Published in final edited form as:
Multimodal Brain Image Anal (2013). 2013 ; 8159: 159–169. doi:10.1007/978-3-319-02126-3_16.

A Graph-Based Integration of Multimodal Brain Imaging Data for the Detection of Early Mild Cognitive Impairment (E-MCI)

Dokyoon Kim[1], **Sungeun Kim**[2,3], **Shannon L. Risacher**[2], **Li Shen**[2,3], **Marylyn D. Ritchie**[1], **Michael W. Weiner**[4,5], **Andrew J. Saykin**[2,6,7,§], **Kwangsik Nho**[2,3,§], and **the Alzheimer's Disease Neuroimaging Initiative (ADNI)**[**]

[1]Center for Systems Genomics, Pennsylvania State University, San Francisco

[2]Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, San Francisco

[3]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, San Francisco

[4]Department of Radiology, Medicine and Psychiatry, University of California, San Francisco

[5]Department of Veterans Affairs Medical Center, San Francisco

[6]Department of Medical and Molecular Genetics, Indiana University School of Medicine, San Francisco

[7]Department of Neurology, Indiana University School of Medicine

Abstract

Alzheimer's disease (AD) is the most common cause of dementia in older adults. By the time an individual has been diagnosed with AD, it may be too late for potential disease modifying therapy to strongly influence outcome. Therefore, it is critical to develop better diagnostic tools that can recognize AD at early symptomatic and especially pre-symptomatic stages. Mild cognitive impairment (MCI), introduced to describe a prodromal stage of AD, is presently classified into early and late stages (E-MCI, L-MCI) based on severity. Using a graph-based semi-supervised learning (SSL) method to integrate multimodal brain imaging data and select valid imaging-based predictors for optimizing prediction accuracy, we developed a model to differentiate E-MCI from healthy controls (HC) for early detection of AD. Multimodal brain imaging scans (MRI and PET) of 174 E-MCI and 98 HC participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort were used in this analysis. Mean targeted region-of-interest (ROI) values extracted from structural MRI (voxel-based morphometry (VBM) and FreeSurfer V5) and PET (FDG and Florbetapir) scans were used as features. Our results show that the graph-based SSL classifiers outperformed support vector machines for this task and the best performance was obtained with 66.8% cross-validated AUC (area under the ROC curve) when FDG and FreeSurfer datasets were

§ Address correspondences to knho@iupui.edu or asaykin@iupui.edu.

** 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/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

integrated. Valid imaging-based phenotypes selected from our approach included ROI values extracted from temporal lobe, hippocampus, and amygdala. Employing a graph-based SSL approach with multimodal brain imaging data appears to have substantial potential for detecting E-MCI for early detection of prodromal AD warranting further investigation.

Keywords

Mild Cognitive Impairment; Multimodal Brain Imaging Data; Data Integration; Graph-based Semi-Supervised Learning; Alzheimer's Disease

1 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease in older adults and at this time, despite incidence rates doubling every 5 years after the age of 65, there is no effective disease modifying treatment for AD to date [1]. AD is predicted to affect 14 million Americans by the year 2050 (www.alz.org) and has become a national priority. The detection and diagnosis of AD at the earliest possible stage is of fundamental importance as early intervention could potentially delay progression to AD and achieve effective disease modification. One of main challenges is to identify and validate biomarkers of AD progression leading to an improved early diagnosis at early symptomatic and especially pre-symptomatic stages. To this end, the concept of mild cognitive impairment (MCI) was introduced [2]. MCI can be classified into early and late stages (E-MCI, L-MCI) based on severity. MCI is thought to be a precursor to the development of early AD, and subjects with late amnestic MCI have a highly elevated probability of developing AD with a conversion rate of approximately 15% per year [3, 4].

New approaches to the search for specific biomarkers to detect MCI/AD compared to healthy controls (HC) have been developed, with neuroimaging (MRI and PET) and cerebrospinal fluid (CSF) biochemical markers showing particular promise [5, 6]. However, in most studies, only patients with L-MCI and AD have been assessed [7]. In order to identify better diagnostic tools that can recognize AD at early symptomatic and especially pre-symptomatic stages, we developed a graph-based semi-supervised learning model to differentiate E-MCI from HC for early detection of AD using multimodal brain imaging scans (MRI and PET) of participants from the ongoing Alzheimer's Disease Neuroimaging Initiative (ADNI).

The semi-supervised learning (SSL) which recently emerged in the machine learning domain, employs a strategy halfway between supervised and unsupervised learning schemes to improve classification performance [8-11]. In particular, the graph-based SSL takes advantage of computational efficiency and representational ease for the biomedical data. Because of the graph structures, it is easy to integrate different types of data for better explaining clinical outcomes [12]. The learning time of graph-based SSL is nearly linear with the number of graph edges while the accuracy remains comparable to the kernel-based methods that suffer from the relative disadvantage of a longer learning time [13, 14]. In addition, the interpretation of biological phenomena can be improved because of the graph structure [15-17], which naturally fits into the graph-based SSL.

2 Materials and Methods

2.1 Data

Samples—The Alzheimer's Disease Neuroimaging Initiative initial phase (ADNI-1) was launched in 2003 to test whether serial magnetic resonance imaging (MRI), position emission tomography (PET), other biological markers, and clinical and neuro-psychological assessment could be combined to measure the progression of MCI and early AD. This multi-site longitudinal study was intended to aid researchers and clinicians develop new treatments for MCI and early AD, monitor their effectiveness, and lessen the time and cost of clinical trials. The ADNI-1 has been extended to its subsequent phases (ADNI-GO and ADNI-2) for follow-up for existing participants and additional new enrollments. Inclusion and exclusion criteria, clinical and neuroimaging protocols, and other information about ADNI have been published previously and can be found at www.adni-info.org [18-20]. Demographic information, raw scan data, *APOE* and GWAS genotypes, neuropsychological test scores, and diagnostic information are available from the ADNI data repository (http://www.loni.ucla.edu/ADNI/). Individuals included in this study were 174 E-MCI (early MCI) and 98 HC (healthy older adults) participants in ADNI-GO or ADNI-2.

Image processing—All available baseline 3T structural brain MRI scans were downloaded from the ADNI database. As detailed in previous studies [18, 19], two widely employed automated MRI analysis techniques were independently used to process MRI scans: whole-brain voxel-based morphometry (VBM) implemented in the Statistical Parametric Mapping 8 (SPM8) software to extract mean grey matter (GM) density for target regions of interest (ROIs) and FreeSurfer version 5.1 to extract mean cortical thickness and volumetric measure for target ROIs. Pre-processed Florbetapir (also known as AV-45 and Amyloid) PET scans to assess brain amyloid β burden were downloaded from the ADNI database. For each scan, mean regional SUVR (standardized uptake value ratio) values were extracted for target ROIs using MarsBaR in SPM8, as detailed in previous study [7, 18]. FDG-PET was used to measure the brain's rate of glucose metabolism with the tracer [^{18}F] Fluorodeoxyglucose. FDG-PET ROI data was downloaded from the ADNI database. All MRI ROI values were adjusted for the baseline age, gender, education, and intracranial volume (ICV) using a regression model, prior to analyses. All the ROI values of Florbetapir and FDG PET were adjusted for the baseline age, gender, and education.

2.2 Classification of Early Mild Cognitive Impairment (E-MCI)

The semi-supervised learning uses both labeled and unlabeled data to improve on the performance of supervised learning. There are several types of SSL algorithms, and the graph-based SSL was used in our study. If two patients' samples were more closely related to others, the algorithm assumed that the diagnosis of E-MCI from those two patients is more likely to be similar. Thus, the classification of E-MCI can be enhanced by considering similarities between patient samples. A natural method of analyzing relationships between entities is a graph, where nodes represent participants and edges show their possible relations. Figure 1 represents an example graph, which was conducted using the brain imaging data. An annotated participant is labeled either by '-1' or '1', indicating the two possible clinical outcomes, either 'healthy older adult' or 'E-MCI'. In order to predict the

label of the unannotated patient '?', the edges connected from/to the patient play an important role in influencing propagation of the relation between the patient and its neighbors. This idea can be easily formulated using a graph-based semi-supervised learning [8]. Edges represent relations, more specifically similarities between participants that may be extracted from different brain imaging data. Different brain imaging data produce different graphs. Thus, the classification of E-MCI can be benefit by integrating diverse graphs from multimodal brain imaging data, i.e., incomplete information and noise. Technically, the data-setup of our experiment for the binary classification can be rephrased as $\{x_n, y_n\}_{n=1}^{N}$ where $x_n \in R^d$ (d is the number of features and N is the number of participants) and $y_n \in \{-1, 1\}$.

Graph-based semi-supervised learning—In the graph-based SSL [8], a participant x_i (i = 1,...,n) is represented as a node i in a graph, and the relationship between participants is represented by an edge. The edge strength from each node j to other node i is encoded in element w_{ij} of a $n \times n$ symmetric weight matrix W. A Gaussian function of Euclidean distance between participants, with length scale hyperparameter σ, is used to specify connection strength:

$$w_{ij} = \begin{cases} exp\left(-\frac{(x_i - x_j)^T (x_i - x_j)}{\sigma^2}\right) & \text{if} \quad i \sim j, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Nodes i and j are connected by an edge if i is in j's k-nearest-neighborhood or vice versa. Thus, nearby participants in Euclidean spaces are assigned large edge weights.

The labeled nodes have labels $y_l \in \{-1, 1\}$, while the unlabeled nodes have zeros $y_u = 0$. The graph-based SSL will output an n-dimensional real-valued vector $f = \left[f_l^{\mathrm{T}} f_u^{\mathrm{T}}\right]^{\mathrm{T}} = (f_1, \ldots, f_l, f_{l+1}, \ldots, f_{n=l+u})^{\mathrm{T}}$, which can be thresholded to make label predictions on $f_l = f_1,...,f_n$ after learning. It is assumed that f_i should be close to the given label y_i in labeled nodes (loss condition), and overall, f_i should not be too different from the f_i of adjacent nodes (smoothness condition). One can obtain f by minimizing the following quadratic functional [8, 9, 11]:

$$\min_f (f - y)^T (f - y) + \mu f^T L f \quad (2)$$

where $y = (y_1,...,y_l, 0,...0)^{\mathrm{T}}$, and the matrix L, called the graph Laplacian matrix [21], is defined as $L = D - W$ where $D = \mathrm{diag}(d_i)$, $d_i = \Sigma_j w_{ij}$. The parameter μ trades off loss versus smoothness. The solution of this problem is obtained as

$$f = (I + \mu L)^{-1} y \quad (3)$$

where I is the identity matrix.

2.3 Integration of Multi-Modal Brain Imaging Dataset

In order to combine the graphs from multimodal brain imaging data, four graphs can be integrated from finding optimum combination coefficients. Information from each graph is regarded as partially independent from and partly complementary to others. Reliability may be enhanced by integrating all available data sources using the graph-based SSL, which has been applied to the extended problem of protein function prediction [22] and clinical outcome prediction using multi-levels of genomic data [12]. Based on the method, the integration of multiple graphs is used to find an optimum value of the linear combination coefficient for the individual graphs (Fig. 2). This corresponds to finding the combination coefficients α for the individual Laplacians of the following mathematical formulation:

$$\min_{\alpha} y^T \left(I + \sum_{k=1}^{K} \alpha_k L_k \right)^{-1} y, \sum_k \alpha_k \leq \mu \quad (4)$$

, where K is the number of graphs and L_k is the corresponding graph-Laplacian of graph G_k. Similar to the output prediction for single graphs, the solution is obtained by

$$f = \left(I + \sum_{k=1}^{K} \alpha_k L_k \right)^{-1} y. \quad (5)$$

3 Results

3.1 Experiment Setting

The receiver operating characteristic (ROC) curve plots sensitivity (true positive rate) as a function of 1-specificity (false positive rate) for a binary classifier system as its discrimination threshold is varied [23]. An ROC score of 0.5 corresponds to random prediction, and an ROC score of 1.0 implies that the model succeeded in putting all of the positive examples before all of the negatives. For each dataset, we calculated area under the curve (AUC) of ROC as a performance measure. In order to avoid the overfitting, five-fold cross-validation was conducted. Since some of the brain imaging dataset is high dimensional and noisy, and contains many redundant features, which may incur computational difficulty and low accuracy, a Student *t*-test based feature selection method was used [24]. Even though there are many feature selection techniques such as filter, wrapper, and embedded method [25], a simple univariate feature selection method was used in order to emphasize not the effect of feature selection but the effect of integration of multimodal brain imaging data. The values of SSL model parameters, k from Equation (1) and μ from Equation (3), were determined by the results of search over $k \in \{3, 4, 5, 6, 7, 8, 9, 10, 20, 30\}$ and $\mu \in \{0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 10, 100, 1000\}$. The optimized combination of model parameters was selected when the greatest AUC was obtained.

3.2 Experiment Results

With multimodal brain imaging data, we provide empirical comparison results about which type of brain imaging data is more informative to a given classification problem for

diagnosis of E-MCI. Figure 3 shows the AUC performance on the classification of E-MCI. The averages of five-fold AUCs from Florbetapir, FDG, FreeSurfer, and VBM are shown in the figure. Among four types of brain imaging dataset, the performance of FreeSurfer dataset showed the best single modality performance with 0.6576 AUC. In Figure 3, AUC increases in the order of the following dataset, FreeSurfer > VBM > FDG > Florbetapir.

3.3 Integration Effects

Since different brain imaging data contain partly independent and partly complementary information content, we integrated across multi-modal brain image datasets for better prediction of E-MCI. We found that multivariate integration across different brain imaging modalities increased the prediction performance for patients with EMCI. Figure 4 shows the results of the integration with all combination of different types of brain imaging dataset. The model combining Florbetapir and VBM (0.6322 AUC) outperformed the model with VBM only (0.609 AUC). In addition, the integration with FDG and FreeSurfer showed the best performance among all combination of four different types of brain imaging dataset with 0.6681 AUC. However, the integration with all four types of brain image data included did not show the best performance.

3.4 Comparison with SVM

The performance from graph-based SSL classifiers was compared with Support Vector Machine (SVM) performance. SVM involves finding an optimal decision boundary, i.e., maximizing the margin by finding the largest achievable distance among the separating hyperplane and the data points on either side. If the data points are separated by a non-linear hyperplane because of some intrinsic property of the problem, it is more appropriate to map the input feature space to a high-dimensional feature space where the data points are separated by a linear hyperplane. This mapping process is conducted by kernel functions. Among kernel functions, the Radial Basis Function (RBF) kernel was used with a wide range of sigma, from 10^{-6} to 1, in order to select the best model. In order to fairly compare the performance, we used the same set of features, which was used in the graph-based SSL. The models from the graph-based SSL outperformed the models from SVM except for Florbetapir data (Table 1). The Wilcoxon signed-rank test was used to assess the significance level of difference in performance between the results of the graph-based SSL and SVM [26]. The model with FreeSurfer dataset from the graph-based SSL showed significantly better than the one from SVM.

4 Discussions and Conclusions

Using automatic whole-brain ROI analysis techniques and a graph-based semi-supervised learning (SSL) method, we developed a classification model to differentiate E-MCI from HC for early detection of AD. In this study, we used MRI (FreeSurfer and VBM) and PET (Florbetapir and FDG) scans from 174 E-MCI and 98 HC in the ADNI cohort. The graph-based SSL technique was used to integrate multi-modal brain imaging data and select imaging-based phenotypes for optimizing E-MCI prediction accuracy. The data integration framework for multimodal brain imaging data has scalability to easily extend to additional types of brain imaging data. In addition, it preserves type-specific properties from the brain

imaging data since the matrices from different types of brain imaging data were not simply merged but combined after conversion into a graph for the integration (Fig. 2).

Our results showed that 1) the graph-based SSL classifiers outperformed support vector machines (SVM) for this task; 2), we obtained the best results when using ROI values extracted by FreeSurfer from structural MRI scans; (3) the overall best performance was obtained with 66.8% cross-validated AUC when FDG PET and FreeSurfer data were combined; (4) the integration with all four types of brain image data included did not show the best performance; and (5) selected imaging-based phenotypes included ROI values extracted from temporal lobe, hippocampus, and amygdala. It has been showed that regional brain atrophy occurs initially and most severely in the entorhinal cortex and hippocampus before spreading throughout the neocortex [27]. These findings suggest that the predictive model may be combined with various data sources from different types of brain imaging data. Integration of independent or complementary information content may improve the chances of successful early diagnosis of AD. The graph-based SSL approach with multimodal brain imaging data has substantial potential for enhanced early detection of AD.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. 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 California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

Samples from the National Cell Repository for AD (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study. Additional support for data analysis was provided by NLM K99 LM011384, NIA R01 AG19771, P30 AG10133, NCI R01 CA101318, NLM R01 LM011360, NSF IIS-1117335, and RC2 AG036535, and U01 AG032984 from the NIH, Foundation for the NIH, and NINDS (R01NS059873).

References

1. Alzheimer's A, Thies W, Bleiler L. 2011 Alzheimer's disease facts and figures. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011; 7:208–244.
2. Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. Archives of neurology. 1999; 56:303–308. [PubMed: 10190820]
3. Stephan BC, Hunter S, Harris D, et al. The neuropathological profile of mild cognitive impairment (MCI): a systematic review. Mol Psychiatry. 2012; 17:1056–1076. [PubMed: 22143004]
4. Petersen RC, Roberts RO, Knopman DS, et al. Mild cognitive impairment: ten years later. Archives of neurology. 2009; 66:1447–1455. [PubMed: 20008648]
5. Wang H, Nie F, Huang H, Kim S, Nho K, et al. Identifying quantitative trait loci via group-sparse multitask regression and feature selection: an imaging genetics study of the ADNI cohort. Bioinformatics. 2012; 28:229–237. [PubMed: 22155867]

6. Meda SA, Narayanan B, Liu J, Perrone-Bizzozero NI, et al. A large scale multivariate parallel ICA method reveals novel imaging-genetic relationships for Alzheimer's disease in the ADNI cohort. NeuroImage. 2012; 60:1608–1621. [PubMed: 22245343]
7. Risacher SL, Kim S, et al. The role of apolipoprotein E (APOE) genotype in early mild cognitive impairment (E-MCI). Frontiers in aging neuroscience. 2013; 5:11. [PubMed: 23554593]
8. Zhou D, Bousquet O, Weston J, Scholkopf B. Learning with local and global consistency. Advances in Neural Information Processing Systems (NIPS). 2004; 16:321–328.
9. Belkin M. Regularization and Semi-supervised Learning on Large Graphs. Proceedings of the 17th Annual Conference on Learning Theory (COLT) 3120. Lecture Notes in Computer Science. 2004:624–638.
10. Zhu, X.; Ghahramani, Z.; Lafferty, J. Semi-supervised learning using Gaussian fields and harmonic functions; Proceedings of the Twenty-first International Conference on Machine Learning (ICML); Washington, DC. AAAI Press; 2003. p. 912-919.
11. Chapelle O, Weston J, Scholkopf B. Cluster kernels for semi-supervised learning. Advances in Neural Information Processing Systems (NIPS). 2003; 15:585–592.
12. Kim D, Shin H, Song YS, Kim JH. Synergistic effect of different levels of genomic data for cancer clinical outcome prediction. J Biomed Inform. 2012; 45:1191–1198. [PubMed: 22910106]
13. Tsuda K, Shin H, Scholkopf B. Fast protein classification with multiple networks. Bioinformatics. 2005; 21(Suppl 2):ii59–65. [PubMed: 16204126]
14. Shin, H.; Tsuda, K. Prediction of Protein Function from Networks.. In: Chapelle, Olivier; Schölkopf, Bernhard; Zien, Alexander, editors. Semi-Supervised Learning. MIT press; 2006. p. 339-352.Chapter 20
15. Spellman PT, Sherlock G. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell. 1998; 9:3273–3297. [PubMed: 9843569]
16. Segal E, Shapira M, et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet. 2003; 34:166–176. [PubMed: 12740579]
17. Ohn JH, Kim J, Kim JH. Genomic characterization of perturbation sensitivity. Bioinformatics. 2007; 23:i354–358. [PubMed: 17646317]
18. Risacher SL, Shen L, West JD, et al. Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. Neurobiology of aging. 2010; 31:1401–1418. [PubMed: 20620664]
19. Risacher SL, Saykin AJ, West JD, et al. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. Current Alzheimer research. 2009; 6:347–361. [PubMed: 19689234]
20. Weiner MW, Veitch DP, et al. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2012; 8:S1–68.
21. Chung FRK. Spectral Graph Theory. Number 92 in Regional Conference Series in Mathematics. 1997
22. Shin H, Lisewski AM, Lichtarge O. Graph sharpening plus graph integration: a synergy that improves protein functional classification. Bioinformatics. 2007; 23:3217–3224. [PubMed: 17977886]
23. Gribskov M, Robinson NL. Use of receiver operating characteristic (ROC) analysis to evaluate sequence matching. Comput Chem. 1996; 20:25–33. [PubMed: 16718863]
24. Jafari P, et al. An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors. BMC Med Inform Decis Mak. 2006; 6:27. [PubMed: 16790051]
25. Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. Bioinformatics. 2007; 23:2507–2517. [PubMed: 17720704]
26. Demsar J. Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research. 2006; 7:1–30.

27. Scahill RI, Schott JM, et al. Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. Proc Natl Acad Sci U S A. 2002; 99:4703–4707. [PubMed: 11930016]

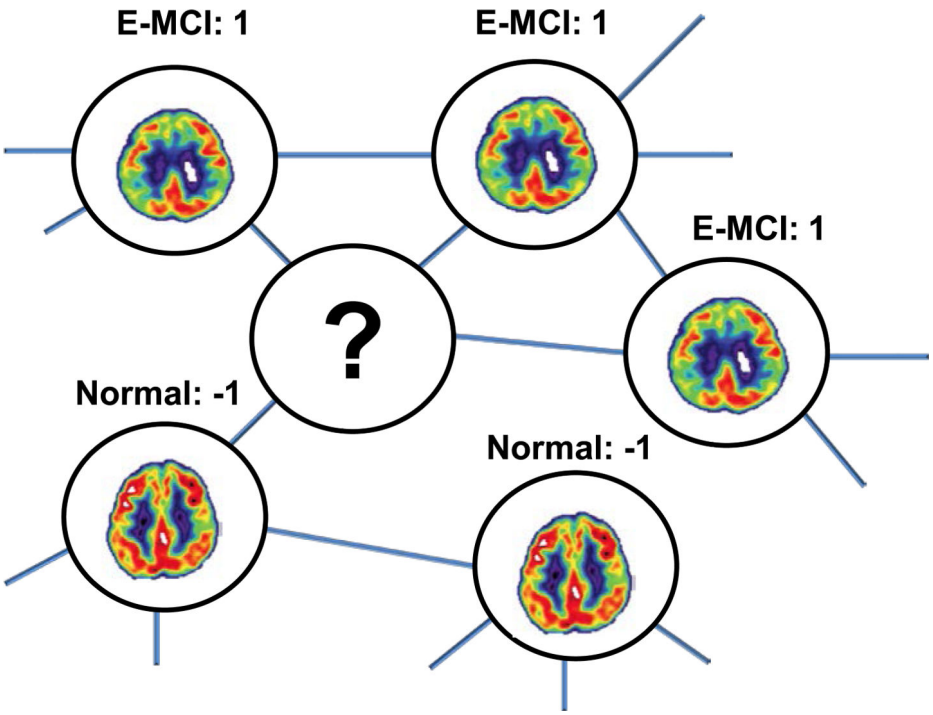

Fig. 1.

Graph representation of brain imaging data between participants. Nodes represent participants and edges depict relations between samples. An annotated sample is labeled either by −1 or +1. In this example, the negative labels indicate samples from ‘healthy older adults’. On the contrary, the positive labels indicate the samples from ‘E-MCI’. The diagnosis of the unannotated sample marked as ‘?’ is predicted by employing the graph-based semi-supervised learning.

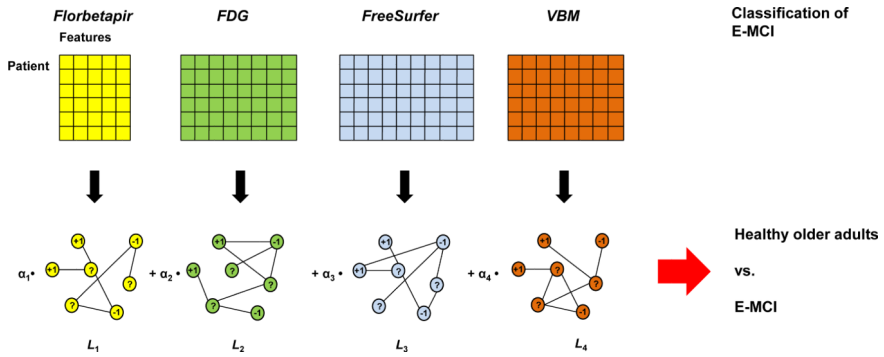

Fig. 2.

Integration scheme of four different types of brain imaging data. Each brain imaging data can be converted into a graph, and then multiple graphs can be combined through finding the optimal value of the combination coefficient.

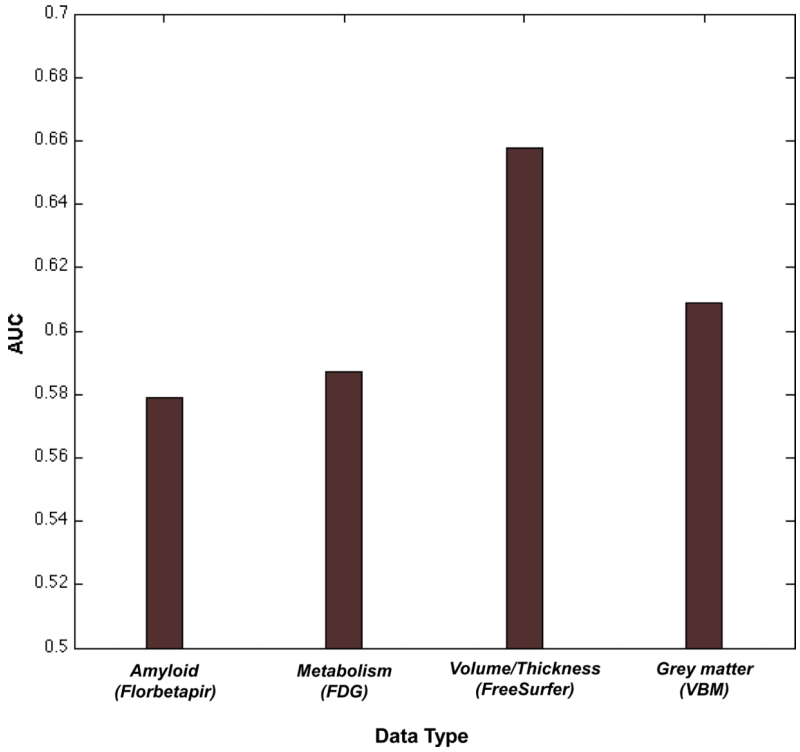

Fig. 3.

Performance comparison among four types of brain imaging dataset. The y axis represents the average AUC and the x axis shows the date type.

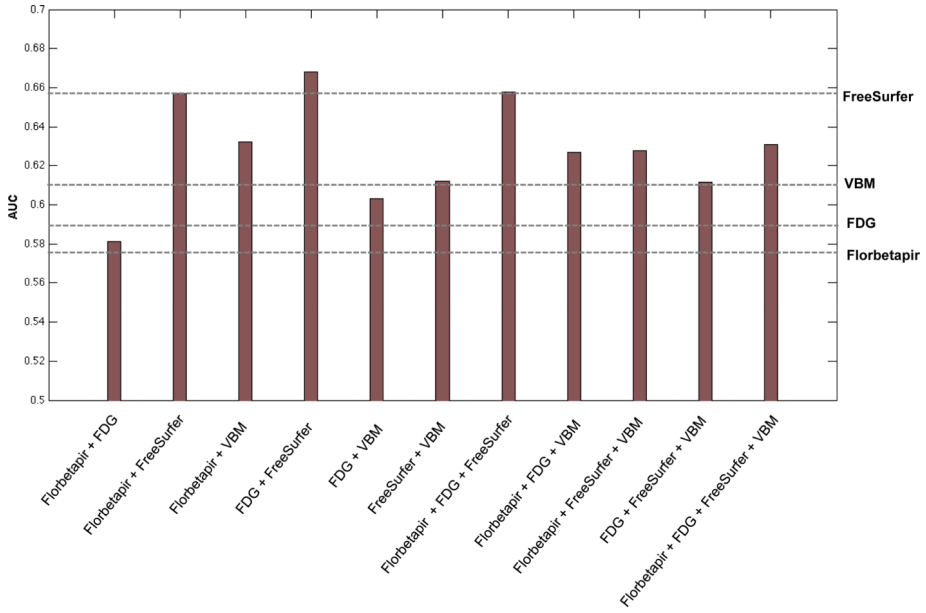

Fig. 4.

Integration of multimodal brain imaging data. The y axis represents the average AUC and the x axis shows the combination of brain imaging datasets.

Table 1

Comparison between the graph-based SSL and SVM. *P*-values were calculated using a Wilcoxon signed-rank test between performance (AUC) of the graph-based SSL and SVM.

Data type	Graph-based SSL	SVM	*P*-value
Florbetapir	0.5789 (± 0.0732)	**0.5825** (± 0.0372)	1.0000
FDG	**0.5873** (± 0.0587)	0.5664 (± 0.0643)	0.8413
FreeSurfer	**0.6576** (± 0.0905)	0.5163 (± 0.0552)	0.0159
VBM	**0.609** (± 0.1059)	0.5709 (± 0.0916)	0.5476