Medical Image Retrieval based on Combination of Visual Semantic and Local Features

Menglin Wu[1, 2], Quansen Sun[2] and Jin Wang[3]

[1]*College of Electronics and Information Engineering, Nanjing University of Technology, Nanjing 211816*

[2]*School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094*

[3]*Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing, 210044*

wumenglin@njut.edu.cn, qssun@126.com, wangjin @nuist.edu.cn

Abstract

Content-based medical image retrieval is an important tool for doctors in their daily activity. In this paper, we propose a novel image retrieval framework to combine visual concept and local features. To obtain visual semantic representation of the image, we first construct a graph model by feature distance and density similarity, and then a graph-based semi-supervised learning method is applied to get the membership degree of query images. Meanwhile, the dense SIFT feature of the image patches is extracted and described by bag of visual words as local features. Besides, we design a similarity measurement based on visual concept and local feature rather than using low level features only. We evaluate the proposed algorithm in ImageCLEFmed dataset. The results demonstrate that our method represents the visual semantic of images effectively, and compares favorably to state-of-the-art approaches based on single low level features in retrieval performance.

Keywords: *content-based medical image retrieval; graph-based semi-supervised learning; visual semantic; bag of words*

1. Introduction

With the development of digit imaging technology in medical domain, hospitals and Medical research institutions produce a large amount of digit images every day. The modes of these images are various: X-ray, computer tomography (CT), magnetic resonance imaging (MRI), ultrasonography (US), etc. Medical image retrieval is an important tool for physicians in Clinical diagnosis and medical research [1, 2]. However, traditional methods based on keywords or have many drawbacks such as inaccurate description and labor-intensive manual annotation. As a complementary search approach, content-based medical image retrieval (CBMIR) has been one of the most active research directions in biomedicine field.

In CBMIR system, feature vectors extracted by low level features, is a basis of similarity measurement in searching procedure. Some methods use global features for retrieval, in [3], Gabor feature was extracted for mammogram retrieval. Besides, other approaches are based

on local features. Greenspan, et al., [4] proposed a Gaussian mixture model-Kullback Leibler (GMM-KL) framework for matching and categorizing x-ray images by body regions feature. Lehman, et al., [5] conducted a comparison of texture feature and multi-scale feature by various classifiers in medical image classification and retrieval. Different from direct low level feature extraction, Avni, et al., [6] subdivided images into local blocks at multiple scales, and use bag of features to describe patch-based image content, while another algorithm in [7] improved the this image representation scheme via multiply assignment and visual words weighting of the patches. Fusion features are also applied to discriminate images of diverse classes. Tommasi, et al., [8] adapted an integration of global and local features for medical image annotation. The local features were randomly sampled modified SIFT descriptors and the global features were downscaling raw pixels, then classification was done by support vector machine (SVM) by three alternative strategies. In spite of these methods demonstrate effective result in medical image retrieval, there is a semantic gap between low level features and high level semantic. As illustrated in Figure 1, images with high variances belong to same class, while others look similar but in different classes. For that reason, the low level features cannot be the complete description of image content.

Semi-supervised learning techniques, which attempt to leverage both labeled and unlabeled data, have been proposed [9]. Lu, et al., [14] developed an image retrieval algorithm based on SVM and local preserving projections (LPP), while Li, et al., [15] used label mean as prior knowledge to improve semi-supervised SVM performance. Xiang, et al., [16] also applied a local spline regression method to semi-supervised classification. As a major family of semi-supervised learning, graph-based methods have attracted increasing research attention. Zhu, et al., [10] introduced an approach based on Gaussian random fields and harmonic functions. The local and global consistency method was proposed in [11], which improved the energy function. Wang and Zhang assumed that each sample point can be linearly reconstructed by its neighborhood and the algorithm named linear neighborhood propagation [12]. Tang, et al., [13] improved this algorithm and used for video annotation. For the pre-processing purpose, a manifold contraction algorithm [17] is proposed to get better classification accuracy.

The major contributions of this paper are following: (1) introduce graph-based semi-supervised learning into medical image retrieval framework, and use sample point density similarity to construct affinity matrix for enhancing the effectiveness of visual concept extraction. (2) To solve the problem that low level features are failed to represent the image content, we apply graph model and label propagation method to obtain membership degree of the query image for semantic representation, and design a similarity measurement which combine local feature and visual semantic.

The rest of the paper is organized as follows. In Section 2, we give an outline of graph-based semi-supervised learning framework and propose a label propagation algorithm improved by adding density similarity measure to get the membership degree of concepts. The local feature of the image is described by Section 3, while in Section 4 we design a similarity measurement based on visual concept and local feature. Experiments and analysis are presented in Section 5. Conclusion and future work are given in final section. An algorithm char flow is showed in Figure 2.

a) High visual variability, but belong to the same class
b) High visual similarity, but belong to different class

Figure 1. X-ray Medical Image

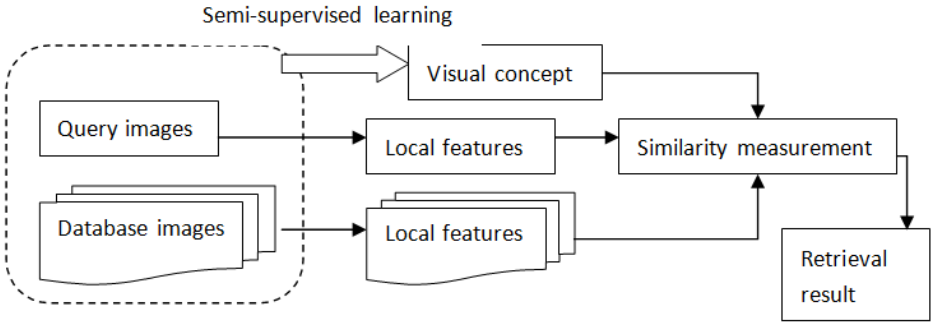


Figure 2. Algorithm Flow Chart

2. Visual Semantic Extraction by Graph-based Semi-supervised Learning

The detail of visual concept extraction is present in this section. First we describe the graph-based semi-supervised learning framework; then we apply the sample density to construct a new affinity matrix and calculate the membership degree of query images as the visual semantic. The iteration algorithm is given in the end of this section.

2.1 Learning Framework

Given a point set $\chi = \{x_1, ..., x_l, x_{l+1}, ...x_n\}$ are n image samples in R^m feature space. The first l points of the sample set are labeled (training images in dataset for our application), while the rest ones are unlabeled (query images for our application). The goal of the learning method is to predict the label of unlabeled point by whole set χ. Let label set $L = \{1, ..., c\}$ be c concept labels, and the first labeled l samples are marked as $y_L = \{y_1^T, y_2^T, ..., y_l^T\}^T$ with $y_i \in R^c$. If a sample y_i has a concept j, y_{ij}=1 and vice-versa. A $n \times c$ matrix $\mathrm{f} = (f_1^T, ..., f_l^T, ... f_n^T)^T$ denotes the membership degree of the each concept for whole sample set. As shown in equation (1), f can be split into two parts after the l-th row, then $\mathrm{f}_L = \mathrm{y}_L$. f_U is the prediction of membership degree for the unlabeled samples, which we can treat as semantic similarity of each medical image class.

$$\mathrm{f} = \begin{pmatrix} \mathrm{f}_L \\ \mathrm{f}_U \end{pmatrix} \quad (1)$$

Construct an undirected connect graph $G=<V,E>$ with the vertex set $V=\chi$. $V=L\bigcup U$, where vertex set $L=\{1,...,l\}$ contains labeled sample points and another set $U=\{l+1,...n\}$ include unlabeled ones. The edge $w_{ij}\in E$ represents the relationship between point i and j:

$$w_{ij}=\exp(-\| x_i-x_j \|^2/(2\sigma^2)) \quad (2)$$

Where σ is bandwidth and x_i is the visual features of sample *i*. According to [13], the concept labels of the unlabeled points are inferred by minimizing the cost function in label propagation procedure:

$$Q(f)=\sum_{1\le i\le j} w_{ij}(f_i-f_j)+\infty\sum_{i\in L}(f_i-y_i) \quad (3)$$

Membership degree prediction function is $f^*=\arg\min_f(Q(f))$. The left term in equation (3) describes the total variation of data labels with respect to neighborhood structure called smoothness term, and the right term called fit term presents the invariant constrain of labeled data. Differentiating $Q(f)$ with respect to f, we have

$$\frac{\partial Q}{\partial f}\Big|_{\mathrm{f}=\mathrm{f}^*}=\mathrm{f}^*-S\mathrm{f}^*+\infty(\mathrm{f}_L^*-y_L)=0 \quad (4)$$

Which can be transform into

$$\begin{aligned}&\mathrm{f}^*=S\mathrm{f}^*\\ &s.t.\quad \mathrm{f}_L^*=\mathrm{y}_L\end{aligned} \quad (5)$$

Here $S=D^{-1}W$ and W is similarity matrix with entry w_{ij}. $D=diag(d_i)$ is a diagonal matrix whose element is $d_i=\sum_{j=1}^{n}w_{ij}$, we can also split Matrix *S* after the *l*-th row and *l*-th column to obtain:

$$S=\begin{bmatrix}S_{LL} & S_{LU}\\ S_{UL} & S_{UU}\end{bmatrix} \quad (6)$$

So equation (5) can convert to:

$$\begin{cases}\mathrm{f}_L=S_{LL}\mathrm{f}_L+S_{LU}\mathrm{f}_U\\ \mathrm{f}_U=S_{UL}\mathrm{f}_L+S_{UU}\mathrm{f}_U\end{cases} \quad (7)$$

Finally, we can get the optimal solution f_U^* with constrain $\mathrm{f}_L^*=\mathrm{y}_L$:

$$\mathrm{f}_U^*=(\mathrm{I}-S_{UU})^{-1}S_{UL}y_L \quad (8)$$

2.2 visual Concept Extract

Since the matrix S defined in (5) is a symmetric matrix, the label information is spread symmetrically, which means the neighborhood points have the same semantic label in feature space. But pairwise similarity measurement in (2) cannot effectively descript the complex structure of the true data set. According to a global consistency assumption, points on the same manifold might have same concept label. Therefore the label propagation in Section 2.1 considers direct contribution from one sample to another through similarity weight, however, it ignores the structural influence.

In this paper, we embed structural similarity model into similarity matrix by sample points' density for improving the performance of our learning method and semantic extraction. Let p_i is the estimation density by Parzen window [18]:

$$p_i = \frac{1}{N_i}\sum_{j=1}^{n} k(x_i - x_j) \tag{9}$$

Where N_i is the number of the sample x_i's neighborhoods, while $k(x)$ is kernel function satisfied $k(x) > 0$ and $\int k(x)dx = 1$, which we choose Gaussian kernel in this paper. The similarity between x_i and x_j is $\tilde{w}_{ij} = w_{ij} \times g_{ij}$, in which density difference g_{ij} is defined:

$$g_{ij} = \exp(-(p_i - p_j)^2 / 2\sigma_p^2) \tag{10}$$

σ_p is the bandwidth parameter controlling the significance of the influence. The formulation (10) indicates that the similarity of two samples not only get smaller by growing the distance of the feature space, but also by the increasing difference of density. Hence, S can be rewritten:

$$\tilde{S} = \tilde{D}^{-1}\tilde{W} \tag{11}$$

Here $\tilde{W}$ is a matrix with element $\tilde{w}_{ij}$, while diagonal matrix $\tilde{D} = diag(\tilde{d}_i)$ and $\tilde{d}_i = \sum_{j=1}^{n} \tilde{w}_{ij}$. Label propagation with density similarity enhances the label information spreading in same structure and suppresses it in different structure.

On the basis of anisotropic diffusion equation [19], formulation (8) equal to an iteration form:

$$\mathrm{f}_U(t+1) = \tilde{S}_{UL}\,\mathrm{f}_L(t) + \tilde{S}_{UU}\,\mathrm{f}_U(t) \tag{12}$$

t is the number of iterations.

For avoiding the image get large membership degree in major class, we define class normalization (13) as class priors to solve this problem. q_j is the image quantity proportions of semantic class j and f_{ij} is the element of row i column j in matrix f_U.

$$f_{ij} = q_j \times f_{ij} / \sum_{j=1}^{c} f_{ij} \tag{13}$$

To sum up, the algorithm for solving the concept membership degree Matrix f_U of unlabeled sample (query images) is following:

Table 1. Algorithm Description

Input: $\chi = \{x_1, ..., x_l, x_{l+1}, ...x_n\}$, $\{x_i\}_{i=1}^{l}$ are labeled by $\{f_i\}_{i=1}^{l} \in R^c$, semantic class $L = \{1, ..., c\}$, and the number of nearest neighborhood k .

1) Calculate the Euclidean distance w_{ij} and difference of density g_{ij} between arbitrary two samples by equation (2) and (10). In practice, we just need to consider k-nearest neighborhood to simplify calculation.
2) Calculate similarity matrix $\tilde{W}$, whose entry is $\tilde{w}_{ij} = w_{ij} \times g_{ij}$.
3) Construct $\tilde{S} = \tilde{D}^{-1} \tilde{W}$ according to (11).
4) Split Matrix $\tilde{S}$ into $\tilde{S}_{LL}$, $\tilde{S}_{LU}$, $\tilde{S}_{UL}$ and $\tilde{S}_{UU}$.
5) Iterate $f_U(t+1) = \tilde{S}_{UL} f_L(t) + \tilde{S}_{UU} f_U(t)$ until convergence, and we get the concept annotations of unlabeled samples $f_U^* = \{f_{l+1}^T, ... f_n^T\}^T$ and $f_i \in R^c (l < i \le n)$, which presents the concept membership degree of c classes.
6) Update f_U^* by class normalization in equation (13).

Output: The membership degree of query images.

3. Local Features Extraction by Patch-based Visual Word

Comparing to normal images, most of the radiographs are gray level images with black background, which have variance imaging mode. Because of noise, artifact and geometric deformable in imaging process, the algorithm based on global features are not suitable for medical image retrieval. On the contrary, we use local features and "bag of features" method to describe the image. The procedures of local feature extraction are shown in Figure 3.

In this paper, we apply SIFT descriptor [20] to extract the local feature, which is robust to scale, rotation and illumination changes so that it can present detail features well. Due to low contrast of the medical images, we cannot use salient point detector directly. Unlike random sampling in [8], we utilize dense SIFT [21] by sampling regular grids of fixed size. Firstly, every image is downscaled to 512×512 and SIFT descriptors of 16×16 pixel patches are computed over a grid with 8 pixels spacing. Secondly, we built the vocabulary and create the visual words using unsupervised K-means clustering algorithm from these descriptors (K is equal to 500 in our paper). At final step, each patch of the image is assigned to the nearest word, and the frequency histogram of the visual words is subsequently normalized for "bag of features" representation of the image.

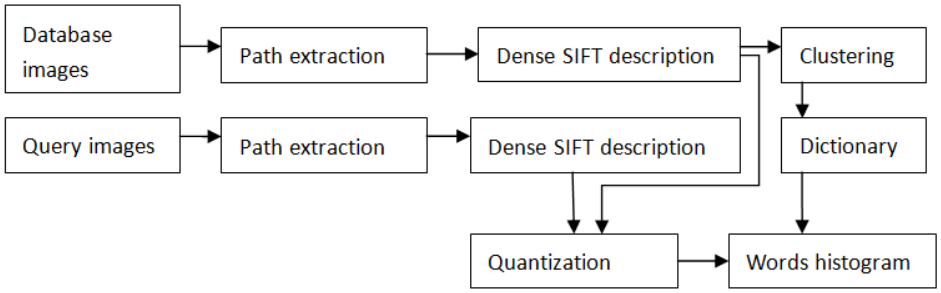


Figure 3. Procedures of Local Feature Extraction

4. Similarity Measurement

In this section, we design a similarity distance metric based on visual concept in Section 2 and local features in Section 3:

$$sim(I,J) = \exp(-d_{JSD}(I,J) / 2\sigma_d^2) \times f_{ij} \quad (14)$$

Where I is a query image and J is a radiograph in dataset respectively. Since the image local features are presented as a kind of histogram, Jensen-Shannon divergence (JSD) is utilized to computer similarity between two visual words histogram:

$$d_{JSD}(I,J) = \sum_{m=1}^{M} I_m \log(\frac{2I_m}{I_m + J_m}) + J_m \log(\frac{2J_m}{J_m + I_m}) \quad (15)$$

In which I_m and J_m are *m-th* bin in histogram. The left multiplier in equation (14) indicates the local feature similarity of two images, which is decreasing with distance increasing, while right multiplier is the membership degree of I relative to the concept class of J. Assuming that the class of the image J is q, and f_U^* is the membership degree matrix calculated in Section 2.2, so f_{ij} is the element of row i column q in matrix f_U^*. The bandwidth σ_d is a tradeoff between visual concept and local features. It means that if σ_d is small, the similarity measurement is more sensitively changed by local features, and vise-versa. Experiment will show the influence of the retrieval performance with σ_d variance in Section 5.3.

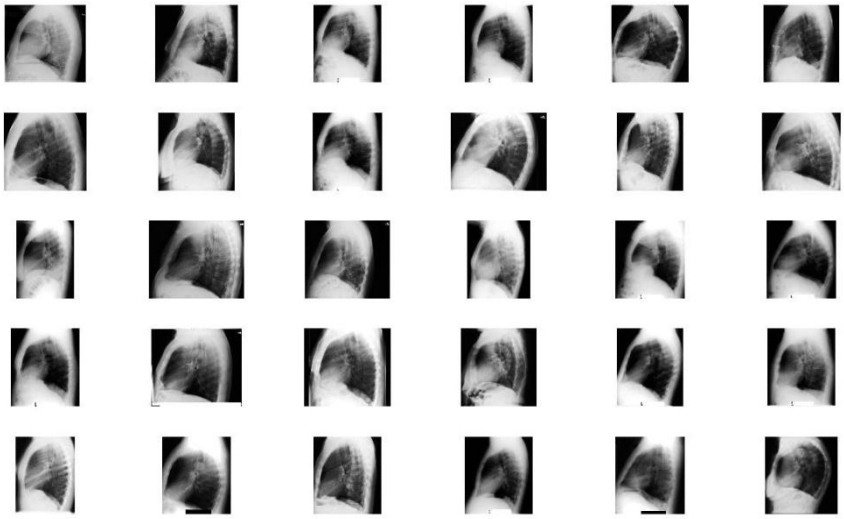
5. Experiments

To evaluate our proposed algorithm for medical image retrieval, we conduct experiments on ImageCLEFmed 2009 dataset [22] from Department of Diagnostic Radiology, Aachen University of Technology. The database contains X-ray images of different ages, genders, view positions and pathologies. Hence, the radiographs quality varies significantly.

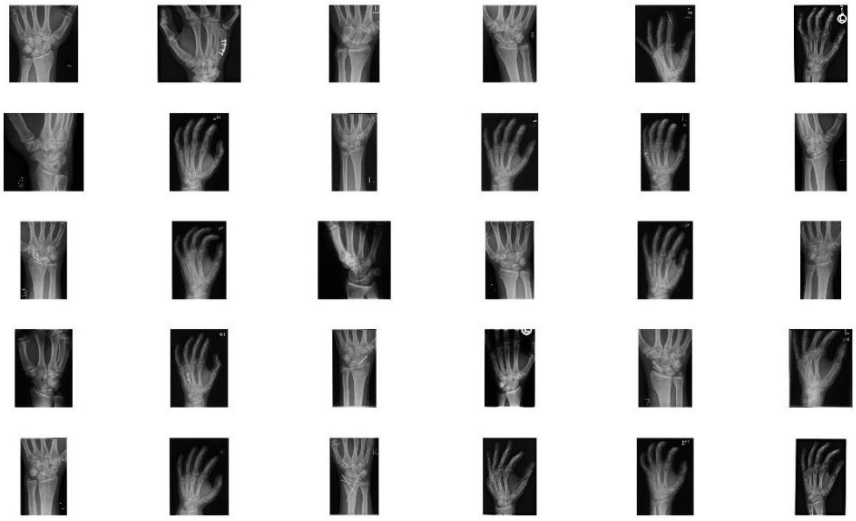
5.1 Results on ImageCLEFmed 2009 Dataset

From this dataset, we choose 4471 images of 50 classes as sub retrieval dataset to ensure that the number of each class is approximately the same, while 1639 radiographs in original testing set are query images. Two retrieval examples is shown in the Figure 4, where the

top-left image is for querying and the rest of ones are searching result. The similarity distance is increasing from left to right and top to bottom.



(a) The query image is a chest image in sagittal view.



(b) The query image is a carpal joint image in coronal view.

Figure 4. Two Examples of Image Retrieval

We compared the proposed retrieval framework with other methods in [3] and [6]. In paper [3], it extracted Gabor features and applied Euclidean distance for similarity measurement referred as Gabor+Euclidean, while patch-based visual words and Jensen-Shannon divergence used in [6] referred as BOW+JSD. We choose parameter σ_d = 0.1 in our method for this comparison. Pension-Recall curve is utilized to test retrieval performance and results are illustrated in Figure 5.

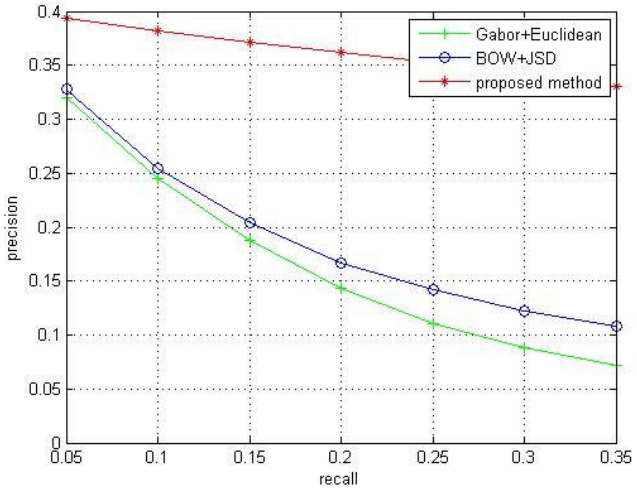


Figure 5. Average Precision-recall Comparison between Three Algorithms

Our method is clearly superior to other two algorithms. More importantly, the precision of the proposed method is decreasing more slowly than other ones as recall number increasing. Therefore, our approach is getting better performance than single low level features. Besides, it also indicates that the local features obtain higher precision than global features for medical image retrieval.

To evaluate the retrieval accuracy of different classes, we use Table 2 to demonstrate the result of 50 classes. Every class is labeled by IRMA [22] code, which contain 'T-D-A-B' four axes: the technical code (T) refers to the image modality; the direction code (D) describes body orientations; the anatomical code (A) models body region examined; the biological code (B) represent the biological system examined.

Table 2. Retrieval Precisions Per-class when the Number of Return Images is 30

category	IRMA code	Gabor	BOW	Proposed method	category	IRMA code	Gabor	BOW	Proposed method
1	11-1-50-0	0.427	0.633	0.881	26	11-1-32-7	0.270	0.288	0.800
2	11-2-50-0	0.197	0.664	0.926	27	11-1-44-7	0.158	0.282	0.386
3	11-1-41-7	0.031	0.073	0.054	28	11-2-44-7	0.117	0.406	0.592
4	11-1-20-7	0.025	0.175	0.218	29	11-4-41-7	0.109	0.248	0.297
5	11-2-23-7	0.022	0.209	0.129	30	11-3-94-7	0.030	0.333	0.030
6	11-1-91-7	0.046	0.146	0.088	31	11-2-43-7	0.080	0.176	0.177
7	11-2-33-7	0.018	0.047	0.100	32	11-2-46-7	0.173	0.380	0.762
8	11-4-21-7	0.249	0.416	0.192	33	11-2-96-7	0.112	0.388	0.667
9	11-1-33-7	0.305	0.267	0.510	34	11-1-95-7	0.057	0.074	0.100
10	11-1-31-7	0.509	0.504	0.573	35	11-1-43-7	0.204	0.280	0.434
11	11-2-31-7	0.082	0.175	0.320	36	11-4-62-6	0.025	0.100	0.025
12	11-1-80-7	0.145	0.281	0.244	37	11-3-62-6	0.027	0.120	0.300
13	11-1-94-7	0.043	0.224	0.368	38	11-4-61-6	0.033	0.092	0.117
14	11-1-92-7	0.260	0.514	0.808	39	11-3-61-6	0.058	0.167	0.217

15	11-1-46-7	0.125	0.293	0.525	40	11-2-91-7	0.043	0.109	0.158
16	11-2-94-7	0.160	0.316	0.715	41	11-2-93-7	0.055	0.114	0.160
17	11-1-70-4	0.147	0.182	0.221	42	11-2-21-7	0.039	0.084	0.241
18	11-2-92-7	0.167	0.411	0.760	43	11-1-93-7	0.119	0.211	0.285
19	11-2-41-7	0.068	0.104	0.220	44	11-1-45-7	0.018	0.049	0.123
20	11-1-51-7	0.152	0.297	0.613	45	11-4-23-7	0.161	0.481	0.753
21	11-2-32-7	0.043	0.060	0.126	46	11-1-71-4	0.165	0.183	0.195
22	11-4-91-7	0.363	0.501	0.747	47	11-2-95-7	0.022	0.033	0.049
23	11-1-96-7	0.215	0.391	0.704	48	11-1-21-7	0.010	0.167	0.021
24	11-2-42-7	0.116	0.246	0.315	49	11-4-31-7	0.067	0.133	0.030
25	11-1-42-7	0.314	0.512	0.688	50	11-2-45-7	0.016	0.046	0.035

The result indicates that our method is superior to two other approaches in most of classes. We also create the Mean Average Precision (MAP) over all the 50 classes for evaluation. The MAP of the three algorithms is 0.254, 0.26 and 0.374 respectively, when the number of return images is 30.

5.2 Comparison with Other Graph-based Semi-supervised Learning Methods

Since the proposed label propagation algorithm is designed for semantic extraction, we test its retrieval performance by comparing with local and global consistency (LGC) method in [10]. Figure 6 demonstrates the comparison of two methods on average precision of retrieval. To fairly evaluate our method, two approaches employ the same feature extraction and similarity measurement described in this paper.

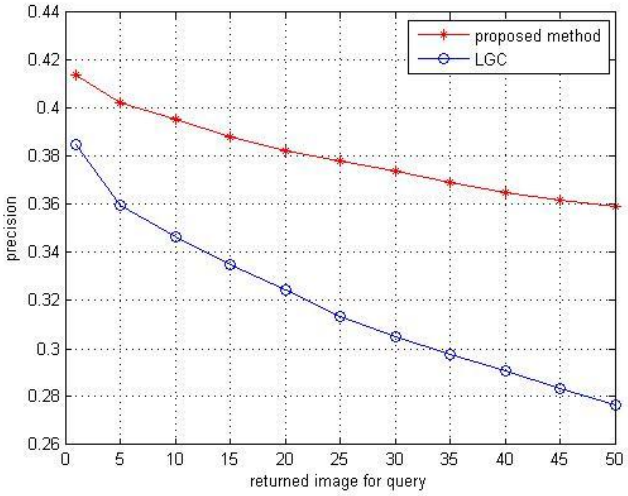


Figure 6. Average Retrieval Precision between Three Algorithms

As a result, the method in our paper obtains higher accuracy than LGC. Since we follow the same steps in retrieval measurement, it is inferred that our approach benefits from the label propagation procedure in semantic extraction. Therefore, embedding structural assumption in label propagation, which means that the similarity between two samples rely on local feature and sample density, can effectively compute the membership degree of the concept classes for query images.

5.3 Results in Different σ_d

In equation (14), bandwidth parameter σ_d is tradeoff between local features and visual concept for similarity measurement. The result changes sharply by the local features varying when σ_d is small. It is indicated that the distance between two images is influence by low level features more than semantics, and vise versa. The average precisions of different σ_d are shown in Figure 7.

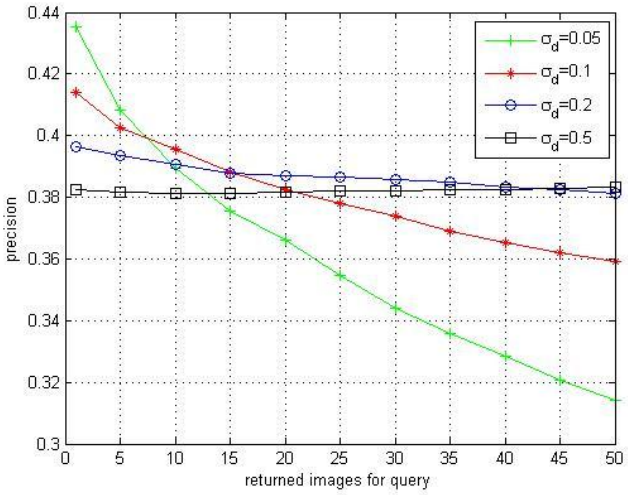


Figure 7. Average Retrieval Precision between Different σ_d

When σ_d = 0.05, initial accuracy is higher than other ones, but the retrieval performance drops down quickly with the number of return image increasing. In contrast to, we get lower accuracy in original, though the precision doesn't change too much. The figure infers that the initial precision is determined by local features while semantic similarity guarantees the stability of the retrieval performance. As aforementioned, we select σ_d = 0.1 in this paper.

6. Conclusions and Future Work

We have demonstrated a novel medical image retrieval framework based on graph-based semi-supervised learning. Unlike the other retrieval method base on low level feature, the proposed algorithm first obtains visual concept by improved learning procedure, and then utilizes dense SIFT and "bag of features" method for local features description. A similarity measurement is also designed combining semantic and local features. Experiment results demonstrate the effectiveness of our method. In the future, we intend to explore the extensions of our algorithms, which allow radiologists to search medical images by interest regions.

Acknowledgements

This work was supported by Nanjing University of Technology Youth Scholars Fund (No. 39709016). The ImageCLEFmed 2009 dataset is courtesy of TM Deserno, Department of Medical Informatics, RWTH Aachen, Germany. It was also supported by the Natural Science Foundation of Jiangsu Province (No. BK2012461) and by the Industrial Strategic Technology Development Program (10041740) funded by the Ministry of Knowledge Economy Korea.

References

[1] Müller H, Michoux N, Bandon D, and Geissbuhler A. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. Medical Informatics. 1,73 **(2004)**.
[2] J. Yao, Z. F. Zhang, S. Antani, R. Long, and G. Thoma. Automatic medical image annotation and retrieval. Neurocomputing. 71, **(2008)**.
[3] C. H. Wei, Y. Li, and C. T. Li. Effective extraction of Gabor features for adaptive mammogram retrieval. Proceedings of IEEE International Conference on Multimedia & Expo,**(2007)** July; Beijing, China
[4] H Greenspan and A. T. Pinhas. Medical image categorization and retrieval for PACS using the GMM-KL framework. IEEE Transactions on Information Technology in Biomedicine. 2, 11**(2007)**.
[5] T. M. Lehmann, M. O. Glüd, T. Deselaers, D. Keysers, H. Schubert, K. Spitzer, H. Ney, and B. B. Wein. Automatic categorization of medical images for content-based retrieval and data mining. Computerized Medical Imaging and Graphics. 2, 29**(2005)**.
[6] U. Avni, E. Konen, M. Sharon, and J. Goldberger. X-ray Categorization and Retrieval on the Organ and Pathology Level, Using Patch-Based Visual Words. IEEE Transactions on Medical Image. 3, 30 **(2011)**.
[7] J. Y. Wang, Y. P. Li, Y. Zhang, C Wang, H. L. Xie, G. L. Chen, and X. Xiao. Bag-of-Features Based Medical Image Retrieval via Multiple Assignment and Visual Words Weighting. IEEE Transactions on Medical Image. 3,30 **(2011)**.
[8] Tatiana T, Francesco O, and Barbara C. Discriminative cue integration for medical image annotation. Pattern Recognition Letters. 15,29 **(2009)**.
[9] X. J. Zhu. Semi-Supervised Learning Literature Survey. Technical Report 1530, Dept. of Computer Sciences, Univ. of Wisconsin Madison **(2005)**.
[10] X. J. Zhu, Z. B. Ghahramani, J Lafferty. Semi-supervised learning using Gaussian fields and harmonic functions. Proceedings of International Conference on Machine Learning, **(2003)** August 21-24; Washington, DC, USA
[11] D. Y. Zhou, O. Bousquet, T. N. Lal, and J. Weston: Learning with Local and Global Consistency. Proceedings of Advances in Neural Information Processing Systems, **(2004)**,December; Belgium
[12] F. Wang, C. S. Zhang. Label Propagation through Linear Neighborhoods. Proceedings of International Conference on Machine Learning. **(2006)**June25-29; Pennsylvania, USA
[13] J. H. Tang, X. S. Hua, G. J. Qi, and X. Q. Wu. Video Annotation Based on Kernel Linear Neighborhood Propagation. IEEE Transactions on Multimedia. 4,10**(2008)**
[14] K. Lu, J. Zhao, and D. Cai. An algorithm for semi-supervised learning in image retrieval. Pattern Recognition. 4,39**(2006)**
[15] Y. F. Li, J. T. Kwok, and Z.-H. Zhou. Semi-supervised learning using label mean. Proceedings of the 26th International Conference on Machine Learning, **(2009)** June 14-18; Montreal, Canada.
[16] S. Xiang, F. Nie, and C. Zhang. Semi-Supervised Classification via Local Spline Regression. IEEE Transactions on Pattern Analysis and Machine Intelligence. 11, 32**(2010)**
[17] E. Hu, S. Chen, X. Yin. Manifold contraction for semi-supervised classification. SCIENCE CHINA Information Sciences, 6, 53**(2010)**
[18] R. Duda, D. Stork, P. Hart. Pattern Classification. 2nd ed. Wiley publishers, New York **(2000)**
[19] J. H. Tang, G. J. Qi, M. Wang, and X. S. Hua. Video semantic analysis based on structure-sensitive anisotropic manifold ranking. Signal Processing. 23,89**(2009)**.
[20] David G. Lowe. Distinctive image features from scale-invariant keypoints. Proceedings of International Journal of Computer Vision. 2,60 **(2004)**.
[21] S Lazebnik, C Schmid, and Jean Ponce. Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories[C]. Proceedings of Computer Vision and Pattern Recognition. **(2006)** June 17-22; New York, USA
[22] T Tommasi, B Caputo, P. Welter, T. M. Deserno. Overview of the CLEF 2009 medical image annotation track. Proceedings of the 10th international conference on Cross-language evaluation forum: **(2010)** September 30 - October 2; Corfu, Greece.

Authors

Menglin Wu

Menglin Wu the B.S. and M.S. degree in the School of Computer Science and Technology, Nanjing University of Science and Technology, china in 2004 and 2006, respectively. Now he is a PhD candidate in the same affiliation, and he is also a lecture in College of Electronics and Information Engineering, Nanjing University of Technology. His research interests mainly include pattern recognition, machine learning and biomedicine image analysis.

QuanSen Sun

Dr. QunSun Sun is a professor in the School of Computer Science and Technology, Nanjing University of Science and Technology. He received his Ph.D. degree in pattern recognition and intelligence system from the same affiliation. His current interests include pattern recognition, image processing, computer vision and data fusion.

Jin Wang

Dr. Jin Wang received the B.S. and M.S. degree in the Electronical Engineering from Nanjing University of Posts and Telecommunications, China in 2002 and 2005, respectively. He received Ph.D. degree in the Ubiquitous Computing laboratory from the Computer Engineering Department of Kyung Hee University Korea in 2010. Now, he is a professor in the Computer and Software Institute, Nanjing University of Information Science and technology. His research interests mainly include routing protocol and algorithm design, performance evaluation and optimization for wireless ad hoc and sensor networks. He is a member of the IEEE and ACM.