Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **6**, November **2018**

Support Vector Machine for Content-based Image Retrieval: A Comprehensive Overview

Dongping Tian

Institute of Computational Information Science
Baoji University of Arts and Sciences
No.1 Hi-Tech Avenue, Hi-Tech District, Baoji, Shaanxi 721013, P.R. China

Institute of Computer Software
Baoji University of Arts and Sciences
No.44 Baoguang Road, Weibin District, Baoji, Shaanxi 721016, P.R. China
tiandp@ics.ict.ac.cn, tdp211@163.com

Received June, 2018; revised September, 2018

Abstract. *Content-based image retrieval (CBIR) has been an active topic of research in computer vision and pattern recognition. In the last two decades, a large number of works on CBIR have been proposed, among which support vector machine (SVM) is one of the most commonly used methods due to its good generalization performance on pattern classification problems without incorporating problem domain knowledge. However, compared to various SVM methods and their corresponding applications in content-based image retrieval, there is almost no review research and analysis about SVM focusing on CBIR. So in the current paper, we mainly revisit some of the important contributions associated with image retrieval based on SVM in the last two decades, spanning 66 references. In the meanwhile, some of the key challenges of SVM involved in the adaptation of existing content-based image retrieval techniques are also discussed so as to explore more powerful strategies and develop more efficient systems to handle the large scale real-world image collections in the future.*
Keywords: SVM, CBIR, Image classification, GMM, Deep learning, Relevance feedback, Active learning

1. **Introduction.** In recent years, content-based image retrieval (CBIR) has been an active research topic in multimedia information processing because of the rapidly increasing requirements in many practical fields, ranging from architectural design, digital libraries and biomedicine to military and education, etc. The main idea of CBIR is to retrieve within large collection images matching a given query based on their visual content analysis. From the literature, it can be easily observed that a great amount of previous researches on CBIR focus on extracting low-level visual features and applying them directly to compute image similarities. Nonetheless, it has shown that low-level image features cannot always capture the semantic concepts in the image [1]. In other words, due to the existence of the two gaps in CBIR, viz., the sensor gap between the object of the world and the information represented by computers as well as the semantic gap between the low-level visual features and high-level human perception and interpretation, the conventional CBIR systems still cannot achieve satisfactory retrieval performance for many practical applications. As a result, much endeavor has been devoted to bridging it in the past few years, including QBIC, Virage, Mars, Netra, PicHunter, Blobworld, Walrus,

Visualseek, SIMPLIcity and others [2]. To be more specific, in a typical CBIR system, features related to visual content such as color, texture, shape and spatial layout are first extracted from a query image, followed by the similarity between the set of features of the query image and that of each target image in a database is computed, and target images are next retrieved which are most similar to the query image.

In addition, it is worth noting that among various machine learning methods, support vector machine has been widely applied in the field of CBIR. SVM is a well-known discriminative approach in the community of machine learning. As a theoretically rich method and because of its advantages such as its use of over-fitting protection 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 comprehensive survey of SVM for content-based image retrieval and focus on the important directions, key solutions and remaining open issues in CBIR. Furthermore, some exciting progresses and new potentials for enhancing the performance of SVM in the area of multimedia 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 retrieval from three aspects, including SVM ensemble for CBIR, hybrid SVM for CBIR, and SVM for some other applications, respectively. Finally, we conclude this paper in section 4 with a summary of some important conclusions and highlight the potential research directions of SVM in CBIR for the future.

2. **Support Vector Machine.** Support vector machine (SVM) is a universal classification algorithm proposed in the middle of the 1990s. 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. Compared with other machine learning algorithms, SVM is considered as a good candidate because of its high generalization performance without the need to add *a priori* knowledge, even when the dimension of the input space is very high [3].

The basic idea of SVM for binary classification is to find an optimal separating hyperplane that maximizes the margin between two classes in a kernel-induced feature space. 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 that is equivalent to nonlinear decision boundaries in the original input space. Given that the labeled training examples are $(x_1, y_1), \cdots, (x_n, y_n)$, where each $x_i \in R^n$ is the i-th input sample and $y_i \in \{+1, -1\}$ is the i-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 Fig. 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. Especially for the non-linear inseparable, the support vector method can be formulated as follows:

$$y(x) = sign(\sum_{i=1}^{n} \alpha_i y_i \psi(x_i, x) + b) \tag{1}$$

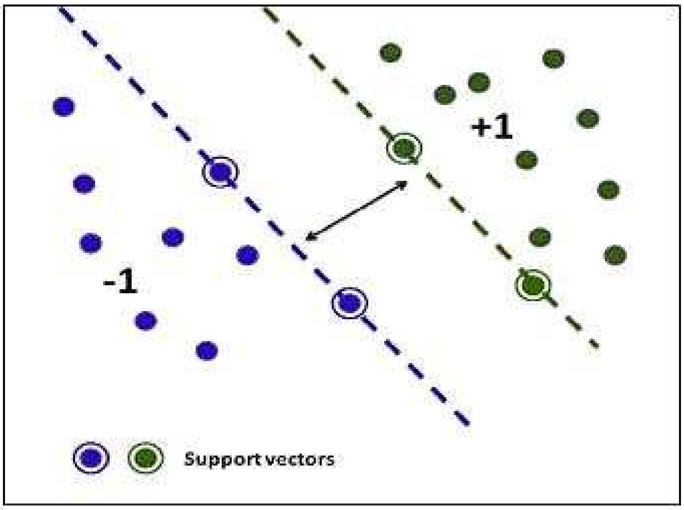

FIGURE 1. SVM for a linearly separable binary classification problem

where α_i represents the support value of each sample, $\psi(x_i, x)$ 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. For more details on the choice of kernel functions, one can refer to the literature [4]. Furthermore, the kernels exploited by SVM can be roughly classified into two categories: global and local kernels. In global kernel, the data points that are far away from the test point have a great effect on the kernel value while in local kernel, only those that are close to the test point have a crucial effect on the kernel value. Alternatively, SVM is able to construct a variety of learning machines by using different kernel functions. The most commonly used kernels are summarized in Table 1. In addition, as for the parameters of SVM, they are usually estimated by n-fold cross-validation and grid-search algorithms respectively. 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 [5].

TABLE 1. Summary of the commonly used kernel functions for SVM

Names	**Expressions**	**Related parameters**
Linear kernel	$k(x_i, x) = x_i \cdot x$	
Polynomial kernel	$k(x_i, x) = (x_i \cdot x + c)^q$	
RBF kernel	$k(x_i, x) = exp(-\gamma \|\|x_i - x\|\|^2), \gamma > 0$	γ, q, c, a and θ
Wavelet kernel	$k(x_i, x) = \prod_{i=1}^{m} (cos(1.75 \frac{x_i - x}{a}) exp(-\frac{\|\|x_i - x\|\|^2}{2a^2}))$	
Sigmoid kernel	$k(x_i, x) = tanh[\gamma(x_i \cdot x) + \theta]$	

3. **SVM for Image Retrieval.** As is well known, we do not yet have a universally acceptable algorithmic means of characterizing human vision, more specifically in the context of image understanding. Hence it is not surprising to see continuing efforts towards it, either building up on prior work or exploring novel directions in the field of computer vision. As is known, content-based image retrieval has been widely studied in the last two decades, whose main idea is to retrieve within large collections images matching a given query thanks to their visual content analysis (as shown in Fig. 2). So in this section, we will focus on some recent literatures about key aspects of CBIR using support vector machine.

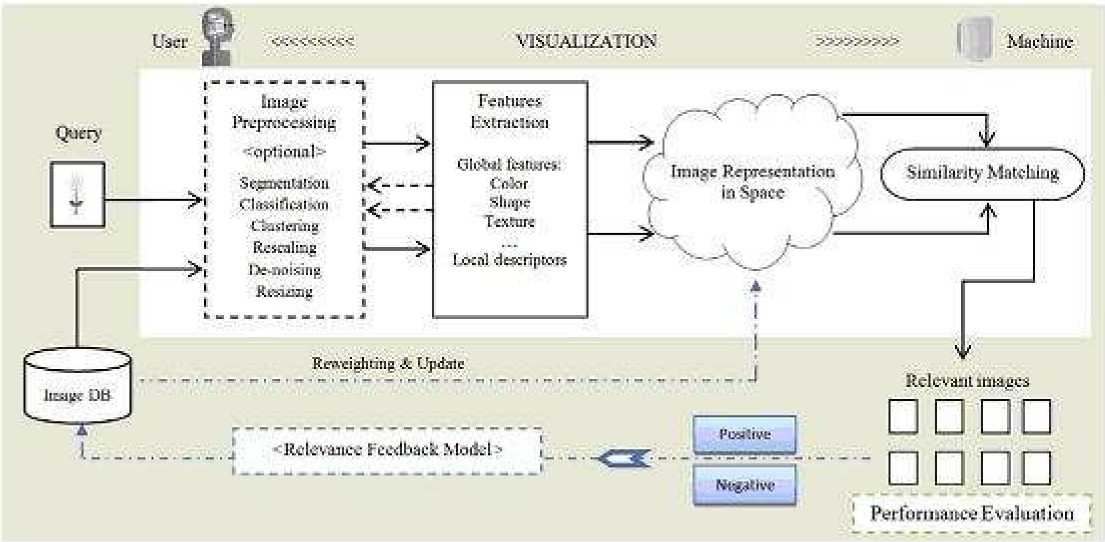

Figure 2. The general framework of CBIR system

3.1. **SVM ensemble for image retrieval.** As a core machine learning technology, SVM has not only strong theoretical foundations but also excellent empirical successes. However, it also has a lot of limitations. First, the regular SVM is originally for binary classification problem, it may not achieve the best performance when applied in multi-class tasks. Moreover, the regular SVM treats fairly with the positive and negative instances. If instances in one of the two-class overnumber another ones, the performance of the regular SVM may suffer dramatically. To overcome these drawbacks, the SVM ensemble technique has been proposed and shown promising improvement over the regular SVM [6]. As is known, ensemble learning aims to mine the complementary information of multiple classifiers to achieve strong generalization ability. In other words, an ensemble of classifiers is a set of classifiers whose individual decisions are combined in some way to classify new examples. Its main advantage lies in that ensembles are often much more accurate than the individual classifiers that make them up [7-9]. In general, an SVM ensemble is a collection of several SVM classifiers in which the decision to classify the test data is made by combining the decision functions of all individual classifiers. Fig. 3 illustrates a general architecture of the SVM ensemble. Note that during the training phase, each individual SVM is trained independently by its own replicated training dataset. All constituent SVMs will be aggregated by various combination strategies. During the testing phase, a test example is applied to all SVMs simultaneously and a collective decision is obtained based on the aggregation strategy.

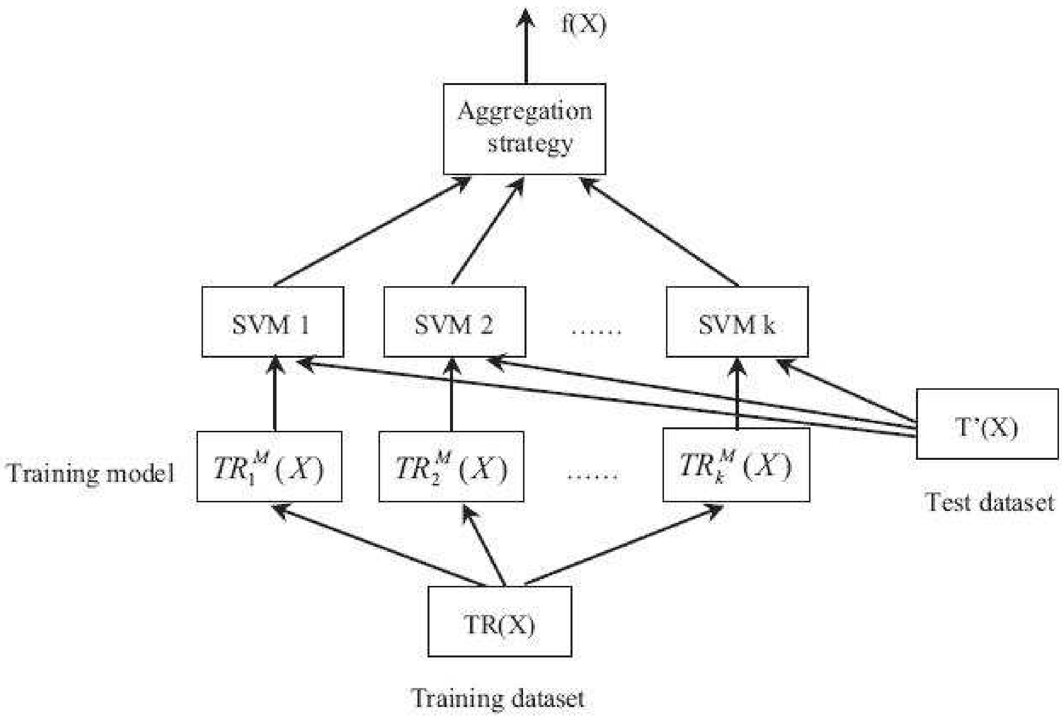

Figure 3. A general architecture of the SVM ensemble

In the early work [10], Zhang et al. propose a support vector machine based learning approach for image retrieval. Based on SVM, a classifier can be learned from training data of relevant images and irrelevant images marked by users, afterwards the model can be used to find more relevant images in the whole database. In [11], one-class SVM is constructed through fitting a tight hypersphere in the nonlinearly transformed feature space to include most of the target images based on the positive examples. It is worth noting that the use of kernel provides an elegant way to deal with nonlinearity in the distribution of the target images while the regularization term in SVM provides good generalization ability. Subsequently, an SVM for region-based image retrieval (RBIR) is presented with a new kernel [12], which is a generalization of Gaussian kernel with the Euclidean distance replaced by the earth mover's distance (EMD). Ko et al.[13] put forward a new method for extraction of salient regions using SVM and learning their importance scores when used for CBIR. Specifically, this method can automatically determine salient regions by SVM within the attention window, along with their different importance scores. In the meanwhile, Gondra et al.[14] propose to apply a generalized SVM (GSVM) based relevance feedback learning algorithm for region-based image retrieval. It should be noted that GSVM does not place restrictions on the kernel, thus any image similarity measure can be used. In particular, the proposed method can use an image similarity measure developed for region-based and variable length representations. Followed by Brown and Pham [15] construct a hierarchical SVM-based image classification system, including methods of training images and the usage of the technique in classification of image contents. In particular, the training phase has three main sub-processes to be performed, i.e., training an SVM to recognize a component and/or sets of components, placing constraints on the organization of the components and objects within an image as well as exportation of the classifier for an operator to employ. Recently, Hoi et al.[16] propose a semi-supervised SVM batch mode active learning for image retrieval, which exploits both semi-supervised kernel learning and batch mode active learning for relevance feedback (RF) in CBIR. Especially a kernel function is first learned from a mixture of labeled and unlabeled examples, subsequently it is used to effectively identify the informative and diverse examples for active learning via a min-max framework. However, one limitation of the current solution for this method is the quadratic programming algorithm, which may not be efficient for large-scale applications. Xu [17] applies support vector machine to image retrieval based on a scaling and rotation invariant encoding scheme for shapes. Wu et al.[18] propose an active learning scheme that deals with SVM ensemble under the semi-supervised setting to address the small sample problems along with a bias-ensemble mechanism is developed to guide the classification model by paying more attention on the positive examples than the negative ones. More recently, an effective CBIR system is developed based on multiple support vector machines ensemble [19]. Experimental results validate its efficiency, accuracy and scalability.

In general, SVM classifier behaves unstable for small-sized training set because its optimal hyperplane is too sensitive to the training examples. Meanwhile the kernel method is not always very effective because the number of feature dimensions in general is much higher than the size of the training set. Based on this recognition, Li et al.[20] present the multi-training SVM (MTSVM) that combines the merits of the co-training technique and a random sampling method in the feature space to alleviate these problems. Besides, by integrating asymmetric bagging-based SVM and random subspace SVM, Tao et al. construct an asymmetric bagging and random subspace SVM (ABRS-SVM) to solve the aforesaid issues and further improve the relevance feedback performance for image retrieval, especially for the hierarchical ABRS-SVM (HABRS-SVM) structure and the parallel ABRS-SVM (PABRS-SVM) structure as shown in Fig. 4. More details on them can be found in reference [21].

3.2. **Hybrid SVM for image retrieval.** Early works on the hybrid SVM with other approaches have also focused on content-based image retrieval, which is one of the basic methods within the machine learning community. As the representative work, Hong et al.[22] incorporate SVM into CBIR with relevance feedback, in which SVM is employed to classify the positive and negative images, and its learning results are used to update the preference weights for the relevant images. In addition, an early approach to image retrieval is the innovative work of Tong and Chang [23] who integrate SVM with active learning to search the data points that can maximally reduce the size of the version space. This method selects the most informative images to query a user and quickly learns a boundary that separates the images that satisfy the user's query concept from the rest of the dataset. However, this method at least has two main drawbacks. First, SVM may fail to learn an accurate classification model from a small number of labeled examples. Given a limited number of training data labeled by the user, directly applying SVM model may not improve the retrieval accuracy significantly although it enjoys excellent generalization performance. Second, unlike traditional pattern classification problem, the labeled positive examples are much less than the labeled negative ones and thus the learned SVM boundary may be biased toward the negative side. As one of the well-known semi-supervised learning techniques, transductive support vector machine (TSVM) has been proposed to incorporate the unlabeled data into image processing. The method of bootstrapping SVM active learning is proposed for CBIR [24], in which the initial SVM classifier is improved by incorporating unlabelled images instead of asking the user to label more randomly selected images. It is, however, observed that the performance of this method may be unstable in some cases. On the other side, it is widely recognized that incorporating prior knowledge in the field of machine learning can help it achieve more accurate generalization, especially when the training sample is scarce. As a practical application, image retrieval has its own characteristics and prior knowledge. In literature [25], to begin with, some prior knowledge of image retrieval is discussed and constructed. Subsequently the knowledge is incorporated into SVM as a constraint and a new target function is formulated. Based on this, a framework of image retrieval with knowledge based SVM is proposed for CBIR. Besides, a hybrid SVM with active learning scheme is also put forward for content-based image retrieval in [26].

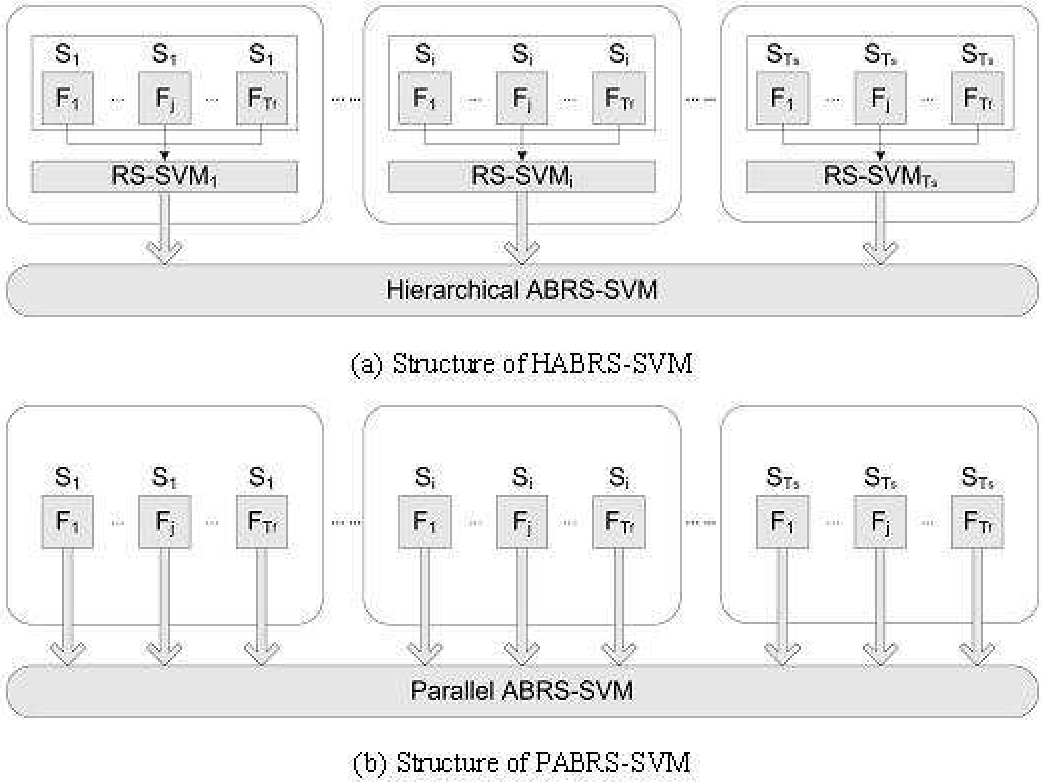

FIGURE 4. Structures of hierarchical ABRS-SVM (a) and parallel ABRS-SVM (b)

Later on, Wang et al.[27] bring forward a new interactive genetic algorithm (IGA) framework incorporating relevant feedback, in which human evaluation is regarded as not only the fitness function of GA, but also the relevant score to instruct interactive machine learning. Different from the approach shown in [11], Zhang et al.[28] leverage one-class support vector machine to solve the multiple instance learning (MIL) problem in region-based image retrieval (RBIR) based on semantic regions instead of the whole image. In particular, one-class SVM is exploited to model the non-linear distribution of image regions and separate positive regions from negative ones. Each semantic region of the test images is given a score by the evaluation function built from the model, and the images with the highest scores are returned to the user as the query results. In the scheme [29], Bhattacharya et al. present a novel approach for CBIR, in which both the probabilistic multi-class SVM and fuzzy c-means (FCM) clustering techniques are utilized for representing images in a semantic space based on category or cluster membership values for each image in the database. It is noted that the small sample problem is a challenging issue in computer vision that will inevitably undermine the retrieval performance. For this purpose, a unified framework that incorporates pseudo-labeling into fuzzy support vector machine is developed in the context of CBIR [30], which takes fully into account of the advantages of pseudo-labeling, active learning and the structure of fuzzy SVM. Liu et al.[31] put forward an SVM-based active feedback in image retrieval using clustering and unlabeled data, in which a new active selection criterion to select images for user's feedback is designed, and unlabeled images are incorporated within the co-training framework. Gao et al.[32] come up with clustering guided SVM (CGSVM) for semantic image retrieval by leveraging the clustering result to select the most informative image samples to be labeled and optimizing the penalty coefficient. In [33], Kim et al. integrate the multi-class SVM learning methods into relevance feedback with adaptive clustering for RBIR. In particular, SVM modeling phase consists of the region-based classification process and the region-merging process respectively. Fig. 5 displays the framework of the RBIR with relevance feedback using multi-class SVM learning.

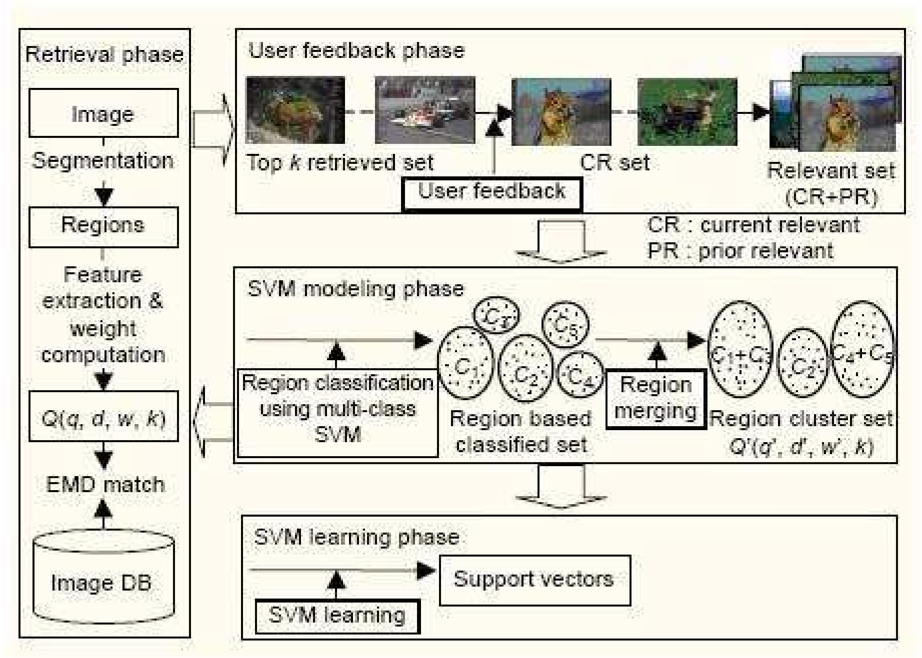

FIGURE 5. The framework of RBIR with RF using multi-class SVM learning

In comparison with CBIR, it should be noted that the region-based image retrieval has been widely investigated over the past several years and much endeavor has been devoted to enhancing its performance [12,34,35]. As the representative work, Jing et al.[12,34] put forward an image retrieval framework by integrating efficient region-based representation and effective online learning capability based on SVM. In addition, a novel generalized SVM algorithm [35] is proposed by taking into account both low-level features and structural information of images to solve the problem of region-based image retrieval via SVM framework. To be specific, for a given image, salient regions are extracted and the concept of salient region adjacency graph (SRAG) is employed to represent the image semantics. Based on SRAG, a novel generalized structure kernel based SVM is constructed for RBIR. Followed by Zeng et al.[36] present an SVM-based relevance feedback technique for region-based image retrieval, which devises a compact representation for image regions and exploits earth mover's distance (EMD) as well as hybrid features to match images. In recent work [37], an integrated approach is proposed to RBIR by using firefly algorithm and support vector machine, in which SVM is exploited to update the weights of preferences for relevant images based on the both relevant and irrelevant feedback images whereas the firefly optimizer is utilized to guide the swarm agents to move towards the cluster of relevant images in the exploration of the search space based on user's feedback. To tackle the overfitting phenomenon and neglect of local correlation, Qi et al.[38] construct a locality-sensitive support vector machine (LSSVM) for image retrieval, which applies locality-sensitive hashing to divide the whole feature space into a number of local regions, on each of them a local model can be better constructed due to the smaller within-class variation on it. Meanwhile, it imposes a global regularizer across local regions so that local classifiers can be smoothly glued together to form a regularized overall classifier. In literature [39], a new content-based image retrieval system is developed based on SVM by using Gaussian mixture models (GMM) as image representations. Recently, Pighetti et al.[40] combine a multi-objective interactive genetic algorithm with a SVM to more precisely retrieve the images of interest with less annotations by allowing a trade-off between image features and user evaluations. In the approach [41], a hybrid particle swarm optimization (PSO) and active learning SVM model is proposed for CBIR, in which the PSO with/without feature selection can optimize the parameters and sub-features in the SVM classifier. On the other side, the active SVM is applied on actively selecting most information images that minimizes the redundancy between candidate images shown to the user. Xing et al.[42] propose a modified AdaBoost-based one-class support vector machine (OCSVM) ensemble for image retrieval. In this method, the weight update formula of training data for AdaBoost is modified to make it fit for combining the results of OCSVM even though OCSVM is regarded as a strong classifier. In the scheme [43], a new interactive CBIR system is developed based on locality-sensitive hashing (LSH) and SVM, in which LSH is adopted to overcome the curse of dimensionality and a SVM-based relevance feedback is introduced to shorten the semantic gap so as to improve the image retrieval performance. Besides, a new technique that learns a boundary to separate the positive and negative examples based on the constrained similarity measure (CSM) is proposed in [44], in which SVM and AdaBoost are used to learn the boundary that is utilized to filter the images in the database for Euclidean similarity measure. Experiments demonstrate the usefulness and effectiveness of the similarity measure for image retrieval.

To further accelerate the convergence of the SVM classifier, Qi and Zhang [45] present a new method for CBIR based on an active learning support vector machine that combines the model selection with the active learning. In more recent work [46], Wang et al. propose a new SVM-based active feedback scheme for CBIR, in which a set of moderate one-class SVM classifiers are first trained separately by using different sub-features vectors, then the weight vector of component SVM classifiers are calculated dynamically by using the parameters for positive and negative samples, finally the results of the component classifiers are combined to form an output code as a hypothesized solution to the overall image retrieval problem. Subsequently, they put forward another SVM-based relevance feedback image retrieval method [47] by using probabilistic feature and weighted kernel function. Besides, an effective image retrieval system is developed by using color, texture and shape information of an image along with the training-based non-linear SVM classifier [48]. In addition, it is worth noting that deep learning, as one of the most currently remarkable machine learning techniques, has been widely used to reduce the semantic gap and solve many real-world problems in the field of computer vision over the last few years, which models the high-level semantics in data by utilizing deep structures that include multiple feature representations and non-linear transformations. In recent years, deep learning has been successfully applied in several CBIR tasks such as image classification [49,50], object recognition [51] and retrieval [52]. Especially in [50], a new deep support vector machine (called DeepSVM) is proposed for image classification (shown in Fig. 6). In each layer of the DeepSVM, Ex-AdaBoost is applied to not only select SVM with the minimal error rate and the highest diversity but also to produce the weight for each feature. For more details of the CBIR systems associated with SVM and deep learning, please refer to the corresponding literature.

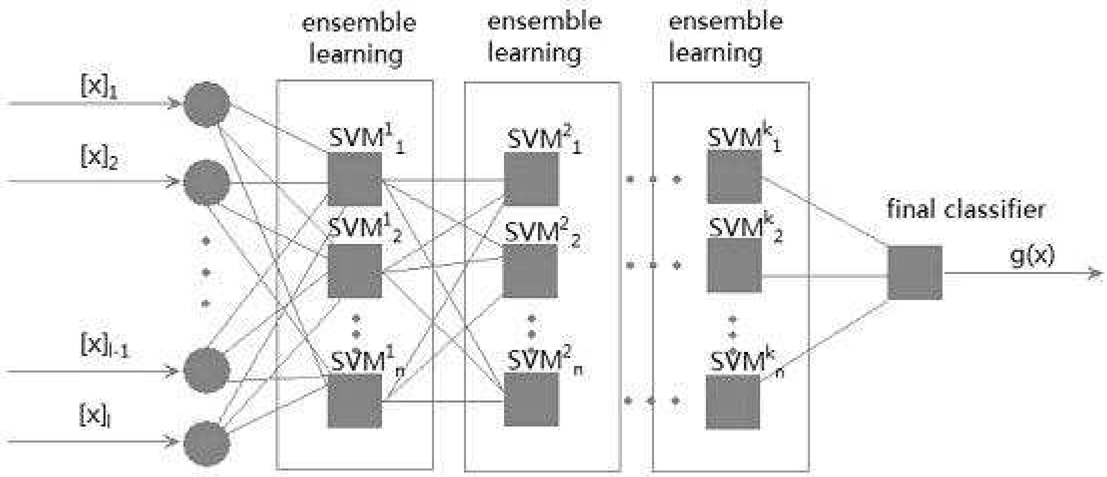

Figure 6. The framework of DeepSVM

3.3. **SVM for other applications.** Aside from the content aforementioned, there are many more other applications of SVM for pattern recognition problems, such as medical data processing [53,54], face recognition [55,56], image classification [57,58,59], audio/video detection [60-62], patent classification [63,64], adult image recognition [65,66] and so on. To summarize, all the SVM related methods discussed in this paper show some promising outcomes. However, they are still in infancy and it is important to avoid a future where the image retrieval community is isolated from real-world interests. We believe that it will have a golden opportunity in the growth of the multimedia search field that is commonly considered the next major frontier of search. In the following, several support vector machine related CBIR methods involved in this paper are concisely summarized in Table 2, mainly involving classifiers and image datasets utilized in the corresponding literature.

4. **Conclusions and Future Work.** Support vector machine is a supervised classifier that 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 achieves optimal class boundaries by finding the maximum distance between classes. In the last two decades, a lot of research has been done on SVM. However, 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 content-based image retrieval have been presented so as to complement the existing surveys in literature. In particular, we put emphasis on SVM for CBIR from two aspects, including SVM ensemble for image retrieval and hybrid SVM for image retrieval, respectively. The primary

Table 2. Summary of SVM related CBIR methods

Sources	Classifiers	Image datasets applied
Zhang et al.[10]	SVM	COREL Dataset
Jing et al.[12]	SVM	COREL Dataset
Ko et al.[13]	SVM	COREL Dataset
Gondra et al.[14]	GSVM	COREL Dataset
Hoi et al.[16]	SVM, Active learning	COREL Dataset
Xu [17]	SVM	Other Dataset
Yildizer et al.[19]	SVM ensemble	COREL Dataset
Li et al.[20]	MTSVM	COREL Dataset
Tao et al.[21]	ABRS-SVM	COREL Dataset
Hong et al.[22]	SVM	COREL Dataset
Tong et al.[23]	SVM, Active learning	COREL Dataset
Wang et al.[24]	TSVM-SAL, Active learning	COREL/VisTex Datasets
Wang et al.[25]	SVM	COREL/Other Datasets
Jing et al.[26]	SVM, Active learning	COREL Dataset
Zhang et al.[28]	SVM, MIL	COREL Dataset
Bhattacharya [29]	SVM, FCM	COREL/Other Datasets
Wu et al.[30]	FSVM, Active learning	COREL Dataset
Gao et al.[32]	CGSVM, GMM	COREL/WWW Datasets
Kim et al.[33]	SVM	COREL Dataset
Jing et al.[34]	SVM	COREL Dataset
Zhang et al.[35]	SVM	COREL Dataset
Kanimozhi et al.[37]	SVM, Firefly algorithm	COREL/Caltech Datasets
Qi et al.[38]	LSSVM	NUS-WIDE Dataset
Marakakis et al.[39]	SVM, GMM, RF	COREL Dataset
Pighetti et al.[40]	SVM, GA	SIMPLIcity Dataset
Ma et al.[41]	SVM, PSO, Active learning	COREL Dataset
Xing et al.[42]	OCSVM, AdaBoost	COREL/Caltech Datasets
Dong et al.[43]	SVM, LSH	COREL Dataset
Guo et al.[44]	SVM, AdaBoost	COREL Dataset
Qi et al.[45]	SVM, Active learning	COREL Dataset
Wang et al.[46]	SVM, Active learning	COREL Dataset
Wang et al.[47]	SVM, RF	COREL/Caltech Datasets
Singh et al.[48]	SVM	COREL/UKbench Datasets

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 content-based image retrieval in the future.

In spite of SVM can obtain a relatively satisfying retrieval performance and seem to be easily for implementation, it still suffers from several issues remain to be solved. Following are some potential directions for SVM research challenges. First, SVM has class-imbalance problem, which means that it has poor performance on imbalanced data. Unfortunately, class-imbalance is a common phenomenon existing in image data, and which will inevitably degrade the performance of the SVM classifier. Second, it is known that kernel function plays an important role in the implementation of SVM. However, in the majority 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 exists 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. Thus 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 image retrieval. Fifth, accurate image segmentation is still beyond current computer vision technique, especially for those approaches heavily depend on local features, which in turn rely on high segmentation accuracy. So to explore more efficient segmentation methods is of great help to boost the performance of CBIR. In the meanwhile, to integrate more sophisticated features including both visual and textual ones into SVM is also a desirable research direction. Sixth, deep learning is a promising approach with the purpose to reduce the semantic gap in CBIR. Selecting appropriate network architectures for deep learning depends on how these architectures will be used in the problem domain. Thus how to employ the nature of SVM to further improve the performance of deep learning is also a very valuable research direction. Last but not the least, to evaluate the image retrieval has been an ongoing challenging problem. Furthermore, in the case of image retrieval, it has been shown that commonly used test datasets such as Corel, Nuswide, Caltech and Mirflickr image collections are not necessarily effective performance indicators for real-world problems.

Acknowledgment. The author would like to sincerely thank the anonymous reviewers for their valuable comments and insightful suggestions that have helped to improve the paper. In addition, the author thanks Prof. Zhongzhi Shi for stimulating discussions and helpful hints. This work is partially supported by the National Program on Key Basic Research Project (No.2013CB329502), the National Natural Science Foundation of China (No.61202212), the Key R&D Program of the Shaanxi Province of China (No.2018GY-037) and the Special Research Project of the Educational Department of Shaanxi Province of China (No.18JK0051).

REFERENCES

[1] Y. Liu, D. Zhang, G. Lu, et al., A survey of content-based image retrieval with high-level semantics, *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, 2007.

[2] R. Datta, D. Joshi, J. Li, et al., Image retrieval: ideas, influences, and trends of the new age, *ACM Computer Survey*, vol. 40, no. 2, pp. 1–60, 2008.

[3] S. Bernhard and J. Alexander, Learning with kernels-support vector machines, regularization, optimization and beyond, *The MIT Press: Cambridge, Massachusetts*, 2003.

[4] V. Vapnik, Statistical learning theory, *New York: Wiley*, 1998.

[5] C. Cortes and V. Vapnik, Support vector networks, *Machine Learning*, vol. 20, no. 3, pp. 1–25, 1995.

[6] F. Yang, G. Xia, G. Liu, et al., Dynamic texture recognition by aggregating spatial and temporal features via ensemble SVMs, *Neurocomputing*, vol. 173, pp. 1310–1321, 2016.

[7] H. Kim, S. Pang, H. Je, et al., Constructing support vector machine ensemble, *Pattern Recognition*, vol. 36, no. 12, pp. 2757–2767, 2003.

[8] R. Wu and W. Chung, Ensemble one-class support vector machines for content-based image retrieval, *Expert Systems with Applications*, vol. 36, no. 3, pp. 4451–4459, 2009.

[9] J. Wu and M. Lu, Asymmetric semi-supervised boosting scheme for interactive image retrieval, *ETRI Journal*, vol. 32, no. 5, pp. 766–773, 2010.

[10] L. Zhang, F. Lin and B. Zhang, Support vector machine learning for image retrieval, *Proc. of the 8th Int'l Conf. on Image Processing (ICIP'01)*, pp. 721–724, 2001.

[11] Y. Chen, X. Zhou and T. Huang, One-class SVM for learning in image retrieval, *Proc. of the 8th Int'l Conf. on Image Processing (ICIP'01)*, pp. 34–37, 2001.

[12] F. Jing, M. Li, H. Zhang, et al., Support vector machines for region-based image retrieval, *Proc. of the Int'l Conf. on Multimedia and Expo (ICME'03)*, pp. 21–24, 2003.

[13] B. Ko, S. Kwak and H. Byun, SVM-based salient region(s) extraction method for image retrieval, *Proc. of the 17th Int'l Conf. on Pattern Recognition (ICPR'04)*, pp. 977–980, 2004.

[14] I. Gondra and D. Heisterkamp, Learning in region-based image retrieval with generalized support vector machines, *Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition Workshop (CVPRW'04)*, pp. 149–156, 2004.

[15] R. Brown and B. Pham, Image mining and retrieval using hierarchical support vector machines, *Proc. of the 11th Int'l Conf. on Multimedia Modeling (MMM'05)*, pp. 446–451, 2005.

[16] S. Hoi, R. Jin, J. Zhu, et al., Semi-supervised SVM batch mode active learning for image retrieval, *Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition (CVPR'08)*, pp. 1–7, 2008.

[17] K. Xu, An application of support vector machines for image retrieval, *Proc. of the 6th Int'l Conf. on Natural Computation (ICNC'10)*, pp. 1528–1532, 2010.

[18] J. Wu, M. Lu and C. Wang, Enhancing SVM active learning for image retrieval using semi-supervised bias-ensemble, *Proc. of the 20th Int'l Conf. on Pattern Recognition (ICPR'10)*, pp. 3175–3178, 2010.

[19] E. Yildizer, A. Balci, M. Hassan, et al., Efficient content-based image retrieval using multiple support vector machines ensemble, *Expert Systems with Applications*, vol. 39, no. 3, pp. 2385–2396, 2012.

[20] J. Li, N. Allinson, D. Tao, et al., Multitraining support vector machine for image retrieval, *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3597–3601, 2006.

[21] D. Tao, X. Tang, X. Li, et al., Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1088–1099, 2006.

[22] P. Hong, Q. Tian and T. Huang, Incorporate support vector machines to content-based image retrieval with relevance feedback, *Proc. of the 7th Int'l Conf. on Image Processing (ICIP'00)*, pp. 750–753, 2000.

[23] S. Tong and E. Chang, Support vector machine active learning for image retrieval, *Proc. of the 9th Int'l Conf. on Multimedia (MM'01)*, pp. 107–118, 2001.

[24] L. Wang, K. Chan and Z. Zhang, Bootstrapping SVM active learning by incorporating unlabelled images for image retrieval, *Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition (CVPR'03)*, pp. 629–634, 2003.

[25] L. Wang, P. Xue and K. Chan, Incorporating prior knowledge into SVM for image retrieval, *Proc. of the 17th Int'l Conf. on Pattern Recognition (ICPR'04)*, pp. 981–984, 2004.

[26] J. Feng, M. Li, H. Zhang, et al., Entropy-based active learning with support vector machines for content-based image retrieval, *Proc. of the Int'l Conf. on Multimedia and Expo (ICME'04)*, pp. 85–88, 2004.

[27] S. Wang, X. Wang and J. Xue, An improved interactive genetic algorithm incorporating relevant feedback, *Proc. of the 4th Int'l Conf. on Machine Learning and Cybernetics (ICMLC'05)*, pp. 2996–3001, 2005.

[28] C. Zhang, X. Chen, M. Chen, et al., A multiple instance learning approach for content-based image retrieval using one-class support vector machine, *Proc. of the Int'l Conf. on Multimedia and Expo (ICME'05)*, pp. 1142–1145, 2005.

[29] P. Bhattacharya, M. Rahman and B. Desai, Image representation and retrieval using support vector machine and fuzzy c-means clustering based semantical spaces, *Proc. of the 18th Int'l Conf. on Pattern Recognition (ICPR'06)*, pp. 929–935, 2006.

[30] K. Wu and K. Yap, Fuzzy SVM for content-based image retrieval: a pseudo-label support vector machine framework, *IEEE Computational Intelligence Magazine*, vol. 1, no. 2, pp. 10–16, 2006.

[31] R. Liu and Y. Wang, SVM-based active feedback in image retrieval using clustering and unlabeled data, *Pattern Recognition*, vol. 41, no. 8, pp. 2645–2655, 2008.

[32] K. Gao, S. Lin, Y. Zhang, et al., Clustering guided SVM for semantic image retrieval, *Proc. of the 2nd Int'l Conf. on Pervasive Computing and Applications (ICPCA'07)*, pp. 199–203, 2007.

[33] D. Kim, J. Song, J. Lee, et al., Support vector machine learning for region-based image retrieval with relevance feedback, *ETRI Journal*, vol. 29, no. 5, pp. 700–702, 2007.

[34] F. Jing, M. Li, H. Zhang, et al., An efficient and effective region-based image retrieval framework, *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 699–709, 2004.

[35] R. Zhang, J. Yuan, J. Huang, et al., A novel generalized SVM algorithm with application to region-based image retrieval, *Proc. of the Int'l Forum on Information Technology and Applications (IFITA'09)*, pp. 280–283, 2009.

[36] Z. Zeng, S. Cai and S. Liu, A novel image representation and learning method using SVM for region-based image retrieval, *Proc. of the 5th Int'l Conf. on Industrial Electronics and Applications (ICIEA'10)*, pp. 1622–1626, 2010.

[37] T. Kanimozhi and K. Latha, An integrated approach to region-based image retrieval using firefly algorithm and support vector machine, *Neurocomputing*, vol. 151, pp. 1099–1111, 2015.

[38] G. Qi, Q. Tian and T. Huang, Locality-sensitive support vector machine by exploring local correlation and global regularization, *Proc. of the Int'l Conf. on Computer Vision and Pattern Recognition (CVPR'11)*, pp. 841–848, 2011.

[39] A. Marakakis, G. Siolas, N. Galatsanos, et al., Relevance feedback approach for image retrieval combining support vector machines and adapted Gaussian mixture models, *IET Image Processing*, vol. 5, no. 6, pp. 531–540, 2011.

[40] R. Pighetti, D. Pallez and F. Precioso, Hybrid content-based image retrieval combining multi-objective interactive genetic algorithm and SVM, *Proc. of the 21st Int'l Conf. on Pattern Recognition (ICPR'12)*, pp. 2849–2852, 2012.

[41] C. Ma, Q. Dai and S. Liu, A hybrid PSO and active learning SVM model for relevance feedback in the content-based images retrieval, *Proc. of the Int'l Conf. on Computer Science & Service System (CSSS'12)*, pp. 130–133, 2012.

[42] H. Xing, J. Wu and X. Chen, Modified AdaBoost based OCSVM ensemble for image retrieval, *Proc. of the 11th Int'l Conf. on Machine Learning and Cybernetics (ICMLC'12)*, pp. 1048–1053, 2012.

[43] T. Dong and Z. Zhao, A novel interactive image retrieval method based on LSH and SVM, *Proc. of the 3rd Int'l Conf. on Network Infrastructure and Digital Content (IC-NIDC'12)*, pp. 482–486, 2012.

[44] G. Guo, A. Jain, W. Ma, et al., Learning similarity measure for natural image retrieval with relevance feedback, *IEEE Transactions on Neural Networks*, vol. 13, no. 4, pp. 811–820, 2002.

[45] Y. Qi and G. Zhang, Strategy of active learning support vector machine for image retrieval, *IET Computer Vision*, vol. 10, no. 1, pp. 87–94, 2016.

[46] X. Wang, H. Yang, Y. Li, et al., A new SVM-based active feedback scheme for image retrieval, *Engineering Applications of Artificial Intelligence*, vol. 37, pp. 43–53, 2015.

[47] X. Wang, L. Liang, W. Li, et al., A new SVM-based relevance feedback image retrieval using probabilistic feature and weighted kernel function, *Journal of Visual Communication and Image Representation*, vol. 38, pp. 256–275, 2016.

[48] C. Singh, E. Walia and K. Kaur, Enhancing color image retrieval performance with feature fusion and non-linear support vector machine classifier, *Optik*, vol. 158, pp. 127–141, 2018.

[49] Y. Li and T. Zhang, Deep neural mapping support vector machines, *Neural Networks*, vol. 93, pp. 185–194, 2017.

[50] Z. Qi, B. Wang, Y. Tian, et al., When ensemble learning meets deep learning: a new deep support vector machine for classification, *Knowledge-Based Systems*, vol. 107, pp. 54–60, 2016.

[51] S. Erfani, S. Rajasegarar, S. Karunasekera, et al., High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning, *Pattern Recognition*, vol. 58, pp. 121–134, 2016.

[52] B. Jiang, J. Yang, Z. Lv, et al., Internet cross-media retrieval based on deep learning, *Journal of Visual Communication and Image Representation*, vol. 48, pp. 356–366, 2017.

[53] Q. Liu, Q. Gu and Z. Wu, Feature selection method based on support vector machine and shape analysis for high-throughput medical data, *Computers in Biology and Medicine*, vol. 91, pp. 103–111, 2017.

[54] S. Souli and Z. Lachiri, Audio sounds classification using scattering features and support vectors machines for medical surveillance, *Applied Acoustics*, vol. 130, pp. 270–282, 2018.

[55] W. Jin, J. Zhang and X. Zhang, Face recognition method based on support vector machine and particle swarm optimization, *Expert Systems with Applications*, vol. 38, no. 4, pp. 4390–4393, 2011.

[56] S. Meshgini, A. Aghagolzadeh and H. Seyedarabi, Face recognition using Gabor-based direct linear discriminant analysis and support vector machine, *Computers & Electrical Engineering*, vol. 39, no. 3, pp. 727–745, 2013.

[57] D. Tian, X. Zhao and Z. Shi, Support vector machine with mixture of kernels for image classification, *Proc. of the 7th Int'l Conf. on Intelligent Information Processing (ICIIP'12)*, pp. 67–75, 2012.

[58] Z. Qi, Y. Tian and Y. Shi, Robust twin support vector machine for pattern classification, *Pattern Recognition*, vol. 46, no. 1, pp. 305–316, 2013.

[59] H. Yin, X. Jiao, Y. Chai, et al., Scene classification based on single-layer SAE and SVM, *Expert Systems with Applications*, vol. 42, no. 7, pp. 3368–3380, 2015.

[60] Y. Zhu, Z. Ming and J. Zhang, Video scene classification and segmentation based on support vector machine, *Proc. of the Int'l Joint Conf. on Neural Networks (IJCNN'08)*, pp. 3571–3576, 2008.

[61] J. Sangeetha and S. Jothilakshmi, A novel spoken keyword spotting system using support vector machine, *Engineering Applications of Artificial Intelligence*, vol. 36, pp. 287–293, 2014.

[62] S. Souli and Z. Lachiri, Audio sounds classification using scattering features and support vectors machines for medical surveillance, *Applied Acoustics*, vol. 130, pp. 270–282, 2018.

[63] C. Wu, Y. Ken and T. Huang, Patent classification system using a new hybrid genetic algorithm support vector machine, *Applied Soft Computing*, vol. 10, no. 4, pp. 1164–1177, 2010.

[64] J. Wu, P. Chang and C. Tsao, A patent quality analysis and classification system using self-organizing maps with support vector machine, *Applied Soft Computing*, vol. 41, pp. 305–316, 2016.

[65] B. Choi, B. Chung and J. Ryou, Adult image detection using Bayesian decision rule weighted by SVM probability, *Proc. of the 4th Int'l Conf. on Computer Sciences and Convergence Information Technology (ICCIT'09)*, pp. 659–662, 2009.

[66] Z. Zhao and A. Cai, Combining multiple SVM classifiers for adult image recognition, *Proc. of the 2nd Int'l Conf. on Network Infrastructure and Digital Content (ICNIDC'10)*, pp. 149–153, 2010.