

Learning with Information Entropy Method for Transportation Image Retrieval

Liu Xiao-jun, Li Qing-ling, Li Yong-jian and Li Jun-yi

Department of Logistics and Information Management, Zhuhai College of Jilin University, Zhuhai 519041, Guangdong, China

Faculty of Business and Administration, University of Macau, Macau 999078, Macau

Department of Logistics and Information Management, Zhuhai College of Jilin University, Zhuhai 519041, Guangdong, China

Electrical and Computer Engineering Department, National University of Singapore, Singapore 119077, Singapore

Lxj02041820@163.com

Abstract

As a new learning framework, Multi-Instance learning is labeled recently and has successfully found application in vision classification. A novel Multi-instance bag generating method is presented in this paper on basis of Gaussian Mixed Model. The generated GMM model composes not only color but also the locally stable unchangeable components. It is frequently named as MI bag by researchers. Besides, another method called Agglomerative Information Bottleneck clustering is applied to replace the MIL problem with the help of single-instance learning ones. Meanwhile, single-instance classifiers are employed for classification. Finally, ensemble learning is adopted to strengthen classifiers' generalization ability of RBM (Restricted Boltzmann Machine) as the base classifier. On the basis of large-scale datasets, this method is tested and the corresponding result shows that our method provides high accuracy and good performance for image annotation, feature matching and example-based object-classification.

Keywords: *Multi-Instance Learning, Image representation, RBM, AIB Clustering, Scene Recognition, Gaussian Mixed Model*

1. Introduction

When people started to study the issue of drug activity prediction, they gradually noticed the term multi-instance learning [1]. The training set in multi-instance framework is mainly made up of labeled bags and meanwhile those labeled bags have many unlabeled instances. As a result, certain concepts could be obtained from the training set so as to label unseen bags accurately. A bag can be correctly labeled if it possesses one positive instance. However, if it contains two or more instances, the circumstances are different. Multi-instance has been adopted in various practical problems since it comes into being. Especially when the object happens to describe something ambiguous, this method is of great use. In the aspect of scene classification, it is necessary for the content of the images to be ambiguous. For example, three components as “beach”, “sky”, “water” are shown in Figure 1. Under such circumstance, if we represent those images on the basis of single-instance, a decline in accuracy of classification is to be generated. However, multi-instance framework can be applied to change this situation. In this instance, a batch of effective feature vectors obtained from each little branch can be used

to build a new instance. In this way, those four instances constitute an integral part of this image. At the same time, the semantics meaning of this image can be expressed.



Figure 1. Example of Ambiguous Images

Image segmentation and clustering are often used to generate multi-instance bag. For example, C. Yang and T. Lozano [10] divide the whole image into many intersecting sub-regions, and then filter sample regions to get feature vectors for instance. All the instances are gathered together to form a multi-instance bag and finally they make use of weighted correlation coefficient to examine the similarity among those images. In 2003, Z.H. Zhou [11] proposed a novel bag generating method based on SOM clustering. The pixel is clustered in terms of the color features and location ones. The results of clustering are fused into a specified number of sub-regions and the color feature of each sub-region could be counted as one instance.

The method of multi-instance bag generation is of great importance to classification performance. Many kinds of algorithms are put forward by researchers in the perspective of multi-instance learning. For instance, scholars like O. Maron *et al* [2,9] come up with a new idea, namely, the algorithms of Diverse Density (DD). Multi-instance learning is used by them to proceed with the classification of images in natural places. In this situation, each image and instance will respectively perform as a bag and the feature of the little branch of the image. Then images are classified by utilizing Diverse Density algorithms. In addition, they have proposed various types of bag generating methods (see Figure.2) and claimed that each method is to play a vital role on the results. Zhou and Zhang (2006) [9] have ever conducted a research on the discrimination of instances which are transmitted to the discrimination of bags. It turns out that the learning algorithms featured with single-instance can be adopted in the multi-instance learning. In fact, people are still trying to make use of those methods nowadays. For example, these methods frequently appear in multi-instance representation. While other scholars hold some unique ideas. For instance, Q. Zhang and SA. Goldman [3] suggest that they could combine EM (Expectation Maximization) algorithm and DD to form a brand-new one, that is, EM-DD algorithm. In terms of object detection, P. Viola and C. Zhang [8] try to integrate multi-instance learning and boosting algorithm. Furthermore, many scholars employ the algorithm of single-instance learning to solve the problems in this field [4-7].

2. Method for Multi-Instance Learning

The novelty of our algorithm is vividly shown in three parts. To start with, a new image multi-instance generation algorithm is put forward. With the help of Gaussian Mixed Model, an image is modeled by multiple Gaussian distribution and the image pixel

is extracted from it. Regarded as a multi-instance bag, generated GMM has the components performing as instances of the corresponding bag. What's more, an entirely different new solution which applies multi-instance representation into algorithms of single-instance learning going is proposed. To be more exact, the instances of all the bags are gathered together and formed several groups. However, as the multi-instance bags generated by step 1 are continuous probability distribution, the traditional clustering method is of no practicability. Therefore our algorithm utilizes agglomerative information bottleneck for clustering MI bag.

The binary features then respectively re-represent each bag. In case that the instance contained in related bag fall into i-th group, its value shall be set to one rather than zero. Thus, one feature vector represents one bag. Besides, the single-instance classifiers can be adopted to distinguish various kinds of bags. Finally, the classifier ensemble comes into being. By using different values of d and repeating the above steps, many classifiers can be produced and accordingly the ensemble learning for prediction can be made. The result of our experiments reveals that the method is of high efficiency and effectiveness. The following steps are to be explained in detail.

2.1. Multi-Instance Bag Generating Via Mixture of Gaussians

Seen from the sub-region of image, each bag possesses the discrete feature while the traditional image bag is produced. However, it is far from being satisfied for image information representation. According to Maron *et al* [12], the multi-instance image bag generating methods are of vital importance to the results of image classification. In previous work, the way of bag generation used to extract color feature information from image sub-region and color feature is the representation of discrete information. However, the image information representation inevitably loses its way. Therefore the inadequacy of image feature representation in previous work shall be fully considered. In the process of studying the probability distribution, an instance is produced by multi-Instance (MI) bag generating method here. We suppose that pixels of images are gained from various Gaussian distribution. The GMM is also called Mi bag which contains instances including locally stable invariant components as spatial location and color [13]. At this point, we also assume that each pixel is represented by color and spatial location. The color variant is concatenated to SIFT feature descriptors computed on RGB/HSV channel. The following formula illustrates how one image is represented as a GMM:

$$f(x; \theta) = \sum_{k=1}^K \alpha_k N(x; \mu_k, \sigma_k) \quad (1)$$

Seen from this formula, x represents a D-dimensional pixel feature (pixel color and spatial location), and θ represents a set of parameters in the aspect of distribution of Gaussian mixture. In addition, this formula (2) demonstrates a Gaussian distribution of D element, α_k means the first k mixture probability and μ_k refers to mean. As for σ_k , it refers to variance,

$$N(x; \mu_k, \sigma_k) = \frac{1}{(2\pi)^{D/2} |\sigma_k|^{D/2}} \exp \left\{ -\frac{1}{2} (x - \mu_k)^T \sigma_k^{-1} (x - \mu_k) \right\} \quad (2)$$

Furthermore,

$$\alpha_k > 0, \sum_{k=1}^K \alpha_k = 1, \mu_k \in R^d. \quad (3)$$

We adopt EM parameter learning algorithm in this paper to check the parameters of these Gaussian mixture model $\theta = (\alpha, \mu, \sigma)$. First of all, Z_n is assumed as a K-dimensional binary. Accordingly responses to n-th pixel are variable. In addition, only one dimension

exists here and the others are zero. $Z_{nk} = 1$ implies that pixel n belongs to the first k Gaussian distribution. To describe this trial more conveniently, z acts as a K -dimensional binary variable which is changeable in accordance with an image pixel. In other words, $P(Z_k = 1) = \alpha_k$ indicates the prior probability of $Z_k = 1$. Regarding the formula $p(x|Z_k = 1) = N(x; \mu_k, \sigma_k)$, it refers to the distribution of x when z functions as the prerequisite. Therefore, the posterior probability of $Z_k = 1$ can be represented as the following formula.

$$\begin{aligned} \gamma(Z_k) &= p(Z_k = 1)p(x|Z_k = 1) \\ &= \frac{p(Z_k = 1)p(x|Z_k = 1)}{\sum_{j=1}^K p(Z_j = 1)p(x|Z_j = 1)} \\ &= \frac{\alpha_k N(x; \mu_k, \sigma_k)}{\sum_{j=1}^K \alpha_j N(x; \mu_j, \sigma_j)} \end{aligned} \quad (4)$$

By virtue of iteration, EM algorithm updates the posterior probability $\gamma(z_k)$ and parameter set $\theta = (\alpha, \mu, \sigma)$. Procedures are listed as follows:

E-step: in formula (3), the posterior probability shall be figured out at first.
 M-step: formula (3) is adopted to replace θ which stands for the set of re-calculated parameters:

$$\mu_k^{new} = \frac{1}{N_k} \sum_{n=1}^n \gamma(z_{nk}) x_n \quad (5)$$

$$\sigma_k^{new} = \frac{1}{N_k} \sum_{n=1}^n \gamma(z_{nk}) (x_n - \mu_k^{new})(x_n - \mu_k^{new})^T \quad (6)$$

$$\alpha_k^{new} = \frac{N_k}{N} \quad (7)$$

$$N_k = \sum_{n=2}^N \gamma(z_{nk}) \quad (8)$$

On basis of the experiment, K ranges from 4 to 7. Information in [14] shows various methods used for evaluating the components' number.

2.2. Conversion of Multi-Instance Problem Based on AIB

In Section 2.1, multi-instance bag is showed in one image whose classification task is adopted in the multi-instance learning problems.

Popular approaches are clustering multi-instance problem into a single instance problem to be solved. This paper has learnt this way of thinking. However, the multi-instance bag generated in Section 2.1 is a Gaussian mixture distribution; the traditional clustering approach implemented on the discrete information feature is no longer applicable while the image bag is in the form of the continuous probability distribution. Therefore this paper uses information bottlenecks theory for clustering [44]. After clustering, the distance from each clustering center to each bag is calculated to form a single instance of feature vectors, which can take advantage of the traditional single-instance learning algorithm for classification.

The term IB (Information Bottleneck) is a newly proposed term by Tishby in 1999. It is an information theoretic principle [15]. According to the IB clustering method, of all the

possible clustering which is given an object, they are made into a certain amount of clusters. As to the target clustering, it can minimize the lost mutual information of the cost of its features extracted from subjects and the target subjects.

Use $P(x, y)$ to function as a joint distribution on the “feature” space Y as well as the “object” space X. As the image clustering is aimed at looking for a clustering with standard

quality, that is $I(x; x)$, on the basis of IB principle, the information loss is minimized. In the

formula $I(x; x)$, between the two variables x , the latter one x is used to represent mutual information:

$$I(x; x) = \sum_{x, \hat{x}} p(x, \hat{x}) \log \frac{p(x, \hat{x})}{p(x)p(\hat{x})} \quad (9)$$

$$I(x; x) = \sum_{x, \hat{x}} p(x)p(x|\hat{x}) \log \frac{p(x|\hat{x})}{p(\hat{x})} \quad (10)$$

To make information loss of $I(X; Y) - I(\hat{X}; Y)$ to be below limit, AIB (Agglomerative Information Bottleneck), a combining algorithm, was put forward by N. Tishby [15]. In AIB clustering algorithm, first there is an ordinary cluster then it comes to interactive selection between the two clusters. After that, $|X|$ -d+1 time merged before finally forming the d clusters. As for $|X|$, it represents size of image sample set. The reason for developing a selection criterion is to minimize the loss of mutual information as much as possible, thus ensure the minimum mutual information loss.

In the formula, c1 and c2 stands for clusters of symbols of X, then after merging of c1 and

c2, the information loss is:

$$d(c_1, c_2) = I(C_{before}, Y) - I(C_{after}, Y) \quad (11)$$

Between the classes and the feature space, $I(C_{before}, Y)$ and $I(C_{after}, Y)$ stand for the mutual information during process of merging of c1 with c2 into the same one. The manipulation of standard information theory which illustrates the distortion function $d(c_1, c_2)$ can be represented as follows:

$$\begin{aligned} d(c_1, c_2) &= \sum_{y, i=1,2} p(c_i, y) \log \frac{p(c_i, y)}{p(c_i)p(y)} \\ &- \sum_y p(c_1 \cup c_2, y) \log \frac{p(c_1 \cup c_2, y)}{p(c_1 \cup c_2)p(y)} \quad (12) \\ &= \sum_{y, i=1,2} p(c_i, y) \log \frac{p(y|c_i)}{p(y|c_1 \cup c_2)} \end{aligned}$$

In this paper, the AIB clustering represents the specific process algorithm [16-18]. The AIB will collect all the bags to d clusters. As the distribution of image can be worked out by putting all the images together and getting the average one, every cluster represents a Gaussian mixture distribution:

$$f(y|c) = \frac{1}{|c|} \sum_{x \in c} f(y|x) = \frac{1}{|c|} \sum_{x \in c} \sum_{k=1}^K \alpha_k N(x, \mu_k, \sigma_k) \quad (13)$$

If the similarity with all clusters is calculated for each pair of image, with the help of semantic concept, an image can be illustrated in the form of a d-dimensional feature vector and each dimension can estimate the similarity of an image with the corresponding semantic concept. The new d-dimensional feature can be trained by the traditional single-instance learning algorithms. As the image models happen to be Gaussian mixture clustering distribution, it is suggested that KL (Kullback-Leibler) divergence measure the similarity among the images.

As the information-theoretic measure, the KL-divergence is particularly used to investigate the distance between continuous and discrete distribution (Kullback [19]). When it comes to a discrete representation, it is easy to gain the KL-measure while it is not so easy in a continuous one. Between two mixtures of Gaussians, it is hard to find suitable description for the KL-divergence, but Monte-Carlo simulation can be adopted to roughly estimate it on the basis of distributions of f and g and then their distance can also be evaluated. The following is the representation:

$$D(f \parallel g) = E_d(f \parallel g) = \int f \log \frac{f}{g} \approx \frac{1}{n} \sum_{i=1}^n \log \frac{f(x_i)}{g(x_i)} \quad (14)$$

x_1, x_2, \dots, x_n is sampled from $f(x)$.

Substitute KL measure (14) in the distortion function (12) we obtain.

$$d(c_1, c_2) = \sum_{i=1,2} p(c_i) D(p(y|c_i) \parallel p(y|c_1 \cup c_2)) \quad (15)$$

Note that in formula (13), as $f(y|x)$ refers to a GMM distribution, c stands for density function of each cluster. Therefore, GMM can be represented as $f(y|c)$.

And $f(y|c_1)$, $f(y|c_2)$ are GMM combined respectively by c_1 and c_2 . Then $c_1 \cup c_2$ is

$$f(y|c_1 \cup c_2) = \frac{1}{|c_1 \cup c_2|} \sum_{x \in c_1 \cup c_2} f(y|x) = \sum_{i=1,2} \frac{|c_i|}{|c_1 \cup c_2|} f(y|c_i)$$

On the basis of (15), we can gain the distance between c_1 and c_2 as follows:

$$d(c_1, c_2) = \sum_{i=1,2} \frac{|c_i|}{|X|} D(f(y|c_i) \parallel f(y|c_1 \cup c_2)) \quad (16)$$

$|X|$ refers to the size of image database. Therefore, to work out how far c_1 and c_2 is, it is necessary for us to work out the KL distance between two GMM distributions. It is as follows:

- 1) Pick when the image is still a cluster and begin;
- 2) Combine cluster c_1 and c_2 so as to gain the minimum information loss $d(c_1, c_2)$ (15) in every step.
- 3) Keep doing step 2) until we obtain the last single cluster.

In Figure 2 the Table shows that the mutual information for image-set decreases in the process of agglomerative algorithm. We selectively pick from COREL databases [23] to create 5 categories, which consist of 600 images. From the result of Figure.4, we can find that if the value of $I(X;Y)$ is high, the correlation of the clustering will be high accordingly. This proves that the clustering quality of the steps is outstanding.

The AIB-HSV-SIFT (128*128) outperforms best.

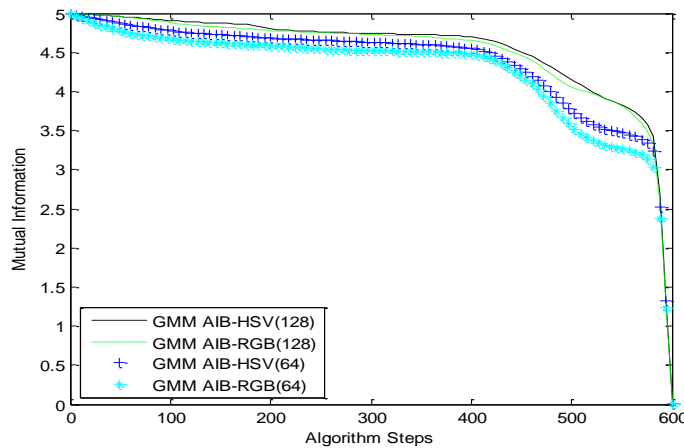


Figure 2. Mutual Information between the Clusters and Categories for each Algorithm Step

2.3. Classifier Ensemble

In Section 2.2, the single instance features of different lengths can be obtained, and single instances of different lengths seem to own various information representations on the basis of the number of different cluster center d . To fully utilize this information, we first use different d values in AIB clustering so that MI bag can be changed into a single instance. In this instance, a single instance which entails different lengths can be used to train different classifiers. The final step is that these classifiers are combined to distinguish classification. Single layer RBM neural network [27-28] is applied as our base classifier in this paper.

Different classifiers are used and combined via evaluating the mean of posterior probability in this paper:

$$p(c_i | x) = \frac{1}{L} \sum_{l=1}^L p_l(c_i | x) \quad (17)$$

In this formula, c_i stands for a category and L for the classifiers' number. As image x is classified to c_i by classifier 1 and the image is classified to c_i by ensemble classifiers. $p_l(c_i | x)$ and $p(c_i | x)$ respectively show the posterior probability of both of them.

Shown in Section 2.1 to 2.3, on the basis of the new bag generation method for the image, we can figure out the multi-instance scheme.

1. EM algorithm is used by us to gain the Gaussian mixture model. And then the $|X|$ Gaussian mixture models for every training image are worked out.
2. By taking advantage of AIB clustering approach, we collect $|X|$ from step 1 and divide them into d class.
3. KL divergence is evaluated and all the cluster centers calculated so as to obtain a new training set.
4. The single-instance classifier is supposed to be cultivated by the generated training set from Step 3.
5. By repeating steps 2-4, different classifiers with different d values shall be collected as many as possible.
6. On the basis of step 5, a number of classifiers can be gained and integrated into an ensemble one.

3. Experiment Results

All the experiments needs to be done 5 times and the results each time should be compared with others by the method of [31-35, 43].

3.1. Caltech-101 Dataset

On the basis of Caltech-101 benchmark, our method has been evaluated for the scene recognition and example-based object. Composed of 101 different types and a background class, Caltech-101 is regarded as an effective benchmark to recognize the targeted object. The class is made up of tens of examples which form a median-scale data set together. As Catech-101 has been employed in many existing algorithms, we use it as a test bed. In particular, three types of local feature are adopted here, pyramid histogram of visual words (PHOW), geometric blur, and PCA-SIFT included.

Table 1. Correctness Rate with 15 or 30 Training Images per Class on Caltech-101

#train	ours	[20]	[28]	[29]	[30]	[31]	[32]
15	62.5	59.05	56.4	52	51	49.52	44
30	71.5	66.23	64.6	N/A	56	58.23	63

Table 1. shows corresponding numbers for 15and 30 training image cases. Seen from the table, we can notice that our algorithm has obtained the better correctness rate in the neighborhood of 62-64% , slightly lower than Li [43], however, in the first period, much improvement over 18% is shown.

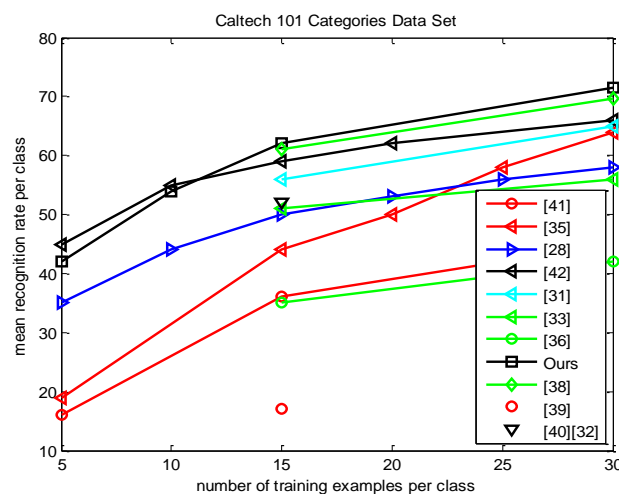


Figure 3. Comparison against Existing Techniques on the Caltech-101

3.2. Local-Patch Indexing

The Photo Tourism project [29-30] is composed of nearly 300k sub images in the size of 32*32. In this case, on the basis of the data obtained from photos of Half Dome, Trevi Fountain and Notre Dame, we test the method on a patch matching task. For Trevi Fountain, we use the n=100000 image patches. It is aimed at evaluating fast algorithms which regain the patches on condition that a query patch is provided. The 128-d SIFT vector is computed for every sub image.

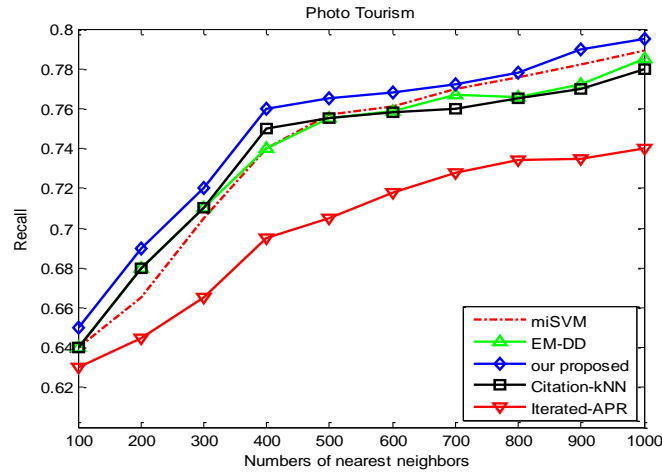


Figure 4. Recall Result on Photo Tourism Data Set

Figure 4. shows that better recall rate performance can be obtained by our algorithm than other multi-instance learning algorithms on Photo Tourism data set. Figure 5. illustrates ROC curve of our proposal method and AUC score is 0.915. As a result, it can be identified as a good classifier.

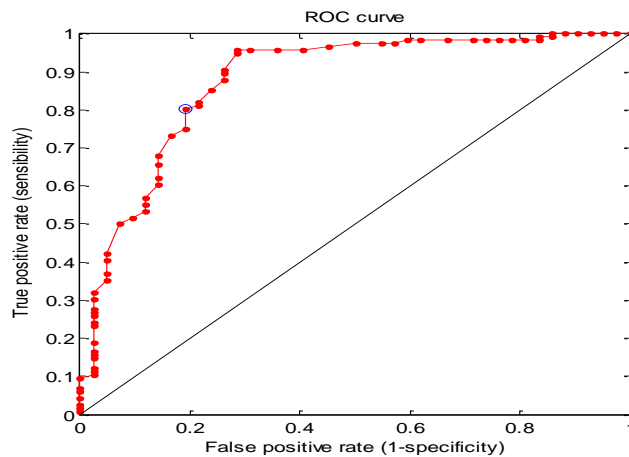


Figure 5. ROC Curve on Photo Tourism Data Set

4. Discussion and Conclusion

In the Multi-instance learning framework, the excellent representation ability for the ambiguity of its object has been successfully demonstrated in image classification task. The MI bag generating method affecting the performance of image classification is also an important factor in image classification. Hence a new multi-instance (MI) bag generating method is put forward for images in this research. It mainly takes advantage of Gaussian Mixed Model (GMM) to model an image. Regarded as a MI bag, the generated GMM is made up of instances including color, spatial location and invariant components which are locally stable. Thus, each instance in a bag can be seen as a Gaussian distribution of one feature. It acts as continuous representation of feature as well and contains more feature information. And we utilize the information bottleneck clustering algorithm to cluster all the images bags into d classes. In this situation, each image can function as a d -dimensional feature vector for another time. Representing the KL

divergence of one image to the i -th cluster, the i -dimension feature represents measures the distance between the image features distribution and overall clustering features distribution, and trains a classifier by a new feature vector. Repeating the above process by different d values, we can obtain several different classifiers; the classification results could be provided by the final ensemble classifier. Through testing five categories from COREL dataset, the experimental results show that on the classification of five categories images, this algorithm is superior to other algorithms such as Citation-kNN, EM-DD, miSVM algorithms and Iterated-discrim APR.

Acknowledgments

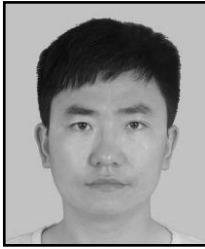
We would like to appreciate Yan Shuicheng and Ni Bingbing from National University of Singapore for helpful comments and discussion on experiments and the paper. To some extent, the accomplishment of this work can not be separated from the help and support of National Program on Key Basic Research Project (973 Program) 2010CB731403/2010CB731406.

References

- [1] T. G. Dietterich, R. H. Lathrop and T. L. Pérez, "J. Solving the multiple-instance problem with axis-parallel rectangles," *Artificial Intelligence Journal*, vol. 1-2, no. 89, (1997), pp. 31-71.
- [2] O. Maron and T. L. Pérez, "A framework for multiple-instance learning," *Advances in Neural Information Processing Systems*, (1998), pp. 570-576.
- [3] Q. Zhang and S. A. Goldman, "EM-DD: An improved multiple-instance learning technique," *Advances in Neural Information Processing Systems*, (2002), pp. 1073-1080.
- [4] Yixin C., Jinbo B. and James Z. W., "MILES: Multiple-Instance Learning via Embedded Instance Selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12 no. 28, (2006), pp. 1931-1947.
- [5] H. T. Pao, S. C. Chuang, Y. Y. Xu and Hsin C. F., "An EM based multiple instance learning method for image classification," *Expert Systems with Applications*, vol. 3 no. 35, (2008), pp. 1468-1472.
- [6] Burr S., Mark C. and Soumya R., "Multiple-instance active learning," *Advances in Neural Information Processing Systems*, (2008), pp. 1073-1080.
- [7] R. Rahmani and S. A. Goldman, "MISSL: Multiple-instance semi-supervised learning," In *Proceedings of the International Conference on Machine Learning (ICML)*, (2006), pp. 705-712.
- [8] P. Viola, J. Platt and C. Zhang, "Multiple instance boosting for object detection," Y. Weiss, B. Schölkopf, J. Platt, Editors *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, vol. 18, (2006), pp. 1419-1426.
- [9] Z. H. Zhou and M. L. Zhang, "J. Solving multi-instance problems with classifier ensemble based on constructive clustering," *Knowledge and Information Systems*, vol. 2 no. 11, (2007), pp. 155-170.
- [10] C. Yang and Lozano P., "Image database retrieval with multiple-instance learning technique. Proceedings of the 16th International Conference on Data Engineering," (2000), pp. 233-243.
- [11] Z. H. Zhou, M. L. Zhang and K. J. Chen, "A novel bag generator for image database retrieval with multi-instance learning," *Proceedings of the 15th International Conference on Tools with Artificial Intelligence*, (2003), pp. 565-570.
- [12] O. Maron and A. L. Ratan, "Multiple-Instance learning for natural scene classification," *Proceedings of the 15th International Conference on Machine Learning (ICML)*, (1998), pp. 341-349.
- [13] D. Lowe, "J. Distinctive image features from scale-invariant key points," *International Journal of Computer Vision*, vol. 2 no. 60, (2004), pp. 91 - 110.
- [14] M. H. Hansen and B. Yu, "J. Model selection and the principle of minimum description length," *Journal of American Statistical Association*, (2001), pp. 746-774.
- [15] N. Slonim and N. Tishby, "Agglomerative information bottleneck," *Advances in Neural Information Processing Systems*, (1999), pp. 617-623.
- [16] Qiyun C., Yuchun F., Jie L. and Wang D., "A novel metrics based on information bottleneck principle for face retrieval," *Proceedings of the 11th Pacific Rim conference on Advances in multimedia information processing*, (2010), pp. 732-741.
- [17] Xue B. and Siwei L., "Combining Bag of Words Model and Information Theoretic Method for Image Clustering," *Advances in Neural Network Research and Application, Lecture Notes in Electrical Engineering*, (2010), pp. 423-430.
- [18] Bardera A., Rigau J., Boada I., Feixas M. and Sbert M., "J. Image Segmentation Using Information Bottleneck Method," *IEEE Transactions on Image Processing*, vol. 7 no. 18, (2009), pp. 1601-1612.

- [19] Greenspan H. and Pinhas A. T., "Medical image categorization and retrieval for PACS using the GMM-KL framework," *IEEE Trans Inf Technol Biomed*, vol. 2 no. 11, (2007), pp. 190–202.
- [20] P. Duygulu, K. Barnard, N. Freitas and D. Forsyth, "Object recognition as machine translation: learning a lexicon for a fixed image vocabulary," In *Proc. ECCV*, (2002).
- [21] S. Feng, R. Manmatha and V. Lavrenko, "Multiple Bernoulli relevance models for image and video annotation. In *Proc. CVPR*, (2004).
- [22] V. Lavrenko, R. Manmatha and J. Jeon, "A model for learning the semantics of pictures," In *NIPS*, (2003).
- [23] "Corel stock photos," <http://elib.cs.berkeley.edu/photos/blobworld/cdlist.html>.
- [24] N. Slonim, N. Friedman and N. Tishby, "Unsupervised document classification using sequential information maximization," presented at the 5th Annual Int. ACM SIGIR Conf. Research and Development in Information Retrieval, (2002).
- [25] N. Slonim, R. Somerville, N. Tishby and O. Lahav, "Objective classification of galaxy spectra using the information bottleneck method," *Monthly Notices Roy. Astron. Soc.*, vol. 323, (2001), pp 270-284.
- [26] L. Hermes, T. Zoller and J. Buhmann, "Parametric distributional clustering for image segmentation," in *Proc. Int. Conf. Computer Vision*, vol. 2, (2002), pp.577-591.
- [27] Andrew K., Kihyuk S. and Honglak L., "Erik Learned-Miller Augmenting CRFs with Boltzmann Machine Shape Priors for Image Labeling," In *Proc. CVPR*, (2013).
- [28] G. E. Hinton, "A practical guide to training restricted boltzmann machines," Technical Report UTML2010-003, University of Toronto, (2010).
- [29] N. Snavely, S. Seitz and R. Szeliski, "Photo Tourism: Exploring Photo Collections in 3D," *Proc. ACM SIGGRAPH*, (2006), pp. 835-846.
- [30] G. Hua, M. Brown and S. Winder, "Discriminant Embedding for Local Image Descriptors," *Proc. IEEE Int'l Conf. Computer Vision*, (2007).
- [31] S. Lazebnik, C. Schmid and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," In *CVPR*, (2006).
- [32] A. C. Berg, "Shape Matching and Object Recognition. PhD thesis," Computer Science Division, University of California, Berkeley, (2005).
- [33] J. Mutch and D. Lowe, "Multi class object recognition using sparse, localized features," In *CVPR*, (2006).
- [34] K. Grauman and T. Darrell, "Pyramid match kernels: Discriminative classification with sets of image features," Technical Report CSAIL-TR-2006-020, MIT, (2006).
- [35] G. Wang, Y. Zhang and L. Fei-Fei, "Using dependent regions for object categorization in a generative framework," In *CVPR*, (2006).
- [36] T. Serre, L. Wolf and T. Poggio, "Object recognition with features inspired by visual cortex," In *CVPR*, (2005).
- [37] Scholkopf B., A. Smola and K. R. Muller, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem," *Neural Computation*, (1998).
- [38] Brain K., Pranteek J. and Krisen G., "Fast Similarity Search for Learned Metrics In *IEEE Trans*," *Pattern Analysis and Machine Intelligence*, (2009).
- [39] L. Fei-Fei, R. Fergus and P. Perona, "Learning Generative Visual Models from Few Training Examples: an Incremental Bayesian Approach Tested on 101 Object Categories," In *Workshop on Generative Model Based Vision*, (2004).
- [40] A. C. Berg, T. L. Berg and J. Malik., "Shape matching and object recognition using low distortion correspondence," In *CVPR*, (2005).
- [41] A. D. Holub, M. Welling and P. Perona, "Combining generative models and fisher kernels for object recognition," In *ICCV*, (2005).
- [42] H. Zhang, A. Berg, M. Maire and J. Malik, "SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition," In *CVPR*, (2006).
- [43] Junyi L. and Jianhua L., "Fast Image Search with Locality-Sensitive Hashing and Homogeneous Kernels Map," *The Scientific World Journal*, (2014).
- [44] Junyi L. and Jianhua L., "A novel semi-supervised Multi-Instance learning approach for scene recognition," *International Conference on Fuzzy Systems and Knowledge Discovery*, (2012).

Author



Liu Xiao-jun, He's a PhD candidate now in Macau University of Science and Technology (MUST) and he's doing his doctoral dissertation on SCM and logistics management information system and information technology. He is currently a lecturer in Department of Logistics and Information Management, Zhuhai College of Jilin University. He has worked on a research program which is A Cloud-based Supply Chain Management System.