

Multi-Phase Feature Representation Learning for Neurodegenerative Disease Diagnosis

Siqi Liu¹, Sidong Liu^{1,2}, Weidong Cai^{1,2}, Sonia Pujol², Ron Kikinis², and
David Dagan Feng^{1,3}

¹ BMIT Research Group, School of Information Technologies, University of Sydney,
NSW, Australia

² Surgical Planning Laboratory, Brigham and Womens Hospital, Harvard Medical
School, Boston, MA, USA

³ Med-X Research Institute, Shanghai Jiaotong University, Shanghai, China

Abstract. Feature learning with high dimensional neuroimaging features has been explored for the applications on neurodegenerative diseases. Low-dimensional biomarkers, such as mental status test scores, gene variations, and protein changes in plasma and cerebrospinal fluid level, are essential in clinical diagnosis of neurological disorders, because they could be simple and effective for the clinicians to assess the disorder’s progression and severity. Rather than only using the low-dimensional biomarkers as inputs for decision making systems, we believe that such low-dimensional biomarkers can be used for enhancing the feature learning pipeline. In this study, we proposed a novel feature representation learning framework, Multi-Phase Feature Representation (MPFR), with low-dimensional biomarkers embedded. MPFR learns high-level neuroimaging features by extracting the associations between the low-dimensional biomarkers and the high-dimensional neuroimaging features with a deep neural network. We validated the proposed framework using the Mini-Mental-State-Examination (MMSE) scores as a low-dimensional biomarker and multi-modal neuroimaging data as the high-dimensional neuroimaging features from the ADNI baseline cohort. The proposed approach outperformed the original neural network in both binary and ternary Alzheimer’s disease classification tasks.

Keywords: Classification, Deep Learning, Neuroimaging

1 Introduction

To aid the clinical decision making, feature learning methods have been applied on learning the correlations between the high dimensional features extracted from neuroimaging data recently [1,2,3]. Low-dimensional biomarkers, such as the Mini-Mental-State-Examination (MMSE) and the cerebrospinal fluid measurements (CSF), are known as supportive diagnostic tools for clinic diagnosis of brain diseases, accompanied with the high-dimensional neuroimaging biomarkers, such as Magnetic Resonance Image (MRI) and Positron Emission Tomography (PET). However unlike the application of high-dimensional neuroimaging

biomarkers [4,5,6,7,8,9,10,11,12], they have not been sufficiently explored in current machine learning based feature representation methods. There have been studies using the low-dimensional biomarkers, such as CSF, as input features for machine learner [13]. However, low-dimensional biomarkers might not be equally sensitive to the cognitive impairment as high-dimensional biomarkers and can be easily overwhelmed when they are fed into the machine learner together. Whereas, other studies attempted to predict low-dimensional clinic assessments, such as Alzheimer’s Disease Assessment Scale-Cognitive Subscale (ADAS-Cog), by building the joint regression models with inputs selected from both high-dimensional and low-dimensional biomarkers [14,15]. However, these methods require the subjects to have paired clinical scores and imaging data, which may lead to shrinkage of training dataset, because of the extra expense on data collection.

In an attempt to solve the above-mentioned problems, we proposed a Multi-Phase Feature Representation (MPFR) framework that have low-dimensional biomarkers embedded in the feature representation learning rather than directly using them as features. MPFR learns the features by assessing the associations between the low-dimensional biomarkers and the high-dimensional biomarkers based on a deep neural network consisting of stacked auto-encoders (SAE), linear regression and softmax regression. The feature representation network is optimised to estimate the low-dimensional biomarkers before it is finally used for classification. The prominent advantage of the proposed framework is the separation of different training phases. Thus, it does not require every training instance to be attached with a diagnostic label nor low-dimensional biomarkers. In addition, the low-dimensional biomarkers will not be overwhelmed by the high-dimensional biomarkers. We validated the proposed framework with the performance of Alzheimer’s Disease (AD) diagnosis with 3 groups from the ADNI cohort and selected the MMSE scores as the low-dimensional biomarkers for feature learning. We modelled the AD diagnosis as a classification problem with 3 classes, including Normal Control (NC), Mild Cognitive Impairment (MCI) and AD. These learnt features revealed the associations between MMSE and the neuroimaging patterns and outperformed the features extracted by the state-of-the-art methods as well as conventional deep-learning-based methods without MPFR in both binary and ternary classification tasks.

2 Multi-Phase Feature Representation

There are three phases of MPFR framework, as shown in Fig. 1. Phase A depicts the unsupervised layer-wised pre-training of the feature representation network which learns a manifold to reconstruct the features at the previous layer. The parameters of the auto-encoders learned by Phase A are unfolded and stacked in Phase B. In Phase B, the feature representation network is enhanced by training to output the low-dimensional features. The outputs of Phase B are the low-dimensional biomarkers estimated with the extracted high-level features. In Phase C, after replacing the output layer with softmax regression, the entire

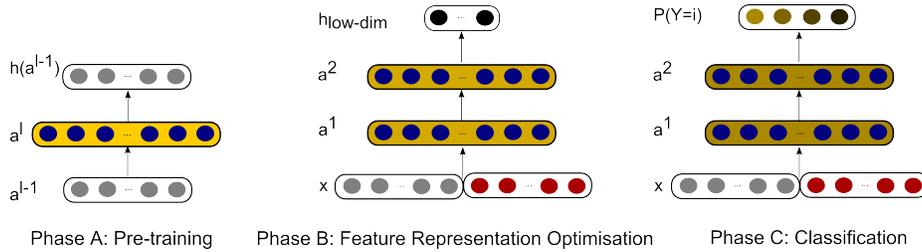


Fig. 1. The illustration of the 3 phases of training the proposed MPFR framework. The grey neurons are the high-dimensional biomarker inputs and the blue neurons are the hidden layer neurons. The neurons with different colors at the bottom layer in Phase B indicate inputs from different neuroimaging modalities, such as MRI and PET, and the black neurons are the estimated low-dimensional biomarker outputs. The darker brown nodes in Phase C indicate more severe stages along the disease progression.

structure is finally fine-tuned for the purpose of classification. The learnt features in Phase C are expected to be more sensitive to the training labels since the hidden layers could learn high-level disorder-specific features.

2.1 Phase A: Pre-Training Stacked Auto-encoders (SAE) with Neuroimaging Biomarkers

The biomarkers extracted from brain images, such as MRI and PET, are initially used as the inputs for the feature representation learner. They are fed into a multi-layered neural network with non-linear activation function. Each hidden layer of the neural net is seen as a higher level of representation of the previous layer. To obtain high-level features that can represent the original inputs, we used stacked autoencoders (SAE) to form the hidden layers of the neural network. The feature representations can be computed by feed-forwarding the activation signals with sigmoid function $S(x) = \frac{1}{1+e^{-x}}$. The parameters of each auto-encoder are optimised to reconstruct the previous input a_{l-1} . The optimisation criterion is to minimise the representation loss:

$$L(W, b, x, z) = \min_{W, b} \|h(W, b, x) - z\|_2^2 + \lambda \|W\|_2^2 \quad (1)$$

where $h(W, b, x)$ is the reconstruction yielded from the auto-encoder; x is the input vector of the auto-encoder; W and b are the weight and bias parameters to be trained; the second term is the weight decay to control over-fitting. Gradient descent based algorithms, such as L-BFGS algorithm [16], can be applied to train auto-encoders, following the greedy layer-wised training strategy by training each hidden layer at once (Phase A in Fig. 1) [17]. To fully extract the synergy between different image modalities, we zero-masked a small proportion of feature parameters from one single modality randomly when training the first hidden layer [18].

2.2 Phase B: Feature Representation Optimisation with Low-Dimensional Biomarkers

Before using the pre-trained network for the classification, the hidden layers are firstly fine-tuned to estimate continuous low-dimensional biomarkers. The network is also expected to extract the correlations between image biomarkers and the low-dimensional biomarkers. For example, if MMSE is chosen to be estimated, the features provided by the fine-tuned feature representation network can be sensitive to the cognitive status of subjects as well as the risk of being diagnosed as AD. However, by observing (a) and (b) in Fig. 2, it is also obvious that neuroimaging data are expected to be more sensitive to the cognitive impairment.

Linear regression is used to estimate the low-dimensional biomarkers. One output layer with linear activation filter is stacked on the top of the pre-trained SAE (Phase B in Fig. 1). The prediction can be demonstrated as

$$h(W, a^{(l)}) = W a^{(l)} + b \quad (2)$$

where $h(W, a^{(l)})$ is the vector of the estimated low-dimensional biomarkers; $a^{(l)}$ is the l -th layer of high-level features obtained from the pre-trained SAE; W and b are the parameters that can be optimised by jointly propagating the error gradients. Since Phase B is designed for augmenting the feature learning network, early stopping is applied in Phase B training to control the over-fitting on the low-dimensional biomarkers in the training set.

2.3 Phase C: Classification with AD Labels

The trained linear regression layer is replaced with a softmax regression output layer (Phase C in Fig. 1). The softmax layer uses a different activation function, which might introduce non-linearity different from the sigmoid function in SAE:

$$P(Y = i|x) = \frac{e^{W^i a + b^i}}{\sum_j e^{W^j a + b^j}} \quad (3)$$

where a is the feature vector obtained by the fine-tuned SAE. Y is assigned to possible stages of a particular disease, such as NC, MCI and AD of AD progression. The prediction $P(Y = i|x)$ with the highest probability is chosen as the final decision. The network is then fine-tuned by back-propagating the classification loss with the pre-labelled training subjects as a supervised classification neural network.

3 Experiments

3.1 Data Acquisition and Feature Extraction

The dataset used in this study was acquired from the ADNI database [20]. We selected 331 subjects from the ADNI baseline cohort, including 77 NC-, 169 MCI-

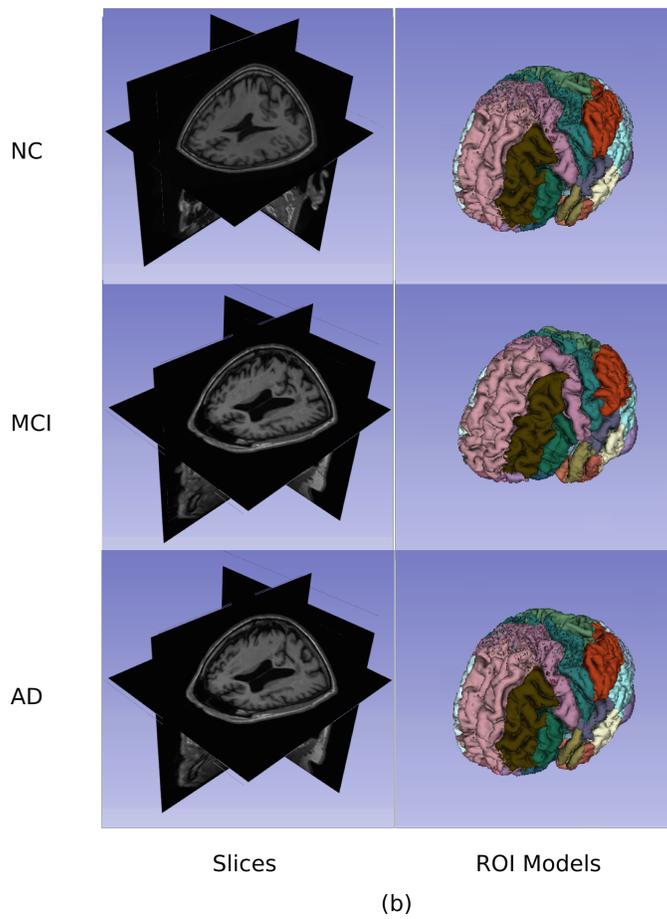
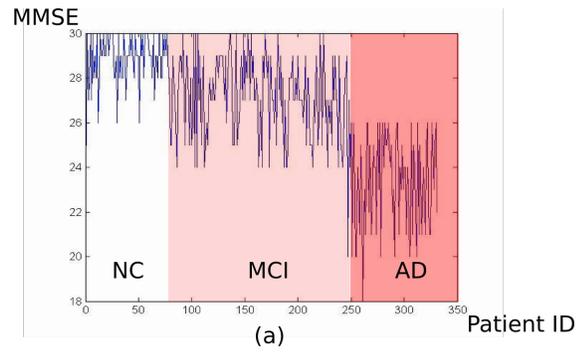


Fig. 2. The comparison between the variations of low-dimensional biomarkers and neuroimaging data. (a) is the plot of MMSE scores. (b) are slices and MAPER whole brain mask models from ADNI baseline cohort, generated using 3D Slicer 4.3.1 [19]. The neuroimaging data contains more information than MMSE. Some sophisticated features, such as ventricle sizes and atrophy, cannot be captured by MMSE, but can be shown in neuroimaging data.

and 85 AD-subjects. For each subject, an FDG-PET image and a T1-weighted volume acquired on a 1.5 Tesla MR scanner were retrieved. All the 3D MRI and PET data were registered following the ADNI image correction protocols [20,21]. The PET images were aligned to the corresponding MRI image using FSL FLIRT [22]. The MRI images were then non-linearly registered to the ICBM 152 template [23] with 83 functional regions using the Image Registration Toolkit (IRTK) [24]. The outputted registration coefficients of MRI images were used to align PET images into the ICBM 152 template. Finally, all registered MRI and PET image were mapped to brain functional regions using the multi-atlas propagation with MAPER approach [25]. Three types of features were extracted from each of the brain regions, including the grey matter volume [25,26,27], solidity [10,28,29] features from MRI data, and the cerebral metabolic rate of glucose consumption (CMRGl_c) from PET data [7,30]. Elastic Net [31,32,33] was applied to choose the most predictive dimensions of features. 85 feature parameters were eventually selected for each subject evaluated in this study.

3.2 Performance Evaluation

The performance on AD diagnosis of the proposed MPFR framework was compared to the state-of-the-art work-flow with multi-kernel SVM (MK SVM) [13] and the conventional deep learning method without embedding MPFR. We chose MMSE scores of patients in the training set to evaluate MPFR in this study, because the variation between the NC/MCI groups and the AD group is obvious (Fig. 2-a). All of the experiments were evaluated with the same features extracted from MRI images and PET images.

MK SVM was implemented using LIBSVM library with the radial basis function (RBF) kernel [34]. We applied the 'one against all' approach to allow MK SVM to perform the trinary classification problem [35]. The proposed deep learning framework was implemented on Matlab 2013b. The approximated random search strategy was applied to choose the hyper-parameters [36].

We used 10-fold cross-validation to evaluate different learning structures. For each fold of cross validation, about 90% subjects were collected for training (including the pre-training, feature optimisation and classification training) and the rest subjects were used for testing. In the experiments of the MPFR framework, the MMSE scores of the testing set were neglected.

3.3 Results

The examination of the high-level features is displayed in Fig 3. Comparing the features learnt with 2 different phases, it is clear that the Phase A only extracted features that represented the neuroimaging data. After Phase B, some dimensions of the learnt features showed variations along the progression of AD.

In Phase B, by comparing the different AD stages, most of high-level features at layer 1 tend to vary according to the severity of AD. Whereas only few dimensions of high-level features at layer 2 showed remarkable variance. It is

reasonable to assume that after the Phase B of MPFR, the higher layers may tend to extract more abstract patterns by flattening the local variations.

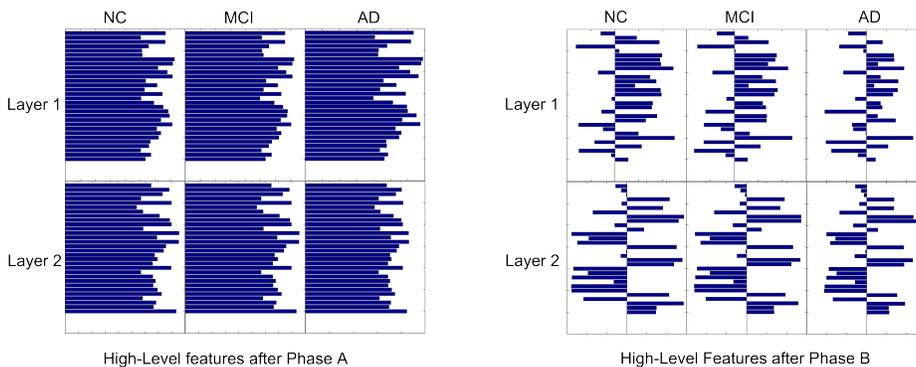


Fig. 3. The examination of the high-level features extracted by Phase A and Phase B (applied MMSE scores) of MPFR. These features were obtained from a network with 2 hidden layers.

Table 1. The performance (%) of the AD binary classification between NC and AD. The first two columns are precisions on each class. The last three columns depict the overall performance (accuracy, sensitivity and specificity).

	NC	AD	ACC	SEN	SPE
MKSVM	89.10 ± 1.26	90.13 ± 1.02	89.60 ± 0.91	89.09 ± 1.39	90.10 ± 1.12
SAE	89.39 ± 10.73	85.67 ± 14.94	88.20 ± 7.68	87.66 ± 9.50	87.50 ± 15.04
Proposed	91.17 ± 8.54	88.35 ± 8.17	90.11 ± 3.06	84.45 ± 10.51	93.89 ± 6.31

The performance of the binary classification between NC and AD is displayed in Table 1. The proposed method outperformed the conventional SAE in both precisions of classifying NC and AD. It also achieved the highest overall accuracy (90.11%) and specificity (93.98%) among the other two methods. The higher sensitivity achieved by MKSVM (89.09%) can be attributed to the limitation on the quantity of the available training data. Parametric models such as SAE and the proposed method tend to have higher standard deviation than non-parametric models like SVM. However, it is notable that the proposed method performed lower standard deviation than SAE. Table 2 shows the performance of trinary classification, including NC, MCI and AD, which was designed for evaluating the effects on the early detection of AD. The proposed method outperformed the conventional SAE and MKSVM in the overall accuracy (59.19%) and the overall sensitivity (50.98%). It is reasonable to convince that the enhancements of classification performance, comparing to the conventional SAE, were primarily benefited from the robust high-level features learnt with MPFR.

Table 2. The performance (%) of the trinary AD classification between NC, MCI and AD. The first three columns are precisions on each class. The last three columns depict the overall performance (accuracy, sensitivity and specificity).

	NC	MCI	AD	ACC	SEN	SPE
MKSVM	47.53 ± 2.59	57.29 ± 0.90	49.74 ± 2.33	53.2 ± 1.27	44.59 ± 2.89	85.00 ± 1.92
SAE	46.97 ± 21.71	61.87 ± 12.26	60.78 ± 15.88	58.57 ± 9.21	49.16 ± 16.60	83.53 ± 6.56
Proposed	49.00 ± 22.42	61.29 ± 12.94	61.52 ± 13.87	59.19 ± 9.20	50.98 ± 16.08	84.36 ± 6.71

4 Conclusions

In this study, we presented a novel framework, the Multi-Phase Feature Representation (MPFR), for the feature representation of neuroimaging data. It differs between the conventional deep learning architecture by learning features to output low-dimensional biomarkers before the deep network is fine-tuned with classification labels. The preliminary results showed that MPFR framework outperformed SAE as well as the state-of-the-art classification method (Multi-Kernel SVM) and had a great potential to embed other low-dimensional biomarkers in feature representation learning without constraining the size of available training data.

5 Acknowledgements

This work was supported in part by the ARC, AADRF, NA-MIC (NIH U54EB005149), and NAC (NIH P41EB015902).

References

1. Suk, H.I., Shen, D.: Deep Learning-Based Feature Representation for AD/MCI Classification. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013. Springer (2013) 583–590
2. Liu, S.Q., Liu, S., Cai, W., Pujol, S., Kikinis, R., Feng, D.: Early Diagnosis of Alzheimer’s Disease with Deep Learning. In: Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on, IEEE (2014) 1015–1018
3. Liu, S.Q., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., Fulham, M.J., Feng, D.: High-level Feature based PET Image Retrieval with Deep Learning Architecture. Journal of Nuclear Medicine **55**(Supplement 1) (2014) 2018
4. Fischl, B., Dale, A.M.: Measuring the Thickness of the Human Cerebral Cortex from Magnetic Resonance Images. Proceedings of the National Academy of Sciences **97**(20) (2000) 11050 – 11055
5. Schaer, M., Cuadra, M.B., Tamarit, L., Lazeyras, F., Eliez, S., Thiran, J.P.: A Surface-Based Approach to Quantify Local Cortical Gyrification. IEEE Transactions on Medical Imaging **27**(2) (2008) 161–170
6. Liu, S., Cai, W., Wen, L., Eberl, S., Fulham, M.J., Feng, D.: A Robust Volumetric Feature Extraction Approach for 3D Neuroimaging Retrieval. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, IEEE (2010) 5657–5660

7. Cai, W., Liu, S., Wen, L., Eberl, S., Fulham, M., Feng, D.: 3D Neurological Image Retrieval with Localized Pathology-Centric CMRGlc Patterns. In: Image Processing (ICIP), 2010 17th IEEE International Conference on. (2010) 3201–3204
8. Liu, S., Cai, W., Wen, L., Eberl, S., Fulham, M.J., Feng, D.: Localized Functional Neuroimaging Retrieval using 3D Discrete Curvelet Transform. In: Biomedical Imaging: From Nano to Macro (ISBI), 2011 IEEE International Symposium on. IEEE (2011) 1877–1880
9. Cash, D.M., Melbourne, A., Modat, M., Cardoso, M.J., Clarkson, M.J., Fox, N.C., Ourselin, S. In: Cortical Folding Analysis on Patients with Alzheimers Disease and Mild Cognitive Impairment. Springer (2012) 289–296
10. Liu, S., Song, Y., Cai, W., Pujol, S., Kikinis, R., Wang, X., Feng, D. In: Multifold Bayesian Kernelization in Alzheimers Diagnosis. Springer (2013) 303–310
11. Liu, S., Cai, W., Wen, L., Feng, D.: Neuroimaging Biomarker Based Prediction of Alzheimer’S Disease Severity with Optimized Graph Construction. In: Biomedical Imaging (ISBI), IEEE 10th International Symposium on, IEEE (2013) 1336–1339
12. Cai, W., Liu, S., Song, Y., Pujol, S., Kikinis, R., Feng, D.: A 3D Difference of Gaussian based Lesion Detector for Brain PET. In: Biomedical Imaging: From Nano to Macro (ISBI), 2014 IEEE International Symposium on. IEEE (2014) 677–680
13. Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D.: Multimodal Classification of Alzheimer’s Disease and Mild Cognitive Impairment. *NeuroImage* **55**(3) (2011) 856 – 867
14. Wang, Y., Fan, Y., Bhatt, P., Davatzikos, C.: High-Dimensional Pattern Regression Using Machine Learning: From Medical Images to Continuous Clinical Variables. *NeuroImage* **50**(4) (2010) 1519 – 1535
15. Zhang, D., Shen, D.: Multi-Modal Multi-Task Learning for Joint Prediction of Multiple Regression and Classification Variables in Alzheimer’s Disease. *NeuroImage* **59**(2) (2012) 895 – 907
16. Ngiam, J., Coates, A., Lahiri, A., Prochnow, B., Ng, A., Le, Q.V.: On Optimization Methods for Deep Learning. In: Proceedings of the 28th International Conference on Machine Learning (ICML-11). (2011) 265–272
17. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., et al.: Greedy Layer-Wise Training of Deep Networks. *Advances in Neural Information Processing Systems* **19** (2007) 153
18. Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., Ng, A.: Multimodal Deep Learning. In: Proceedings of the 28th International Conference on Machine Learning (ICML-11). (2011) 689–696
19. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., et al.: 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* **30**(9) (2012) 1323–1341
20. Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., et al.: The Alzheimer’s Disease Neuroimaging Initiative (ADNI): MRI Methods. *Journal of Magnetic Resonance Imaging* **27**(4) (2008) 685–691
21. Jagust, W.J., Bandy, D., Chen, K., Foster, N.L., Landau, S.M., Mathis, C.A., Price, J.C., et al.: The Alzheimer’s Disease Neuroimaging Initiative Positron Emission Tomography Core. *Alzheimer’s & Dementia* **6**(3) (2010) 221–229
22. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage* **17**(2) (2002) 825–841

23. Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., et al.: 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) (2001) 1293–1322
24. Schnabel, J.A., Rueckert, D., Quist, M., Blackall, J.M., et al.: A Generic Framework for Non-Rigid Registration Based on Non-Uniform Multi-Level Free-Form Deformations. In Niessen, W.J., Viergever, M.A., eds.: *Medical Image Computing and Computer-Assisted Intervention MICCAI 2001*. Springer (2001) 573–581
25. Heckemann, R.A., Keihaninejad, S., Aljabar, P., Gray, K.R., Nielsen, C., Rueckert, D., Hajnal, J.V., Hammers, A.: Automatic Morphometry in Alzheimer’s Disease and Mild Cognitive Impairment. *Neuroimage* **56**(4) (2011) 2024–2037
26. Liu, S., Cai, W., Song, Y., Pujol, S., Kikinis, R., Wen, L., Feng, D.: Localized Sparse Code Gradient in Alzheimer’s Disease Staging. In: *Engineering in Medicine and Biology Society (EMBC), 2013 Annual International Conference of the IEEE. IEEE (2013) 5398–5401*
27. Che, H., Liu, S., Cai, W., Pujol, S., Kikinis, R., Feng, D.: Co-neighbor Multi-View Spectral Embedding for Medical Content-based Retrieval. In: *Biomedical Imaging: From Nano to Macro (ISBI), 2014 IEEE International Symposium on. IEEE (2014) 911–914*
28. Batchelor, P.G., Castellano Smith, A.D., Hill, D.L.G., Hawkes, D.J., Cox, T.C.S., Dean, A.: Measures of Folding Applied to the Development of the Human Fetal Brain. *IEEE Transactions on Medical Imaging* **21**(8) (2002) 953–965
29. Liu, S., Zhang, L., Cai, W., Song, Y., Wang, Z., Wen, L., Feng, D.: A Supervised Multiview Spectral Embedding Method for Neuroimaging Classification. In: *Image Processing (ICIP), 2013 IEEE International Conference on. IEEE (2013) 601–605*
30. Liu, S., Cai, W., Wen, L., Eberl, S., Fulham, M.J., Feng, D.: Generalized Regional Disorder-Sensitive-Weighting Scheme for 3D Neuroimaging Retrieval. In: *Engineering in Medicine and Biology Society (EMBC), 2011 Annual International Conference of the IEEE. IEEE (2011) 7009–7012*
31. Zou, H., Hastie, T.: Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **67**(2) (2005) 301–320
32. Liu, S., Cai, W., Wen, L., Feng, D.: Multi-Channel Brain Atrophy Pattern Analysis in Neuroimaging Retrieval. In: *Biomedical Imaging: From Nano to Macro (ISBI), 2013 IEEE International Symposium on. IEEE (2013) 206–209*
33. Liu, S., Cai, W., Wen, L., Feng, D.D., Pujol, S., et al.: Multi-Channel Neurodegenerative Pattern Analysis and Its Application in Alzheimer’s Disease Characterization. *Computerized Medical Imaging and Graphics* **38**(4) (2014) 436–444
34. Chang, C.C., Lin, C.J.: LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* **2**(3) (2011) 27:1–27:27
35. Liu, Y., Zheng, Y.: One-Against-All Multi-Class SVM Classification Using Reliability Measures. In: *Neural Networks (IJCNN 2005), IEEE International Joint Conference on. Volume 2. (2005) 849–854*
36. Bergstra, J., Bengio, Y.: Random Search for Hyper-Parameter Optimization. *The Journal of Machine Learning Research* **13** (2012) 281–305