Genetic, Structural and Functional Imaging Biomarkers for Early Detection of Conversion from MCI to AD

Nikhil Singh, Angela Y. Wang, Preethi Sankaranarayanan, P. Thomas Fletcher, and Sarang Joshi for the Alzheimer's Disease Neuroimaging Initiative*

University of Utah, Salt Lake City, UT
nikhil@cs.utah.edu

Abstract. With the advent of advanced imaging techniques, genotyping, and methods to assess clinical and biological progression, there is a growing need for a unified framework that could exploit information available from multiple sources to aid diagnosis and the identification of early signs of Alzheimer's disease (AD). We propose a modeling strategy using supervised feature extraction to optimally combine high-dimensional imaging modalities with several other low-dimensional disease risk factors. The motivation is to discover new imaging biomarkers and use them in conjunction with other known biomarkers for prognosis of individuals at high risk of developing AD. Our framework also has the ability to assess the relative importance of imaging modalities for predicting AD conversion. We evaluate the proposed methodology on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database to predict conversion of individuals with Mild Cognitive Impairment (MCI) to AD, only using information available at baseline.

1 Introduction

Mild cognitive impairment (MCI) is an intermediate stage between healthy aging and dementia. Patients diagnosed with MCI are at high risk of developing Alzheimer's disease (AD), but not everyone with MCI will convert. Accurate prognosis for MCI patients is an important prerequisite for providing the optimal treatment and management of the disease. The complex anatomical shape changes that occur during disease progression can be extracted from magnetic resonance images (MRI) of the brain. Decreased synaptic response and brain

* Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.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.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

N. Ayache et al. (Eds.): MICCAI 2012, Part I, LNCS 7510, pp. 132–140, 2012.

function can be measured using functional imaging modalities, such as [^{18}F]-fluorodeoxyglucose Positron Emission Tomography (FDG-PET). Additional potential risk biomarkers include blood and cerebrospinal fluid (CSF) markers, including genetic susceptibility assessed by apolipoprotein E (APOE) genotype and plaque deposition assessed by concentration of $A\beta$-42 and ptau$_{181}$. The challenge for predicting conversion is to combine these heterogeneous data sources, some of which are high-dimensional (MRI and PET) and some low-dimensional (clinical, CSF, APOE carrier), by selecting features that optimally weight the relative contribution from each modality.

Recent studies have examined the role of different classes of biomarkers, cognitive measures, and genetic risk factors either in combination with a single imaging modality or independently for predicting conversion from MCI to AD [1,2]. Weiner et. al [3] offer a comprehensive review of this ongoing research. Despite evidence for the predictive capability of individual biomarkers, cognitive measures, or neuroimaging data, relatively little attention has been given to combining information available from multiple imaging modalities with the biomarkers [4]. In one such study, Kohannim et. al [4] combine FDG-PET-derived numerical summaries, MRI-derived volume measures, CSF biomarkers, APOE genotype, and subject demographics for the task of discriminating MCI from AD. However, their work did not address prediction of conversion to AD.

In this article, we present a unified framework to combine the high-dimensional information available from multiple imaging modalities, anatomical shape atrophy (derived from MRI) and neuronal hypometabolism (derived from FDG-PET), with other low-dimensional biomarkers, such as APOE carrier status, $A\beta$-42 and ptau$_{181}$ concentration. We use Partial Least Squares as a supervised dimensionality-reduction technique to fuse the weighted combination of the two imaging modalities together with the clinical information. This data-driven formulation finds the optimal combination of these high-dimensional modalities that best characterize the disease progression. The focus of this work is to assess the combined predictive capability of this model for early detection of conversion of MCI to AD by using only the information available at baseline.

2 Methodology

We use the general framework of computational anatomy [5] to characterize the anatomical shape variation. Since the anatomical shape and neuronal metabolic activity are two separate measures obtained from independent imaging modalities, we combine the two to form a product space of the joint imaging modalities. To make pattern analysis robust, we propose a supervised dimensionality reduction to represent this high-dimensional data in terms of a few features, specifically selected to best explain factors relevant to dementia. Further, the extracted imaging features are used in conjuntion with APOE genotype and/or CSF biomarkers for assessing the risk of conversion of an MCI individual to AD. Fig. 1 summarizes our feature selection and classification framework.

Anatomical Shape Variations—Deformation Momenta: We follow the now well-established framework of large deformation diffeomorphic

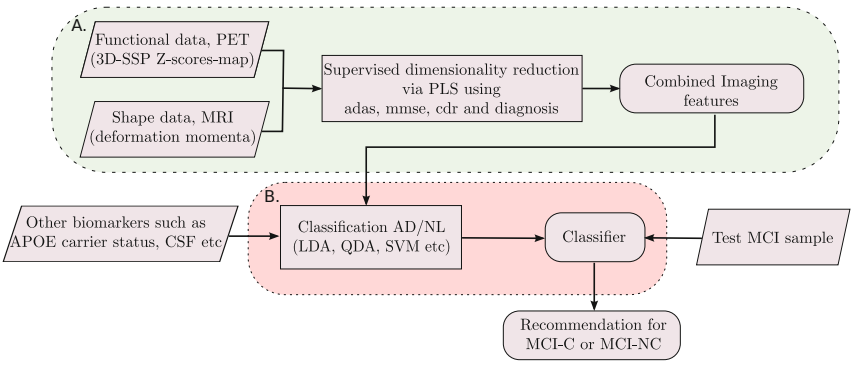

Fig. 1. MCI-C/MCI-NC prediction framework. Block A: Feature extraction process from high-dimensional imaging data. Block B: Classification.

transformations (LDDMM [5]) for capturing structural variation. A convenient and natural machinery for generating diffeomorphic transformations is by the integration of ordinary differential equations (ODE) on underlying coordinate space, Ω defined via the smooth time-indexed velocity vector fields $v(t, y) : (t \in [0, 1], y \in \Omega) \rightarrow \mathbb{R}^3$. The function $\phi^v(t, x)$ given by the solution of the ODE $\frac{dy}{dt} = v(t, y)$ with the initial condition $y(0) = x$ defines a diffeomorphism of Ω.

Following [6], we use a group-wise approach and build the mean population-based atlas, $\bar{I}$. We quantify the structural variability of the individual by registering the atlas to each image via estimating geodesic diffeomorphic transformations. Given a collection of anatomical images $\{I^i, i = 1, \cdots, N\}$, the minimum mean squared energy atlas construction problem is that of jointly estimating an image $\bar{I}$ and N individual deformations:

$$\bar{I} = \arg\min_{I,\phi_i} \frac{1}{N} \sum_{i=1}^{N} \int_{\Omega} ||I \circ \phi_i^{-1} - I^i||^2 dx + d(id, \phi_i)^2, \tag{1}$$

where d is the Riemannian metric defined on the space of diffeomorphisms and id is the identity diffeomorphism. For each of the N LDDMM image matching problems the geodesic evolution are given in terms of deformation momenta, $\alpha^i(t)$, by:

$$v^i(t) + K \star \nabla I_t \alpha^i(t) = 0, \;\; \partial_t \alpha^i(t) + \nabla \cdot (\alpha^i(t) v^i(t)) = 0 \text{ and } \partial_t I_t + \nabla I_t \cdot v^i(t) = 0$$

where K is the kernel associated with the metric d. The second equation is the conservation of momenta while the third is the infinitesimal action of the velocity field, v on the image. This results in the estimate of N geodesics emanating from the atlas towards each image. The geodesic equations are completely determined via the scalar initial momenta, $\alpha^i(0)$ in the atlas space corresponding to each individual image deformation direction. The LDDMM image matching problem is solved using the iterative backward-integration based gradient descent algorithm. The gradient of the energy functional in (1) is expressed in terms

of time-dependent Lagrangian multiplier or adjoint variables over the path of geodesics resulting in a set of adjoint equations (details in [7]).

FDG-PET Metabolism Activity—SSP: As the disease advances the progressive neurodegeneration is accompanied by reduced neuronal metabolism and increased synaptic dysfunction. This results in decreased uptake of [^{18}F]-fluorodeoxyglucose (FDG) measured by Positron Emission Tomography (PET) functional imaging. ADNI FDG-PET images are co-registered to the talaraich atlas space using Neurostat [8]. Peak pixel values are selected and 3D-stereotactic surface projection (3D-SSP) maps of glucose metabolism are computed relative to pons. Corresponding statistical maps of Z-scores, $p_i (i = 1, \cdots, N)$, are generated in comparison to cognitively normal control subjects ($\mu_{\text{age}} = 69.6 \pm 7.7$).

Combining Structure & Function: The shape space represented by the space of deformation momenta, $\mathcal{S}$, and the space of neuronal metabolic activity represented by 3D-SSP, $\mathcal{P}$, are both high-dimensional spaces. Since the anatomical shape and metabolic activity are two separate measures obtained from independent imaging modalities, we combine the two spaces to form a product space that defines the combined space of imaging modalities, $\mathcal{M}$ such that: $M = \mathcal{S} \times \mathcal{P}$. Inner product between a pair $m_i = (\alpha_i, p_i) \in \mathcal{M}$ and $m_j = (\alpha_j, p_j) \in \mathcal{M}$ is defined via a their convex combination as: $\langle m_i, m_j \rangle_{\mathcal{M}} = \eta \langle \alpha_i, \alpha_j \rangle_{\mathcal{S}} + (1 - \eta) \langle p_i, p_j \rangle_{\mathcal{P}}$. The factor, η is interpretable as a relative weight when both the modalaties are normalized to have unit variance.

Supervised Dimensionality Reduction via Partial Least Squares: The structural and functional information extracted from two imaging modalities results in a feature space with much higher dimension than the population size. Although classifiers utilizing kernel approaches such as Support Vector Machines (SVM) could work in the high-dimensional imaging feature space, for Linear Discriminant Analysis (LDA), dimensionality reduction has to be performed. We adopt a well known methodology for regression called Partial Least Squares (PLS) [9][10]. The Partial Least Squares can be interpreted as a supervised dimensionality reduction technique based on latent decomposition model. We adapt the PLS methodology for the purpose of extracting relevant features from the combination of shape and 3D-SSP data supervised by the clinical scores such as MMSE, ADAS, CDR and clinical cognitive status that are treated as global measures of dementia. We find directions $\hat{m}$ in the combined product space of imaging modalities, $\mathcal{M}$, and directions $\hat{y}$ in the clinical response space, $\mathcal{Y}$, that explain their association in the sense of their common variance. The projections of shape and pet data along the directions, $\hat{m}_i$ are treated as the features for the classifier. The PLS problem is given by:

$$\max \ \text{cov}(\langle \hat{m}, m^i \rangle, \langle \hat{y}, y^i \rangle) \text{ subject to } \|\hat{m}\| = 1 \ , \ \|\hat{y}\| = 1 \tag{2}$$

The subsequent directions are found by removing the component extracted (deflating the data) both in space, $\mathcal{M}$ and the clinical response space, $\mathcal{Y}$ as:

$$m^i \leftarrow m^i - \langle \hat{m}, m^i \rangle_{\mathcal{M}} \hat{m} \text{ and } y^i \leftarrow y^i - \langle \hat{y}, y^i \rangle_{\mathcal{Y}} \hat{y}$$

The solution to this covariance maximization problem is the Singular Value Decomposition (SVD) of the cross covariance matrix. The corresponding direction vectors $\hat{m}$'s and $\hat{y}$'s are the respective left and right singular vectors. The maximum number of possible latent vectors are limited by the inherent dimensionality of the two spaces, i.e., by $\min(\dim(\mathcal{M}), \dim(\mathcal{Y}))$.

Note that the efficient implementations of solution to the PLS via SVD uses the Gram matrix of inner products of the data. If we denote the Gram matrix of momenta by G_S and that of 3D-SSP by G_P, the fused Gram matrix for the product space weighted by η can be written as: $G_M = \eta G_S + (1 - \eta)G_P$. The projection scores, thus obtained by PLS, have combined information of anatomical shape and glucose metabolic activity that is used as features together with low-dimensional modalities such as genetic biomarkers of APOE carrier status and/or CSF biomarker available from spinal tap tests.

APOE Carrier Status—Genetic Biomarker: A confirmed risk factor for Alzheimer's disease is the status of apolipoprotein E (APOE) gene in an individual. APOE exhibit polymorphisms with three major isomorphisms or alleles: APOE $\varepsilon 2$, APOE $\varepsilon 3$ and APOE $\varepsilon 4$. Majority of the population with late-onset of AD is found to be dominant in APOE $\varepsilon 4$ allele. APOE carrier status is computed based on the allele copy inherited from parents in an individual. We consider the binary status for APOE genetic risk based on whether the individual has at least one copy of allele $\varepsilon 4$ and treat those subjects as APOE-carrier.

Prediction of Conversion to AD: Distinguishing the probable convertors from the population of MCI is a binary classification problem. While there are several ways to look at this problem, we present here a formulation of the classifier supervised by the AD group and healthy control group (NL). In other words, the classifier is trained on the AD and NL but is used as a "recommender" for the test MCI subject. Based on the classification score obtained on the MCI subject, the prediction of the classifier is interpreted. We denote the test MCI subject as "AD-like" when the classifier recommends AD and treated as predicted MCI-C otherwise termed as "Stable-MCI" or predicted MCI-NC. The classifier accuracy is assessed by comparing the predicted MCI-C or MCI-NC status with the conversion status from the follow-up study for that test MCI subject. The proposed methodology is evaluated using the LDA, its quadratic variant–Quadratic Discriminant Analysis (QDA), and SVM as binary classifiers.

3 Results and Discussion

Data Preprocessing: All the baseline and screening T1 weighted, bias-field-corrected and N3 scaled structural Magnetic Resonance Images were downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Preprocessing the MRI involved skull stripping and registration to talairach coordinates as a part of the ADNI preprocessing pipeline. Tissue-wise intensity normalization for white matter, gray matter, and cerebrospinal fluid was performed using the

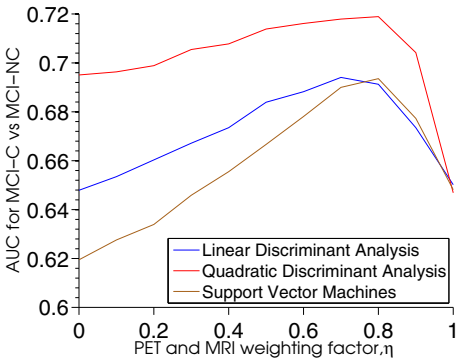

Fig. 2. Shape and PET weighting factor, η for different classifiers based on AUC

Table 1. ADNI data details

Diagnosis	54 Stable NL controls, 127 MCI, 61 AD
Education	$\mu = 15.27$ and $\sigma = 3.23$
Age	$\mu = 75.56$ and $\sigma = 6.65$
Gender	98 Females and 144 Males
Handedness	229 Right and 13 Left
APOE positive	13 NL's, 70 MCI's, 41 AD's
Follow-up	From baseline upto 48 months
MCI-C/NC status	54 out of 127 MCI converted to AD

Table 2. MCI-C vs. MCI-NC classification results for η_{OPT}

	AUC	Acc (%)	Sen(%)	Spec(%)	η
QDA	0.72	66.14	64.81	67.12	0.7
LDA	0.69	63.78	74.07	56.16	0.8
SVM	0.69	64.57	72.20	58.90	0.8

expectation maximization based segmentation followed by the piecewise polynomial histogram matching algorithm. The FDG-PET data was processed to get 3D-SSP as detailed in Section 2. The corresponding clinical test score, the CSF-biomarker data and the APOE genotype information were also retrieved. The baseline subjects that had all the clinical, APOE genotyping, FDG-PET imaging and MRI imaging data from the ADNI database comprised of a total of 242 individuals. Table 1 reports the details about the subject demographics, diagnosis, apoe carrier status and future conversion status.

To extract the anatomical shape features, the unbiased atlas, $\bar{I}$ is constructed from the preprocessed baseline MR brain images on the Graphical Processing Unit (GPU) [6]. The geodesics emanating from this estimated atlas towards each subject are estimated by warping $\bar{I}$ to each of the baseline subjects to give initial deformation momenta, $\alpha^i(0)(i = 1, \cdots, N)$ [7]. The corresponding 3D-stereotactic surface projection (3D-SSP) maps, $p_i(i = 0, \cdots, N)$, of glucose metabolism from FDG-PET are computed using Neurostat [8] to give Z-score maps. The supervised PLS dimensionality reduction is applied on combined imaging data of AD and NL subjects. Since the response is 4-D, the resulting feature space is 4-D and is represented by $\hat{m}_i$ $(i = 1, \cdots, 4)$. The imaging features are then combined together with low-dimensional biomarkers such as APOE carrier status to train the binary classifier for AD/NL classification.

The independent test MCI subject is projected into the shape and PET feature space defined by the training AD and NL group in terms of $\hat{m}_i$'s. The imaging

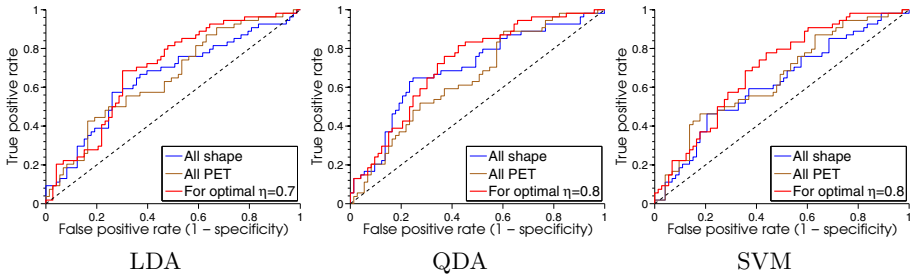

Fig. 3. Receiver operating characterstic curves (ROC) for MCI-C/MCI-NC classification with only shape information, only PET information and optimal combination of shape and PET as per η_{OPT}

features for the test MCI subject are combined with its APOE carrier status. The trained AD/NL-classifier's prediction on MCI baseline features is then used as a recommendation for future conversion to AD. Note that, for the test MCI subject, no clinical scores such as ADAS, MMSE, CDR or diagnostic information in any form is used during feature extraction from imaging data or classifier prediction. The accuracy of prediction is evaluated by comparing against the actual conversion status using the follow-up diagnosis data.

Fig. 2 shows area under the receiver operating characterstic curve (AUC) as a function of the weighting factor, η, for the three separate classifiers discriminating MCI-C vs MCI-NC. The accuracy of prediction of MCI to AD conversion and the associated η is given in Table 2. The reported numbers correspond to optimal η, based on AUC. QDA performed the best with accuracy of 66% and AUC of 0.72 at $\eta = 0.8$. Also, the optimal combination of PET and shape performed much better as compared to only using PET or anatomical shape information irrespective of the choice of classifier used (Fig. 3). Besides APOE carrier status, the above analysis was also done after adding log transformed CSF-biomarkers: $A\beta$-42 and ptau_{181} concentration, which reduced the study sample-size to only: 29 NL, 36 AD and 59 MCI. With CSF-biomarkers, a slight increase in accuracy was observed for QDA: accuracy=68% and AUC= 0.72 ($\eta = 0.8$).

The log Jacobians of the deformation, overlayed on atlas image $\bar{I}$, resulting from evolving $\bar{I}$ along the geodesic represented by the classifier weights are shown in Fig. 4. The selected slices from this 3D overlay shown here capture relevant regions of the neuro-anatomical structures, such as hippocampus, pertinent to cognitive impairment in Alzheimer's and related dementia. Similarly, the PET classifier weights are translated back in the Z-score space of 3D-SSP (Fig. 5).

The major contribution of this article is the ability to extract, in order of relevance, the disease-characterizing patterns from multiple imaging modalities. The presented framework has broad applicability to data analysis studies involving heterogeneous data sources, both in terms of modalities and dimensions. We observed that the shape component dominated the model with up to 80% contribution compared to only 20% contribution from the PET component, irrespective of the classifier used. The spatial patterns of anatomical shape changes

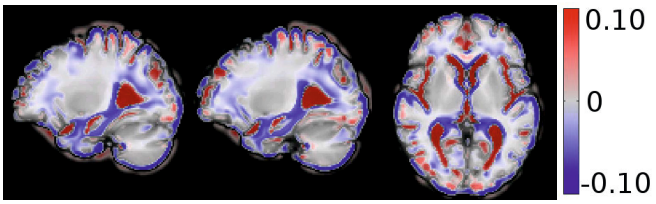

Fig. 4. Shape: Discriminating regions obtained from classifier weights for prediction of MCI conversion to AD. Log of Jacobians overlayed on atlas. Red denotes regions of local expansion and blue denotes regions of local contraction.

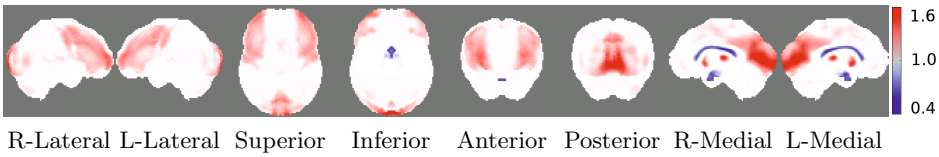

Fig. 5. FDG-PET: Discriminating regions obtained from classifier weights for prediction of MCI conversion to AD in 3D-SSP Z-score space.

were primarily the expansion of lateral ventricles and CSF, together with the shrinkage of the cortical surface.Another critical observation was the clearly evident shrinkage of the hippocampus and cortical and sub-cortical gray matter along the discriminating directions. Such patterns of atrophy are well known to characterize the disease progression in AD and related dementia.

Acknowledgements: Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (NIH Grant U01 AG024904). The research in this paper was supported by NIH grant 5R01EB007688, the University of California, San Francisco (NIH grant P41 RR023953), NSF grant CNS-0751152), and NSF CAREER Grant 1054057.

References

1. Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.N., Trojanowski, J.Q.: Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiology of Aging 32(12), 2322.e19–2322.e27 (2011)
2. Lemoine, B., Rayburn, S., Benton, R.: Data Fusion and Feature Selection for Alzheimer's Diagnosis. In: Yao, Y., Sun, R., Poggio, T., Liu, J., Zhong, N., Huang, J. (eds.) BI 2010. LNCS, vol. 6334, pp. 320–327. Springer, Heidelberg (2010)
3. Weiner, M.W., et al.: The Alzheimers Disease Neuroimaging Initiative: A review of papers published since its inception. Alzheimer's and Dementia, S1–S68 (2012)
4. Kohannim, O., et al.: Boosting power for clinical trials using classifiers based on multiple biomarkers. Neurobiology of Aging 31(8), 1429–1442 (2010)

5. Younes, L., Arrate, F., Miller, M.: Evolutions equations in computational anatomy. NeuroImage 45(1S1), 40–50 (2009)
6. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 23, 151–160 (2004)
7. Vialard, F.X., et al.: Diffeomorphic 3D image registration via geodesic shooting using an efficient adjoint calculation. IJCV, 1–13 (2011)
8. Minoshima, S., Frey, K.A., Koeppe, R.A., Foster, N.L., Kuhl, D.E.: A diagnostic approach in Alzheimers disease using three-dimensional stereotactic surface projections of Fluorine-18-FDG PET. J. of Nuclear Medicine 36(7), 1238–1248 (1995)
9. Bookstein, F.L.: Partial Least Squares: A dose-response model for measurement in the behavioral and brain sciences. Psycoloquy 5(23) (1994) (revised)
10. Singh, N., Fletcher, P.T., Preston, J.S., Ha, L., King, R., Marron, J.S., Wiener, M., Joshi, S.: Multivariate Statistical Analysis of Deformation Momenta Relating Anatomical Shape to Neuropsychological Measures. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 529–537. Springer, Heidelberg (2010)