Published in final edited form as:

Inf Process Med Imaging. 2017 June ; 10265: 158–169. doi:10.1007/978-3-319-59050-9_13.

A Novel Dynamic Hyper-Graph Inference Framework for Computer Assisted Diagnosis of Neuro-Diseases

Yingying Zhu, Xiaofeng Zhu, Minjeong Kim, and Guorong Wu

Department of Radiology and BRIC, the University of North Carolina at Chapel Hill

Abstract

Recently hyper-graph learning gains increasing attention in medical imaging area since the hyper-graph, a generalization of a graph, opts to characterize the complex subject-wise relationship behind multi-modal neuroimaging data. However, current hyper-graph methods are re-strained with two major limitations: (1) The data representation encoded in the hyper-graph is learned only from the observed imaging features for each modality separately. Therefore, the learned subject-wise relation-ships are neither consistent across modalities nor fully consensus with the clinical labels or clinical scores. (2) The learning procedure of data representation is completely independent to the subsequent classification step. Since the data representation optimized in the feature domain is not exactly aligned with the clinical labels, such independent step-by-step workflow might result in sub-optimal classification. To overcome these limitations, we propose a novel dynamic hyper-graph inference frame-work. Working in a semi-supervised manner, it iteratively estimates and adjusts the subject-wise relationship from multi-modal neuroimaging data until the learned data representation (encoded in the hyper-graph) achieves the largest consensus with the observed clinical labels and scores. It is worth noting that our inference framework is also flexible to integrate classification (identifying individuals with neuro-disease) and regression (predicting the clinical scores). We have demonstrated that the performance of our proposed dynamic hyper-graph inference framework renders more accurate diagnosis result in identifying MCI (Mild Cognition Impairment) subjects and the fine-grained recognition of different progression stage of MCI compared to conventional counterpart methods.

1 Introduction

There is an overwhelming consensus that the morphological patterns from neuroimaging data, either structural image (e.g., MRI) or functional image (e.g., PET/SPECT), are highly correlated with neuro-diseases, such as Alzheimer's disease (AD). Large inter-subject variations and complex disease pathology make it challenging to develop imaging-based diagnosis system that can identify an individual patient or high risk subjects, in vivo, with high sensitivity and specificity.

Since different imaging modalities convey complementary information, many state-of-the-art machine learning approaches have been proposed to improve the diagnosis accuracy using multi-modal imaging. To name a few works, multi-task learning by ensemble SVM and SVR are used for AD and Mild Cognitive Impairment (MCI) classification and regression in [1], Deep Boltzmann machine is used to fuse multiple modality imaging data into a high-level representation for classification [2], etc. Showing the advantage of

integrating multi-modal imaging information in assisting diagnosis, those learn models have a deficiency in the generality and accuracy due to the limited number of training samples with both multi-modal imaging data and clinical labels. Hence, semi-supervised learning approaches come to the stage which leverage the data distribution of the entire population (regardless with or without label) to overcome the lack of labeled subjects. In general, most semi-supervised methods represent data distribution using graph techniques.

Due to large inter-subject variations, pairwise relationship encoded in the simple graph is not sufficient to characterize such complex and diverse subject-to-subject relationship. The emerging hyper-graph technique shows potential to solve this problem as it has been successfully applied in various medical imaging applications including image segmentation [3] and classification [4]. Particularly, hyper-graph has been used to combine multi-modal neuroimaging information for the identification of MCI subjects in [4]. The main idea is to construct several hyper-graphs, with each measuring the relevance among a group of subjects based on one modality. Compared to conventional simple graph technique, hyper-graph allows connecting more than two nodes (subjects) in each hyper-edge. Thus, hyper-graph intuitively conveys more information than conventional simple graph through a set of hyper-edges, which makes it more powerful in capturing high-order relationship among different subjects. Given the hyper-graph, the MCI classification is eventually a label propagation procedure to disseminate the clinical labels throughout the hyper-graph structure, following the principle that subjects falling in the same hyper-edge should bear similar clinical labels.

Although promising results have been achieved in [4], there are still several issues. First, the hyper-graph is built for each modality independently. For example, there are two hyper-graphs separately built from morphological features extracted from MR and PET images, as shown by blue and purple solid ellipses in the middle of Fig. 1. The final classification result is the linear combination of label likelihood scores by different hyper-graphs. Although the weights for combination can be learned from training data, the learned prior weights cannot guarantee to be optimal for each testing subject. Second, the construction of hyper-graphs is exclusively driven by morphological feature representations and completely independent to the subsequent classification and regression task with the morphological patterns already aligned with clinical labels. More critically, the hyper-graphs are fixed once built upon the observed morphological feature only. Due to the possible noisy and redundant patterns existing in the observed high-dimensional imaging features, the learned hyper-graph might be sub-optimal for label propagation and thus misguide the propagation of clinical labels to each individual testing subject.

To address these issues, we propose a novel dynamic hyper-graph inference framework for multi-modal diagnosis of neuro-diseases. Instead of constructing separate hyper-graphs for different modal imaging data, we propose to learn a unified dynamic hyper-graph for all modal imaging data aligned with phenotype data (e.g., clinical labels scores), as shown in the right of Fig. 1. The insight of our work is that all subjects share a common hyper-graph disregard of various appearances of morphological patterns across different modalities. The dynamic hyper-graph is designated to (1) respect the subject-wise relationship measured by the observed multi-modal imaging data (2) automatically learn the optimal weights for

dynamical hyper-graph on multi-modal imaging data (3) bring minimal discrepancy after the verification between the estimated data labels and clinical labels assigned in the training dataset. To achieve these goals, we present an iterative solution which alternatively estimates and adjusts the subject-wise relationship from multi-modal neuroimaging data until the learned data representation (encoded in the hyper-graph) achieves the largest consensus with the observed clinical labels. Meanwhile, we notice that the clinical scores such as MMSE and CDR scores in AD are widely used in current learning-based diagnosis methods since these phenotype data has a higher correlation with the clinical labels compared to imaging data. To that end, we further extend our hyper-graph inference framework to integrate both classification (identifying individuals with neuro-disease) and regression (predicting the clinical scores). We apply our proposed hyper-graph based diagnosis method to identify MCI subjects and classify MCI subjects at different progression stages. The experiments result demonstrated that our method outperforms conventional hyper-graph methods about 1.8% increase in identifying MCI subjects from NC and AD cohorts and 3.1% increase in fine grained recognition of recognizing different stage of MCI respectively.

2 Methods

We will first briefly review the standard hyper-graph learning techniques and then introduce the proposed dynamic hyper-graph learning framework.

2.1 Conventional Hyper-Graph Learning

Conventional hyper-graph for one modality—A hyper-graph is denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the vertex set and each vertex represent one subject here, $\mathcal{E}$ is the hyper-edge set. In a conventional graph, one edge connects a pair of vertex, however, in a hyper-graph, one hyper-edge connects a set of vertexes as shown in Fig. 2. Given n subjects with imaging features, we first compute a subject-wise distance matrix $S \in \mathcal{R}^{n \times n}$ using imaging features, then, we construct the hyper-graph incidence distance matrix $\mathbf{H} = \{\mathbf{0}, \mathbf{1}\} \in \mathcal{R}^{|\mathcal{V}| \times |\mathcal{E}|}$ to model the relationship among different vertices and edges as,

$$\mathbf{H}(v, e) = \mathbf{1}, S(v, e) < \theta, \mathbf{H}(v, e) = \mathbf{0}, \text{otherwise}, \quad (1)$$

where θ is the predefined threshold. The (v, e) – th entry of the incidence matrix indices whether the vertex is connected via the hyper-edge e to other vertices (1 represent connected and 0 represent un-connected). Two diagonal matrices $\mathbf{D}_v$ and $\mathbf{D}_e$ to represent the vertex degrees and hyper-edge degrees are also constructed from the incidence matrix $\mathbf{H}$. The vertex degree and edge degree are computed as,

$$\mathbf{D}_v(v, v) = \sum_{e \in \mathcal{E}} \mathbf{H}(v, e), \mathbf{D}_e(e, e) = \sum_{v \in \mathcal{V}} \mathbf{H}(v, e),$$

After constructed the incidence matrix, vertexes degree matrix and hyper-edge degree matrix, denote $\mathbf{I}$ as the identity matrix, one can compute the hyper-graph Laplacian matrix $\mathbf{L}$ as,

$$\mathbf{L} = \mathbf{I} - \mathbf{D}_v^{-\frac{1}{2}} \mathbf{H} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-\frac{1}{2}}, \quad (2)$$

Given known data clinical labels $\mathbf{Y}$ and learned hyper-graph Laplacian matrix, conventional methods solve the following function for classification labels propagation on hyper-graph,

$$\begin{aligned} \arg \min_{\mathbf{F}} \quad & \text{tr}(\mathbf{F}^T \mathbf{L} \mathbf{F}) \\ \text{s.t.} \quad & \mathbf{A} \mathbf{F} = \mathbf{Y}, \end{aligned} \quad (3)$$

where $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_n]$ are the labels for all subjects, $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_q]$ are the known clinical labels, $\mathbf{A} \in \mathbb{R}^{q \times n}$ is a selection matrix to select the known label subjects from all subjects. Eq. (2) can be solved using Augmented Lagrange Methods (ALMs) [6].

Multiple hyper-graphs for multiple modalities—Eq. (2) only considers one modality imaging data, in order to apply hyper-graph for multiple modality imaging data, Gao et.al. [5] proposed to combine multiple hyper-graph models by a linear model and estimating the unknown data labels on the combined multiple hyper-graph structures. Suppose there are M modalities data, they construct M different hyper-graph Laplacian matrixes $\mathbf{L}_m$, $m = 1, \dots, M$. A weight matrix $\mathbf{u} = [u_1, \dots, u_M]$ for different modalities is learned to combine M hyper-graphs Laplacian matrixes. The labels are propagated on the combined hyper-graph by solving

$$\begin{aligned} \arg \min_{\mathbf{u}, \mathbf{F}} \quad & \text{tr}(\mathbf{F}^T \sum \mathbf{u}_m \mathbf{L}_m \mathbf{F}) \\ \text{s.t.} \quad & \mathbf{A} \mathbf{F} = \mathbf{Y}, \mathbf{u} > \mathbf{0}, \mathbf{u}^T \mathbf{1} = 1. \end{aligned} \quad (4)$$

Eq. (3) also can be solved using Augmented Lagrange Methods (ALMs) [6] efficiently.

Conventional hyper-graph issues—The hyper-graph construction is only based on the imaging morphological features, which is independent of the clinical label and score propagation process as shown in Fig. 1. There is no guarantee that the learned hyper-graph is optimal to model subject-wise relationship among clinical labels. Therefore, this learned sub-optimal hyper-graph structure can lead to a sub-optimal solution in the process of propagating clinical labels and clinical scores from known subjects to unknown subjects.

2.2 Joint Data Representation Learning and Classification on Dynamic Hyper Graph

To address these above issues in the conventional hyper-graph learning, we propose a novel method to solve the hyper-graph learning and classification problems simultaneously. Here, we define an intrinsic subject-wise distance matrix $\mathbf{S}$, where s_{ij} is the similarity between subjects i and j . We learn a dynamic hyper-graph Laplacian matrix $\mathbf{L}$ from the subject-wise similarity matrix $\mathbf{S}$ by: 1) construct incidence matrix, vertex degree and hyper-edge degree matrixes $\mathbf{H}$, $\mathbf{D}_v$, $\mathbf{D}_e$ from the estimated similarity matrix $\mathbf{S}$ 2) compute different hyper-graph

Laplacian matrix using Eq. (2). We then propose a dynamic hyper-graph classification framework which learns the similarity matrix $\mathbf{S}$ on clinical label domain and feature domain, then update the data clinical labels on the hyper-graph simultaneously,

$$\begin{aligned} \arg \min_{\mathbf{S}, \mathbf{F}, \mathbf{L}, \boldsymbol{\alpha}} \frac{1}{2} \sum_{ij}^n \left(\|\mathbf{f}_i - \mathbf{f}_j\|_2^2 s_{ij} + \sum_{m=1}^M \alpha_m \|\mathbf{x}_i^m - \mathbf{x}_j^m\|_2^2 s_{ij} \right) + \gamma \text{tr}(\mathbf{F}^T \mathbf{L} \mathbf{F}) + \|\boldsymbol{\alpha}\|_2^2 \\ \text{s.t. } \forall i, m, \mathbf{A} \mathbf{F} = \mathbf{Y}, \mathbf{s}_i^T \mathbf{1} = 1, \mathbf{s}_i \geq 0, \boldsymbol{\alpha}^T \mathbf{1} = 1, \boldsymbol{\alpha} \geq 0, \end{aligned} \quad (5)$$

where $\mathbf{x}_i^m$ is the feature vector of subject i from modality m , $\mathbf{f}_i$ is the class label for subject i , $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_n]$, $\mathbf{s}_i = [s_{i1}, \dots, s_{im}]^T \in \mathcal{R}^{n \times 1}$, $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_m]^T \in \mathcal{R}^{M \times 1}$ is the weights for each modality. An intrinsic subject distance matrix $\mathbf{S}$ is learned using clinical labels and different modality imaging features from both training and testing subject in this model. Furthermore, the data labels propagation on hyper-graph is performed simultaneously in Eq. (5). Eq. (5) is can be considered to be convex problem with respect to each parameter: $\mathbf{S}, \mathbf{F}, \mathbf{L}, \boldsymbol{\alpha}$, so we can break Eq. (5) into several convex problems and solve each sub-convex problem iteratively until converge.

2.3 Joint Classification and Regression on Dynamic Hyper-Graph

The above dynamic hyper-graph model is able to learn hyper-graph representation on imaging data and clinical labels, however, the clinical score are not used yet. Convergent evidence shows that clinical score (such as cognitive score) prediction can be reliably used to evaluate and analysis the progression of the disease clinically. Furthermore, the clinical score is an additional useful information for disease identification. Therefore, it is straightforward to integrate the clinical score for disease diagnosis. We proposed to adopt a joint clinical score regression model into the dynamic hyper-graph model using conventional graph regression model.

The basic idea of conventional graph regression model is to estimate the value of each vertex on a graph as a linear combination of other vertexes in this graph. Let $\mathbf{C} \in \mathcal{R}^{n \times d}$ denote the clinical score of all subjects on a graph, the clinical score $\mathbf{c}_i$ of subject i can, be simply computed as, $\mathbf{c}_i = \mathbf{w}_i^T \mathbf{C}$, where $\mathbf{w}_i \in \mathcal{R}^{n \times 1}$ is the linear regression weights for subject i [7].

Denote the regression weight matrix for all subjects as $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_n]$, $\mathbf{Q}$ as the known clinical score for training subjects, $\mathbf{C}$ as the clinical score for all subjects, $\mathbf{A}$ is the selection matrix which select the training subjects from all subjects, we have $\mathbf{Q} = \mathbf{A} \mathbf{C}$. We propose a dynamic hyper-graph which learns the hyper-graph from the imaging features, clinical labels and clinical scores, which jointly perform clinical label and score estimation as,

$$\begin{aligned}
& \arg \min_{\mathbf{S}, \mathbf{F}, \mathbf{L}, \mathbf{C}, \mathbf{W}, \boldsymbol{\alpha}} \frac{1}{2} \sum_{ij} \left(\|\mathbf{f}_i - \mathbf{f}_j\|_{2^{s_{ij}}}^2 + \sum_{m=1}^M \alpha_m \|\mathbf{x}_i^m - \mathbf{x}_j^m\|_{2^{s_{ij}}}^2 + \beta \|\mathbf{w}_i^T \mathbf{C} - \mathbf{w}_j^T \mathbf{C}\|_{2^{s_{ij}}}^2 \right) \\
& + \gamma \text{tr}(\mathbf{F}^T \mathbf{L} \mathbf{F}) + \|\boldsymbol{\alpha}\|_2^2, \\
& \text{s.t. } \forall i, m, \mathbf{A} \mathbf{F} = \mathbf{Y}, \mathbf{A} \mathbf{C} = \mathbf{Q}, \mathbf{s}_i^T \mathbf{1} = 1, \mathbf{s}_i \geq 0, \boldsymbol{\alpha}^T \mathbf{1} = 1, \boldsymbol{\alpha} \geq 0.
\end{aligned} \tag{6}$$

where β is the weight for clinical score term. In Eq. (5), the subject-wise distance matrix $\mathbf{S}$ is learned on clinical label domain, imaging feature domain and clinical score domain in order to better represent the subject-wise relationship. The similarity matrix $\mathbf{S}$, the regression matrix $\mathbf{W}$ and the clinical labels and scores are optimized simultaneously in Eq. (5). Eq. (5) can be efficiently solved with respect to each parameters iteratively using Augmented Lagrange Methods (ALMs) [6, 8–10].

Discussion—Recall the conventional hyper-graph methods, they learn a *fixed hyper-graph model* on imaging data only, then propagates the subject clinical labels on the pre-learned hyper-graph model as shown in Fig. 2 right. The hyper-graph model is fixed during the label propagation process. Our model learns an *optimal dynamical hyper-graph* representation for multiple modality imaging features, clinical labels and clinical scores as shown in Fig. 2 left. The learned hyper-graph is based on the imaging features, the subject labels, and the clinical score information. Furthermore, the learned hyper-graph is optimized on both the training subjects with known clinical data and the testing subjects which only have imaging data. There are three major contributions in the proposed dynamical hyper-graph model as shown in Fig. 2 right, 1) imaging features of both training and testing subjects are used to construct the hyper-graph, which is especially helpful for real clinical practice with sufficient imaging data but very limited clinical labels; 2) Besides multiple modality imaging features, the clinical labels and scores are also employed in hyper-graph construction in order to make the learned hyper-graph structure in census with not only imaging features but also clinical labels and scores; 3) hyper-graph structure learning, clinical label propagation, and clinical score prediction are solved in a joint framework with a global optimal solution.

3 Experiments

We apply our method for Alzheimer’s Disease diagnosis, which is one of the most common neuro-degenerative diseases with memory loss, cognition defects, altered behavior, and as AD progresses, it can lead death ultimately [11, 12]. A high risk stage before AD is called MCI with noticeable and measurable cognitive ability changes [5, 13, 14]. MCI is a very important stage in the progression of Alzheimer’s disease (AD). Convergent evidence shows that MCI subjects have a high risk to progress towards a prodromal stage of AD, although some subjects can remain staying in the MCI stage for a long time or even get better [13, 15], which are referred as MCI Non-Converters (MCI-NC). Other MCI subjects eventually developed to AD finally are referred as MCI Converters (MCI-C). Thus, identifying the heterogeneous subtypes such as AD, MCI-C, MCI-NC and predict the clinical score in the

potential AD patients is of high impact in clinical practice. We evaluate the performance of the model on identifying the heterogeneous subtypes in AD and also predict the clinical score using multiple modality imaging data.

3.1 Image Processing

In our experiments, we use the baseline MRI, PET images from the ADNI dataset (www.adni-info.org) including 81 AD subjects and 90 NC (Normal Control). In the MCI subjects, there are 50 MCI Converter (MCI Convert) and 60 MCI Non-Converter (MCI Non-Convert) subjects. For each subject, we first register its MR image to the PET image. Then, we apply skull strip, bias correction, and segmentation [16]. After that, we obtain the parcellation of 90 anatomical structures by warping the template to the underlying subject. We use AAL template which has manually labeled 90 regions of interest (ROI). We use HAMMER for registration (<http://www.nitrc.org/projects/hammer/>). The volume percentile of gray matter and the average intensity of PET image in each ROI are used as the morphological features.

3.2 Experimental Setting

We have evaluated the performance of our method on two multiple class classification problems: AD/MCI/NC classification and AD/MCI-C/MCI-NC/NC classification tasks. We use a 10-fold cross validation strategy. We randomly partition the subjects into 10 non-overlapping approximately equal size subsets. Each time one subset is used for testing and all rest subsets are used for training. The optimal parameters are learned by grid search strategy in the training set using five-fold inner cross-validation. The search range for parameters γ and β is $10^{-5} - 10^5$. This procedure is repeated 10 times to compute the overall cross-validation classification performance. Three statistical measures are used for classification performance evaluation, including accuracy (ACC), sensitivity (SEN) and specificity (SPEC). We also evaluate the regression performance of our model for clinical score prediction. The Root Mean Square Error (RMSE) and Correlation Coefficients (CC) are used here to evaluate the regression performance.

Performance Evaluation—We compare our method with the recent static hyper-graph method [5] for multi-modality data classification, which combines several hyper-graphs learned on different modalities linearly for the challenging MCI detection (Denoted as static hyper-graph (Gao et.al.)). We also compare our method with the ensemble SVM [2], which is the state of art methods for multiple modality data fusion and classification. The recent work on the counterpart method: transductive graph learning in [17] is also compared to our model. This model learns a conventional graph model on different modalities and clinical labels for AD diagnosis. Our model is different to transductive graph learning models in several aspects: first, we employ a hyper-graph model which is able to model more complex relationship; secondly, our model is for multiple-class problem; third, we learned the weights among different modalities automatically; last, both the clinical score and labels are jointly optimized in our model. In order to make the static hyper-graph method and transductive graph learning method suitable for regression task, we add graph based regression term and jointly optimize the label propagation and regression on the

conventional static hyper-graph model and transductive graph learning model in our experiments.

3.3 Classification Results

Fig. 3. shows the MCI classification results of all compared methods using multiple modality imaging data. Our method outperforms all competing for state-of-the-art methods. Specifically, by using all modality imaging data, our method achieves an improvement of **1.8%** for the NC/MCI/AD classification problem compared to static hyper-graph model and **1.6%** for the NC/MCI-C/MCI-NC/AD classification problem in terms of accuracy.

Different Sizes of Training Data for Classification—To evaluate the performance of our method on different training data sizes, we vary the size of training data (the data with known clinical labels and clinical scores) and compare the classification performance of our method with static hyper-graph (Gao et. al.) method. Fig. 5. right shows the accuracy of the competing methods on different data size for three class classification (AD/MCI/NC). Fig. 4. left shows the accuracy of competing methods on different data size for four class classification problem (AD/MCI-C/MCI-NC/NC). The performance of our method decreases gracefully when the training dataset size decreases. However, the competing method decreases more than our method when training dataset decreases. This result suggests that our method is able to learn a more optimal hyper-graph model than the conventional method (Gao et. al.) after using the label information in the hyper-graph learning process.

3.4 Regression Results

Fig. 5. shows the clinical score regression results of all compared methods using multiple modality imaging data. Our method outperforms all competing for state-of-the-art methods. Specifically, as shown in the left of Fig. 5., by using all modality imaging data, our method achieves an improvement of **1.8%** for the ADAS-COG regression problem and **2.8%** for the MMSE clinical score regression problem compared to the modified static hyper-graph regression model in terms of accuracy. We furthermore show the performance of competing methods by measuring the Regression Root Mean Square (RMSE) in the right Fig. 5. Our method achieves error decrease at **0.8** on ADAS-COG score and **0.9** on MMSE score compared to the static hyper-graph regression model.

Different Sizes of Training Data for Regression—We also evaluate the performance of our method on different training data sizes for clinical score regression task. We change the size of training data (the data with known clinical labels and clinical scores) and compare the classification performance of our method with the modified static hyper-graph method from Gao et. al.'s method. Fig. 6. right shows the ADAS-COG clinical score errors measured by the Correlation Coefficients (CC) by all competing methods on different data size. Fig. 6. left shows the MMSE clinical score regression error in terms of Correlation Coefficients (CC) by all competing methods on different data size. The performance of our method decreases gracefully when the training dataset size decreases. However, the competing method decreases dramatically with training dataset decreases. This result suggests that our method is able to learn a more optimal hyper-graph model than the

conventional method (Gao et. al.) after using the label and clinical score information in the hyper-graph learning and regression process.

Discussion—The better performance can be dedicated to the following reasons. First, different from conventional hyper-graph methods that learn fixed hyper-graph from imaging data independently from the classification process, our method uses both the imaging data, the clinical data to learn a dynamic hyper-graph to explore the underlying intrinsic relationship among multiple modality data in a semi-supervised way. Thus, the solution space for our method is more optimized on clinical data and imaging data compared with conventional hyper-graph methods. Second, the hyper-graph is capable of modeling complex relationship among different subjects and different modality imaging data, therefore, the employed hyper-graph structure is superior on formulating the joint relationship among multiple vertices compared with simple graph model.

4 Conclusions

In this work, we proposed a dynamic hyper-graph method for joint learning, classification, and regression on multiple classification problems (MCI/NC/AD) using multiple modal imaging data. We use a dynamic hyper-graph learned from multi-modal imaging data, the clinical labels and clinical scores on both training dataset and testing dataset to guide the classification of new subjects. We evaluated our method on MCI-C/MCI-NC/NC/AD diagnosis and clinical score prediction. The results demonstrated that our method could improve the performance significantly compared to the conventional multi-modality static hyper-graph learning method on both diagnosis and clinical score prediction.

References

1. Zhu X, Suk H, Thung K, Zhu Y, Wu G, Shen D. Joint Discriminative and Representative Feature Selection for Alzheimer's Disease Diagnosis. International Workshop on Machine Learning in Medical Imaging. 2016
2. Zhang D. Multi-modal multi-task learning for joint feature prediction of multiple regression and classification variables in Alzheimer's Disease. Neuroimage. 2012; 59:895–907. [PubMed: 21992749]
3. Suk H, Shen D. Hierarchical Feature Representation and Multimodal Fusion with Deep Learning for AD/MCI Diagnosis. Neuroimage. 2015; 1
4. Dong P, Guo Y, Shen D, Wu G. Patch-Based Techniques in Medical Imaging. Springer; 2015. Multi-atlas and Multi-modal Hippocampus Segmentation for Infant MR Brain Images by Propagating Anatomical Labels on Hypergraph; 188–196.
5. Gao Y, Wee C, Kim M, Giannakopoulos P, Montandon M, Haller S. , et al. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015. Springer; 2015. MCI Identification by Joint Learning on Multiple MRI Data; 78–85.
6. Boyd S, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends® in Machine Learning. 2011; 3:1–122.
7. SMITH A. Nonparametric Regression on a Graph. Journal of Computational and Graphical Statistics. 2011; 20
8. Zhu Y. Convolutional Sparse Coding for Trajectory Reconstruction. TPAMI. 2015; 37:529–540.
9. Zhu Y, Cox M, Lucey S. 3D motion reconstruction for real-world camera motion. CVPR. 2011:1–8.
10. Zhu Y, Huang D, Torre FDL, Lucey S. Complex non-rigid motion 3D reconstruction by union of subspaces. CVPR. 2014:23–34.

11. Thompson P, et al. Tracking Alzheimer's Disease. *Annals of the New York Academy of Sciences*. 2007; 1097:183–214. [PubMed: 17413023]
12. Risacher S, Saykin A. Neuroimaging Biomarkers of Neurodegenerative Diseases and Dementia. *Seminars in neurology*. 2013; 33:386–416. 11/14. [PubMed: 24234359]
13. Reisberg B, Ferris S, Kluger A, Franssen E, Wegiel J, De Leon MJ. Mild cognitive impairment (MCI): a historical perspective. *International Psychogeriatrics*. 2008; 20:18–31. [PubMed: 18031593]
14. Zhang D, Wang Y, Zhou L, Yuan H, Shen D, I Alzheimer's Disease Neuroimaging. Multimodal Classification of Alzheimer's Disease and Mild Cognitive Impairment. *NeuroImage*. 2011; 55:856–867. 01/12. [PubMed: 21236349]
15. Ganguli M, Dodge H, Shen C, DeKosky ST. Mild cognitive impairment, amnesic type An epidemiologic study. *Neurology*. 2004; 63:115–121. [PubMed: 15249620]
16. Guo X, Wang Z, Li K, Li Z, Qi Z, Jin Z, et al. Voxel-based assessment of gray and white matter volumes in Alzheimer's disease. *Neuroscience letters*. 2010; 468:146–150. [PubMed: 19879920]
17. Wang Z, Adeli E, Zhu Y, Zu C, Nie F, Shen D, Wu G. Progressive Graph-Based Transductive Learning for Multi-Modal Classification of Alzheimer's Disease. *MICCAI*. 2016

Fig. 1.

The comparison between conventional hyper-graph learning (left) and our joint hyper-graph learning, classification, and regression method (right) using multiple modal imaging data. The main improvements in our method are (1) hyper-graph across multiple modalities and clinical scores, (2) refinement of hyper-edges dynamically among labelled subjects and unlabeled subjects (designated by the dash ellipses, instead of solid ellipses in the left).

Fig. 2.

Comparison of conventional hyper-graph and our method. The conventional hyper-graph is learned using the image feature only and the classification process is performed by propagating the clinical labels on the fixed hyper-graph in conventional hyper-graph method (shown in the left). Our method optimizes both the hyper-graph structure, the subject labels and the subject clinical score dynamically and simultaneously (shown in the right).

Fig. 3. Multiple class (NC/MCI-C/MCI-NC/AD) classification performances using multiple modality imaging data by ensemble SVM, Static Hyper-Graph and Our Dynamic Hyper-Graph model.

Fig. 4.

Classification performance of our method vs. Gao et. al.'s method on various sizes of training data. Our method is able to achieve high performance using small amount of labeling data in a semi-supervised manner.

Fig. 5. Clinical score regression performance using multiple modality imaging data by ensemble SVR, Static Hyper-Graph and Our Dynamic Hyper-Graph model. We measure the regression error by the Correlation Coefficients (CC) and Root Mean Square Error (RMSE) term.

Fig. 6.

Regression performance of our method vs. modified hyper-graph regression method from Gao et. al.'s method on various sizes of training data. Our method is able to achieve high performance using small amount of labeling data in a semi-supervised manner.