Deep spectral-based shape features for Alzheimer's Disease classification

Mahsa Shakeri[1,2], Herve Lombaert[3], Shashank Tripathi[1],
Samuel Kadoury[1,2]
for the Alzheimer's Disease Neuroimaging Initiative*

[1] MEDICAL, Polytechnique Montreal, Montreal, Canada,
[2] CHU Sainte-Justine Research Center, Montreal, Canada,
[3] INRIA Sophia-Antipolis, France

Abstract. Alzheimer's disease (AD) and mild cognitive impairment (MCI) are the most prevalent neurodegenerative brain diseases in elderly population. Recent studies on medical imaging and biological data have shown morphological alterations of subcortical structures in patients with these pathologies. In this work, we take advantage of these structural deformations for classification purposes. First, triangulated surface meshes are extracted from segmented hippocampus structures in MRI and point-to-point correspondences are established among population of surfaces using a spectral matching method. Then, a deep learning variational auto-encoder is applied on the vertex coordinates of the mesh models to learn the low dimensional feature representation. A multi-layer perceptrons using softmax activation is trained simultaneously to classify Alzheimer's patients from normal subjects. Experiments on ADNI dataset demonstrate the potential of the proposed method in classification of normal individuals from early MCI (EMCI), late MCI (LMCI), and AD subjects with classification rates outperforming standard SVM based approach.

Keywords: classification, spectral matching, variational autoencoder, Alzheimer's disease.

1 Introduction

Alzheimer's disease (AD) is characterized by progressive impairment of cognitive and memory functions in elderly population. Considering its worldwide prevalence, early diagnosis of this disease might have a huge impact on the overall well-being of the population, and the burden to caregivers, as well as the associated financial costs to the world's health system. Studies reported that AD can be diagnosed by clinical assessments in most of the cases [1], while by the time the patient is diagnosed the disease

*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`

progression may have deteriorated. Therefore, early diagnosis of this neuropathology is of special interest.

Mild cognitive impairment (MCI) is considered as a transition state between normal aging and dementia [2]. The cognitive deficits in MCI patients are not as severe as those seen in individuals with AD. However, studies have suggested that about $10-12\%$ of subjects with MCI progress to AD per year [2]. Therefore, these individuals with milder degrees of cognitive and functional impairment than AD patients are particularly interesting subjects, since biomarker manifestation could potentially be different at such an early stage of the disease.

Studies have shown that the neuropathological changes in AD and MCI affect the hippocampus structure, which is a brain region crucial to various cognitive functions [3]. Neuroimaging datasets for AD including magnetic resonance imaging (MRI) and other types of biomarkers have shown considerable promise to detect longitudinal changes in subjects [4], by offering rich information on the patients morphometric and anatomical profiles. Their use stems from the premise that morphological changes may be more reproducible and more precisely measured with MRI than other parameters such as clinical scores, cerebrospinal fluid (CSF), or proteomic assessments.

Recent advances in medical imaging and classification techniques have led to a better discrimination between Alzheimer's disease and healthy aging. Because of the high dimensionality of medical image, various dimensionality reduction approaches have been developed to facilitate and enhance classification accuracy. A simple method is principal components analysis (PCA) [5], which finds the directions of greatest variance in the dataset and represents each data point by its coordinates along each of these directions. A nonlinear generalization of PCA is multi-layer autoencoders (AE) [6], which is a feedforward neural network to encode the input into a more compact from and reconstruct the input with the learned representation. Among available AE architectures, the deep variational autoencoder (VAE) [7] method has recently become popular in computer vision due to its capability to learn a manifold without the assumption of linearity in addition to its generative property.

With respect to surface representation, recent studies have shown the advantage of spectral shape description compared to Euclidean surface representation [8] [9] [10]. The use of eigenvalues have led to interesting results for AD classification in [11], where Laplace-Beltrami spectrum on the intrinsic geometry of the structural meshes was computed to define the shape descriptors.The spectral coordinates, which were derived from the Laplacian eigenfunctions of shapes have been used in [8] to parametrize surfaces explicitly. The authors applied a Random Decision Forest classifier on spectral representation of surfaces and achieved a significant improvement on cortical parcellations. Also, in [9] and [10], the eigendecomposion of the surfaces in the spectral domain were used to provide pointwise information on meshes and establish accurate point-to-point correspondences across surfaces.

In this work, we present a surface-based classification technique based on classification of spectral features using variational stacked auto-encoders. We first extract 3D surface meshes of hippocampus structures from segmented binary MR images. Then, the point-to-point surface correspondences is established across populations (NC, AD, EMCI, LMCI) using a spectral matching approach. In spectral based shape matching approach, relationships are modeled as graphs and an eigendecomposition on these graphs enables us to match similar features. Once the matched surfaces are created, the vertex coordinates are used as shape feature descriptors. Then, variational autoencoder (VAE) obtains the non-linear low-dimensional embedding of the shape features. A multi-layer perceptron (MLP) classifier is simultaneously trained to model the non-linear decision boundaries between classes.

The work follows on the prior work of [12], which used a Stacked Auto-Encoder (SAE) to discover the latent representation from the grey matter (GM) tissue densities and voxel intensities. Unlike Suk et al. [12], which selects intensity and volume based features from MRI and PET modalities, we create the feature descriptors from matched hippocampi surfaces extracted from MRI. Moreover, instead of training a separate classifier on the low dimensional features as in [12], we add a softmax multi-layer perceptron on top of our variational autoencoder network to obtain both dimensionality reduction and the classification output at the same time.

The rest of the paper is organized as follows. In Section 2, we present the morphological feature extraction method using spectral shape matching, as well as the feature representation and classification method based on variational autoencoder and multi-layer perceptron. Section 3 includes the description of the dataset, experiments and discussion. Our conclusions are presented in Section 4, along with envisioned future research directions.

2 Methodology

Given MR images along with their corresponding hippocampus segmentations (produced manually or automatically), we first extract features from MRI as explained in Section 2.1. Then, we use a deep variational autoencoder (VAE) to learn a latent feature representation from the low-level features and train a multi-layer perceptron (MLP) for classification purposes in Section 2.2.

2.1 Shape feature extraction using spectral matching

Given a reference surface mesh S_r and a population of n surfaces $\{S_i\}_{i=1..n}$, the spectral matching between each surface meshe S_i and S_r is done in a two step process. First, an initial map is calculated between the two surfaces [9]. This initial map is then used in the second step to establish a smooth map between the two meshes [10].

Here, we consider vertices and neighbouring points in each surface mesh as nodes and edges of a graph. Then a laplacian graph is created for each surface graph from

the set of vertices and edges of each mesh. The general Laplacian operator L_i [13] is defined on each surface as following:

$$L_i = G_i^{-1} (D_i - W_i) \tag{1}$$

where W_i is the weighted adjacency matrix, which is created based on a distance between connected nodes. The term D_i is a diagonal matrix, in which the elements are set by the degree of vertices. G_i is a node weighting matrix created based on the mean curvature at each node as described in [14].

The eigendecomposition of Laplacian matrix L_i provides its spectral components. After reordering the spectral components by finding the optimal permutation of components between the pair of meshes, regularization is performed by matching the spectral embeddings. The correspondence initial map c between each pair of vertices on S_i and S_r is established with a simple nearest-neighbour search between their spectral representations.

In the next step, given initial map c, the final smooth map between two surfaces S_i and S_r is obtained. In this process, an association graph is defined as the union of the set of vertices and edges of two surfaces with an initial set of correspondence links c between both surfaces. Then, a Laplacian matrix is created for the association graph, and the spectral decomposition is computed to produce a shared set of eigenvectors that enables a direct mapping between two meshes S_i and S_r.

Once all 3D meshes are matched to the reference, the vertices of all surfaces are rearranged to create the new reconstructed meshes with consistent vertex ordering. Now, the shape descriptor x_i will be created for the surface S_i as a vector of (X, Y, Z) coordinate of all vertices.

2.2 Feature learning and classification

In this work we use a deep learning-based feature representation method to improve the classification accuracy. Here, we take inspiration from the variational autoencoder network, which learns the low-dimensional manifold without the linearity assumption and has a generative model. In this section, we explain the proposed network architecture, which is a combination of a variational autoencoder network (VAE) and a softmax multi-layer perceptron (MLP). The combined VAE-MLP network architecture is shown in Figure 1.

Deep variational autoencoder and MLP classifier:

Auto-encoders are a type of deep neural networks structurally defined by input, hidden, and output layers. Given the input data $x \in R^D$ defined from the spectral representation of mesh shapes, an auto-encoder maps it to a latent representation $z \in R^d$

(encoding), which could be used for unsupervised learning or for feature extraction. The representation z from the hidden layer is then mapped back to a vector $y \in R^D$ (decoding), which approximately reconstructs the input vector x. The hidden layer in the middle, i.e., z, can be constrained to be a bottleneck to learn compact representations of the input data.

Variational autoencoder (VAE) assumes that data is generated by a directed graphical model with a latent variable z. VAE uses the encoder network to map the input x into the continuous latent variables ($q_\phi(z|x)$) and uses decoder network to map latent variables to reconstructed data ($p_\theta(x|z)$), where ϕ and θ are the parameters of the encoder (recognition model) and decoder (generative model), respectively.

The lower bound VAE loss function of the variational autoencoder for individual datapoint x_i has the following form:

$$L_{VAE}(\theta, \phi; x_i) = -D_{KL}\left(q_\phi\left(z|x_i\right)||p_\theta\left(z\right)\right) + E_{q_\phi(z|x_i)}\left[\log p_\theta\left(x_i|z\right)\right] \quad (2)$$

The first component is the regularization term, which is the KL divergence of the approximate posterior from the prior, while the second term is the expected reconstruction error. As shown in [7], we assume both $p_\theta\left(z\right)$ and $q_\phi\left(z|x_i\right)$ as Gaussian. Given J as the dimensionality of z and K as the number of samples per datapoint, the resulting estimator for x_i will be as follows:

$$L_{VAE}(\theta, \phi; x_i) = -\frac{1}{2}\sum_{j=1}^{J}\left(1 + \log\left(\sigma_j^2\right) - \mu_j^2 - \sigma_j^2\right) + \frac{1}{K}\sum_{k=1}^{K}\log p_\theta\left(x_i|z_{i,k}\right) \quad (3)$$

where, $z_{i,k} = \mu_i + \sigma_i \odot \epsilon_k$ and $\epsilon_k \sim N\left(0, I\right)$.

Here, μ and σ can be computed using the deterministic encoder network. The reconstruction (decoding) term of $\log p_\theta\left(x_i|z_{i,k}\right)$ could be set as a bernoulli cross-entropy loss function.

The low dimensional features $z_i = \mu_i + \sigma_i$ from the latent layer are fed to an MLP classifier for solving the classification problem. For the last layer, we use the cross entropy loss function and the softmax activation function, which is standard for classification problems [15]. The softmax function ensures that the network outputs are all between zero and one, and that they sum to one on every time step. Therefore, they can be interpreted as the posterior probabilities, given all the inputs up to the current one. We set the number of units in the classification output layer to be equal to the number of classes of interest (i.e., two).

The network architecture:

Annotated medical image datasets tend to be small and generally hard to obtain. This increases the risk of network overfitting in medical applications. Therefore, we make a series of design choices for our network to avoid overfitting. Our network includes L_2 regularization at each layer to penalize the squared magnitude of all parameters directly in the objective function. That is, for every weight w in the network, we add the term $\frac{1}{2}\lambda w^2$ to the cost function, where λ is the regularization strength.

We also add a drop out layer with the probability of 0.5 after each dense layer. During training, dropout is implemented by only keeping a neurone active with some probability p, or setting it to zero otherwise. Network weights are set based on the uniform initialization scaled by the square root of the number of inputs.

We train the network for 100 epochs with batch size of 28 starting with a learning rate of 0.00001 and dropping it at a logarithmic rate to 0.000001. For the deep learning library, we use Keras and Theano. We determine the number of hidden units based on the classification results. The optimal structure of the network is shown in Figure 1.

3 Experiments

We evaluate the performance of our approach on a popular brain imaging dataset in Alzheimer's disease, namely the Alzheimer's Disease Neuroimaging Initiative (ADNI). The ADNI database (adni.loni.usc.edu) was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org. The database of ADNI consists of cross-sectional and longitudinal data including 1.5 or 3.0 T structural MR images. The detailed description of the MRI protocol of ADNI is provided in [16].

For this study, a subset of latest 1.5 T MR images is used including 150 normal controls (NC), 90 AD patients, 160 early MCI (EMCI), and 160 individuals with late MCI (LMCI). ADNI performed additional post-processing steps on MR images to correct certain image artifacts and to enhance standardization across sites and platforms [16]. The post-processing steps include gradient non-linearity correction, intensity inhomogeneity correction, bias field correction, and phantom-based geometrical scaling to remove calibration errors. In this work, we use these processed images. Here, hippocampi was segmented using FSL-FIRST automatic segmentation software package [17] and visual inspection was performed on the output binary masks to ensure the quality of the automatic segmentation.

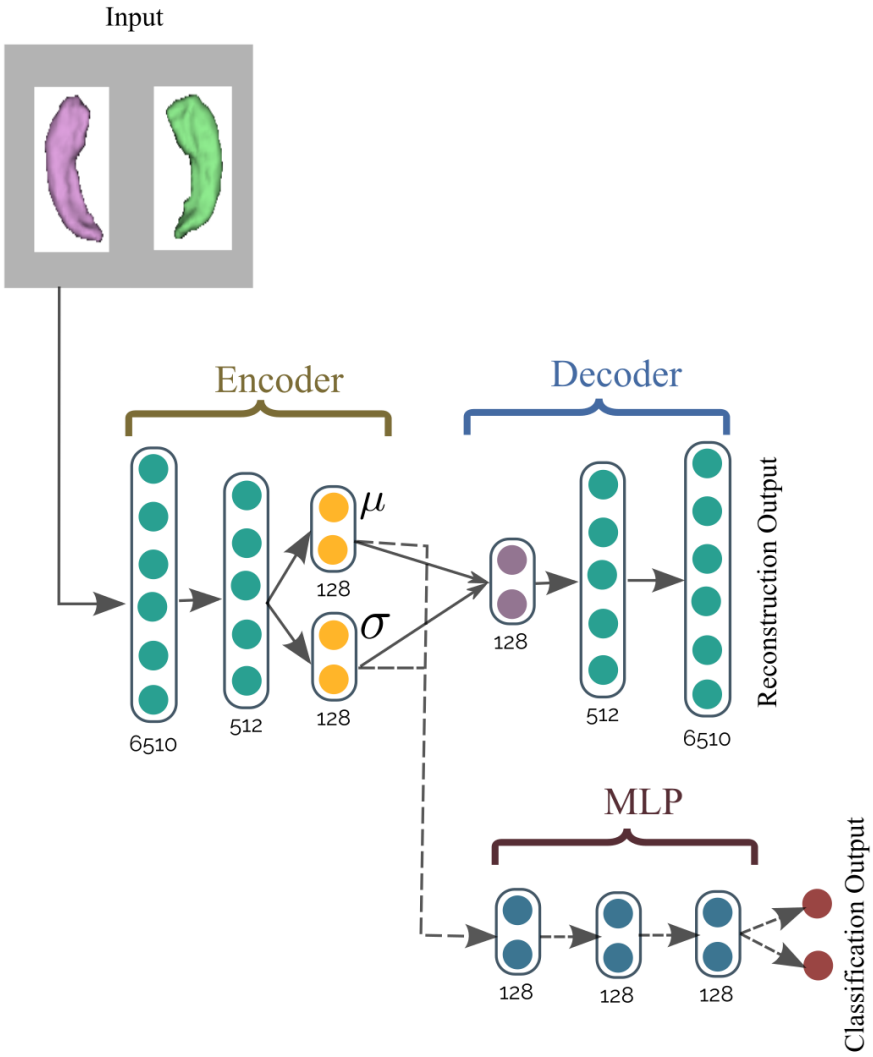

Fig. 1. The architecture of our proposed network. The numbers mentioned under each layer correspond to the layer's dimension.

Here we consider six binary classification problems: AD vs. NC, NC vs. EMCI, NC vs. LMCI, AD vs. EMCI, AD vs. LMCI, and EMCI vs. LMCI. We consider 20% of data for test and the rest for train. Each time 20% of train set is left out and used for validation.The whole process is repeated five times for unbiased evaluation. The regularization strength λ is set as 0.05 based on experimental results.

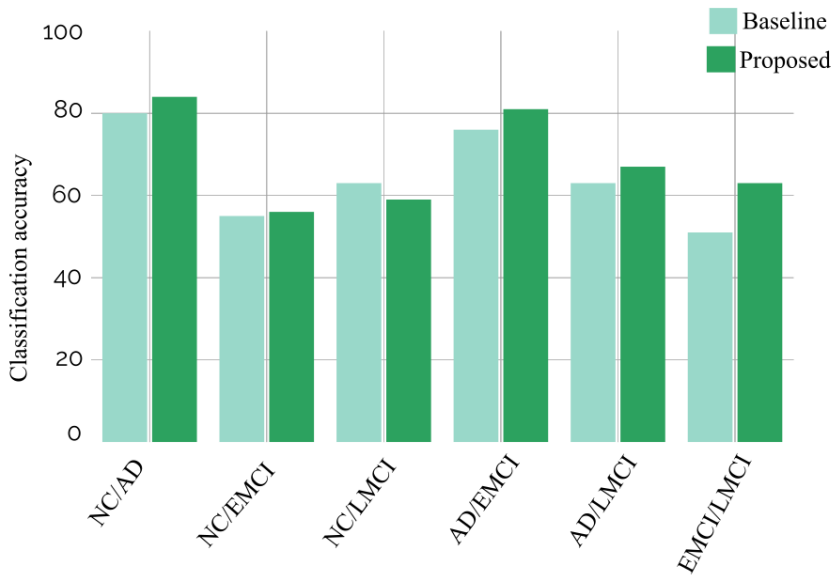

Fig. 2. Comparison of the classification accuracy with a baseline approach using the same spectral-based shape feature representation. The VAE-based method achieved higher accuracy in most of the cases.

We tested different network architectures and realized that going deeper than the proposed model in Figure 1 would not help improving the classification accuracy, however the dimensionality of the hidden and the latent unit had direct effect on the classification performance.

In the analysis of the results, the performance of the classifier are measured by its sensitivity (SE), specificity (SP) and accuracy (AC). Sensitivity, which is the ability of the classifier to correctly identify positive results, is defined as TP/(TP+FN). Specificity refers to the ability to correctly identify negative results and is formulated as TN/(FP+ TN). Accuracy is defined as (TP+ TN)/(TP+TN+FN+FP).

As baseline, we train a linear Support Vector Machines (SVM) on the same dataset after applying principle components analysis (PCA) for dimensionality reduction. The features are extracted from 3D surface meshes after applying spectral matching in the same way as our proposed method. The classification accuracy for the proposed and the baseline methods is illustrated in Figure 2. We summarize the classification accuracy along with the sensitivity (SE), and specificity (SP) measures in Table 1.

Table 1. Comparison of the classification accuracy (AC%), sensitivity (SE%), and specificity (SP%) with a baseline method using the same spectral-based shape feature descriptor. The proposed method achieved higher accuracy in most of the cases.

	NC/AD			NC/EMCI			NC/LMCI			AD/EMCI			AD/LMCI			EMCI/LMCI		
	AC	SE	SP	AC	SE	SP	AC	SE	SP	AC	SE	SP	AC	SE	SP	AC	SE	SP
Baseline	80	70	86	55	52	58	**63**	**56**	**75**	76	65	71	63	58	66	51	50	52
Proposed	**84**	**73**	**89**	**56**	**52**	**60**	59	52	64	**81**	**70**	**82**	**67**	**58**	**73**	**63**	**62**	**66**

These results show that our method produces higher accuracy in most of the cases. As expected, the best classification accuracies are those obtained for groups, which are well separated diagnostically. For instance, 84% and 81% for the classification of NC versus AD and EMCI versus AD, respectively. The computational time of both methods is around 60 *sec* for training on 300 surfaces and less than 5 *ms* for testing on one surface.

In addition, the obtained results is comparable to the previously proposed approaches that have used MRI based features. For instance, Suk et al. [12] and Goryawala et al. [18] found the accuracy of 85% and 84%, respectively for the classification of NC versus AD. These method have also included additional information from PET modality or neuropsychological test to improve the classification performance. One future direction of our proposed approach would be to include a combination of informative features to reach a higher accuracy.

4 Conclusions

In this paper we have proposed a deep learning method based on a spectral feature representation using hippocampus morphology for the classification of Alzheimer's Disease. The morphological features were extracted as 3D surface meshes from MR image and spectral matching process was used to establish point-to-point correspondences in mesh vertices. A variational autoencoder was trained to find the latent feature representation from hippocampus morphological variations. A softmax classifier was applied to differentiate between NC, EMCI, LMCI, and AD.

Experimental evaluation on the ADNI dataset demonstrates the effectiveness of our approach especially in classifying AD vs. NC and AD vs. EMCI. This work shows the importance of the VAE-based morphological feature representation in improving the diagnosis accuracy in different stages of dementia. Future research directions include adding other informative features, such as cognitive information and multimodal data (e.g., PET) to increase the classification accuracy.

Acknowledgements

Funding was provided by the Canada Research Chairs and from the CHU Sainte-Justine Hospital's Research Center, Montreal, Canada.

ADNI data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimers Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.

The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

1. Ranginwala, N.A., Hynan, L.S., Weiner, M.F., White, C.L.I.: Clinical criteria for the diagnosis of alzheimer disease: Still good after all these years. The American Journal of Geriatric Psychiatry **16**(5) (2008) 384 – 388
2. Petersen, R., Smith, G., Waring, S., Ivnik, R., Tangalos, E., Kokmen, E.: Mild cognitive impairment: clinical characterization and outcome. Arch. Neurol **56**(3) (1999) 303 – 308
3. Du AT, Schuff N, e.a.: Magnetic resonance imaging of the entorhinal cortex and hippocampus in mild cognitive impairment and alzheimer's disease. Journal of neurology, neurosurgery, and psychiatry **71** (2001)
4. Wyman, B., Harvey, D., Crawford, K., Bernstein, M., Carmichael, O., Cole, P., Crane, P., Decarli, C., Fox, N., Gunter, J., Hill, D., Killiany, R., Pachai, C., Schwarz, A., Schuff, N., Senjem, M., Suhy, J., Thompson, P., Weiner, M., Jack, C.: Standardization of analysis sets for reporting results from adni mri data. Alzheimer's and Dementia **9**(3) (2013) 332–337
5. Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S.: Alzheimers disease via pattern classification of mri. Neurobiology of aging **29**(4) (2008) 514–523
6. Bengio, Y.: Learning deep architectures for ai. Found. Trends Mach. Learn. **2**(1) (January 2009) 1–127

7. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In: International Conference on Learning Representations (ICLR). (2013)
8. Lombaert, H., Criminisi, A., Ayache, N. In: Spectral Forests: Learning of Surface Data, Application to Cortical Parcellation. Springer International Publishing, Cham (2015) 547–555
9. Lombaert, H., Grady, L., Polimeni, J.R., Cheriet, F.: FOCUSR: feature oriented correspondence using spectral regularization-a method for precise surface matching. IEEE Trans. Pattern Anal. Mach. Intell. **35**(9) (2013) 2143–2160
10. Lombaert, H., Sporring, J., Siddiqi, K.: Diffeomorphic spectral matching of cortical surfaces. In: Information Processing in Medical Imaging - 23rd International Conference, IPMI 2013, Asilomar, CA, USA, June 28-July 3, 2013. Proceedings. (2013) 376–389
11. Wachinger, C., Reuter, M.: Domain Adaptation for Alzheimer's Disease Diagnostics. Neuroimage **139** (2016) 470–479
12. Suk, H.I., Shen, D. In: Deep Learning-Based Feature Representation for AD/MCI Classification. Springer Berlin Heidelberg, Berlin, Heidelberg (2013) 583–590
13. Grady, L.J., Polimeni, J.R.: Discrete Calculus. (2010)
14. Shakeri, M., Lombaert, H., Datta, A.N., Oser, N., Ltourneau-Guillon, L., Lapointe, L.V., Martin, F., Malfait, D., Tucholka, A., Lippe, S., Kadoury, S.: Statistical shape analysis of subcortical structures using spectral matching. Computerized Medical Imaging and Graphics (2016)
15. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY, USA (1995)
16. Jack, C., Bernstein, M., Fox, N., et al.: The alzheimer's disease neuroimaging initiative (adni): Mri methods. Journal of magnetic resonance imaging: JMRI **27**(4) (2008) 685–691
17. Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M.: A bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage (2011)
18. Goryawala, M., et al.: Inclusion of neuropsychological scores in atrophy models improves diagnostic classification of alzheimers disease and mild cognitive impairment. Computational Intelligence and Neuroscience (2015)