LETTER

MR-MIL: Manifold Ranking Based Multiple-Instance Learning for Automatic Image Annotation

Yufeng ZHAO[†a)], Yao ZHAO[†], Zhenfeng ZHU[†], *Nonmembers, and* Jeng-Shyang PAN[††], *Member*

SUMMARY A novel automatic image annotation (AIA) scheme is proposed based on multiple-instance learning (MIL). For a given concept, manifold ranking (MR) is first employed to MIL (referred as MR-MIL) for effectively mining the positive instances (i.e. regions in images) embedded in the positive bags (i.e. images). With the mined positive instances, the semantic model of the concept is built by the probabilistic output of SVM classifier. The experimental results reveal that high annotation accuracy can be achieved at region-level.

key words: *automatic image annotation, multiple-instance leaning, manifold ranking, SVM*

1. Introduction

In [1], Chang et al. presented a MIL algorithm, called as sequential point-wise diverse density (SPWDD), to extract only one representative instance for a given concept. Tang et al. treated natural scene classification as ranking the typicality of images via semi-supervised multiple-instance typicality ranking (SSMITR) [2]. However, although some endeavors have been made, the annotation performance is still unsatisfactory. First, there may be some concepts not captured by a single region even if image segmentation is ideal. Second, due to lack of the prior knowledge of instances in MIL, the annotation accuracy is not satisfied. To concentrate on these problems, MR-MIL is proposed for effectively mining the positive instances embedded in the positive bags. With the mined positive instances, a probabilistic SVM classifier is captured for a concept.

2. MR-MIL Based Annotation Scheme

As the application of MIL in AIA, each image is taken as a labeled bag with multiple instances and the segmented regions in the image correspond to the instances in the bag. A bag is labeled positive as long as one of its instances is positive. A bag is labeled negative only if its instances are all negative. Given a concept c, let $\bar{X} = \{X_i \,|i = 1, \cdots, N\}$ be the set of labeled positive and negative bags, where N is the total number of labeled bags. Let $\tilde{x} = \left\{x_p \,|p = 1, \cdots, M\right\}$ be the set of instances contained in $\bigcup_{i=1}^{N} X_i$, where M is the total number of instances. Let $\bar{Y} = \{Y_i \,|i = 1, \cdots, N\}$ be the label set of bags and $\tilde{y} = \left\{y_p \,|p = 1, \cdots, M\right\}$ be the label set of instances. In the case of MIL, the labels of bags need to be known beforehand and the labels of instances are generally unknown. MR-MIL based annotation scheme is composed of the following three steps.

Relevance score propagation of instance: For a given concept c, the relevance score to it of each instance, which will act as trusty measurement for positive instance selection, can be obtained via MR. Instead of initializing $\tilde{y}$ with the bag label [2], the prior relevance score of each instance is assigned with its normalized region area, which is concordant with human's perception to some extent. Table 1 shows the propagation algorithm for the relevance scores of instances.

In Table 1, S denotes the normalized adjacency graph matrix with the heat kernel being used for measuring the pair-wise similarity. By Eq. (1), the initial relevance score of each instance to the concept c is propagated to its neighbors iteratively until a steady state $R^* = (1 - \alpha)(I - \alpha S)^{-1} Y$ is reached. See [3] for the rigorous proof. Here, the element r_p of R^* represents the relevant degree of the instance x_p to the concept c.

Positive instance selection: Based on the relevance scores of instances, the positive instances embedded in the positive bags can be mined accurately. On the assumption that there is at least one positive instance in each positive bag, Table 2 reports a novel strategy for positive instances selection.

Concept learning: The SVM classifier for the concept c is trained with the selected positive instance set PI and some negative instances sampled randomly from the nega-

Table 1 Propagation algorithm for the relevance scores of instances.

Input: Labeled bag set $\bar{X}$, instance set $\tilde{x}$ and its area set $\tilde{y}$.

1. Form the prior relevance score vector $Y_{M\times 1} = \left[y_1, \cdots, y_M\right]^T$.
2. Compute the pair-wise similarity among instances and stack them into the normalized matrix $S_{M\times M}$.
3. Set $R(0) = Y$, perform Eq.1 and converge to R^*.

$$R(t+1) = \alpha \times S \times R(t) + (1-\alpha) \times Y \quad (1)$$

where t is the iteration number, $\alpha \in [0,1]$ is the parameter.

Outputs: Relevance score vector $R^*_{M\times 1} = \left[r_1, \cdots, r_M\right]^T$.

Manuscript received February 25, 2008.
Manuscript revised June 19, 2008.
[†]The authors are with the Institute of Information Science, Beijing Jiaotong University, P.R. China.
[††]The author is with the Department of Electronic Engineering, Kaohsiung University of Applied Sciences, Taiwan.
a) E-mail: snowmanzhao@163.com
DOI: 10.1093/ietfec/e91–a.10.3088

Table 2 Selection algorithm for the positive instances.

Input: Positive bags $X_i^+, i = 1, \cdots, N^+$. Relevance score vector R^*.

1. Compute the variance $var\left(X_i^+\right), i = 1, \cdots, N^+$ of relevance scores of instances contained in each positive bag.
2. Set the threshold $\xi = \frac{1}{2}\left[\max_i \left\{var(X_i^+)\right\} + \min_i \left\{var(X_i^+)\right\}\right]$.
3. Pick out all positive bags X'^+ satisfying $var\left(X'^+\right) \geq \xi$.
4. Select only one instance with the maximum relevance score from each picked-out positive bag X'^+.

Output: A set of selected positive instances $PI = \left\{x_1^*, \cdots, x_n^*\right\}$.

tive bags. The simplified probabilistic SVM modifies the binary output of the standard SVM as follows:

$$p(x|c) = \frac{1}{1 + \exp(-yf(x))} \tag{2}$$

where $f(x)$ is the distance of x to the super-plane through SVM. For a test bag $X = \{x_1, \cdots, x_m\}$, the probability of the given concept c to be annotated with is determined by:

$$p(c|X) \propto \prod_{i=1}^{m} p(x_i|c) \tag{3}$$

According to Eq. (3), the top-five concepts with maximum probabilities are selected to annotate the test bag.

3. Experimental Results

We verify the performance of the proposed AIA scheme on the same Corel dataset with [1]. Furthermore, we also measure the image annotation effectiveness by using the annotation recall and precision defined in [1]. Table 3 shows the annotation qualities of SPWDD and MR-MIL on the total 374 concepts.

Clearly, the proposed AIA scheme achieves competitive results. The average recall and precision are increased by 10% and 15%, respectively. The major reason for the great improvement is that the semantic model for each concept is characterized more accurately than SPWDD.

Both SSMITR and MR-MIL are to mine the positive instances embedded in the positive bags even if SSMTR is not originally proposed for AIA. The evaluation on them for ten most frequent concepts is presented in Fig. 1. For each concept, some instances are selected by SSMITR and MR-MIL, respectively. We define the precision as the ratio of the number of ground truth to the total number in the selected instances.

As shown in Fig. 1, the performance is promoted largely when the prior areas of instances are concerned. Over ten most frequent concepts, the average precision of

Table 3 Comparison with other related annotation models.

Models	***SPWDD***	***MR-MIL***
Average Recall	0.09	0.19
Average Precision	0.07	0.22

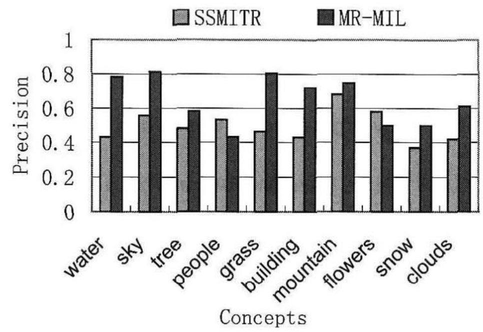

Fig. 1 Precisions of ten most frequent concepts.

MR-MIL goes up by 16% compared with SSMITR. The potential reason for that is the prior knowledge of instances take on some important clues for associating the bag with the positive instances contained in it.

4. Conclusions

Mining of positive instances is of great importance for MIL based AIA. In this paper, two aspects are proposed to focus on it. With the help of prior knowledge of instances, MR is applied to mutually propagate the relevance scores associated with a concept. Meanwhile, a novel strategy based on the local statistic of instances' relevance scores from each bag is proposed to pick out positive instances for learning concept model. The final experimental results demonstrate that quality of MR-MIL based annotation scheme is greatly improved.

Acknowledgments

This work was funded in part by NSFC (No. 60602030, No. 90604032), 863 Program (No. 2007AA01Z175), PCSIRT (No. IRT707) and Research Foundation of BJTU (No. 2005SM013, No. 2005SZ005).

References

[1] C. Yang, M. Dong, and F. Fotouhi, "Region based image annotation through multiple instance learning," Proc. ACM, MM, pp.435–438, Singapore, 2005.

[2] J. Tang, X. Hua, G. Qi, and X. Wu, "Typicality ranking via semi-supervised multiple-instance learning," Proc. ACM, MM, pp.297–300, Augsburg, Germany, 2007.

[3] D. Zhou, O. Bousquet, T.N. Lal J. Weston, and B. Scholkopf, "Learning with local and global consistency," Proc. IEEE, NIPS, Canada, 2003.