

Medical Image Retrieval Using Multi-graph Learning for MCI Diagnostic Assistance

Yue Gao¹, Ehsan Adeli-M.¹, Minjeong Kim¹,
Panteleimon Giannakopoulos², Sven Haller³, and Dinggang Shen¹

¹ Department of Radiology and BRIC,
University of North Carolina at Chapel Hill, NC 27599, USA
² Division of Psychiatry, Geneva University Hospitals, Switzerland
³ Department of Neuroradiology, University Hospitals of Geneva
and Faculty of Medicine of the University of Geneva, Switzerland

Abstract. Alzheimer’s disease (AD) is an irreversible neurodegenerative disorder that can lead to progressive memory loss and cognition impairment. Therefore, diagnosing AD during the risk stage, a.k.a. Mild Cognitive Impairment (MCI), has attracted ever increasing interest. Besides the automated diagnosis of MCI, it is important to provide physicians with related MCI cases with visually similar imaging data for case-based reasoning or evidence-based medicine in clinical practices. To this end, we propose a multi-graph learning based medical image retrieval technique for MCI diagnostic assistance. Our method is comprised of two stages, the query category prediction and ranking. In the first stage, the query is formulated into a multi-graph structure with a set of selected subjects in the database to learn the relevance between the query subject and the existing subject categories through learning the multi-graph combination weights. This predicts the category that the query belongs to, based on which a set of subjects in the database are selected as candidate retrieval results. In the second stage, the relationship between these candidates and the query is further learned with a new multi-graph, which is used to rank the candidates. The returned subjects can be demonstrated to physicians as reference cases for MCI diagnosing. We evaluated the proposed method on a cohort of 60 consecutive MCI subjects and 350 normal controls with MRI data under three imaging parameters: T1 weighted imaging (T1), Diffusion Tensor Imaging (DTI) and Arterial Spin Labeling (ASL). The proposed method can achieve average 3.45 relevant samples in top 5 returned results, which significantly outperforms the baseline methods compared.

1 Introduction

Alzheimer’s disease (AD) is an irreversible neurodegenerative disorder found in elderly over 65 years of age, accounts for 60% to 80% of age-related dementia cases [10]. AD can lead to progressive memory loss and cognition impairment. The number of AD patients has reached 26.6 million and is expected to double in the next two decades [1]. Therefore, accurate diagnosis of AD during the risk stage, a.k.a. Mild Cognitive Impairment (MCI), is important. In recent years, extensive research efforts have been dedicated to MCI identification using different imaging data, such as MRI [2], positron emission tomography (PET) [5], and Cerebrospinal fluid (CSF) [3], which aims to

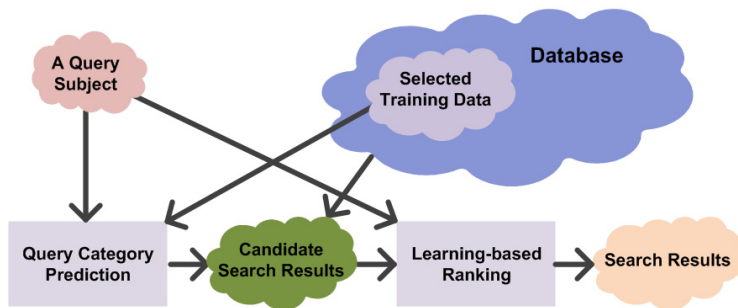


Fig. 1. The framework of the proposed medical image retrieval method for MCI diagnostic aid.

provide the physicians with brain structural and functional information of the brain for the patient’s condition.

In addition to automatic MCI classification based on imaging data, providing physicians with cases of similar visual appearances and corresponding treatment records can indeed facilitate clinical decisions. It can supply references for physicians to perform case-based reasoning or evidence-based medicine with even more confidence. Therefore, medical image retrieval has attracted much more attention in recent years [11, 12, 7, 4]. We notice that most works target at retrieving similar objects in image content [11, 12] or the same imaging modalities [7, 4] for a given query image. However, for the purpose of MCI diagnostic aid, it should retrieve subjects from the database with similar brain patterns across all the imaging modalities. To assist MCI diagnosis, our goal is to identify similar brain patterns from imaging data. Given a **query**, as a set of imaging data belonging to one subject, our goal is to find the subjects with similar brain patterns from a **database**. The database contains subjects with clinical treatment records and the same imaging data types as the query.

Thus, in this work, we propose a medical image retrieval technique for application to MCI diagnosis assistance. The proposed method is composed of two main stages: *query category prediction* for candidate selection and *ranking*, as shown in Figure 1. In the first stage, we locate the most relevant subjects from the database to the query subject. In this step, a group of subjects from the database¹ are selected and then used to predict the category of the query subject, in a supervised manner. In this case, the categories are MCI and NC, while this framework could be adopted for all other neurodegenerative brain diseases. Given the selected supervised data from the database, for each imaging data modality, a graph is constructed based on the pair-wise subject distances. Then, graphs from different imaging data are combined into a single multi-graph to formulate the overall similarity, which includes both the selected subjects and the query. A learning procedure is conducted on the multi-graph to jointly estimate the relevance of the query subject to all predefined categories and to learn the optimal combination weights of the multiple graphs, which leads to a category prediction for the query subject. Then, all the subjects belonging to the same category are selected as candidate subjects.

¹ Note that, for all the subjects in the database, we have their medical records and therefore the category information for these subjects are available.

In the second stage, all these candidate subjects and the query subject are again modeled in a new multi-graph, which is built by combining the obtained weights from the previous stage of query category prediction, but only on the selected candidate subjects. A learning procedure on the graph reveals the relevance of these candidates to the query subject, which would lead to a ranking of candidate subjects from the database on how relevant they are compared to the query subject. The proposed method is evaluated on a cohort of 60 consecutive MCI subjects and 350 normal controls (NC) containing MRI data with three imaging parameters, including T1 weighted imaging (T1), Diffusion Tensor Imaging (DTI), and Arterial Spin Labeling (ASL).

2 The Method

In principle, our methods works in two-stage, i.e., a query category prediction stage to select the potential candidates along with the original query for retrieval and the ranking stage built based upon a multi-graph ranking scheme to find similar brain patterns to the query.

Query Category Prediction: Given the query imaging data, the first step is to predict the query category, which estimates the relevant data category in the database.

Graph Construction- Here, a group of subjects with annotations from the database are selected, for the purpose of query category prediction. Let \mathcal{Q} denote the query subject and $\mathcal{M} = \{M_1, M_2, \dots, M_N\}$ denote the N selected training subjects in the database (N is set equal to 100 in our experiments). The relationship among these subjects (with multimodal imaging data) could be formulated in a multi-graph structure as below. Let N_{Mod} denote the number of imaging modalities in total. For the i^{th} imaging data, a graph $\mathcal{G}_i = \{\mathcal{V}_i, \mathcal{E}_i, \mathbf{W}_i\}_{N+1}$ is constructed by using all the N annotated subjects and the query \mathcal{Q} , where \mathcal{V}_i is the vertex set containing $N + 1$ samples, and \mathcal{E}_i is the edge set. $\mathcal{E}_i(v_s, v_t)$ is the edge connecting the s^{th} and the t^{th} vertices in \mathcal{G}_i corresponding to the edge weight $\mathbf{W}_i(v_s, v_t)$. Here, $\mathbf{W}_i(v_s, v_t)$ is defined as:

$$\mathbf{W}_i(v_s, v_t) = \exp\left(-\frac{d^2(v_s, v_t)}{\sigma_i^2}\right), \quad (1)$$

where $d(v_s, v_t)$ is the distance between v_s and v_t , which is calculated as the Euclidean distance between two corresponding features (feature extraction will be introduced in Section 3). Let $\Theta_i = \mathbf{D}_i^{-1/2} \mathbf{W}_i \mathbf{D}_i^{-1/2}$ denote the normalized weight matrix of the graph \mathcal{G}_i , where \mathbf{D}_i is a diagonal matrix defined as $\mathbf{D}_i(s, s) = \sum_{v_t} \mathbf{W}_i(v_s, v_t)$.

Objective Function and Solution- Multi-graph learning has been applied in many applications, such as video annotation [9] and affective image analysis [13]. Given the N_{Mod} graphs generated from multimodal data, the query category prediction task can be formulated as a multi-graph learning framework, where the cost function is:

$$\mathcal{Q}(\mathbf{F}, \boldsymbol{\omega}) = \left\{ \sum_{i=1}^{N_{\text{Mod}}} \omega_i \boldsymbol{\Omega}_i(\mathbf{F}) + \mu \mathcal{R}(\mathbf{F}) + \eta \|\boldsymbol{\omega}\|_2^2 \right\} \quad s.t. \quad \sum_{i=1}^{N_{\text{Mod}}} \omega_i = 1, \quad (2)$$

where $\boldsymbol{\omega} = \{\omega_1, \omega_2, \dots, \omega_{N_{\text{Mod}}}\}$ is the weighting parameter for each graph, $\boldsymbol{\Omega}_i$ is the regularizer on the i^{th} graph, \mathbf{F} is the relevance matrix, $\mathcal{R}(\mathbf{F})$ is the empirical loss,

$\|\boldsymbol{\omega}\|_2^2 = \sum_{i=1}^{N_{\text{Mod}}} \omega_i^2$ is the squared ℓ_2 norm of $\boldsymbol{\omega}$, and μ and η are the parameters to balance different components in the cost function. ω_i can be initialized as $1/N_{\text{Mod}}$. We define the regularizer term, $\boldsymbol{\Omega}_i$, as:

$$\boldsymbol{\Omega}_i = \frac{1}{2} \sum_{v_s, v_t} \mathbf{W}_i(v_s, v_t) \left\| \frac{\mathbf{F}(v_s, \cdot)}{\sqrt{\mathbf{D}_i(v_s, v_s)}} - \frac{\mathbf{F}(v_t, \cdot)}{\sqrt{\mathbf{D}_i(v_t, v_t)}} \right\|^2. \quad (3)$$

$\mathcal{R}(\mathbf{F})$ is defined as $\mathcal{R}(\mathbf{F}) = \sum \|\mathbf{F}_i - \mathbf{Y}_i\|^2$, where $\mathbf{Y} \in \mathbb{R}^{n \times 2}$ is the labeled matrix and initialized as follows. If the a^{th} vertex belongs to the k^{th} category, $\mathbf{Y}(a, k) = 1$; otherwise $\mathbf{Y}(a, k) = 0$.

The learning task on \mathbf{F} and $\boldsymbol{\omega}$ can be written as

$$\arg \min_{\mathbf{F}, \boldsymbol{\omega}} \mathcal{Q}(\mathbf{F}, \boldsymbol{\omega}), \quad s.t. \sum_{i=1}^{N_{\text{Mod}}} \omega_i = 1. \quad (4)$$

where the relevance among vertices and the multi-graph combination weights can be simultaneously optimized.

To solve the above optimization task, we alternately optimize \mathbf{F} and $\boldsymbol{\omega}$, respectively. In the first step, $\boldsymbol{\omega}$ is fixed and \mathbf{F} is optimized, which leads to the following problem:

$$\arg \min_{\mathbf{F}} \left\{ \sum_{i=1}^{N_{\text{Mod}}} \omega_i \boldsymbol{\Omega}_i(\mathbf{F}) + \mu \mathcal{R}(\mathbf{F}) \right\}. \quad (5)$$

According to [14], it can be efficiently solved in an iterative procedure by

$$\mathbf{F}(t+1) = \frac{1}{1+\mu} \sum_{i=1}^{N_{\text{Mod}}} \omega_i \boldsymbol{\Theta}_i \mathbf{F}(t) + \frac{\mu}{1+\mu} \mathbf{Y}, \quad (6)$$

where $\mathbf{F}(t+1)$ is the \mathbf{F} in the $(t+1)^{\text{th}}$ iteration and $\mathbf{F}(0) = \mathbf{Y}$. This procedure is iterated until the convergence.

In the second step, \mathbf{F} is fixed and $\boldsymbol{\omega}$ is optimized. The objective function in Eq. (4) can be rewritten as:

$$\arg \min_{\boldsymbol{\omega}} \left\{ \sum_{i=1}^{N_{\text{Mod}}} \omega_i \boldsymbol{\Omega}_i(\mathbf{F}) + \eta \|\boldsymbol{\omega}\|_2^2 \right\}, \quad s.t. \sum_{i=1}^{N_{\text{Mod}}} \omega_i = 1. \quad (7)$$

which can be solved using the Lagrangian method.

The above alternating optimization process is repeated until convergence. Here let $\mathbf{F}(q, \cdot)$ denote the corresponding relevance vector of \mathcal{Q} . Then, \mathcal{Q} can be classified into a predefined category by $\arg \max(\mathbf{F}(q, \cdot))$. Then, all the data belonging to the same category in the database are selected as the candidate retrieval results, denoted by $\boldsymbol{\Gamma} = \{S_1, S_2, \dots, S_{N_c}\}$. Here, each S_i is one candidate subject belonging to the query's category and N_c is the number of candidates.

Learning-based Ranking: Although the subjects in $\boldsymbol{\Gamma}$ are relevant to the query based on the category information, they could still be different from the point-of-view of imaging appearance or patterns present in the image. In this step, we need to rank all these

selected candidates based on the imaging data to retrieve the most relevant subjects. Here, the N_c candidates Γ and the query \mathcal{Q} are formulated in a new multi-graph structure and N_{Mod} graphs $\{\hat{\mathcal{G}}_1, \hat{\mathcal{G}}_2, \dots, \hat{\mathcal{G}}_{N_{\text{Mod}}}\}$ are generated, using these $N_c + 1$ subjects in a similar way as in the query category prediction step. Since the optimized combination weights ω are previously learned, in this step we only need to estimate the relevance between each subject in the candidate set and the query. So, the the optimization task on $\hat{\mathbf{f}}$ can be written as

$$\arg \min_{\hat{\mathbf{f}}} \hat{\mathcal{Q}}(\hat{\mathbf{f}}) = \arg \min_{\hat{\mathbf{f}}} \left\{ \sum_{i=1}^{N_{\text{Mod}}} \omega_i \hat{\Omega}_i(\hat{\mathbf{f}}) + \hat{\lambda} \hat{\mathcal{R}}(\hat{\mathbf{f}}) \right\}, \quad (8)$$

where $\hat{\mathbf{f}}$ is the to-be-learned relevance vector, $\hat{\Omega}_i$ is the graph regularizer on the i^{th} graph $\hat{\mathcal{G}}_i$, $\hat{\mathcal{R}}(\hat{\mathbf{f}}) = \|\hat{\mathbf{f}} - \hat{\mathbf{y}}\|^2$ is the empirical loss, and $\hat{\lambda}$ is the parameter to balance the graph regularizer and the empirical loss. Here, $\hat{\mathbf{y}}$ is the labeled vector, in which all the elements are 0 except that the value corresponding to the query is 1. Similar to the solution of Eq. (5), we can solve $\hat{\mathbf{f}}$ by iterating Eq. (9) until convergence, where $\hat{\mathbf{f}}(0) = \mathbf{y}$ and $\hat{\Theta}_i = \hat{\mathbf{D}}_i^{-1/2} \hat{\mathbf{W}}_i \hat{\mathbf{D}}_i^{-1/2}$ is the normalized graph Laplacian of $\hat{\mathcal{G}}_i$.

$$\hat{\mathbf{f}}(t+1) = \frac{1}{1 + \hat{\lambda}} \sum_{i=1}^{N_{\text{Mod}}} \omega_i \hat{\Theta}_i \mathbf{f}(t) + \frac{\hat{\lambda}}{1 + \hat{\lambda}} \hat{\mathbf{y}}. \quad (9)$$

Based on $\hat{\mathbf{f}}$, all the candidates can be ranked in a descending order to demonstrate the retrieval results.

3 Validation

Experimental Settings: To evaluate the performance of the proposed medical imaging retrieval approach for MCI diagnosis assistance, a dataset containing 60 MCI patients and 350 normal controls was collected. Generally, T1, DTI and ASL were collected for each subject. For T1, the anatomical automatic labeling atlas, which are parcellated with 90 predefined regions-of-interest (ROIs), was registered to the native space of each subject. The white matter (WM) and gray matter (GM) tissue volumes in these ROIs are calculated to generate a 180-dimensional feature vector, (i.e., 90 WM and 90 GM features). For DTI and ASL, after the 90 ROIs parcellation, two 90×90 connectivity matrices are computed. Then, the local clustering coefficients are calculated to quantify the cliquishness of the nodes [8], which generate two 90-dimensional feature vectors for these two types of imaging data, respectively. In the retrieval task, 50 MCI patients and 50 NCs are employed as the queries. In the query category prediction procedure, the training samples are randomly selected. The parameters in the proposed method are selected by a grid search via a 5-fold cross validation in the category prediction stage, and kept consistent in the ranking stage.

The evaluation criteria include #Correct@K and Normalized discounted cumulative gain value in the top K results (NDCG@K) [6]. Here, #Correct@K measures the average number of correct results in top K returned results. NDCG measures the performance of a ranking list based on the graded relevance of the results and ranges from

0 to 1. Generally, the most relevant results should be ranked at top-most positions leading to a higher NDCG value. NDCG has been commonly used in information retrieval to evaluate the performance of web search engines. In our experiments, the Mini Mental State Examination (MMSE) score, the most commonly used test for complaints of memory problems and dementia diagnosis in clinical practice, is used as the criteria to define the subject relevance. Generally, closer MMSE scores indicate high similarity between two subjects. Given a query \mathcal{Q} , the relevance of \mathcal{T} to \mathcal{Q} is generated by:

$$r(\mathcal{Q}, \mathcal{T}) = \begin{cases} 3 & \text{if } \mathcal{Q} \text{ and } \mathcal{T} \text{ belong to the same category and } dif_{MMSE}(\mathcal{Q}, \mathcal{T}) = 0 \\ 2 & \text{if } \mathcal{Q} \text{ and } \mathcal{T} \text{ belong to the same category and } dif_{MMSE}(\mathcal{Q}, \mathcal{T}) = 1 \\ 1 & \text{if } \mathcal{Q} \text{ and } \mathcal{T} \text{ belong to the same category and } dif_{MMSE}(\mathcal{Q}, \mathcal{T}) > 1 \\ 0 & \text{if } \mathcal{Q} \text{ and } \mathcal{T} \text{ do not belong to the same category} \end{cases},$$

where $dif_{MMSE}(\mathcal{Q}, \mathcal{T})$ is the difference between the MMSE scores of \mathcal{Q} and \mathcal{T} .

Three baseline methods are implemented for comparison:

- Mean distance. To compare two subjects, the distances by using different imaging modalities are calculated first, and the mean distance is used.
- Minimal distance. To compare two subjects, the distances by using different imaging modalities between these two subjects are calculated and the minimal distance is selected as the pair-wise subject similarity.
- Graph learning fusion. In this method, a graph is first constructed for each imaging modality, using all the subjects in the database together with the query, based on the distances from that imaging modality. The learning is conducted on each graph to estimate the pair-wise subject relevance. The relevance from different modalities are combined together as the similarity between the two subjects. In this method, the category information in the database is not used.

Experimental Results: Figure 2 shows the performance comparisons of different methods, when using different imaging data and all the three types of imaging data in terms of #Correct@K and NDCG@K, respectively. It is noted that in many cases, the performance curves of the three baseline methods overlap each other, such as in all the #Correct@K curves. We can observe from the results that the proposed method can achieve better results compared to three baseline methods. For example, when all the three imaging data are employed, the proposed method can return 3.45 relevant subjects in top 5 returned results and 6.90 relevant subjects in top 10 returned results. The proposed method can improve #Correct@10 by 1.90 ($p = 1.2e^{-5}$), 1.89 ($p = 4.5e^{-5}$), and 1.84 ($p = 1.6e^{-5}$) compared with the three baseline methods, respectively. The improvements on NDCG@10 are 0.10 ($p = 9.5e^{-3}$), 0.10 ($p = 2.0e^{-2}$), and 0.12 ($p = 1.6e^{-3}$), respectively. The better performance on #Correct@K indicates that the proposed method is able to return more relevant medical records, and a higher NDCG@K value means that highly relevant subjects are located at top-most positions. The better performance comes from the accurate query category prediction and the following ranking procedure, which can precisely locate possible relevant subjects and then finally provide a fine ranking list. It is possible for the physicians to obtain useful reference from existing patient’s database given the imaging data of the query subject.

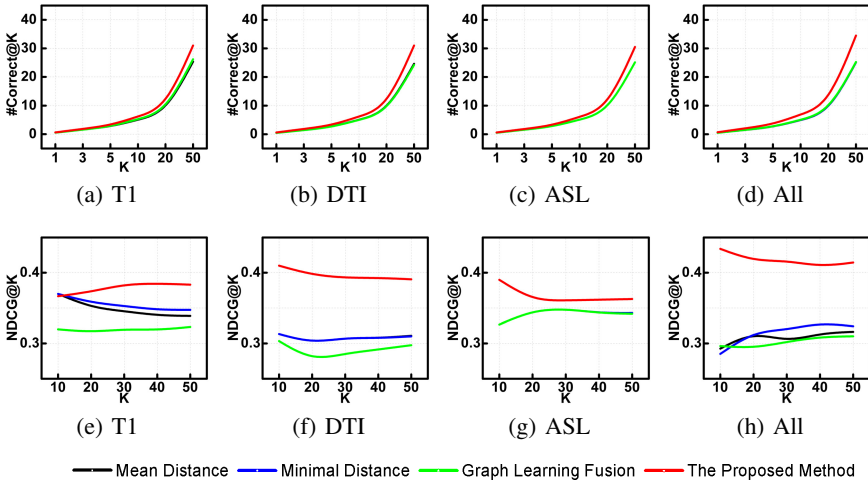


Fig. 2. The retrieval results comparison among different methods using different modalities in terms of #Correct@K and NDCG@K, where (d) and (h) refer to the using of all the three types of imaging data.

Table 1. The comparison among the proposed method using different data combinations.

Data Combination	#Correct@3	#Correct@5	#Correct@10	NDCG@10	NDCG@20
T1	1.830	3.050	6.100	0.390	0.361
DTI	1.860	3.100	6.200	0.367	0.373
ASL	1.860	3.100	6.200	0.410	0.397
T1+DTI	1.830	3.050	6.100	0.350	0.368
T1+ASL	1.650	2.750	5.500	0.337	0.346
DTI+ASL	1.860	3.100	6.200	0.340	0.334
T1+DTI+ASL	2.070	3.450	6.900	0.417	0.417

We further investigate the impact of multimodal data in the retrieval task. Table 1 demonstrates the comparisons of our method when different multimodal data combinations are employed. We can observe from the results that more imaging data from the query can lead to better results. The best performance comes from the using of T1, DTI and ASL together, which can be dedicated to the fact that different imaging data can represent the subject through different views and can be complementary for each other. This result also suggests that having more imaging data for a patient can be helpful for diagnosis assistance.

4 Conclusion

In this paper, we proposed to retrieve relevant medical records using imaging data to assist MCI diagnosis. Different from existing medical image retrieval tasks, which focus

on similar image object content or the same type of imaging modalities as the retrieval target, our objective is to find disease-related imaging data, with similar brain patterns in the same imaging modalities. Our two-stage retrieval method is able to locate candidate results and then rank them to obtain relevant imaging data for the query. The results show that the top returned subjects are highly relevant to the query, which demonstrates the effectiveness of the proposed method on MCI diagnosis assistance.

References

- [1] Brookmeyer, R., Johnson, E., Ziegler-Graham, K., Arrighi, M.H.: Forecasting the global burden of Alzheimer's disease. *Alzheimer's & Dementia: the Journal of the Alzheimer's Association* 3(3), 186–191 (2007)
- [2] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M.O., Chupin, M., Benali, H., Colliot, O.: Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *NeuroImage* 56(2), 766–781 (2011)
- [3] Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M.: The Alzheimer's Disease Neuroimaging Initiative: CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *The Journal of Neuroscience* 30(6), 2088–2101 (2010)
- [4] de Herrera, A.G.S., Kalpathy-Cramer, J., Fushman, D.D., Antani, S., Müller, H.: Overview of the imageclef 2013 medical tasks. In: *Working Notes of CLEF 2013*, pp. 1–15 (2013)
- [5] Hinrichs, C., Singh, V., Mukherjee, L., Xu, G., Chung, M.K., Johnson, S.C.: Spatially augmented l₁-boosting for AD classification with evaluations on the ADNI dataset. *NeuroImage* 48(1), 138–149 (2009)
- [6] Järvelin, K., Kekäläinen, J.: Ir evaluation methods for retrieving highly relevant documents. In: *ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 41–48. ACM (2000)
- [7] Müller, H., de Herrera, A.G.S., Kalpathy-Cramer, J., Demner-Fushman, D., Antani, S., Eggel, I.: Overview of the imageclef 2012 medical image retrieval and classification tasks. In: *CLEF (Online Working Notes/Labs/Workshop)*, pp. 1–16 (2012)
- [8] Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52(3), 1059–1069 (2010)
- [9] Wang, M., Hua, X.S., Hong, R., Tang, J., Qi, G.J., Song, Y.: Unified video annotation via multigraph learning. *IEEE Transactions on Circuits and Systems for Video Technology* 19(5), 733–746 (2009)
- [10] Ye, J., Wu, T., Li, J., Chen, K.: Machine learning approaches for the neuroimaging study of Alzheimer's disease. *Computer* 44(4), 99–101 (2011)
- [11] Zhang, X., Liu, W., Dundar, M., Badve, S., Zhang, S.: Towards large-scale histopathological image analysis: Hashing-based image retrieval. *IEEE Transactions on Medical Imaging* 34(2), 496–506 (2015)
- [12] Zhang, X., Yang, L., Liu, W., Su, H., Zhang, S.: Mining histopathological images via composite hashing and online learning. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *MICCAI 2014, Part II. LNCS*, vol. 8674, pp. 479–486. Springer, Heidelberg (2014)
- [13] Zhao, S., Yao, H., Yang, Y., Zhang, Y.: Affective image retrieval via multi-graph learning. In: *ACM International Conference on Multimedia*, pp. 1025–1028. ACM (2014)
- [14] Zhou, D., Bousquet, O., Lal, T.N., Weston, J., Schölkopf, B.: Learning with local and global consistency. *Advances in Neural Information Processing Systems* 16(16), 321–328 (2004)