Minimizing the Semantic Gap in Biomedical Content-Based Image Retrieval

Haiying Guan[a], Sameer Antani[a], L. Rodney Long[a], and George R. Thoma[a]

[a]Lister Hill National Center for Biomedical Communications
National Library of Medicine, National Institutes of Health
guanh, santani, rlong, gthoma@mail.nih.gov

ABSTRACT

A major challenge in biomedical Content-Based Image Retrieval (CBIR) is to achieve meaningful mappings that minimize the semantic gap between the high-level biomedical semantic concepts and the low-level visual features in images. This paper presents a comprehensive learning-based scheme toward meeting this challenge and improving retrieval quality. The article presents two algorithms: a learning-based feature selection and fusion algorithm and the Ranking Support Vector Machine (Ranking SVM) algorithm. The feature selection algorithm aims to select 'good' features and fuse them using different similarity measurements to provide a better representation of the high-level concepts with the low-level image features. Ranking SVM is applied to learn the retrieval rank function and associate the selected low-level features with query concepts, given the ground-truth ranking of the training samples. The proposed scheme addresses four major issues in CBIR to improve the retrieval accuracy: image feature extraction, selection and fusion, similarity measurements, the association of the low-level features with high-level concepts, and the generation of the rank function to support high-level semantic image retrieval. It models the relationship between semantic concepts and image features, and enables retrieval at the semantic level. We apply it to the problem of vertebra shape retrieval from a digitized spine x-ray image set collected by the second National Health and Nutrition Examination Survey (NHANES II). The experimental results show an improvement of up to 41.92% in the mean average precision (MAP) over conventional image similarity computation methods.

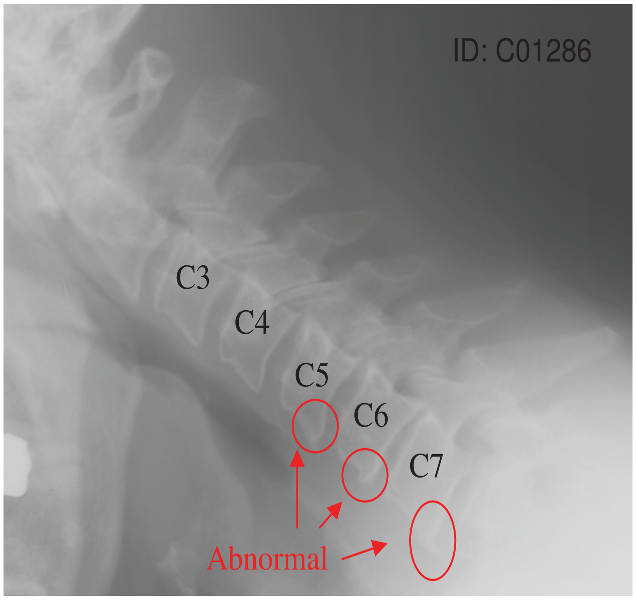
Keywords: Content-Based Image Retrieval, NHANES II database, Ranking SVM, semantic gap, shape retrieval, digital radiography

1. INTRODUCTION

In recent years, there has been a significant increase in the use of images in clinical medicine and biomedical research. Content-Based Image Retrieval (CBIR) has been discussed in the technical literature as a method that may develop into an efficient image search and retrieval technique. Prior work on medical image retrieval has mainly focused on extracting low-level visual features (e.g., color, texture, shape, spatial layout) and then using them directly to compute image similarity. Extensive experiments have shown, however, that low-level image features cannot always capture the biomedical *semantic* concepts in the image.[1] This poses a serious shortcoming in applying CBIR to routine clinical use, where image content is defined in terms of biomedical concepts. In general, it is challenging to link high-level semantic concepts and automatically-extracted, low-level image features. Therefore, to support query by semantic concept, there is a compelling need for CBIR systems to provide maximum support towards bridging the 'semantic gap' between the low-level visual features and the semantics in biomedical concepts.

In this paper we focus on the retrieval of vertebral shapes from the cervical spine x-ray images collected in the second National Health and Nutrition Examination Survey (NHANES II) conducted by the National Center for Health Statistics (NCHS). In this set of x-ray images, osteophytes (bone "spurs") are a visual feature of interest to the osteoarthritis community. In our work, we have specifically focused on anterior osteophytes (AO), rather than those located on the vertebra posterior, which are less frequently found in the collection. The x-ray image collection consists of 10,000 cervical spine images and 7,000 lumbar spine images, all digitized from film. Abnormal vertebrae exhibiting AO are identified by shape variation in the anterior "corner" of the vertebral outline seen in the sagittal view. Macnab's osteophyte classification[2] (claw and traction) has been

Medical Imaging 2010: Advanced PACS-based Imaging Informatics and Therapeutic Applications,
edited by Brent J. Liu, William W. Boonn, Proc. of SPIE Vol. 7628, 762807 · © 2010 SPIE
CCC code: 1605-7422/10/$18 · doi: 10.1117/12.844470



	Expert 1	Expert 2	Expert 3
C3	Normal	Normal	Moderate, Claw
C4*	Normal	Normal	Normal
C5	Severe, Claw	Severe, Claw , Traction	Severe, Claw
C6	Severe, Traction	Moderate , Traction	Moderate, Traction
C7*	Severe, Claw	Severe, Claw	Severe, Claw

*Used for training and testing

Figure 1. Cervical spine x-ray with AO labeled by three experts for the C_3-C_7 vertebrae

used as pathology criteria for image retrieval. Combining three severity grades (slight, moderate, and severe) with Macnab's categories,[3] the vertebra shapes may be organized into ten groups: normal, slight claw, slight traction, slight claw and traction, moderate claw, moderate traction, moderate claw and traction, severe claw, severe traction, severe claw and traction. We defined different levels of relevancy in each group, and the task we undertook was to (1) retrieve the vertebra images having the same category label (claw or traction) as that of the query image, and (2) to sort these retrieved images by the similarity of the severity grades to the query image.

In the technical literature, Content-Based Image Retrieval methods that use shape have been applied to trademark data, fish images, and silhouettes.[4–6] However, not all shape retrieval algorithms are suitable for biomedical applications due to (i) high similarity across anatomical shapes, (ii) deficiencies in features for representing subtle boundary differences (which could be indicative of pathology), and (iii) inadequacy of shape description for supporting pathology-specified queries. In addition, images of the same anatomy but with different pathologies often exhibit very subtle differences that lead to disagreements even among experts. Figure 1 shows an example cervical spine x-ray image. A pathological or abnormal vertebra is identified by the abnormal, subtle variation of the shapes of the anterior vertebrae corners. The table on the right describes the AO types and severity levels of C_3-C_7, as labeled by three board-certified radiologists. Due to the inter- and intra-personal variability of the medical experts, the "ground truth" of the vertebra corner type and grade is problematic to some degree. Thus, it is a very challenging problem to minimize the gap between the semantics expressing pathology and the low-level features; that is, given the low-level shape features, how do we retrieve a semantically meaningful shape, considering the subtle variations among many similar shapes?

This paper presents a learning-based retrieval scheme that attempts to bridge this gap using machine learning techniques. First, we extract a number of global and local shape features, what we expect to be an "abundant set", as the low-level feature representation. The global features include those related to geometry (elongation, eccentricity, roughness, and compactness), Fourier descriptors with complex coordinates,[7] Fourier descriptors with Centroid Contour Distance Curve (CCDC),[7] Coefficient of Fourier Expansion of Bent function,[8] and moment invariants.[9] The local shape features are turn angle[10] and Distance Across the Shape (DAS).[11] When standard measures of similarity are applied to these low-level features, however, we find that they do not directly reflect high-level biomedical concepts (for example, the pathological type or grade of the vertebra shapes) of ultimate interest to us. In order to bridge these gaps between the image features and the biomedical concepts, we propose a two-step learning scheme: (1) learning-based feature selection based on the *retrieval accuracy performance of* *individual features in a training data set* to choose 'good' semantically representative features and the fusion of the similarity measurements of the best-performing features to create a new similarity score based on multiple features; and (2) the use of this similarity score in a state-of-the-art ranking algorithm. This ranking algorithm is learning-based rank function optimization approach, the Ranking SVM, which we apply to CBIR for these images. The Ranking SVM algorithm was proposed by Herbrich et al. in 2000.[12] Joachims et al. has applied it to search engine optimization for text retrieval,[13] and we are introducing it to the CBIR domain for the retrieval of medical images. This method predicts ranking function as a means of bridging the gap between the image pathology (expert-marked ground truth data) and the low-level image features. The retrieval function is learned by optimizing a set of inequalities using SVM techniques that improve retrieval effectiveness and efficiency. We demonstrate our approach for retrieval of vertebral shapes segmented from digitized spine x-ray images from health survey data archived by the National Library of Medicine. Our experiment shows that the proposed scheme significantly improves retrieval accuracy.

2. METHODS

In this section, we introduce our learning-based scheme for image retrieval. The scheme consists of two steps: (i) a learning-based feature selection and fusion algorithm; and (ii) the Ranking SVM algorithm to optimize the rank function for better retrieval accuracy. Our objective is to bridge the semantic gap between high-level concepts and low-level image features by a machine learning approach, given the ground-truth ranking of training samples.

2.1 Learning-based feature selection and fusion

A large number of "low-level" visual features may be extracted from images, but not all features are representative of the semantic concepts important for a particular task. It has been pointed out that for a fixed sample size, increasing the number of features (with a corresponding increase in the unknown parameters) adversely affects the reliability of parameter estimates,[14] commonly known as the "curse of dimensionality", which may result degrade classifier performance. Thus, it may not be a good choice to include all the features or feature components in the rank function optimization step in image retrieval. An "intelligent" feature selection and fusion approach is necessary to provide representative features with high discriminative power to achieve better retrieval performance. Some CBIR retrieval methods heuristically design or choose features to minimize such problems, but it is a very challenging task when there is a large semantic gap between high-level human concepts and low-level features.

In this section, we propose a learning-based feature selection approach, which exploits the measured performance of individual features against a set of training data. By then selecting and combining features according to their retrieval performance on the training data set, we develop a learning-based feature fusion approach based on the similarity measurements. First, according to each feature's retrieval accuracy on the training data, we select the features with high retrieval performance.

The scheme we have developed is as follows. Let MAP^m denote the *mean average precision* (MAP)[15] achieved over our set of training queries for feature m ($m = 1, 2, 3, ..., M$, where M is the number of features); then after sorting all the MAP^m, we choose the first K features with highest precision rates, $MAP^{m_k}(k = 1, ..., K)$, where k corresponds to the kth largest mean average precision of a single feature. Then, we define a weight w^{m_k} for each selected feature as follows:

$$w^{m_k} = MAP^{m_k} / \sum_{i=1}^{K} MAP^{m_i} \tag{1}$$

In order to be able to measure similarities between shapes with respect to more than one feature, we define the *fusion score* f_{qi}^K with K features for the query shape q and retrieved shape i as follows:

$$f_{qi}^K = \sum_{k=1}^{K} w^k |\vec{n}_q^k - \vec{n}_i^k| \tag{2}$$

where w^k is the weight of the selected feature, as defined above; $\vec{n}_q^k$ is the value of feature vector m_k (with performance MAP^{m_k}) of the query, and $\vec{n}_i^k$ is the value of feature vector m_k of the ith item in the data set. (Both vectors are usually normalized. In the equation above, the "||" notation denotes L1-norm.) For convenience, we name this fusion score similarity measurement the *Picked-Weighted Score (PWS)*. It should be noted that the number K of features selected may significantly affect the retrieval performance. If the retrieval performance (MAP) of the training data using the fusion score in Eq. 2 is denoted as MAP^{f_K}, then $K_{optimal}$, the optimal number of features, is determined by the following equation,

$$K_{optimal} = \arg \max_{K=1}^{M} MAP^{f_K} \tag{3}$$

where M is the total number of features. In the experiment, we exhaustively calculated the MAP across the entire set of features to determine $K_{optimal}$.

2.2 The Ranking SVM algorithm

The Ranking SVM algorithm was originally proposed by Joachims[13] for search engine optimization in the domain of document retrieval. Ranking SVM learns the retrieval function by optimizing a set of inequalities. Unlike the standard SVM approach for classification, the Ranking SVM is specifically formulated to effectively adapt the retrieval function to a specified partial ranking order, instead of simply assigning samples to different classes. The training samples used include expert ranking information, which shows the *degree of relevance* of the sample to the query, instead of simple class information. This algorithm, with the available training data, is particularly well-suited for the retrieval problem and so we have applied Ranking SVM to image retrieval.

We formulate the ranking problem for image retrieval as follows: For a given image query q and an image database collection $D = \{d_1, ..., d_m\}$, the optimal retrieval system should return an optimal ranking r^* that orders the database elements in D according to their relevance to the query. The objective of the retrieval algorithm is to find a retrieval function f, whose ordering $r_{f(q)}$ approximates the optimum ranking r^*, according to a target standard of optimality.

Typically, retrieval systems do not achieve an optimal ordering r^*. Instead, a retrieval function f is evaluated according to how closely its ordering $r_{f(q)}$ approximates the optimum, given an independently and identically distributed training sample S of size n containing queries q with their target rankings r^*, $(q_1, r_1^*), (q_2, r_2^*), ..., (q_n, r_n^*)$. The retrieval problem may be posed as a maximization problem in terms of Kendall's rank correlation coefficient τ; In the formulation,[16] the *learner* L selects a ranking function f from a family of ranking functions F that maximizes the empirical τ on the training sample. More precisely, the objective is to learn a scoring function for ranking; that is, to learn to accurately rank a set of objects by combining a given collection of ranking or preference functions.

The input to the algorithm is a list of "order preferences" (e.g., d_i should be ranked above d_j), and a list of "base ranking functions" from a family of ranking functions F. The goal is to find a function $f \in F$ that maximizes Kendall's τ and generalizes beyond the training data. Consider the class of linear ranking functions,

$$(d_i, d_j) \in f_{\overrightarrow{w}}(q) \Longleftrightarrow \overrightarrow{w}^T \Phi(q, d_i) > \overrightarrow{w}^T \Phi(q, d_j), \tag{4}$$

where, $\overrightarrow{w}$ is a $(p \times 1)$ weight vector that is adjusted by learning. $\Phi(q, d)$ is a mapping, $\Phi : Q \times D \to R^p$, onto features that describe the match between query q and the image d, where $Q = \textit{possible query images}$, and $D = \textit{database images}$. A possible formulation of the retrieval problem is finding the weight vector, $\overrightarrow{w}$, that minimizes the following equation (see Joachims's paper[13] for details),

$$V(\overrightarrow{w}, \overrightarrow{\xi}) = \frac{1}{2} \overrightarrow{w}^T \cdot \overrightarrow{w} + C \sum \xi_{i,j,k}, \tag{5}$$

subject to the constraints of the inequalities:

$$\forall (d_i, d_j) \in r_1^* : \overrightarrow{w}^T \Phi(q_1, d_i) \geq \overrightarrow{w}^T \Phi(q_1, d_j) + 1 - \xi_{i,j,1} \tag{6}$$

$$...$$

$$\forall (d_i, d_j) \in r_n^* : \overrightarrow{w}^T \Phi(q_n, d_i) \geq \overrightarrow{w}^T \Phi(q_n, d_j) + 1 - \xi_{i,j,n} \tag{7}$$

$$\forall i \forall j \forall k : \xi_{i,j,k} \geq 0.$$

C is a parameter that allows trading-off margin size against training error. ξ is introduced as a non-negative slack variable, such that the upper bound of $\sum \xi_{i,j,k}$ is minimized. The optimization problem is equivalent to that of a classification SVM on pairwise difference vectors $\Phi(q_k, d_i) - \Phi(q_k, d_j)$. It can be solved using decomposition algorithms similar to those used for SVM classification. An adaptation of the $SVM^{light} algorithm$[17] is used in our experiment.

In our two-step learning-based scheme, instead of using the original low-level image features as the input feature to the Ranking SVM algorithm, we use the weighted selected features, which are obtained by the learning-based feature selection algorithm described in section 2.1, and their similarity measurements based on commonly used Minkowski metric distance, such as Normalized Euclidean Distance (NED) measure, Manhattan distance (SUM), Chebychev distance (MAX), Mahalanobis distance (MAH), and *Picked-Weighted Score (PWS)* (described in section 2.1) to serve as the input features to the Ranking SVM algorithm.

3. FEATURE EXTRACTION

In this section, we focus on shape-based features for image retrieval. For a given vertebra shape contour, we use both global and local shape features. We choose features that are invariant with respect to rotation, scale, and translation.

Firstly, the vertebra shapes were segmented using an active contours method modified to constrain evolving contour points to follow "orthogonal curves",[18] to avoid convergence to a self-intersecting solution contour at vertebra corners.[19] We preprocess the contour shapes by curve smoothing (to reduce noise), fitting (for smoothness), interpolation, and re-sampling (to obtain evenly distributed points) to obtain the final shape contour description. The curve fitting and interpolation are done with the natural cubic spline algorithm. Then the shape contour is resampled by equal arc length sampling. Finally, the whole vertebra shape is represented by two boundary point sets with different resolutions.

To capture the characteristics of shape in local regions, we use two different local features. The first is Turn Angle (TA), also called Turning Angle or Bent Angle.[10] For an arbitrary shape, the Turn Angle feature could be calculated from the approximating polygon for that shape. Turn Angle for a polygon with n vertices, then, may be represented as a vector in R^n . The second feature for capturing local shape is Distance Across the Shape.[11] DAS is defined for each vertex P in a polygon, as the length of the angle bisector at P, measured as the line segment from P to the intersecting side of the polygon. If the bisector intersects the shape multiple times, the distance to the closest intersection is used.

We also extract some global features, such as the geometric polygon features (elongation, eccentricity, compactness, and roughness[20]) and Hu moments.[9] Hu moments is a region-based feature. Although region-based shape representation may be expected to capture more global properties, contour-based shape representation is also useful in some cases.[7] For example, the Fourier Descriptor (FD) is a commonly used contour-based feature for shape representation. Many FD methods have been discussed in the literature, and various *shape signatures* have been exploited to obtain FD. However, FD derived from different signatures result in significantly different performance for shape retrieval. In this work, we use FD derived from three shape signatures, namely, central distance, complex coordinates, and the Fourier series coefficient of the Bent function. (The reader should refer to Guan[21] for additional details.)

4. EXPERIMENT AND EVALUATION

In the experiments, we randomly selected 200 cervical spine (C-spine) x-ray images from a collection of 17,000 spine x-rays from the NHANES II[22] survey. Due to the inter- and intra-observer variability of medical experts who interpreted these images, our ground truth of the vertebra corner type and grade for these 200 images confounds the class categories to some degree, and reduced the original image set to 117 vertebrae that had complete agreement among the radiologists; these 117 were used for training and testing. The vertebra shapes were segmented using an Active Contour Segmentation algorithm.[19] After curve fitting, smoothing, and re-sampling, each vertebra was represented by 180 boundary points. Finally, the global and local vertebra shape features were extracted and normalized using Gaussian normalization. The global shape features used were geometric (elongation, eccentricity, roughness, and compactness), Fourier descriptors with complex coordinates,[23] Fourier descriptors with Centroid Contour Distance Curve,[23] Fourier Coefficient of Fourier Expansion of Bent function,[8] and moment invariants. The local shape features used were Turn Angle and Distance Across the Shape.[11]

We applied Ranking SVM with the learning-based feature fusion scheme to vertebra shape retrieval for these images. We trained and tested on this data set using the leave-one-out procedure. Specifically, each vertebra shape was selected in turn as the test image and the remaining images were used for training. The test results were then averaged over all runs to compute overall performance. We evaluated the results quantitatively by using the average precision-recall graph. As described in the Introduction, the segmented vertebra shapes were classified into pathology types according to Macnab's system and into three severity grades, and grouped into ten classes. We considered an image to be relevant to a query if the retrieved images were in the same group (i.e. both have the same type and grade) as the query image.

We performed retrieval for each query image in six different ways: (1)-(4) by calculating image similarities to the query image using four conventional Minkowski metric-based similarity measurements (Euclidean, Mahalanobis, Manhattan, and Chebychev distances), (5) by calculating the image similarities to the query image using the fused score described in Eq. 2 (Picked-Weighted Score, PWS), (6) by using Ranking SVM with all the original low-level image features (without feature selection and fusion), and (7) by using Ranking SVM with the learning-based feature selection and fusion approach. We compared them using the mean average precision (MAP). The resulting MAPs of MAH (Mahalonobis), MAX (Chebychev), NED (Normalized Euclidean), SUM (Manhattan), PWS, the ranking SVM algorithm without the feature selection, and the ranking SVM with the feature selection are 10.38%, 14.40%, 25.51%, 25.80%, 38.81%, 49.65%, and 52.30% respectively, which indicate the MAP of proposed scheme is $26.50\% - 41.92\%$ higher than the conventional Minkowski similarity measurements, 13.49% better than the PWS algorithm alone, and 2.65% better than the ranking SVM algorithm without the feature selection.

We compared the retrieval performance using the average precision-recall graphs. Figure 2 shows the performance comparison of these seven retrieval methods, and demonstrates the favorable performance of Ranking SVM with the fusion score, as compared to simple database ranking by the four Minkowski metrics[24] using low-level image features, and also as compared to the Ranking SVM algorithm without the fusion score. In summary, the experiments demonstrates that our proposed scheme is the best among all the algorithm we tested.

Unlike the existing approaches to CBIR, which are typically based on conventional Minkowski distance measures for image similarity, our scheme is an off-line learning based approach, which seeks to predict the measures of image similarity that are implicit in expert-labeled training data. Our method treats the learning of the similarity function as an optimization problem and discovers a mapping between similarity ranking using high-level human perception to ranking using low-level image features. In this way our approach models the relation of high-level semantic concepts to the actual image data presented to the algorithm.

5. CONCLUSION

In this paper, we have introduced a learning-based scheme to model biomedical semantic concepts using low-level image features for CBIR. The results demonstrate that for vertebra shape retrieval, our method may be able to approximate the retrieval rank function implicit in an expert-labeled training data set, which expresses pathological semantic concepts, and serves as a basis for retrieving visually similar vertebrae. Our results also show that the precision of our method significantly outperforms commonly used simple distance-based similarity metrics for the experimental data set.

6. ACKNOWLEDGEMENT

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC).

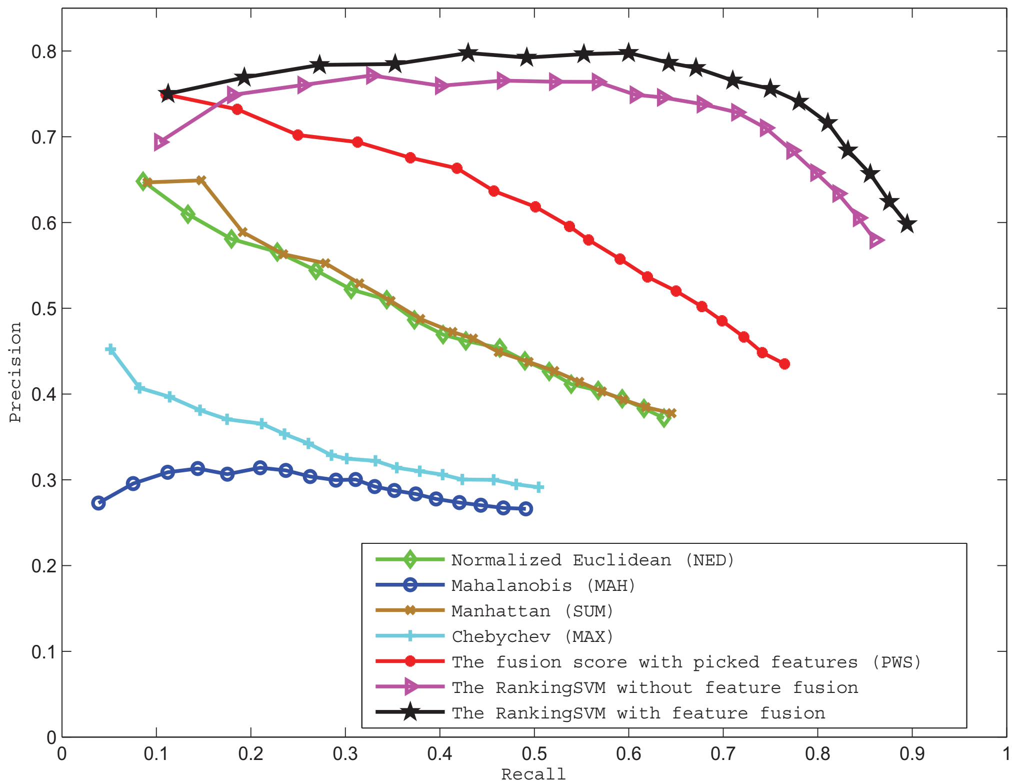


Figure 2. The Average Precision-Recall graph

REFERENCES

[1] Y. Liu, D. Zhang, G. Lu, and W. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition* **40**(1), pp. 262–282, 2007.

[2] M. Heggeness and B. Doherty, "Morphologic study of lumbar vertebral osteophytes," *Southern Medical Journal* **91**, pp. 187–189, February 1998.

[3] J. Shou, S. Antani, L. Long, and G. R. Thoma, "Evaluating partial shape queries for pathology-based retrieval of verterbra," in *Proc. 8th World Multiconference on Systemics, Cybernetics and Informatics*, pp. 155–60, July 2004.

[4] M. Pelillo, K. Siddiqi, and S. Zucker, "Matching hierarchical structures using association graphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(11), pp. 1105–1120, 1999.

[5] T. Sebastian, P. Klein, and B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(5), pp. 550–571, 2004.

[6] D. Zhang and G. Lu, "A comparative study on shape retrieval using fourier descriptors with different shape signatures," *Journal of Visual Communication and Image Representation* (14 (1)), pp. 41–60, 2003.

[7] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition* , pp. 1–19, 2004.

[8] D.-J. Lee, P. Zhan, A. Thomas, and R. B. Schoenberger, "Shape-based human detection for threat assessment," *Visual Information Processing XIII* **5438**(1), pp. 81–91, SPIE, 2004.

[9] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory* **8**, pp. 179–187, February 1962.

[10] E. Arkin, L. Chew, D. Huttenlocher, K. Kedem, and J. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(3), pp. 209–216, 1991.

[11] L. Chen, R. Feris, and M. Turk, "Efficient partial shape matching using Smith-Waterman algorithm," in *CVPR Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment*, pp. 23–28, (Anchorage, Alaska), 2008.

[12] R. Herbrich, T. Graepel, and K. Obermayer, *Large Margin Rank Boundaries for Ordinal Regression.* MIT Press, Cambridge, MA, 2000.

[13] T. Joachims, "Optimizing search engines using clickthrough data," in *KDD '02: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 133–142, ACM Press, (New York, NY, USA), 2002.

[14] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, pp. 4–37, January 2000.

[15] C. D. Manning, P. Raghavan, and H. Schtze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.

[16] A. Mood, F. Graybill, and D. Boes, *Introduction to the Theory of Statistics*, McGraw-Hill Companies, 3 ed., 1974.

[17] T. Joachims, *Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms*, Kluwer Academic Publishers, April 2002.

[18] H. Tagare, "Deformable 2-d template matching using orthogonal curves," *IEEE Transactions on Medical Imaging* **16**, pp. 108–117, February 1997.

[19] L. Long, S. Antani, and G. Thoma., "Image informatics at a national research center," *Computerized Medical Imaging and Graphics* **29**, pp. 171–193, 2005.

[20] F. Aghdasi, R. Ward, J. Morgan-Parkes, and B. Palcic, "Feature selection for classification of mammographic microcalcification clusters," in *Engineering in Medicine and Biology Society, 15th Annual International Conference of the IEEE*, pp. 58–59, 1993.

[21] H. Guan, S. Antani, L. R. Long, and G. R. Thoma, "Comparative study of spine vertebra shape retrieval using learning-based feature selection," in *special track on Biomedical Image Applications in Research and Clinical Medicine, the 22nd IEEE International Symposium on Computer-Based Medical Systems*, Aug. 2009.

[22] L. Long, S. Antani, D. Lee, D. Krainak, and G. Thoma, "Biomedical information from a national collection of spine x-rays: Film to content-based retrieval," in *Proc. of SPIE Medical Imaging: PACS and Integrated Medical Information Systems*, pp. 5033:70–84, February 2003.

[23] D. Zhang and G. Lu, "Content-based shape retrieval using different shape descriptors: A comparative study," *IEEE International Conference on Multimedia and Expo* **00**, pp. 1139– 1142, 2001.

[24] G. L. Naber, *The Geometry of Minkowski Spacetime: An Introduction to the Mathematics of the Special Theory of Relativity*, Dover Publications, 2003.