International Journal of Hybrid Information Technology
Vol.8, No.11 (2015), pp.435-446
http://dx.doi.org/10.14257/ijhit.2015.8.11.39

Support Vector Machine for Automatic Image Annotation

Dongping Tian[1,2]

[1]*Institute of Computer Software, Baoji University of Arts and Sciences, Baoji, Shaanxi, 721007, China*
[2]*Institute of Computational Information Science, Baoji University of Arts and Sciences, Baoji, Shaanxi, 721007, China*
tdp211@163.com

Abstract

Automatic image annotation (AIA) is an active topic of research in computer vision and pattern recognition. In the last two decades, large amount of researches on AIA have been proposed, mainly including classification-based methods and probabilistic modeling methods. As one of the most common methods for AIA, support vector machine (SVM) has been widely applied in the multimedia research community, especially for image classification, image annotation and retrieval. However, compared with various SVM methods and their corresponding applications in the literature, there is almost no review research and analysis about SVM related studies. So the current paper, to start with, elaborates the basic principles of SVM. Followed by it summarizes SVM with applications to image annotation from three aspects of SVM ensemble for AIA, SVM with mixture of kernels for AIA and hybrid SVM for AIA respectively. In addition, SVM exploited in several other applications are also briefly reviewed. Finally, we end this paper with a summary of some important conclusions and highlight the potential research directions of SVM in automatic image annotation for the future.

Keywords: *SVM, Image classification, Image annotation, Image retrieval*

1. Introduction

Automatic image annotation (AIA) has become one of the core research topics in computer vision and multimedia areas due to its potential impact on both image understanding and semantic based image retrieval. In general, AIA refers to the process of learning statistical models from a training set of pre-annotated images in order to generate annotations for unseen images using visual feature extracting technology. The goal of AIA is to find suitable annotation words to represent the visual content of an untagged or noisily tagged image. In other words, the correlation between images and annotation words is a central problem in view of technical solutions. As can be seen from the literature, the state-of-the-art research on AIA has proceeded along two categories. The first one poses image annotation as a super- vised classification problem [1, 2], which treats each semantic keyword or concept as an independent class and assigns each keyword or concept one classifier. To be more specific, such kind of approaches predicts the annotations for a new image by computing the similarity at the visual level and propagating the corresponding keywords subsequently. In contrast, the second category treats the words and visual tokens in each image as equivalent features in different modalities. Followed by image annotation is formalized via modeling the joint distribution of visual and textual features on the training data and predicting the missing textual features for a new image. The representative work includes the translation model (TM) [3], cross-media relevance model (CMRM) [4], continuous space relevance model (CRM) [5], multiple Bernoulli relevance model (MBRM) [6], dual cross-media relevance model (DCM- RM) [7], pLSA-words [8], SGMM-RW [9], pLSA-RW [10], pLSA-MB

ISSN: 1738-9968 IJHIT

[11], regularized latent Dirichlet allocation (rLDA) [12] and TagProp [13], *etc*. By comparison, the former approach is relatively direct and natural to be understood. However, its performance is limited with the increase of the number of the semantic concepts and explosive multimedia data on the web. On the other hand, the latter often requires large-scale parameters to be estimated and the accuracy is strongly affected by the quantity and quality of the training data available. Furthermore, it has been argued that the discriminative approach can result in superior performance than that by generative method which is more powerful in capturing decision boundaries among different classes and more widely applied in recognition tasks. For this reason, it is highly desirable to develop approaches with the flexibility of generative learning and the performance of discriminative methods. At present, several hybrid generative discriminative schemes have been proposed to combine the strengths of these two kinds of models in a number of applications, from scene classification [14-16], object recognition [17, 18], image annotation [19], image classification [20-22], speech recognition [23] to biological sequence analysis [24], resulting in state-of-the-art performance.

Support vector machine (SVM) is a well-known example of discriminative methods. As a theoretically rich method and because of its advantages such as its use of over-fitting protect- ion independently from the number of features and its effectiveness in the case of sparse data, SVM has been extensively applied to a wide variety of domains such as computer vision, pattern recognition, object detection and function estimation, etc. In its basic form, an SVM creates a hyperplane as the decision plane, which separates the positive and negative classes with the largest margin. This paper aims to conduct a survey of SVM for automatic image annotation and focus on the important directions, key solutions and remaining open issues in AIA. In addition, some exciting progresses and new potentials for enhancing the performance of SVM in the field of multimedia area are also observed. The remainder of this paper is organized as follows. In section 2, we introduce the basic principle of support vector machine. Section 3 elaborates the SVM for image annotation from three aspects, including SVM ensemble for AIA, SVM with mixture of kernels for AIA and hybrid SVM for AIA, respectively. In section 4, SVM for some other applications are briefly reviewed. Finally, we conclude this paper in section 5 with a summary of some important conclusions and highlight the potential research directions of SVM in AIA for the future.

2. Support Vector Machine

Support vector machine (SVM) is a kind of machine learning algorithm based on the statistical learning theory that works according to the principle of structural risk minimization (SRM) rather than the empirical risk minimization of large samples. SVM has good generalization performance without the need to add a priori knowledge, even when the dimension of the input space is very high [25].

To be specific, SVM works by mapping the training data into a high dimensional feature space. After that it separates the two classes of data with a hyperplane and maximizes the distance which is called the margin. By introducing kernels into the algorithm, it is possible to maximize the margin in the feature space, which is equivalent to nonlinear decision bound- aries in the original input space. Given that the labeled training examples are $(x_1,y_1),...,(x_n,y_n)$, where each $x_k \in R^n$ is the k-th input sample and $y_k \in \{+1,-1\}$ is the k-th output pattern. In their simplest form, SVM can find out the hyperplanes that separate the training data by a maximal margin. All vectors lying on one side of the hyperplane are labeled as -1, and all vectors lying on the other side are labeled as +1. The training instances that locate closest to the hyperplane are called support vectors, as a linearly separable binary classification problem shown in Figure 1.

The goal of SVM is to produce a model that predicts target value of data instances only with the attributes in the testing set. Mathematically, the support vector method can be formu- lated as follows:

$$y(x) = sign(\sum_{k=1}^{n} \alpha_k y_k \psi(x_k, x) + b) \tag{1}$$

where α_k represents the support value of each sample, $\Psi(x_k,x)$ denotes the kernel function that satisfies Mercer's condition, and $y(x)$ stands for the class label predicted by the SVM model. It should be noted that a challenging problem for SVM is the choice of kernel functions which is actually a measure of similarity between two vectors. Gaussian radial basis function (RBF) and polynomial kernel are two commonly used kernel functions listed as follows:

$$K(x_k, x)_{RBF} = \exp(-\gamma \| x_k - x \|^2), \gamma > 0 \tag{2}$$

$$K(x_k, x)_{Poly} = [(x_k \times x) + c]^q \tag{3}$$

As for the parameters of SVM, they are usually estimated by n-fold cross-validation and grid-search algorithms respectively. In addition, since image annotations are not likely to be linearly separable in the projected space, it needs to be allowed for some training errors. This need gives rise to the soft-margin SVM algorithm, which can be formulated as a special case of the hard margin version with the modified kernel by adding a factor to penalize training errors. More details can be gleaned from reference [26].

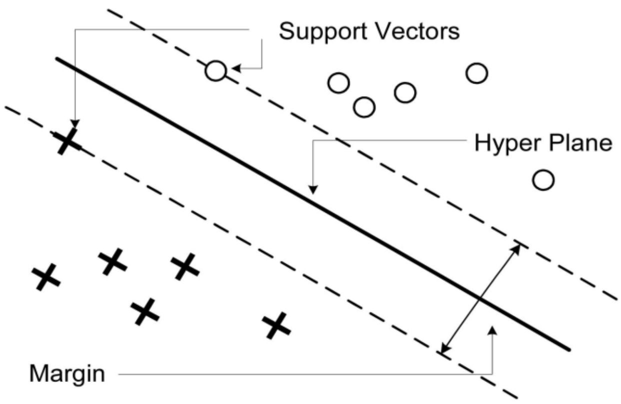

Figure 1. SVM for a Linearly Separable Binary Classification Problem

3. SVM for Image Annotation

As is well known, support vector machine is a supervised classifier, which has been shown with high effectiveness in high dimensional data classifications, especially when the training dataset is small. SVM can classify both linear and non-linear data due to the use of kernel mapping. The advantage of SVM over other classifiers is that it can achieve optimal class boundaries by finding the maximum distance between classes. So far, it has been successfully applied to a number of classification problems, such as text classification, object recognition and image annotation, etc. It should be noted that this paper mainly focuses on SVM with application to automatic image annotation. Furthermore, it will be revisited from three aspects of SVM ensemble for AIA, SVM with mixture of kernels for AIA and hybrid SVM for AIA, respectively.

3.1. SVM Ensemble for AIA

Compared to common classifiers, ensemble scheme is a more flexible and cheaper approach when applied to the machine learning community. Some results report that it is more reliable than most one-level classifier in the task of automatic image annotation. In general, ensemble method can be divided into two categories based on their design goals. The first one is to reduce prediction variance caused by training data selection, including the well-known bagging, arcing and boosting, *etc*. In contrast, the goal of the second ensemble category is to reduce prediction error by using decomposition and reconstruction methods with good error- correction capability, such as one per class, pair-wise coupling, error-correcting output coding and so on [27]. It is worth noting that to construct ensemble classifiers for image annotation and retrieval has attracted much more attentions recently.

SVM, in actual fact, is a binary classifier and mainly applied to two-class problems, but it can be adapted to multi-class problems by ensemble schemes. It is known that automatic image annotation needs multi-class classifier. The most common approach is to train a separate SVM for each concept with each SVM generating a probability value. During the testing phase, the decisions from all classifiers are fused to get the final class label of a test image. Figure 2 illustrates the process. Note that the base level consists of multiple binary classifiers and the second level fuses the decisions from the base level classifiers.

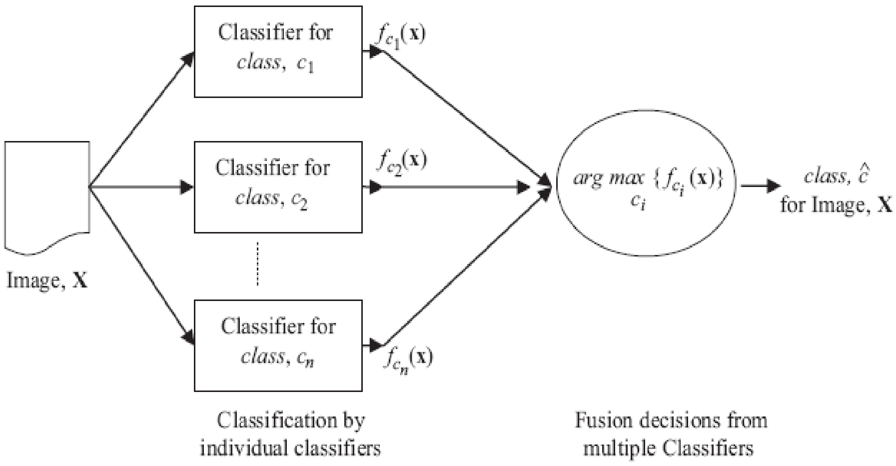

Figure 2. Multi-class Classifier using Multiple Binary SVM Classifiers

Based on the basic framework shown in Figure 1, Cusano, *et al.,* [28] employ SVM to do image segmentation and classification simultaneously. More specifically, an image is first partitioned into overlapping tiles that are sampled at fixed interval. Each pixel is covered by a number of tiles. The tiles are classified independently into one of the seven predefined concepts. The concept of the pixel is determined by the majority voting from the classes of its parent tiles. Pixels belonging to the same concept constitute a segment. Thus the approach can simultaneously segment images into regions and annotate the corresponding regions. An early approach to ensemble classifiers for AIA is the innovative work of Chang, *et al.,* [29], who present CBSA to semi-automate the process of annotating unlabeled images with multiple soft labels. Based on the scenario of ensemble SVMs, Li, *et al.,* [30] present a confidence-based dynamic ensemble by bagging SVMs to improve the performance of image annotation based on traditional static classifiers, in which color and texture features are used to train the ensemble SVMs for image annotation. In the approach [31], images are segmented using k-means algorithms and 23 SVM classifiers are trained to learn 23 region level concepts for semantic image annotation. Goh, *et al.*, [32] construct a three-level multiple sets of SVMs for multi-class image annotation by fusing one-class, two-class and multi-class SVMs. To be specific, level 1 consists of several sets of classifiers trained by different subset of training samples, and its probabilistic outputs are normalized by each set of classifiers before using them in level 2. Level 2 is the fusion process used to find a confidence factor in addition to the highest probabilistic decision. At level 3, the confidence factors of same concept are added together, and the concept with the maximum cumulative confidence is the final decision corresponding to the final annotation. Afterwards a combination of multiple SVM classifiers is leveraged for AIA [33], which is constructed by combining the output of several effective weak classifiers using a boosting technique. In [34], the CLAIRE image annotation system is developed based on the idea of stacked generalization, in which the first-level classifiers focus on classifying color and texture features respectively and the outputs of the classifiers are fed into the second-level classifier for the final classification. This system shows a promising way to assign keywords to images due to it avoids the direct mapping of the low-level features to high-level semantic concepts. Subsequently, Qi and Han [35] leverage a similar framework to [32] by combining MIL-based SVMs and global-feature-based SVMs for image annotation, but they fuse the decisions classifier by classifier instead of set by set. Meanwhile, they use both global and local features in two different sets of SVMs, which can effectively compensate the limitations of one type of feature by the other. However, this approach has a high time complexity and it doesn't take into account the semantic relationship between blocks. For more details and a more complete explanation please refer to the corresponding literatures.

3.2. SVM with Mixture of Kernels for AIA

Support vector machine is a kind of machine learning algorithm. By introducing kernels into SVM, it is possible to maximize the margin in the feature space that is equivalent to the nonlinear decision boundaries in the original space. The design of SVM classifier is very simple and mainly requires the choice of the kernel function. However, it has to be chosen carefully since an inappropriate kernel can lead to poor performance. To this end, much research effort has been devoted to the study of kernel function [36-38]. In addition, it is reported that kernels used by SVM can be roughly divided into two categories: global and local kernels. In global kernels, points far away from the test point have a great effect on kernel values. But in local kernels, only those close to the test point have a great impact on kernel values. Among the existing kernels, the radial basis function (RBF) and polynomial kernel function are two typical local and global kernels respectively. Specifically, the former can provide good fitting performance for SVM while the latter can restrain the fluctuation and keep a stable prediction performance. Furthermore, research shows that non-linear SVM with the Gaussian radial basis function (RBF) kernel is able to yield excellent results compared with linear and polynomial kernels respectively. Hence Scholkopf, *et al.*, [36] capitalize on the Gaussian RBF kernel in their nonlinear SVM systems.

In early work [37], Chapelle, *et al.*, compare SVMs based on polynomial and Gaussian RBF kernels learned by global RGB and HSV color histograms individually to annotate images. In particular, this paper details how to choose the distance used in a RBF kernel that affects the performance on histogram classification and how to find Laplacian RBF kernels to be superior to the standard Gaussian RBF kernels. Followed by Smits and Jordaan [38] define the mixtures of the RBF and polynomial kernels for SVM. Similarly, a new kind of SVM method based on the weighted linear combination of RBF and polynomial kernel is proposed in [39]. Simulation results validate the good performance of the mixtures of kernels for SVM in modeling analysis compared to any single kernel. Besides, Zheng *et al.*, [40] put forward a star cluster grouping algorithm based on the second order directional derivative operators deduced from the mapped least squares support vector machine (LS-SVM) with mixtures of RBF and polynomial kernels. The experimental results demonstrate that the proposed app- roach with the mixtures of RBF and polynomial kernels provides more optimal performance than that with any single kernel and has high robustness and efficiency. Recently, Bouguila *et al.*, [41] bring forward a classification scheme that incorporates both finite multinomial Dirichlet mixture models and SVM in a way that combines their respective advantages. Experiments show that SVM with the proposed multinomial Dirichlet mixture (MDM) kernel significantly outperforms those with polynomial, Gaussian, RBF, Sigmoid and finite multi- nomial mixture kernels respectively. More recently, Wei *et al.*, [42] propose an AIA approach by using multi-class SVM with ontology to achieve a higher accuracy, in which the kernel is formed with RBF and polynomial kernels to realize advantageous complementarities. Especially in our previous work [43], a novel SVM with mixture of kernels (SVM-MK) for automatic image annotation is presented. On the one hand, the combined global and local block- based image features are extracted in order to reflect the intrinsic content of images as complete as possible. SVM-MK, on the other hand, is constructed based on the mixture of kernels to shoot for better annotation performance. Figure 3 illustrates the framework of SVM-MK for automatic image annotation.

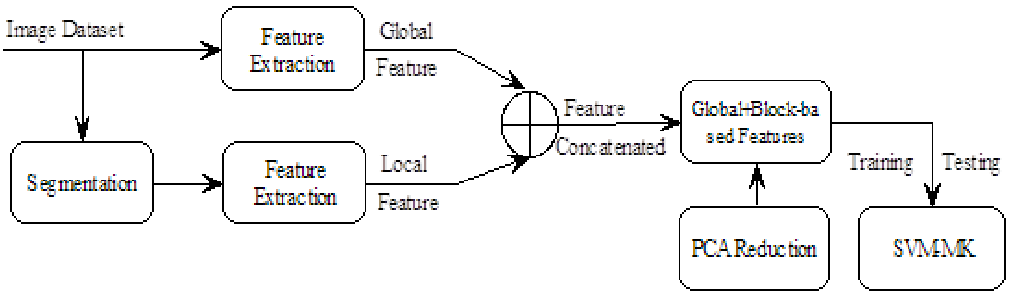

Figure 3. Framework of SVM-MK for AIA

3.3. Hybrid SVM for AIA

In recent years, hybrid SVM has become a desirable research direction in the field of computer vision. An early approach to AIA is the innovative work of Andrews, *et al.*, [44] who propose the MI-SVM method. The instance-based (*i.e.*, region-based) image features are iteratively fed into SVMs until no updates for all the positive training images. The converged instance-based features are then used to annotate an image. Followed by Chen, *et al.*, [45] propose diverse density support vector machine method (DD-SVM), which combines EM-DD with SVMs, to construct bag-based image features using multiple local maxima instead of one global maximum. These features are thereafter fed into SVMs to form the hyperplanes for image classification and annotation. However, this method is limited by region naming and learning certain concepts. Experiments validate that the DD-SVM method can achieve the best annotation accuracy while the MI-SVM method can achieve better annotation accuracy than both EM-DD and DD methods. In addition, Han and Qi [46] integrate multiple instance learning (MIL) based SVM with global feature based SVM for effective image annotation. Note that MIL is applied to the image blocks to obtain bag features to be input to a set of SVMs for finding the optimum hyperplanes for categorizing training images, while the global features are fed into another set of SVMs to categorize training images. For each test image, two sets of image features are constructed and sent to the respective set of SVMs, and the outputs from these SVMs are incorporated to obtain the final image annotation results. Yang, *et al.*, [47] propose asymmetrical support vector machine based multiple-instance learning (MIL) algorithm to extend the conventional support vector machine through MIL so as to train the SVM in a new space for region-based image annotation. Afterwards Zhao and Zhu [48] construct TSVM-HMM for automatic image annotation, which integrates the discriminative classification with the generative model to mutually complete their advantages for better annotation performance. Lu, *et al.*, [49] exploit SVM to annotate main concepts of five kinds of animals in images.

Recently, Jiang, *et al.*, [50] adopt learning vector quantization (LVQ) technique to optimize low-level features extracted from given images, and then select some representative vectors with LVQ to train SVM classifiers instead of using all feature data for semantic image annotation. In [51], a hierarchical annotation scheme is presented by considering that human's visual identification to a scenery object is a rough-to-fine hierarchical process. To be specific, the input image is first segmented into multiple regions and each segmented region is roughly labeled with a general keyword using the multi-classification SVM, then an active semi- supervised expectation-maximization (EM) algorithm is employed to find the representative pattern of each fine class and classify the roughly labeled regions into corresponding fine classes, finally the contextual relationship is applied again to revise the improper fine labels. In the meanwhile, Huang, *et al.*, [52] put forward a novel approach for AIA by extracting color, texture and shape features from a set of training images to build the main object classifier as well as background object models by using support vector machine. Particularly, the combination of active contour model (ACM) and JSEG is leveraged to improve the system performance, and Gaussian mixture model (GMM) is employed to explore the relationship between image classes and image backgrounds based on the built association knowledge base. Similar to the ideas of [34], Lei, *et al.*, [53] present a two-stage mapping AIA technique by combining both hidden Markov model (HMM) and support vector machine (HMM-SVM). In the first stage, two HMMs are constructed separately from color and texture features of images for mapping the low-level features to mid-level features. In the second stage, SVM is applied to map the so-called mid-level features to high-level concepts. Figure 4 illustrates the framework of HMM-SVM for automatic image annotation.

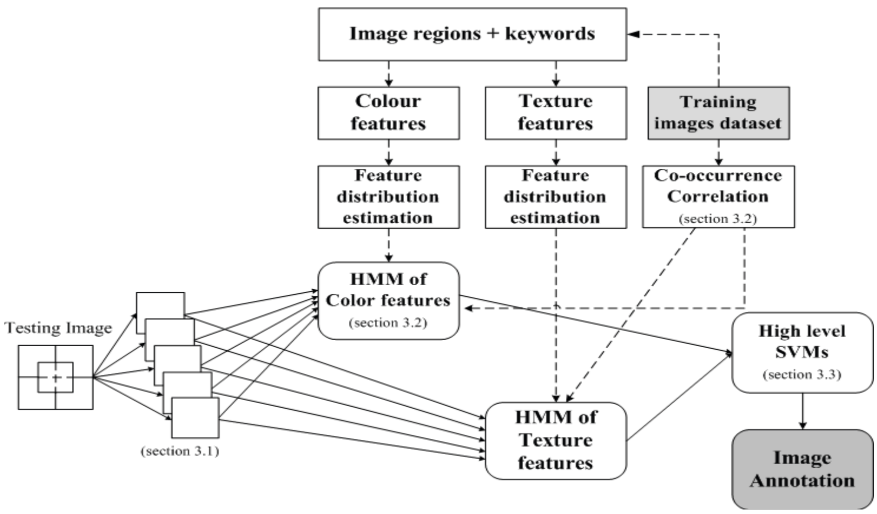

Figure 4. Block Diagram of HMM-SVM Approach

More recently, Alham, *et al.*, [54] present a distributed SVM algorithm for large-scale image annotation (MRSVM), which partitions the training dataset into smaller subsets and trains SVM in parallel using a cluster of computing nodes. To be specific, MRSVM builds on the sequent minimal optimization algorithm for high efficiency in training and employs Map- reduce for parallel computation across a cluster of computers. Subsequently, they propose a Mapreduce-based distributed SVM ensemble algorithm

(MRESVM) for scalable image annotation [55] that is designed based on the bagging architecture by training multiple SVMs on bootstrap training datasets and combining the output in an appropriate manner. In parti- cular, a balanced sampling strategy for bootstrap is introduced to increase the classification accuracy. More details can be gleaned from the corresponding literature.

4. SVM for Other Applications

Aside from the content aforementioned, there are many more other applications of SVM for pattern recognition problems, such as medical image annotation [56,57], face detection [58], image classification [59, 60], facial expression recognition [61], video detection [62-64] and adult image recognition [65,66], *etc.* Such as work [59], Serrano et al. exploit color and texture SVMs to classify color and texture features of 16 blocks per image into indoor and outdoor classes individually. Then, two decision values associated with each block of an image assigned by the color and texture SVMs are further used to train a new SVM. The classification result is better than the majority voting suggested in other literatures. Besides, Li, *et al.*, [60] present a multi-label SVM active learning method for multi-label image classi- fication. Especially the max loss strategy and mean max loss strategy are employed as the optimization of selection strategies in their approach. In the approach [57], Amaral, *et al.*, address the problem of hierarchical medical image annotation by building a content-based image retrieval system aiming to explore the combination of three different methods using SVMs. Experimental results show that even if almost all fusion methods result in an improvement over standalone classifications, none clearly outperforms each other. In the following, several SVM related automatic image annotation methods involved in this paper are concisely summarized in Table 1, mainly including the classifiers and image datasets adopted in the corresponding literature.

Table 1. Summary of SVM Related Automatic Image Annotation Models

Sources	Classifiers	Image Datasets
Cusano, *et al.*, [28]	SVM	WWW Dataset
Li, *et al*., [30]	Ensemble SVMs	COREL/WWW Datasets
Shi, *et al.*, [31]	SVM	COREL/Other Datasets
Goh, *et al.*, [32]	Ensemble SVMs	COREL Dataset
Tsai, *et al.*, [34]	Ensemble SVMs	COREL Dataset
Qi, *et al.*, [35]	SVM, MIL	COREL Dataset
Chapelle, *et al.*, [37]	SVM	COREL14/COREL7 Datasets
Wei, *et al.*, [42]	SVM	COREL Dataset
Tian, *et al.*, [43]	SVM-MK	COREL Dataset
Chen, *et al.*, [45]	DD-SVM	COREL/MUSK Datasets
Han and Qi [46]	MIL, SVM	COREL Dataset
Yang, *et al.*, [47]	SVM, MIL	COREL Dataset
Zhao, *et al.*, [48]	TSVM, HMM	COREL Dataset
Lu, *et al.*, [49]	SVM	COREL Dataset
Jiang, *et al.*, [50]	SVM, LVQ	COREL Dataset
Gao, *et al.*, [51]	SVM, Semi-supervised EM	COREL Dataset
Huang, *et al.*, [52]	SVM, GMM, ACM	COREL Dataset
Lei, *et al.*, [53]	HMM-SVM	COREL Dataset
Alham, *et al.*, [54]	MRSVM, SMO, MapReduce	COREL Dataset
Alham, *et al.*, [55]	MRESVM, SMO, MapReduce	COREL Dataset
Qiu [56]	SVM	ImageCLEF2006 Dataset
Serrano, *et al.*, [59]	Ensemble SVMs	Other Dataset
Feng, *et al.*, [67]	Ensemble SVMs	COREL Dataset
Boutell, *et al.*, [68]	SVM	COREL/Other Datasets
Fan, *et al.*, [69]	SVM	COREL/WWW Datasets

5. Conclusions

Support vector machine is a core machine learning technology that has strong theoretical foundations and excellent empirical successes. In recent years, SVM has been applied to a wide variety of domains such as pattern recognition, object detection and so on. Compared with large amount of SVMs and their corresponding applications, there is almost no review research and analysis about SVM related studies. So in this paper, a number of encouraging SVM approaches for AIA have been presented so as to complement the existing surveys in literature. Specifically, we spotlight on SVM for automatic image annotation from three aspects, including SVM ensemble for AIA, SVM with mixture of kernels for AIA and hybrid SVM for AIA, respectively. The primary purpose of this paper is to illustrate the pros and cons of SVM combined with a great deal of existing works as well as to point out the promising research directions of SVM for automatic image annotation in the future.

In most cases, although SVM can obtain a satisfying annotation performance and seem to be relatively easily for implementation, it still suffers from several issues remain to be investigated. First, SVM has class-imbalance problem, which means that it has poor performance on imbalanced data. Unfortunately, class-imbalance is a common phenomenon in image data, and which will inevitably degrade the quality of the SVM classifiers. Second, it is known that kernel function plays an important role in the implementation of SVM. However, in most of the applications, the intrinsic structure of the image data has been ignored by these standard kernels. Moreover, it is shown that the kernel function should be generated directly from data which gives better results. So how to formulate appropriate kernels in terms of the image data is a worthy research direction. Third, as described thus far, the most obvious drawback to SVM algorithm is that it apparently only handles binary classification problems. So how to extend SVM for multi-class classification problems by considering the compromise of efficiency and accuracy is a valuable research direction in the future. Fourth, as for large-scale image annotation, there is a common phenomenon that many conceptually different categories are visually similar in the feature space which may cause feature overlapping and thus degrade the generalization performance. So how to integrate the contextual and correlation information of candidate annotations into the process of image annotation for SVM can be of great help to improve the performance of AIA. Last but not the least, image segmentation remains an open research problem in computer vision, especially for those approaches rely very much on local features, which in turn rely on high segmentation accuracy. However, segmentation can hardly be done reliably, particularly on compressed images. So to explore more efficient image segmentation methods is helpful to boost the overall annotation accuracy. Furthermore, image segmentation itself is a worthy of further research direction.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (No.61379030), the Natural Science Basic Research Plan in Shaanxi Province of China (No.2014JQ2-6036) and the Science and Technology R&D Program of Baoji City (No.203020013, No.2013R2-2).

References

[1] J. Li, J. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, **(2003)**, pp.1075- 1088.

[2] G. Carneiro, A. Chan, P. Moreno and N. Vasconcelos, "Supervised learning of semantic classes for image annotation and retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 3, **(2007)**, pp. 394-410.

[3] P. Duygulu, K. Barnard and N. de Freitas, “Object recognition as machine translation: learning a lexicon for a fixed image vocabulary”, In Proceedings of ECCV, **(2002)**, pp. 97-112.
[4] L. Jeon, V. Lavrenko and R. Manmantha, “Automatic image annotation and retrieval using cross-media relevance models”, In Proceedings of SIGIR, **(2003)**, pp. 119-126.
[5] R. Manmatha, V. Lavrenko and J. Jeon, “A model for learning the semantics of pictures”, In Proceedings of NIPS, **(2003)**, pp. 553-560.
[6] S. Feng, R. Manmatha and V. Lavrenko, “Multiple Bernoulli relevance models for image and video annotation”, In Proceedings of ICPR, **(2004)**, pp. 1002-1009.
[7] J. Liu, B. Wang and M. Li, “Dual cross-media relevance model for image annotation”, In Proceedings of ACM MM, **(2007)**, pp. 605-614.
[8] F. Monay and D. Gatica-Perez, “pLSA-based image auto-annotation: constraining the latent space”, In Proceedings of ACM MM, **(2004)**, pp. 348-351.
[9] D. Tian, “Semi-supervised learning for refining image annotation based on random walk model”, Knowledge-Based Systems, vol. 72, no. 12, **(2014)**, pp. 72-80.
[10] D. Tian, X. Zhao and Z. Shi, “An efficient refining image annotation technique by combining probabilistic latent semantic analysis and random walk model”, Intelligent Automation & Soft Computing, vol. 20, no. 3, **(2014)**, pp. 335-345.
[11] D. Tian, W. Zhang, X. Zhao and Z. Shi, “Employing PLSA model and max-bisection for refining image annotation”, In Proceedings of ICIP, **(2013)**, pp. 3996-4000.
[12] H. Xu, J. Wang, X. Hua and S. Li, “Tag refinement by regularized LDA”, In Proceedings of ACM MM, **(2009)**, pp. 573-576.
[13] M. Guillaumin, T. Mensink, J. Verbeek and C. Schmid, “TagProp: discriminative metric learning in nearest neighbor models for image auto-annotation”, In Proceedings of ICCV, **(2009)**, pp. 309-316.
[14] A. Bosch, A. Zisserman and X. Muoz, “Scene classification using a hybrid generative / discriminative approach”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, **(2008)**, pp. 712-727.
[15] A. Perina, M. Cristani and U. Castellani, “A hybrid generative /discriminative classification framework based on free-energy terms”, In Proceedings of ICCV, **(2009)**, pp. 1-8.
[16] X. Li, T. Lee and Y. Liu, “Hybrid generative-discriminative classification using posterior divergence”, In Proceedings of CVPR, **(2011)**, pp. 2713-2720.
[17] D. Zhang and S. Chang, “A generative-discriminative hybrid method for multi-view object detection”, In Proceedings of CVPR, **(2006)**, pp. 1-8.
[18] A. Holub, M. Welling and P. Perona, “Hybrid generative-discriminative visual categorization”, International Journal of Computer Vision, vol. 77, no. 1, **(2008)**, pp. 239-258.
[19] S. Yang, J. Bian and H. Zha, “Hybrid generative / discriminative learning for automatic image annotation”, In Proceedings of UAI, **(2010)**, pp. 1-8.
[20] Y. Li, L. Shapiro and J. Bilmes, “A generative/discriminative learning algorithm for image classification”, In Proc. ICCV, **(2005)**, pp. 1605-1612.
[21] M. Enzweiler, D. Gavrila, “A mixed generative-discriminative framework for pedestrian classification”, In Proceedings of CVPR, **(2008)**, pp. 1-8.
[22] M. Zare, A. Mueen and M. Awedh, “Automatic classification of medical X-ray images: hybrid generative-discriminative approach”, IET Image Processing, vol. 7, no. 5, **(2013)**, pp. 523-532.
[23] N. Smith and M. Gales, “Speech recognition using SVMs”, In Proceedings of NIPS, **(2002)**, pp. 117-129.
[24] K. Tsuda, T. Kin and K. Asai, “Marginalized kernels for biological sequences”, Bioinformatics, vol. 18, no. 1, **(2002)**, pp. 268-275.
[25] S. Bernhard and J. Alexander, “Learning with kernels-support vector machines, regularization, optimization and beyond”, The MIT Press, Cambridge, Massachusetts, **(2003)**.
[26] C. Cortes and V. Vapnik, “Support vector networks. Machine Learning”, vol. 20, no. 3, **(1995)**, pp. 1-25.
[27] K. Goh, E. Chang and K. Cheng, “SVM binary classifier ensembles for image classification”, In Proceedings of CIKM, **(2001)**, pp. 395-402.
[28] C. Cusano and G. Ciocca, “Image annotation using SVM”, In Proceedings of II, SPIE, vol. 5304, **(2004)**.
[29] E. Chang, K. Goh, G. Sychay and G. Wu, “CBSA: content-based soft annotation for multimodal image retrieval using Bayes point machines”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 1, **(2003)**, pp. 26-38.
[30] B. Li and K. Goh, “Confidence-based dynamic ensemble for image annotation and semantics discovery”, In Proceedings of ACM MM, **(2003)**, pp. 195-206.
[31] R. Shi, H. Feng, T. Chua and C. Lee, “An adaptive image content representation and segmentation approach to automatic image annotation”, In Proceedings of CIVR, **(2004)**, pp. 545-554.
[32] K. Goh, E. Chang and B. Li, “Using one-class and two-class SVMs for multiclass image annotation”, IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 10, **(2005)**, pp. 1333-1346.
[33] Y. Gao, J. Fan and H. Luo, “Automatic image annotation by incorporating feature hierarchy and boosting to scale up SVM classifiers”, In Proceedings of ACM MM, **(2006)**, pp. 901-910.

[34] C. Tsai and K. McGarry, "CLAIRE: a modular support vector image indexing and classification system", ACM Transactions on Information Systems, vol. 24, no. 3, **(2006)**, pp. 353-379.
[35] X. Qi and Y. Han, "Incorporating multiple SVMs for automatic image annotation", Pattern Recognition, vol. 40, no. 2, **(2007)**, pp. 728-741.
[36] B. Scholkopf, K. Sung and C. Burges, "Comparing support vector machines with Gaussian kernels to radial basis function classifiers", MIT, A.I. Memo, vol. 1599, **(1996)**.
[37] O. Chapelle and P. Haffner, "Support vector machines for histogram-based image classification", IEEE Transactions on Neural Networks, vol. 10, no. 5, **(1999)**, pp. 1055-1064.
[38] G. Smits and E. Jordaan, "Improved SVM regression using mixtures of kernels", In Proceedings of IJCNN, **(2002)**, pp. 2785-2790.
[39] Y. Zhu, L. Tian and W. Li, "Mixtures of kernels for SVM modeling", In Proceedings of ICNC, **(2005)**, pp. 601-607.
[40] S. Zheng, J. Liu and J. Tian, "An efficient star acquisition method based on SVM with mixtures of kernels", Pattern Recognition Letters, vol. 26, no. 2, **(2005)**, pp. 147-165.
[41] N. Bouguila, "A discrete mixture-based kernel for SVMs: application to spam and image categorization", Information Processing and Management, vol. 45, no. 6, **(2009)**, pp. 631- 642.
[42] Z. Wei, X. Luo and F. Zhou, "Ontology based automatic image annotation using multi-class SVM", In Proceedings of ICIG, **(2013)**, pp. 434-438.
[43] D. Tian, X. Zhao and Z. Shi, "Support vector machine with mixture of kernels for automatic image annotation", High Technology Letters, vol. 19, no. 3, **(2013)**, pp. 295-300.
[44] S. Andrews, I. Tsochantaridis, T. Hofmann, "Support vector machines for multiple-instance learning", In Proceedings of NIPS, **(2002)**, pp. 561-568.
[45] Y. Chen and J. Wang, "Image categorization by learning and reasoning with regions", The Journal of Machine Learning Research, vol. 5, **(2004)**, pp. 913-939.
[46] Y. Han and X. Qi, "A complementary SVMs-based image annotation system", In Proc. ICIP, **(2005)**, pp. 1185-1188.
[47] C. Yang, M. Dong and J. Hua, "Region-based image annotation using asymmetrical support vector machine-based multiple-instance learning", In Proceedings of CVPR, **(2006)**, pp. 2057-2063.
[48] Y. Zhao and Y. Zhao, "TSVM-HMM: transductive SVM based hidden Markov model for automatic image annotation", Expert Systems with Applications, vol. 36, no. 6, **(2009)**, pp. 9813-9818.
[49] H. Lu, Y. Zheng, X. Xue and Y. Zhang, "Content and context-based multi-label image annotation", In Proceedings of CVPRW, **(2009)**, pp. 61-68.
[50] Z. Jiang, J. He and P. Guo, "Feature data optimization with LVQ technique in semantic image annotation" In Proceedings of ISDA, **(2010)**, pp. 906-911.
[51] Y. Gao, Y. Yin and T. Uozumi, "A hierarchical image annotation method based on SVM and semi-supervised EM", Acta Automatica Sinica, vol. 36, no. 7, **(2010)**, pp. 960-967.
[52] Y. Huang and H. Lu, "Automatic image annotation using multi-object identification", In Proceedings of PSIVT, **(2010)**, pp. 386-392.
[53] Y. Lei, W. Wong, W. Liu and M. Bennamoun, "An HMM-SVM-based automatic image annotation approach", In Proceedings of ACCV, **(2010)**, pp. 115-126.
[54] N. Alham, M. Li and Y. Liu, "A distributed SVM for scalable image annotation", In Proceedings of FSKD, **(2011)**, pp. 2655-2658.
[55] N. Alhama and M. Li, "A MapReduce-based distributed SVM ensemble for scalable image classification and annotation", Computers & Mathematics with Applications, vol. 66, no. 10, **(2013)**, pp. 1920-1934.
[56] B. Qiu, "A refined SVM applied in medical image annotation", In Proceedings of CLEF, **(2007)**, pp. 690-693.
[57] I. Amaral, F. Coelho, J. Costa and J. Cardoso, "Hierarchical medical image annotation using SVM-based approaches", In Proceedings of ITAB, **(2010)**, pp. 1-5.
[58] C. Waring and X. Liu, "Face detection using spectral histograms and SVMs", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 35, no. 3, **(2005)**, pp. 467-476.
[59] N. Serrano, A. Savakis and J. Luo, "A computationally efficient approach to indoor/outdoor scene classification", In Proceedings of ICPR, **(2002)**, pp. 146-149.
[60] X. Li, L. Wang and E. Sung, "Multi-label SVM active learning for image classification", In Proceedings of ICIP, **(2004)**, pp. 2207-2210.
[61] I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines", IEEE Transactions on Image Processing, vol. 16, no. 1, **(2007)**, pp. 172-187.
[62] Y. Zhu and Z. Ming, "SVM-based video scene classification and segmentation", In Proceedings of MUE, **(2008)**, pp. 407-412.
[63] Y. Zhu, Z. Ming and J. Zhang, "Video scene classification and segmentation based on support vector machine", In Proceedings of IJCNN, **(2008)**, pp. 3571-3576.
[64] Y. Hu and G. Zhai, "Automatic classification system of wushu video based on SVM", In Proceedings of FSKD, **(2009)**, pp. 124-128.
[65] Z. Zhao and A. Cai, "Combining multiple SVM classifiers for adult image recognition", In Proceedings of ICNIDC, **(2010)**, pp. 149-153.

[66] B. Choi, B. Chung and J. Ryou, "Adult image detection using Bayesian decision rule weighted by SVM probability", In Proceedings of ICCIT, **(2009)**, pp. 659-662.
[67] H. Feng and T. Chua, "A bootstrapping approach to annotating large image collection", In Proceedings of ACM MIR, **(2003)**, pp. 55-62.
[68] M. Boutell, J. Luo and X. Shen, "Learning multi-label scene classification. Pattern Recognition", vol. 37, no. 9, **(2004)**, pp. 1757-1771.
[69] J. Fan, Y. Gao, H. Luo and G. Xu, "Automatic image annotation by using concept-sensitive salient objects for image content representation", In Proceedings of SIGIR, **(2004)**, pp. 361-368.

