An Effective Image Representation Method using Kernel Classification

Haoxiang Wang
School of Electronic Engineering
Xidian University
Xi' an 710126, China
Email: wanghaoxiang1102@hotmail.com

Jingbin Wang
National Time Service Center
Chinese Academy of Sciences
Xi' an 710600 , China

Abstract—The learning of image representation is always the most important problem in computer vision community. In this paper, we propose a novel image representation method by learning and using kernel classifiers. We firstly train classifiers using the one-against-all rule, then use them classify the candidate images, and finally using the classification responses as the new representations. The Euclidean distance between the classification response vectors are used as the new similarity measure. The experimental results from a large scale image database show that the proposed algorithm can outperform the original feature on image retrieval problem.

Keywords-Image Retrieval; Image Representation; Kernel Classification

I. INTRODUCTION

In modern times, countless digital photos are being taken everyday, and the problem of how to store, represent and retrieval them raises [1]–[5]. In recent years, the digital management system of images has been proposed to effectively handle this problem [6], [7]. Many online systems are developed to store, manage and share the images. For example the google image engine [8], [9] can that allows users to search the Web for image content. Using such services, we can classify the images correctly according to the concept of them. In figure 1, we shown some examples of classified image using google image engine. These classified images can provide effective information of image concept, by containing more themes information. Furthermore, Flickr groups can represent richer themes [10], [11].

In this paper, an effective representation is learned for an image by using these Google image class memberships. If two images share a set of similar class memberships, we regard them as a pair of similar images. Thus we use the class memberships to construct the new representation of the image. Based on this new presentation, we could gain an accurate retrieval of images from a image database [12]–[14]. A kernel classifier is used to measure the membership of an image to a special image class. The number of images in each class is always tremendous, thus the training of kernel classifier may takes a long time [15]–[18]. To solve this problem, an efficient algorithm is needed. In this paper, we develop a new iterative algorithm to learn the

kernel classifier, which is an efficient fast implementation of support vector machine (SVM) [19]–[21]. We try to develop the fast algorithm by using both the stochastic learning strategy [22]–[24] and the support vector approximation technology [25], [26]. The developed classifier is fast and robust, which makes the learning from a large number of images possible. Because the training images are input into the algorithm one by one, thus it does not takes much memory, and it is also a online algorithm. Moreover, we use the Histogram Intersection Kernel (HIK) [27]–[29], which is also easy and fast to calculate for histogram features vectors of multi-instance based image representation [30]–[32].

Using the proposed kernel classifiers, and with a candidate image, we firstly input it into the classifiers of different classes and obtain the classification responses. Then the classification responses are connected to form a classification response vector. Finally, a pair of images are compared by consulating the Euclidean distance [33]–[36] between
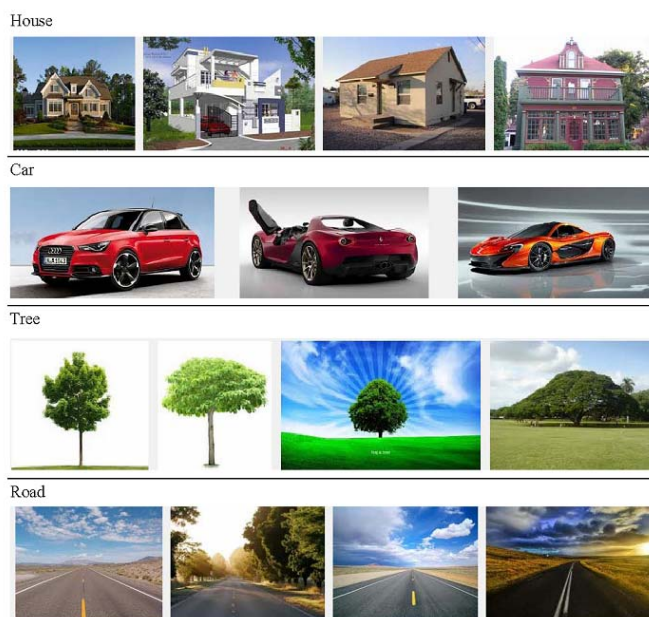

Figure 1.   Classified Images of Google Image Engine

their classification response vectors. The experiment results show that this method is both efficient and effective for the kernel classifier learning. However, the goal of kernel classification is not to learn good classifiers themselves, but to use the classifier for the representation of images. We also show that the proposed kernel classification response vector performs better than the original histogram features. We use the proposed method to retrieve images from large scale images. It is shown that it outperforms other methods. The contribution of this paper is summarized as follows:

- In this paper, we demonstrate a new approach to use the image of different classes to represent a new image.
- Moreover, we develop a novel iterative algorithm to learn the kernel classifier by using the training samples one by one. It could also be used to learn other types of classifiers.

II. REPRESENTATION BY KERNEL CLASSIFICATION

A. Overview

The overview and framework of the proposed method is shown in figure 2. We first download many images of different classes from internet and then train SVM classifiers to separate the images of each class to other classes using the one-against-all rule [37]–[40]. Given a test image, we apply the SVM classifiers to compute its response, and then the response vector is used as the new representation. Finally, the image dissimilarity is measured using the Euclidean distance between the response vectors [41].

B. Kernel Classifier Learning

For each image class, we train a kernel classifier to distinguish the images of this class against the images of other classes. To this end, we first download a large number of images of this class, and also a great deal of images of different classes. Then, we extract the histogram feature vector from each image. The set of extracted image histogram vectors are presented as $\{\mathbf{x}_i\}_{i=1}^N$, where $\mathbf{x}_i = [x_{i1}, \cdots, x_{id}]^\top \in R^d$ is the d dimensional vector of the i-th image, and N is the number of images. A HIK kernel is used to match a pair of histograms as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \sum_k min(x_{ik}, x_{jk}) \tag{1}$$

Finally, based on this kernel, we design a classifier to predict if an image $\mathbf{x}$ belongs to this class or not.

$$
\begin{aligned}
f(\mathbf{x}) &= \sum_{i=1}^N \alpha_i \phi(\mathbf{x}_i)^\top \phi(\mathbf{x}) \\
&= \sum_{i=1}^N \alpha_i K(\mathbf{x}_i, \mathbf{x}) \\
&= \boldsymbol{\alpha}^\top K(\bullet, \mathbf{x})
\end{aligned}
\tag{2}
$$

where $\phi(\mathbf{x})$ is an implicit nonlinear mapping function so that $\phi(\mathbf{x}_i)^\top \phi(\mathbf{x}) = K(\mathbf{x}_i, \mathbf{x})$, α_i is the weight for i-th training image, $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_N]^\top$, and $K(\bullet, \mathbf{x}) = [K(\mathbf{x}_1, \mathbf{x}), \cdots, K(\mathbf{x}_N, \mathbf{x})]^\top$. The kernel classifier parameter $\boldsymbol{\alpha}$ is learned by minimizing the objective function as

$$min_{\boldsymbol{\alpha}} \frac{1}{2} \|\boldsymbol{\alpha}\|_2^2 + C\frac{1}{N} \sum_{i=1}^N \triangle(f(\mathbf{x}_i), y_i) \tag{3}$$

where $\triangle$ is a loss function for the classifier, $y_i \in \{+1, -1\}$ is the class label of i-th image. Here we use the hinge-loss

$$\triangle(f(\mathbf{x}_i), y_i) = max(0, 1 - y_i f(\mathbf{x}_i)) \tag{4}$$

$\|\boldsymbol{\alpha}\|_2^2$ is introduced to control the complexity of the classifier, so that the structural loss can also be minimized [42].

We write the objective function in (3) as

$$
\begin{aligned}
O(\boldsymbol{\alpha}) &= \frac{1}{2} \|\boldsymbol{\alpha}\|_2^2 + C\frac{1}{N} \sum_{i=1}^N \triangle(f(\mathbf{x}_i), y_i) \\
&= \frac{1}{2} \boldsymbol{\alpha}^\top \boldsymbol{\alpha} + C\frac{1}{N} \sum_{i=1}^N \beta_i \left(1 - y_i \boldsymbol{\alpha}^\top K(\bullet, \mathbf{x}_i)\right)
\end{aligned}
\tag{5}
$$

β_i is defined as

$$\beta_i = \begin{cases} 1, & if \ 1 - y_i \boldsymbol{\alpha}^\top K(\bullet, \mathbf{x}_i) \geq 0 \\ 0, & else \end{cases} \tag{6}$$

We try to use the gradient descent method [43]–[45] to update the classifier parameter $\boldsymbol{\alpha}$ in an iterative algorithm. However, for the hinge loss has no close form of derivative with regarding to $\boldsymbol{\alpha}$, thus we fix β_i as the previous one. In the t-th iteration, we first compute β_i^{t-1} as

$$\beta_i^{t-1} = \begin{cases} 1, & if \ 1 - y_i \boldsymbol{\alpha}^{t-1^\top} K(\bullet, \mathbf{x}_i) \geq 0 \\ 0, & else \end{cases} \tag{7}$$

then fix it and compute the derivative of $O(\boldsymbol{\alpha})$ with regarding to $\boldsymbol{\alpha}$, as

$$\frac{\partial O(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \boldsymbol{\alpha} - C\frac{1}{N} \sum_{i=1}^N \beta_i^{t-1} y_i K(\bullet, \mathbf{x}_i) \tag{8}$$

Then $\boldsymbol{\alpha}$ is updated as follows

$$
\begin{aligned}
\boldsymbol{\alpha}^t &= \boldsymbol{\alpha}^{t-1} - \tau^t \left. \frac{\partial O(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^{t-1}} \\
&= \boldsymbol{\alpha}^{t-1} - \tau^t \left(\boldsymbol{\alpha}^{t-1} - C\frac{1}{N} \sum_{i=1}^N \beta_i^{t-1} y_i K(\bullet, \mathbf{x}_i) \right) \\
&= \left(1 - \tau^t\right) \boldsymbol{\alpha}^{t-1} - \tau^t C\frac{1}{N} \sum_{i=1}^N \beta_i^{t-1} y_i K(\bullet, \mathbf{x}_i)
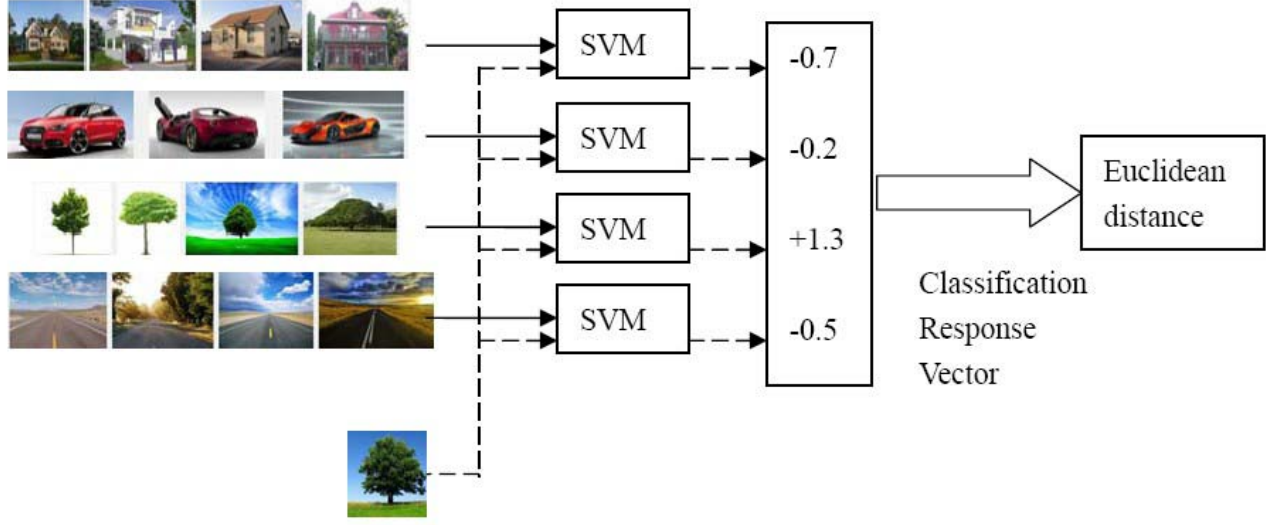\end{aligned}
\tag{9}
$$

854

Figure 2. Overview of the proposed method.

where τ^t is the descent step at t-th iteration. It can be concluded that in each iteration, all the training images are used to compute $\frac{\partial O(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^{t-1}}$. When the training image number N is large, the computational cost could be very expensive.

1) Fast updating method: To accelerate the updating procedure, we develop a novel and efficient algorithm for this problem. The training images are given to the algorithm one by one. In each iteration, only one image is used. We assume that in t-th iteration, $\mathbf{x}_{i^t}$ is used, and the objective function is constructed using only $\mathbf{x}_{i^t}$ as

$$
\begin{aligned}
O(\boldsymbol{\alpha}) &= \frac{1}{2}\left\|\boldsymbol{\alpha}\right\|_2^2 + C\triangle(f(\mathbf{x}_{i^t}), y_{i^t}) \\
&= \frac{1}{2}\boldsymbol{\alpha}^\top\boldsymbol{\alpha} + C\beta_{i^t}\left(1 - y_{i^t}\boldsymbol{\alpha}^\top K(\bullet, \mathbf{x}_{i^t})\right)
\end{aligned}
\tag{10}
$$

The derivative of $O(\boldsymbol{\alpha})$ with regarding to $\boldsymbol{\alpha}$ is as

$$
\frac{\partial O(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = \boldsymbol{\alpha} - C\beta_{i^t}y_{i^t}K(\bullet, \mathbf{x}_{i^t})
\tag{11}
$$

where β_{i^t} is also computed using the previous $\boldsymbol{\alpha}^{t-1}$ as

$$
\beta_{i^t} = \begin{cases} 1, & if\ 1 - y_{i^t}{\boldsymbol{\alpha}^{t-1}}^\top K(\bullet, \mathbf{x}_{i^t}) \geq 0 \\ 0, & else \end{cases}
\tag{12}
$$

In the t-th iteration, $\boldsymbol{\alpha}$ is updated as

$$
\begin{aligned}
\boldsymbol{\alpha}^t &= \boldsymbol{\alpha}^{t-1} - \tau^t\left.\frac{\partial O(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^{t-1}} \\
&= \boldsymbol{\alpha}^{t-1} - \tau^t\left(\boldsymbol{\alpha}^{t-1} - C\beta_{i^t}y_{i^t}K(\bullet, \mathbf{x}_{i^t})\right) \\
&= \left(1 - \tau^t\right)\boldsymbol{\alpha}^{t-1} - \tau^t C\beta_{i^t}y_{i^t}K(\bullet, \mathbf{x}_{i^t})
\end{aligned}
\tag{13}
$$

C. Classification Response Vector

Given M classes of images, we train M SVM classifiers denoted as $\{f^m(\mathbf{x})\}_{m=1}^M$ using the developed efficient algorithm. When a test image comes, its feature vector is extracted as $\mathbf{x}$, and its response to M classifiers are given as $f^m(\mathbf{x}), m = 1\cdots M$. These classification responses are further organized in a vector $\mathbf{f} = [f^1(\mathbf{x}), \cdots, f^M(\mathbf{x})]^\top$. When a pair of images $\mathbf{x}_i$ and $\mathbf{x}_j$ are available, they are firstly represented as $\mathbf{f}_i = [f^1(\mathbf{x}_i), \cdots, f^M(\mathbf{x}_i)]^\top$ and $\mathbf{f}_j = [f^1(\mathbf{x}_j), \cdots, f^M(\mathbf{x}_j)]^\top$, then we compute the Euclidean distance to compare the similarity between them

$$
Dist(\mathbf{f}_i, \mathbf{f}_j) = \sqrt{\sum_m (f^m(\mathbf{x}_i) - f^m(\mathbf{x}_j))^2}
\tag{14}
$$

III. EXPERIMENTS

Moreover, we collected a large scale database of 10,000 images from google image engine which are classified according to google keywords. Some example images of our database are given in figure 3.

We test our algorithm on the database retrieval task of the collected data set. We randomly selected 1,000 images as queries and leave the remaining ones as the database images. We define an image in the database as a relevant image to the query image if they belong to the same class. We compare our method to the method based on original histogram feature and HIK as similarity measure. The compared method is denoted as "Histogram+HIK". Our method is denoted as "Kernel Classification Representation".

855

Figure 3. Example images of collected image data set.

Figure 5. The nearest neighbors of an query image.

The average retrieval precision (ARP) [46]–[48] of different retrieved images is given in figure 4.

Figure 4. The retrieval ARP of the proposed method and compared method.

Figure 6. The training time of the updating algorithms.

In the experiment, a precision value is calculated for each query image, an then they are averaged to obtain the final ARP. The original histogram feature, the precision value at one retrieved image is 0.4620. But when the kernel classifier is used to represent the images, it is improved to 0.7164. Such a significant improvement is archived by including the discriminant information provided by the classifier. A much more accurate representation of image is provided by using the kernel classifier trained from the images of different classes.
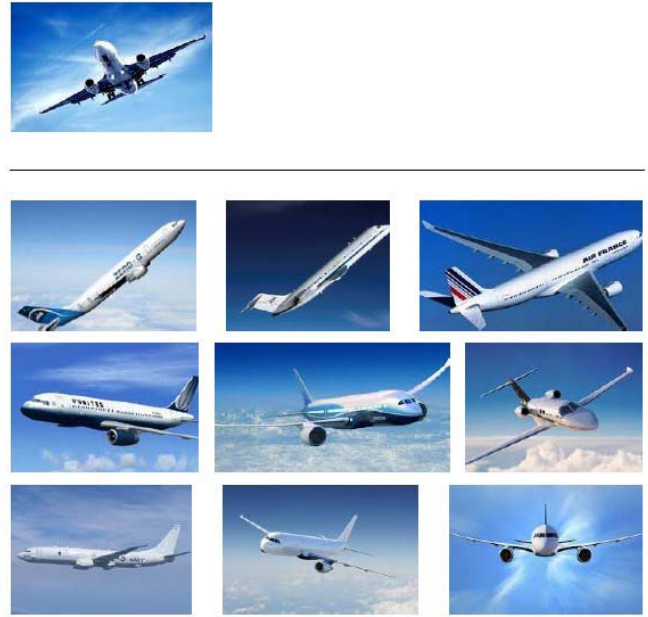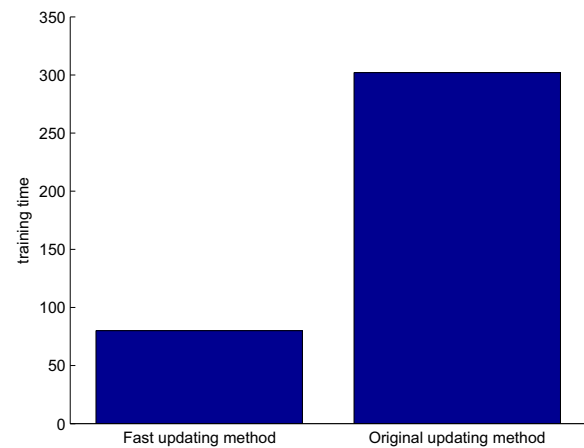
In figure 5, the top nine nearest neighbor images of a query image is shown. On the top of the figure, the query image is presented, and the retrieved nearest neighbors are presented below. It is clear that the proposed method can relevant images from the database.

We also compare the proposed fast updating algorithm with the original updating one. The training time is given in figure 6. It is indicated that the proposed algorithm can

reduce the training time significantly.

IV. CONCLUSION

In this paper, we propose a novel method to represent image using the kernel classifiers. The motivation behind this is that two images are regarded as similar to each other if their response to a classifier are similar. We developed a fast and effective iterative algorithm to train the kernel classifier when the training image number is tremendous. The presented experiment results show the proposed presentation method outperforms the original features. In addition, our methodfig2 reduce the training time eminently.

REFERENCES

[1] H. Liu and S. Li, "Decision fusion of sparse representation and support vector machine for sar image target recognition," *Neurocomputing*, vol. 113, pp. 97–104, 2013.

[2] P. Vundela and V. Sourirajan, "A robust multiwavelet-based watermarking scheme for copyright protection of digital images using human visual system," *International Arab Journal of Information Technology*, vol. 10, no. 6, 2013.

[3] L. Dai, X. Gao, Y. Guo, J. Xiao, Z. Zhang, *et al.*, "Bioinformatics clouds for big data manipulation," *Biology direct*, vol. 7, no. 1, p. 43, 2012.

[4] J. J.-Y. Wang, X. Wang, and X. Gao, "Non-negative matrix factorization by maximizing correntropy for cancer clustering," *BMC bioinformatics*, vol. 14, no. 1, p. 107, 2013.

[5] J. J.-Y. Wang, H. Bensmail, and X. Gao, "Multiple graph regularized nonnegative matrix factorization," *Pattern Recognition*, vol. 46, no. 10, pp. 2840–2847, 2013.

[6] A. Bosco, D. Giacalone, A. Bruna, S. Battiato, and R. Rizzo, "Signal activity estimation with built-in noise management in raw digital images," in *VISAPP 2013 - Proceedings of the International Conference on Computer Vision Theory and Applications*, vol. 1, 2013, pp. 118–121.

[7] L. Ruan and A. Groves, "Illinois firefighter line of duty deaths digital image collection database: A knowledge management initiative," *Fire Technology*, vol. 49, no. 1, pp. 163–173, 2013.

[8] X. Liu, Z.-P. Shi, and Z.-Z. Shi, "A co-boost framework for learning object categories from google images with 1st and 2nd order features," *Visual Computer*, pp. 1–13, 2013.

[9] Y. Zhou, K. Chen, and X. Yang, "Google image search refinement: Finding text in images using local features," in *CSAE 2012 - Proceedings, 2012 IEEE International Conference on Computer Science and Automation Engineering*, vol. 1, 2012, pp. 98–101.

[10] E. Hoque, O. Hoeber, and M. Gong, "Cider: Concept-based image diversification, exploration, and retrieval," *Information Processing and Management*, 2013.

[11] L. Nie, S. Yan, M. Wang, R. Hong, and T.-S. Chua, "Harvesting visual concepts for image search with complex queries," in *MM 2012 - Proceedings of the 20th ACM International Conference on Multimedia*, 2012, pp. 59–68.

[12] J. J. Wang, H. Bensmail, and X. Gao, "Multiple graph regularized protein domain ranking," *BMC bioinformatics*, vol. 13, no. 1, p. 307, 2012.

[13] C. Yan, "Accurate image retrieval algorithm based on color and texture feature," *Journal of Multimedia*, vol. 8, no. 3, pp. 277–283, 2013.

[14] J. Wang, Y. Li, Q. Wang, X. You, J. Man, C. Wang, and X. Gao, "ProClusEnsem: Predicting membrane protein types by fusing different modes of pseudo amino acid composition," *Computers in biology and medicine*, vol. 42, no. 5, pp. 564–574, 2012.

[15] X.-F. Liu, "Face recognition with texture facial features classification using kernel fisher classifier," *International Journal of Advancements in Computing Technology*, vol. 5, no. 8, pp. 259–266, 2013.

[16] Z. Fu, G. Lu, K. Ting, and D. Zhang, "Learning sparse kernel classifiers for multi-instance classification," *IEEE Transactions on Neural Networks and Learning Systems*, 2013.

[17] S. Tabibian, A. Akbari, and B. Nasersharif, "Extension of a kernel-based classifier for discriminative spoken keyword spotting," *Neural Processing Letters*, pp. 1–24, 2013.

[18] N. Inoue and Y. Yamashita, "Simultaneous learning of localized multiple kernels and classifier with weighted regularization," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7626 LNCS, pp. 354–362, 2012.

[19] P. Papadakis, "Enhanced pose normalization and matching of non-rigid objects based on support vector machine modelling," *Pattern Recognition*, 2013.

[20] S. Asha, C. Sudalaimani, P. Devanand, T. Thomas, and S. Sudhamony, "Automated seizure detection from multichannel eeg signals using support vector machine and artificial neural networks," in *Proceedings - 2013 IEEE International Multi Conference on Automation, Computing, Control, Communication and Compressed Sensing, iMac4s 2013*, 2013, pp. 558–563.

[21] A. Hipni, A. El-shafie, A. Najah, O. Karim, A. Hussain, and M. Mukhlisin, "Daily forecasting of dam water levels: Comparing a support vector machine (svm) model with adaptive neuro fuzzy inference system (anfis)," *Water Resources Management*, vol. 27, no. 10, pp. 3803–3823, 2013.

[22] T. Silva and L. Zhao, "Uncovering overlapping cluster structures via stochastic competitive learning," *Information Sciences*, 2013.

[23] I. Chattopadhyay and H. Lipson, "Abductive learning of quantized stochastic processes with probabilistic finite automata," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1984, 2013.

[24] Z. Liu, A. Abbas, B.-Y. Jing, and X. Gao, "Wavpeak: picking nmr peaks through wavelet-based smoothing and volume-based filtering," *Bioinformatics*, vol. 28, no. 7, pp. 914–920, 2012.

[25] J. Park, N. Kim, M. Jeong, and K. Shin, "Multiphase support vector regression for function approximation with breakpoints," *Journal of the Operational Research Society*, vol. 64, no. 5, pp. 775–785, 2013.

[26] D. Geebelen, J. Suykens, and J. Vandewalle, "Reducing the number of support vectors of svm classifiers using the smoothed separable case approximation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 4, pp. 682–688, 2012.

[27] A. Freytag, E. Rodner, P. Bodesheim, and J. Denzler, "Rapid uncertainty computation with gaussian processes and histogram intersection kernels," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7725 LNCS, no. PART 2, pp. 511–524, 2013.

[28] A. Freytag, B. Frohlich, E. Rodner, and J. Denzler, "Efficient semantic segmentation with gaussian processes and histogram intersection kernels," in *Proceedings - International Conference on Pattern Recognition*, 2012, pp. 3313–3316.

[29] Y. Usui and K. Kondo, "The sift image feature reduction method using the histogram intersection kernel," in *ISPACS 2009 - 2009 International Symposium on Intelligent Signal Processing and Communication Systems, Proceedings*, 2009, pp. 517–520.

[30] T. Chen, H. Deng, and Y. Lan, "A novel image annotation method using region-based multi-instance semi-supervised learning," *Journal of Computational Information Systems*, vol. 9, no. 9, pp. 3667–3674, 2013.

[31] J.-R. Ding, J.-H. Huang, J.-F. Liu, and Y.-T. Zhang, "Combining local features and multi-instance learning for ultrasound image classification," *Zidonghua Xuebao/Acta Automatica Sinica*, vol. 39, no. 6, pp. 861–867, 2013.

[32] J. J.-Y. Wang, H. Bensmail, and X. Gao, "Joint learning and weighting of visual vocabulary for bag-of-feature based tissue classification," *Pattern Recognition*, vol. 46, no. 12, pp. 3249–3255, 2013.

[33] R. Hibi and N. Takahashi, "A modified multiplicative update algorithm for euclidean distance-based nonnegative matrix factorization and its global convergence," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7063 LNCS, no. PART 2, pp. 655–662, 2011.

[34] P. Nilsson and E. Hertz, "Ultra low power hardware for computing squared euclidean distances," in *2011 20th European Conference on Circuit Theory and Design, ECCTD 2011*, 2011, pp. 580–583.

[35] A. Sony, K. Ajith, K. Thomas, T. Thomas, and P. Deepa, "Video summarization by clustering using euclidean distance," in *2011 - International Conference on Signal Processing, Communication, Computing and Networking Technologies, ICSCCN-2011*, 2011, pp. 642–646.

[36] H. Wang, F. Shkjezi, and E. Hoxha, "Distance metric learning for multi-camera people matching," in *Advanced Computational Intelligence (ICACI), 2013 Sixth International Conference on*. IEEE, 2013, pp. 140–143.

[37] X. Yang, Q. Yu, L. He, and T. Guo, "The one-against-all partition based binary tree support vector machine algorithms for multi-class classification," *Neurocomputing*, vol. 113, pp. 1–7, 2013.

[38] M. Kumar and M. Gopal, "Reduced one-against-all method for multiclass svm classification," *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 238–14 248, 2011.

[39] T. Mota and A. Thomé, "One-against-all-based multiclass svm strategies applied to vehicle plate character recognition," in *Proceedings of the International Joint Conference on Neural Networks*, 2009, pp. 2153–2159.

[40] G.-S. Chen, H.-Y. Tang, D.-Q. Zhang, Y.-X. Jiao, and H.-X. Wang, "Existence of three solutions for a nonlocal elliptic system of-kirchhoff type," *Boundary Value Problems*, vol. 2013, no. 1, pp. 1–9, 2013.

[41] J. Wang, X. Gao, Q. Wang, and Y. Li, "ProDis-ContSHC: learning protein dissimilarity measures and hierarchical context coherently for protein-protein comparison in protein database retrieval," *BMC bioinformatics*, vol. 13, no. Suppl 7, p. S2, 2012.

[42] M. Banitt, L. Ventura, W. Feuer, E. Savatovsky, G. Luna, O. Shif, B. Bosse, and V. Porciatti, "Progressive loss of retinal ganglion cell function precedes structural loss by several years in glaucoma suspects," *Investigative Ophthalmology and Visual Science*, vol. 54, no. 3, pp. 2346–2352, 2013.

[43] H. Lai, Y. Pan, Y. Tang, and N. Liu, "Efficient gradient descent algorithm for sparse models with application in learning-to-rank," *Knowledge-Based Systems*, vol. 49, pp. 190–198, 2013.

[44] M. Orouskhani, M. Mansouri, Y. Orouskhani, and M. Teshnehlab, "A hybrid method of modified cat swarm optimization and gradient descent algorithm for training anfis," *International Journal of Computational Intelligence and Applications*, vol. 12, no. 2, 2013.

[45] H. Xiong, S. Chen, X. Peng, and S. Wu, "An approximate gradient descent algorithm for wireless sensor signal reconstruction," *Journal of Information and Computational Science*, vol. 10, no. 8, pp. 2437–2444, 2013.

[46] H. Carroll, M. Kann, S. Sheetlin, and J. Spouge, "Threshold average precision (tap-k): A measure of retrieval designed for bioinformatics," *Bioinformatics*, vol. 26, no. 14, pp. 1708–1713, 2010.

[47] L. Ru, S. Ma, and J. Lu, "Feature fusion based on the average precision in image retrieval," *Jisuanji Yanjiu yu Fazhan/Computer Research and Development*, vol. 42, no. 9, pp. 1640–1646, 2005.

[48] K. Kishida, "Property of average precision and its generalization: An examination of evaluation indicator for information retrieval experiments," *NII Technical Reports*, vol. 2005, no. 14, pp. 1–19, 2005.