^{11}C]PIB maps: high (negative) weights were identified in the PCC, precuneus, bilateral parietal and anterior cingulate cortex. This is in line with the proposed vulnerability of the DMN for AD pathology [40]. Common univariate group fMRI analyses often identify only posterior regions of the DMN in AD [8,61]. We know from a previous study that EC is less sensitive for detecting classic AD regions such as the precuneus, but rather represents a shift in total network organization, where frontal regions become more prominent in AD [62]. When examining discrimination EC maps, frontal brain regions with strong positive weights seem more important for correct classification than lower centrality values in occipital cortex. Increases in centrality and functional connectivity are believed to reflect a compensatory mechanism in AD [63,64].

Limitations and future recommendations

A limitation of this study is that all subjects were only included who had received PET and MRI scanning less than 1 year apart, resulting in a small number of subjects included in the study. Future studies of more subjects across different stages of the disease are required for more robust classifier training and validation. Testing of our classifiers was not performed on a separate data set, but validation of the classifiers was done with cross-validation and permutation testing. Existing software packages were used for both regional and voxelwise analysis, which facilitates comparison of results and is user friendly. Default settings of the software packages include certain assumptions, for example regarding the precise definition of the hyperplane, which could have affected our results. For the clinical diagnosis of AD no PET data was used and thus some amyloid-negative AD patients were included in the study, which may have influenced classifier accuracy. Sophisticated feature selection and kernel weighting on a voxel level should be further explored. Exploration of voxelwise weight maps will provide information on spatiotemporal distribution of biomarkers, characterizing in detail the pathological processes in different stages of the disease.

Other fMRI measures, similar to Dai et al. [49] could enhance classification. fMRI is especially important in early stages of AD [11,64]. The current limitation of the regional approach is that importance of separate brain regions is not taken into account. An event-based model (EBM) has been proven successful in ordering biomarker transitions across the spectrum of AD disease progression [65]. Combining EBM with regional MKL may provide opportunities for incorporating regional importance in multi-modal classification across the disease stages.The application of multi-modality classifiers will lead to better modeling of AD progress in imaging data, and will enable earlier, more accurate diagnosis of AD, by making more rational use of imaging.

Conclusion

This paper demonstrates the feasibility of reliable image-based classification between AD patients and controls that does not require PET scanning. Multi-modality pattern classification of brain MRI scans is currently more successful at the regional level. To move to a higher resolution in the feature space (e.g., the voxel level) in multi-modality classifiers for diagnostic support, will require more research and innovation.

References

[1] Hardy J, Selkoe DJ, 2002. The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. Science, 297(5580), pp.353–356.

[2] Klunk WE, et al., 2004. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Ann Neurol, 55(3), 306–319.

[3] Cuingnet R, et al., 2011. Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. Neuroimage, 56(2), 766–781.

[4] Henneman WJP, et al., 2009. Hippocampal atrophy rates in Alzheimer disease Added value over whole brain volume measures. Neurology, 72(11), 999–1007.

[5] Klöppel S, et al., 2008. Automatic classification of MR scans in Alzheimer's disease. Brain, 131(3), 681–689.

[6] Fox NC, Schott JM, 2004. Imaging cerebral atrophy: normal ageing to Alzheimer's disease. Lancet, 363(9406), 392–394.

[7] Agosta, F. et al., 2012. Resting state fMRI in Alzheimer's disease: beyond the default mode network. Neurobiology of Aging, 33(8), 564–1578.

[8] Binnewijzend MAA, et al., 2012. Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. Neurobiol Aging, 33(9), 2018–2028.

[9] Chen Y, et al., 2011. Voxel-level comparison of arterial spin-labeled perfusion MRI and FDG-PET in Alzheimer disease. Neurology, 77(22), 1977–1985.

[10] Tolboom N, et al., 2009. Detection of Alzheimer Pathology In Vivo Using Both 11C-PIB and 18F-FDDNP PET. J Nucl Med, 50(2), 191–197.

[11] Filippini N, et al., 2009. Distinct patterns of brain activity in young carriers of the APOE-ε4 allele. Proceedings of the National Academy of Sciences. Available at: http://www.pnas.org/content/early/2009/04/08/0811879106 [Accessed February 1, 2013].

[12] Fan Y, Resnick SM, Wu X, Davatzikos C, 2008. Structural and functional biomarkers of prodromal Alzheimer's disease: A high-dimensional pattern classification study. Neuroimage, 41(2), 277-85.

[13] Cortes C, Vapnik V, 1995. Support-vector networks. Machine Learning, 20(3), 273–297.

[14] Bennett KP, Campbell C, 2000. Support vector machines: hype or hallelujah? SIGKDD Explor Newsl, 2(2), 1–13.

[15] Pareira F, Mitchell T, 2009. Machine learning classifiers and fMRI: A tutorial overview. Neuroimage, 45, S199–S209.

[16] Marquand AF, et al., 2013. Automated high accuracy classifcation of Parkinsonian disorders: a pattern recognition approach. PLoS One, 15, 8(7):e69237.

[17] Bateman RJ, et al., 2012. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. New England Journal of Medicine, 367(9), 795–804.

[18] Ewers M, et al., 2011. Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia. Trends in Neurosciences, 34(8), 430–442.

[19] O'Dwyer L, et al., 2012. Using Support Vector Machines with Multiple Indices of Diffusion for Automated Classification of Mild Cognitive Impairment. PLoS ONE, 7(2), e32441.

[20] Plant C, et al., 2010. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. Neuroimage, 50(1), 162–174.

[21] Zhang D, et al., 2011. Multimodal classification of Alzheimer's disease and mild cognitive

impairment. Neuroimage, 55, 856–867.

[22] Zhang D, Shen D, 2012. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. Neuroimage, 59, 895–907.

[23] Bron EE, et al., 2014. Diagnostic classification of arterial spin labeling and structural MRI in presenile early stage dementia. Hum Brain Mapp, 35, 4916-4931.

[24] van der Flier WM, et al., 2014. Optimizing patient care and research: the Amsterdam Dementia Cohort. J Alzheimers Dis, 41, 313-327.

[25] Ossenkoppele R, et al., 2012. Longitudinal imaging of Alzheimer pathology using [11C] PIB, [18F]FDDNP and [18F]FDG PET. Eur J Nucl Med Mol Imaging, 39(6), 990–1000.

[26] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM, 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. Neurology, 34(7), 939–944.

[27] Folstein MF, et al., 1975. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res, 12(3), 189-198.

[28] Brix G, et al., 1997. Performance evaluation of a whole-body PET scanner using the NEMA protocol. J Nucl Med, 38(10), 1614–1623.

[29] Svarer C, et al., 2005. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. Neuroimage, 24, 969–979.

[30] Wu Y, Carson RE, 2002. Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging. J Cereb Blood Flow Metab, 22, 1440–1452.

[31] Yaqub M, Tolboom N, Boellaard R, van Berckel BN, van Tilburg EW, Luurtsema G, Scheltens P, Lammertsma AA (2008): Simplified parametric methods for [11C]PIB studies. Neuroimage 42:76–86.

[32] Yamaguchi H, Hirai S, Morimatsu M, Shoji M, Nakazato Y, 1989. Diffuse type of senile plaques in the cerebellum of Alzheimer-type dementia demonstrated by beta protein immunostain. Acta Neuropathol, 77(3), 314-319.

[33] Mazziotta J, et al., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 356(1412), 1293–1322.

[34] Jenkinson M, et al., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage, 17(2), 825–841.

[35] Smith SM, et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23(S1), 208-219.

[36] Andersson JLR, Jenkinson M, Smith S, 2010. Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2

[37] Rueckert D, et al., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging, 18(8), 712 –721.

[38] Smith SM, et al., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A, 106(31), 13040-13045.

[39] Adriaanse, S.M. et al., 2012. Amyloid and its association with default network integrity in Alzheimer's disease. Hum Brain Mapp, 35(3), 779-791.

[40] Buckner RL, et al., 2009. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. J Neurosci, 29(6), 1860–1873.

[41] Rombouts SARB, et al., 2005. Altered resting state networks in mild cognitive impairment

and mild Alzheimer's disease: an fMRI study. Hum Brain Mapp, 26(4), 231–239.

[42] Rombouts SARB, et al., 2009. Model-free group analysis shows altered BOLD FMRI networks in dementia. Hum Brain Mapp, 30(1), 256–266.

[43] Wink AM, et al. 2012. Fast Eigenvector Centrality Mapping of VoxelwiseConnectivity in Functional Magnetic Resonance Imaging: Implementation, Validation, and Interpretation. Brain Connect, 2(5), 265-274.

[44] Gong G, et al., 2009. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. Cereb Cortex, 19(3), 524–536.

[45] Tzourio-Mazoyer N, et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage, 15, 273–289.

[46] Chang C, Lin C. 2001. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27.

[47] Burges C, 1998. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov, 2, 121–167.

[48] Agresti A, 2002. Categorical Data Analysis. Hoboken, New Jersey: John Wiley and Sons.

[49] Dai Z, et al, 2012. Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). Neuroimage, 59, 2187–2195.

[50] Dukart J, et al., 2013. Relationship between imaging biomarkers, age, progression and symptom severity in Alzheimer's disease. Neuroimage Clin, 3, 84-94.

[51] Silverman DH, et al., 2001. Positron emission tomography in evaluation of dementia: Regional brain metabolism and long-term outcome. JAMA, 286(17), 2120-7.

[52] Chetelat G, Desgranges B, de la Sayette V, Viader F, Eustache F, Baron J-C, 2002. Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. NeuroReport, 13, 1939–1943.

[53] Dickerson BC, et al., 2001. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiol Aging, 22(5), 747-754.

[54] Wang L, et al., 2014. Multi-subject brain decoding with multi-task feature selection. Biomed Mater Eng, 24, 2987–2994.

[55] Gong Q, et al. 2011. Prognostic prediction of therapeutic response in depression using high-field MR imaging. Neuroimage, 55, 1497-1503.

[56] Karas GB, et al., 2004. Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. NeuroImage, 23, 708–716.

[57] Rabinovici GD, Jagust WJ, 2009. Amyloid imaging in aging and dementia: testing the amyloid hypothesis in vivo. Behav Neurol, 21, 117–28.

[58] Drzezga A, et al. 2003. Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. Eur J Nucl Med Mol Imaging, 30, 1104–1113.

[59] Li Y, et al. 2008. Regional analysis of FDG and PIBPET images in normal aging, mild cognitive impairment, and Alzheimer's disease. Eur J Nucl Med Mol Imaging, 35, 2169–81.

[60] Mosconi L, et al. 2006. Hypometabolism exceeds atrophy in presymptomatic early-onset familial Alzheimer's disease. J Nucl Med, 47, 1778–86.

[61] Greicius MD, Srivastava G, Reiss AL, Menon V, 2004. Default mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. Proc Natl Acad Sci USA, 101, 4637–4642.

[62] Binnewijzend MAA, et al., 2013. Brain network alterations in Alzheimer's disease measured by eigenvector centrality in fMRI are related to cognition and CSF biomarkers. Hum Brain Mapp, 35(5), 2383-2393.

[63] Mormino EC, et al., 2012. Not quite PIB-positive, not quite PIB-negative: Slight PIB elevations in elderly normal control subjects are biologically relevant. Neuroimage, 59(2), 1152–1160.

[64] Hedden T, et al., 2009. Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. J Neurosci, 29, 12686–12694.

[64] Sanz-Arigita EJ, et al., 2010. Loss of 'smallworld' networks in Alzheimer's disease: Graph analysis of fMRI resting-state functional connectivity. PLoS One, 5, e13788.

[65] Young AL, et al, 2014. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain, 137, 2564-2577.