BrainPrint in the Computer-Aided Diagnosis of Alzheimer’s Disease

C. Wachinger\textsuperscript{1,2}, K. Batmanghelich\textsuperscript{1}, P. Golland\textsuperscript{1}, M. Reuter\textsuperscript{1,2}

\textsuperscript{1}Computer Science and Artificial Intelligence Lab, MIT
\textsuperscript{2}Massachusetts General Hospital, Harvard Medical School

Abstract. We investigate the potential of shape information in assisting the computer-aided diagnosis of Alzheimer’s disease and its prodromal stage of mild cognitive impairment. We employ BrainPrint to obtain an extensive characterization of the shape of brain structures. BrainPrint captures shape information of an ensemble of cortical and subcortical structures by solving the 2D and 3D Laplace-Beltrami operator on triangular and tetrahedral meshes. From the shape descriptor, we derive features for the classification by computing lateral shape differences and the projection on the principal component. Volume and thickness measurements from FreeSurfer complement the shape features in our model. We use the generalized linear model with a multinomial link function for the classification. Next to manual model selection, we employ the elastic-net regularizer and stepwise model selection with the Akaike information criterion. Training is performed on data provided by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) and testing on the data provided by the challenge. The approach runs fully automatically.

1 Introduction

This paper describes our method submitted to the challenge on Computer-Aided Diagnosis of Dementia based on structural MRI data held in conjunction with MICCAI 2014 in Boston. The task of the challenge is to differentiate between patients with Alzheimer’s disease (AD), mild cognitive impairment (MCI), and healthy controls (CN) given T1-weighted MRI data. Alzheimer’s disease leads to structural changes in the brain that are visible in T1 images. The challenge provides multi-center data from hospitals in the Netherlands, including 30 subjects for validation and 354 subjects for testing. The objective of the study is to provide a large-scale objective validation of methods for computer-aided diagnosis of dementia and therefore promoting their application in clinical practice.

We propose to augment the commonly used volume and thickness measurements for AD classification with shape information. Shape information can contribute valuable information to the characterization of brain structures, which is only coarsely represented by its volume. For quantifying shape, we use the recently introduced BrainPrint [22], which is a holistic representation of the brain anatomy, containing the shape information of an ensemble of cortical and subcortical structures. Previous studies that employed shape information for the
classification focused on single structures, e.g. the hippocampus [8, 6, 3, 21], while BrainPrint provides an extensive characterization of the brain anatomy. Such a holistic characterization seems promising in diagnosing Alzheimer’s disease, which is associated with global atrophy across the brain. BrainPrint naturally extends the ROI-based analysis in FreeSurfer that provides volume and thickness measurements with shape information.

BrainPrint quantifies the shape information by calculating the spectrum of the Laplace-Beltrami operator on both triangular meshes that represent boundary surfaces, e.g. the white matter surface, and tetrahedral meshes for volumetric representations. The meshes are constructed from the segmentation provided by FreeSurfer. From the detailed, high-dimensional characterization with BrainPrint, we compute lateral shape differences and the projection on the principal component to obtain shape features that can easily be integrated in the classification. These shape features complement volume and thickness measurements in our analysis, which are appropriately normalized to account for head size and gender. The classification is performed with the generalized linear model and a multinomial link function. Next to manual model selection, we employ the elastic-net regularizer and stepwise model selection with the Akaike information criterion. Training is performed on data provided by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) and testing on the data provided by the challenge.

1.1 Related Work

A recent review of machine learning approaches in Alzheimer’s disease [5], refers to only two articles that employ shape features [6, 3], illustrating that shape is not commonly used. A recent comparison of methods for AD classification [4] included one shape-based approach [8]. The focus of these shape-based approaches is the hippocampus. More precisely, [8] approximate the hippocampal shape by a series of spherical harmonics. [6] use permutation tests to extract surface locations that are significantly different among patients with AD and controls. [3] incorporate shape information by deriving thickness measurements of the hippocampus from a medical representation. [21] use statistical shape models to detect hippocampal shape changes. Other structures of interest for shape analysis were the cortex and ventricles. [11] use multi-resolution shape features with non-Euclidean wavelets for the analysis of cortical thickness. [12] analyze the fractal dimension of the cortical ribbon. [10] model surface changes of the ventricles in a longitudinal setup with a medial representation.

The shape characterization in BrainPrint [22] builds upon shapeDNA [20], which contains the spectral information of objects. Spectral signatures have previously been investigated for AD classification [1], with a focus on right hippocampus, right thalamus and right putamen. In contrast, we characterize cortical structures and the wide range of subcortical structures. Moreover, a 3D object can be represented by its volume (e.g. tetrahedra mesh) or its boundary (e.g. triangle meshes), where the spectra of 3D solid objects and their 2D boundary surfaces contain complementary information [20]. Most prior work computes
the shapeDNA for triangular surface meshes [1, 2, 16], while we also work with
spectra of tetrahedral volume tessellations [19]. Given that the Laplace spectra
are isometry invariant, the 2D boundary representation alone may yield a weaker
descriptor, due to the large set of potential (near-) isometric deformations. E.g. a
closed 2D surface with a protrusion pointing inwards yields the same descriptor
as one with the protrusion pointing outwards, while the spectra of the enclosed
3D solids differ. Furthermore, it has been shown in [17] that spectra of the 2D
boundary surface are capable of distinguishing two isospectral 3D solids (GWW-
prisms). For these reasons we combine the information from both the 3D solid
and 2D boundary shape representations in the BrainPrint.

2 Shape Features for Classification

2.1 Shape Descriptor BrainPrint

The shape description is based on the automated segmentation of anatomical
brain structures with FreeSurfer [7]. In this work we use the shapeDNA [20] as
shape descriptor, which performed among the best in a comparison of methods
for non-rigid 3D shape retrieval [14]. The ShapeDNA is computed from the
intrinsic geometry of an object by calculating the Laplace-Beltrami spectrum.
It provides a compact shape representation that facilitates the further analysis.
Considering the Laplace-Beltrami operator $\Delta$, we obtain the spectrum by solving
the Laplacian eigenvalue problem (Helmholtz equation)

$$\Delta f = -\lambda f$$

using the finite element method. The solution consists of eigenvalue $\lambda_i \in \mathbb{R}$ and
eigenfunction $f_i$ pairs (sorted via their eigenvalues: $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots$). The first
$l$ non-zero eigenvalues form the shapeDNA: $\lambda = (\lambda_1, \ldots, \lambda_l)$. Next to this direct
definition of shapeDNA, we also consider the normalization of the eigenvalues
to make the representation independent of the objects’ scale

$$\lambda' = \text{vol}^\frac{2}{D} \lambda,$$

where $\text{vol}$ is the Riemannian volume of the $D$-dimensional manifold (i.e. the area
for 2D surfaces) [20].

The eigenvalues are isometry invariant with respect to the Riemannian mani-
fold, meaning that length-preserving deformations will not change the spectrum.
Isometry invariance includes rigid body motion and therefore permits the com-
parison of subjects by directly comparing the shapeDNA, without the need for
alignment. A second property is that the spectrum continuously changes with
topology-preserving deformations of the boundary of the object. Fig. 1 illustrates
the eigenfunctions calculated on the surface of the cerebral cortex. The eigen-
functions show natural vibrations of the shape when oscillating at a frequency
specified by the square root of the eigenvalue.

We compute the spectra for all cortical and subcortical structures on the
2D boundary surfaces (triangle meshes) and additionally for cortical structures
Fig. 1: Left cerebral cortex and first eigenfunctions of the Laplace-Beltrami operator calculated on the surface (yellow positive, red negative, and green zero).

(white and pial surfaces in both hemispheres) also on the full 3D solid (tetrahedra meshes), forming the BrainPrint $A = (\lambda_1, \ldots, \lambda_\eta)$. Triangle meshes of the cortical surfaces were obtained automatically for each hemisphere using FreeSurfer. Surface meshes of subcortical structures were constructed via marching cubes from the FreeSurfer subcortical segmentation. To construct tetrahedral meshes, first handles were removed from the surface meshes, which were then uniformly resampled to 60K vertices. The gmsh package [9] was used to compute the tetrahedra volume meshes. For the computation of the spectra we used the linear finite element method [20] with Neumann boundary condition (zero normal derivative) for tetrahedra meshes.

2.2 Shape Features from BrainPrint

BrainPrint contains shape information from $\eta = 44$ brain structures, each one described by $l = 50$ eigenvalues. This results in a characterization of a subject’s brain shape by over 2000 variables, which can easily cause overfitting of the model. To decrease the number of variables and increase robustness, we therefore (i) include a 1D asymmetry measure (the distance of BrainPrint across hemispheres), and (ii) employ principal component analysis. When working with BrainPrint, the potential switching of eigenfunctions across shape deformations as discussed in [18] are not problematic as we compare the sorted sequences of eigenvalues (shapeDNA). In order for eigenfunctions to switch, their eigenvalues have to be very close initially, so a switch has only a limited effect on the BrainPrint. Another aspect of shapeDNA is that the eigenvalues form an increasing sequence with the variance increasing as well. Depending on the distance measure, this can cause higher eigenvalues to dominate the similarity measure between two shapes, although these components do not necessarily contain the most important geometric information. Additionally, the increased impact of higher eigenvalues puts a particular emphasis on number of selected eigenvalues $l$. To account for these issues we normalize the BrainPrint and employ appropriate distance computations as described next.

Lateral differences: As a first shape feature, we compute lateral (left/right) shape differences of brain structures. More precisely, we compute differences for white and gray matter surfaces with triangular and tetrahedral meshes, as well
as triangular meshes for cerebellum white matter and gray matter, striatum, lateral ventricles, hippocampus and amygdala. We calculate the Mahalanobis distance between a brain structure $s$

$$d_s = \|\lambda_s^{\text{left}} - \lambda_s^{\text{right}}\| \Sigma_s,$$

with $\Sigma_s$ the covariance matrix across all subjects for structure $s$.

Alternative distance functions that have been proposed for shapeDNA are the Euclidean distance (or any p-norm), some Hausdorff distances, the Euclidean distance on re-weighted eigenvalues $\hat{\lambda}_i = \lambda_i / i$ [20, 17], and the weighted spectral distance [13]. The weighted distances (latter two approaches) are motivated by the need to reduce the impact of higher eigenvalues on the distance. The linear re-weighting is based upon the observation that the eigenvalues demonstrate a linear growth pattern (Weyl’s law) and therefore yields an approximately equal contribution of each eigenvalue. The weighted spectral distance is similar to a division by the squared eigenvalue number and therefore functions like a low-pass filter. The proposed Mahalanobis distance accounts for the covariance pattern in the data and supports an equal contribution of all eigenvalues in the sequence.

**Principal Component Analysis:** A second set of features that we derive from BrainPrint is the calculation of the principal component analysis for each of the 44 brain structures. Projecting the shapeDNA on the principal component retains most of the variance in the dataset, while reducing the dimensionality. Problematic in this regard is once again that higher eigenvalues show most variance, so that they will dominate the identification of the principal component. We have experimented with (i) linear re-weighting, $\hat{\lambda}_i = \lambda_i / i$, and (ii) the normalization of each eigenvalue to unit variance across the dataset. Evaluation of both approaches yielded similar results, so that we employ the simpler linear re-weighting.

### 3 Alzheimer Classification with BrainPrint

We use the generalized linear model (GLM) for the classification with a multinomial link function. We employ linear models in this study because we expect non-linear methods in combination with the large number of features to be more susceptible to overfitting.

#### 3.1 Features

Each subject is characterized by 163 features. These split up in 70 thickness and 39 volume measurements provided by FreeSurfer, together with 10 lateral shape differences, and 44 PCA shape variations. According to the recent analysis of the normalization of variables for AD classification [23], we normalize volumetric measures by ICV but do not normalize cortical thickness measures. Linear regression with respect to age is performed for each feature to remove the confounding effect of age in the analysis. After the normalization, we exclude age and gender from the variables.
3.2 Model Selection

We use three different approaches for model selection: manual, stepwise selection, and elastic-net. For the manual model, we choose volume and thickness measurements that have previously been reported important for AD classification and add shape features. For volume, we use hippocampus and amygdala. For thickness, we use entorhinal cortex, middle temporal lobe, parahippocampal gyrus, and banks of the superior temporal sulcus. For shape, we use the lateral differences of hippocampus, amygdala, and lateral ventricles. Experiments show a clear improvement in classification accuracy with the additional shape features.

As second approach, we employ stepwise model selection in the GLM with the Akaike information criterion (AIC). The AIC considers model complexity in terms of the number of variables and the likelihood of the data given the data. The model with the best trade-off between high likelihood and few variables is selected by the AIC. We performed the model selection once on the ADNI training data, yielding a model with 25 thickness, 15 volume, and 13 shape features, referred to as “stepwise I”. In a second run, we added the challenge training data to the model selection, leading to a model with 30 thickness, 11 volume, and 19 shape features, referred to as “stepwise II”.

In the final approach for model selection, we use elastic-net regularized general linear models provided by the R package glmnet. The regularizer in elastic-net combines lasso and ridge-regression penalties, modulated by the parameter $\alpha$

$$P_\alpha(\beta) = (1 - \alpha) \frac{1}{2} \| \beta \|^2_2 + \alpha \| \beta \|_1,$$

with coefficients $\beta$. Our experiments indicated best results for the lasso penalty ($\alpha = 1$). An additional parameter $\lambda$, balancing the data fit and penalty term, was determined with cross-validation, see Fig. 2. The lowest multinomial deviance of the model corresponds roughly to $\lambda = \exp(-4.5)$. 

Fig. 2: Multinomial deviance of elastic-net computed with cross-validation for different parameters $\lambda$ (bottom) and the corresponding number of features (top). The plot shows the mean deviance together with upper and lower standard deviation.
Subjects Diagnosis Gender Age quantiles
(CN/MCI/AD) Male Female (1st/2nd/3rd)
ADNI 751 (213/364/174) 437(58%) 314(42%) (71.1/75.3/79.8)
Challenge-Validation 30 (12/9/9) 17(57%) 13(43%) (59.3/65.0/68.0)
Challenge-Test 354 213(60%) 141(40%) (59.0/64.0/71.0)

Table 1: Statistics of datasets used for training and testing.

4 Results

We processed the non-uniformity corrected brain scans distributed by the challenge with FreeSurfer v5.3. We processed the baseline scans from ADNI-1 [15] with FreeSurfer v5.1. Table 1 lists statistics about diagnosis and demographics of both datasets. Notable is the age difference between the challenge and ADNI data. The first quantile of ADNI is higher than the third quantile of the challenge data meaning that 75% of ADNI cases are older than 75% of challenge cases. This mismatch may have a detrimental effect on the classification. All results reported in this paper are obtained by training on the ADNI data and predicting on the challenge-validation data. The results on the challenge-test data will be announced by organizers later. For the final training with the submission of the results to the challenge, we also include the validation data to the training set because of the potential benefit of including samples from the target dataset to training.

The processing of the challenge data with FreeSurfer failed for 3 subjects. These subjects are labeled as AD without further inspection, since we assume a relation between processing difficulties and severe atrophy or increased subject motion in AD patients. The BrainPrint calculation caused errors on triangular meshes for 1 subject and on tetrahedral meshes on 5 subjects. The errors are related to FreeSurfer producing a mesh that is a non-manifold (the mesh has edges sharing more than two triangles) and problems in the volumetric mesh creation with gmsh. We impute missing values for these 6 subjects by the population mean in the statistical analysis. The mean runtime of FreeSurfer was 16.8h on the challenge data. The calculation time for BrainPrint is about 0.6h per subject, where most time is spent on mesh processing. The statistical analysis runs in less than a minute. All processing steps run fully automatically.

Our approach depends on a number of design choices and variables. One is the distance function for the calculation of lateral shape differences. The results across the model selection schemes for different distance functions are not unanimous but the Mahalanobis distance achieved competitive results. We choose a linear re-weighting of the eigenvalues in PCA. Another important parameter is the number of eigenvalues $l$ forming the shapeDNA. Fig. 3 illustrates the classification accuracy on the challenge-validation data as a function of the number of eigenvalues. We compare the manual model selection, the two stepwise approaches, and the elastic-net approach. The plot also shows the impact of
Fig. 3: Classification accuracy on challenge-validation data. The plots contrast the impact of the volume normalization in Eq.(2). Results are shown for different model selection schemes and are plotted as function of the number eigenvalues $l$.

the volume normalization of shapeDNA with Eq.(2) on the classification performance.

Based on the results, we select $l = 40$ eigenvalues for the further analysis. Table 2 shows additional evaluation measures for this case. The scripts to calculate these measures were provided by the challenge organizers. For the final prediction on the challenge-test data, we also added the challenge-validation data to the training set. We measure the consistency of the predicted classes on the challenge-test data for the changed training set and list it in the table. We observe a tendency for less consistent results with more sophisticated methods. Since the number of results submitted per group was limited to 5, we select the methods with best performances, highlighted in bold face in table 2. We also included the manual model, although it performs slightly worse than the elastic-net.

5 Conclusion

We investigated the potential of BrainPrint for classifying between AD, MCI, and CN in the MICCAI dementia challenge. We derived shape features from BrainPrint and used them together with volume and thickness measures in generalized linear models for classification. Three different approaches for model selection were presented. We trained our methods on the ADNI dataset. Our results indicate the good performance in classification when adding the ensemble of shape information from BrainPrint.

Acknowledgements: This work was supported in part by the Humboldt foundation, the Martinos Center for Biomedical Imaging (P41-RR014075, P41-EB015896), the National Alliance for Medical Image Computing (U54-EB005149)
Table 2: Evaluation of classification results for different models calculated on the challenge-validation data. Confidence intervals are shown in parenthesis. All values are in %. (Norm = volume normalization, TPF = true positive fraction, AUC = area under the curve, CON = consistency of prediction). Methods submitted to the challenge are highlighted with bold face.

and the NeuroImaging Analysis Center (P41-EB015902). We thank the Alzheimer’s Disease Neuroimaging Initiative (ADNI) for image data.

References