International Refereed Journal of Engineering and Science (IRJES)
ISSN (Online) 2319-183X, (Print) 2319-1821 *Volume 1, Issue 1 (September 2012), PP. 022-030*
www.irjes.com

Web Based Image Retrieval Incorporatinggeneralized Multi-Instance (Gmi) Learning And Bag-Based Reranking

R.SOWMYA*, SHARAD KULKARNI**

*(*Department of Electronics and communication engineering, Sri KottamTulasi Reddy Memorial College, JNTU-H university,Kondair. AP)*
*(**Department of Electronics and communication engineering, Sri KottamTulasi Reddy Memorial College, JNTU-H university,Kondair. AP)*

ABSTRACT----*This paper incorporates two newschemes called Generalized Multi-Instance (GMI) Learning and Bag-Based Rerankingfor large-scale TBIR. It has two key steps: first, cluster relevant images using both textual and visual features. Instead of directly reranking the relevant images by using traditional image reranking methods, we have partitioned the relevant images into clusters. By treating eachcluster as a "bag" and the images in a bag as "instances." second, To facilitate (G)MI learning in our framework, we have proposed an automatic bag annotation method to automatically find positive and negative bags for training classifiers. To demonstrate the effectiveness of the proposed method, wecompare the performance of the proposed method with other existing ones like SIL-SVM, mi-SVM. The experimentalresults show that the proposed method is usually better than the others.*

Keywords:*Generalized Multi-Instance(GMI) Learning and Bag-Based Reranking, weak bag annotation, bag ranking score, text-based image retrieval (TBIR).*

I. INTRODUCTION

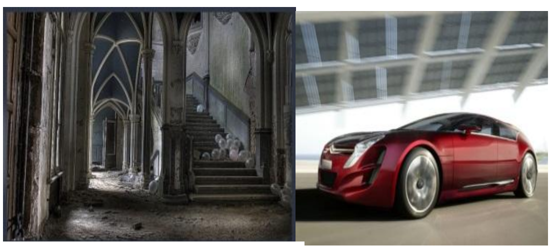
Internet makes it possible for human to access huge amount of information. The great paradox of the World Wide Web is that the more information available about a given topic, the more difficult it is to locate the accurate and relevant information. Most of the users know what information they need, without knowing where to get it from. Some of the users know what the information they are looking for and where to get it from; and they get it by following appropriate links. But these users usually miss the relevant information available on the web which is far from their known links. Search engines can facilitate all users to locate such relevant information.

Many information retrieval systems appeared in recent years. Text retrieval systems satisfy users with sufficient success. Google and Yahoo! are two examples of the top retrieval systems which have billions of hits a day. Even though Internet contains media like images, audio and video, retrieval systems for these types of media are rare and have not achieved success as that of text retrieval systems.

Nowadays, web image search engines (e.g. *Google, yahoo*) rely almost purely on surrounding text features. This leads to ambiguous and noisy results. Image search reranking methods usually fail to capture the user's intention when the query term is ambiguousas shown in Fig 1.

To address this issue, many image reranking methods havebeen developed [5], [12]–[14], [31], [32], [36], [42] to rerank theinitially retrieved images using visual features.

To improve the retrieval performance, in this paper, we introduce a new framework, referred to as the bag-based image reranking framework, for large-scale TBIR. We first partition the relevant images into clusters by using visual and textual features. We treat each cluster of imagesas a "bag" andthe images inside the cluster as "instances."



The main Ground floor hall. Mission – Chateau Noisy, Belgium We set

Edward's Volvo, the silver S60R: Carlisle's MercedesS55 AMG

Fig. 1.Web images with noisy tags.

In traditional MI learning methods, if a bag contains at least one relevant instance, this bag is labeled as positive; if the instances in a bag are all irrelevant, this bag is labeled as negative.In our image retrieval application, we observe that it isVery likely that multiple relevant images are clustered in a positive bag while a few relevant images may be clustered with irrelevant images in a negative bag. Different from traditional MIlearning, we propose a generalized MI (GMI) setting for this application in which at least a certain portion of a positive bag is of positive instances, while a negative bag might contain atmost a few positive instances. In this case, the traditional MI methods may not be effective to address the ambiguities on the instance labels in both positive and negative bags. Therefore, we

Propose a new GMI learning algorithm using SVM, referred toGMI-SVM, which uses the recently proposed "Label Generation"strategy [18] and maximum margin criterion to effectivelyrerank the relevant images by propagating the labels from thebag level to the instance level.

Inoursetting,each bag(cluster)canhavearoughestimateoftheproportionofpositiveinstances(images).Forexample,the positivebagsconsist of$atleast\mu = 10\%$ positiveinstances,whereasthenegative bagshave$atmost\gamma = 2\%$positiveinstances.NotethatournewassumptionisdifferentfromtheconventionalMIassumption intwoaspects:1) itremovesthe strictassertionof the negative bagsand2)itprovides more information for positive bags.To address the ambiguities on the instance label sin both positive and negative bags,wethengeneralizetheMIlearningproblem underthenewsettinganddevelopaGMI-SVM algorithm for labelprediction oninstances (images) toenhance there trieval performance.

To facilitate (G)MI learning in our framework, we conducta so-called *weak bag annotation* process to automatically findpositive and negative bags for training classifiers. First, weintroducean *instance ranking score* defined by the similarity betweenthe textual query and each relevant image. Then, we obtaina *bag ranking score* for each bag by averaging the instanceranking scores of the instances in this bag. Finally, we rank allbags with the bag ranking score. In our automatic bag annotationmethod, the top ranked bags are used as the pseudopositivebags, and pseudonegative bags are obtained by randomly samplinga few irrelevant images that are not associated with thetextual query. After that, these bags are used to train a classifierthat is then used to rerank the database images.

Fig. 2 showsthe overall flowchart of our proposed bag-based framework forthe TBIR. We will show in the experiments that our frameworkwith the automatic bag annotation method performs much betterthan the existing image reranking methods [12], [42]. Moreover,users are also allowed to manually annotate positive/negativebags during the RF process, and our experiments show that theretrieval performance of GMI-SVM can be further improved byusing the manually labeled training bags.

II. BAG-BASEDWEBIMAGERERANKING FRAMEWORK

Here,wepresentourproposedbag-basedrerankingframe-workforlarge-scale TBIR. Ourgoalisto improve the Web imageretrievalin Internetimage databases, suchas $\mathtt{Flickr}i$. TheseWebimages areusuallyaccompaniedbytextualdescriptions.For the thWebimage,thelow-levelvisualfeature

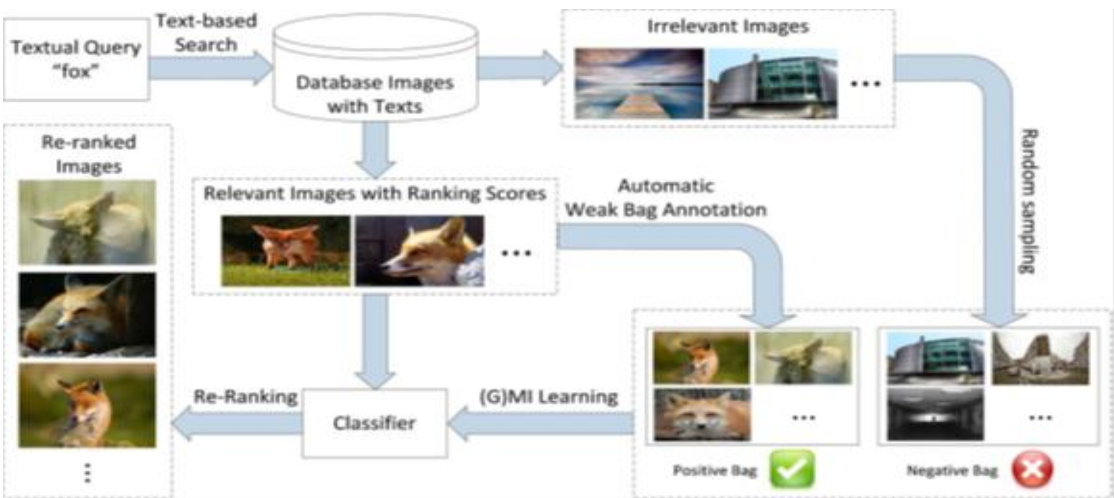


Fig: 2. Bag-based image reranking framework for large-scale TBIR.

$\mathbf{v}_i$ (e.g.,color,texture,andshape)andthetextualfeature$\mathbf{t}_i$
(e.g.,termfrequency)canbeextracted.Wefurtheraggregatethemintoasinglefeature vector forsubsequentoperations, namely, $\mathbf{x}_i = \lambda\mathbf{v}_i', \mathbf{t}_i']'$, where$\lambda$is aweightparameter.

A. Initial Reranking

Aftertheuserprovidesatextualquery q(e.g.,"fox"),our systemexploitstheinverted-file method[19]toautomatically findrelevantWebimageswhosesurroundingtextcontainsthe textualquery tag q,aswellasirrelevant Webimages whose surroundingtextdonotcontain q .Foreachretrievedrelevant image x,aninstancerankingscorecanbedefinedasfollows[3]:

$$r(\mathbf{x}) = -\tau + \frac{1}{\delta} \quad (1)$$

whereδisthetotalnumberoftagsinimage x andτisthe rankpositionofthequerytag q inthetaglistofimage x.If$\tau_i < \tau_j$ and$i \neq j$,thenwehave$r(\mathbf{x}_i) > r(\mathbf{x}_j)$.Inotherwords,whenonerelevantimagecontainsthetextualquery qatthetoppositioninitstaglist,thisimagewillbeassigned a higherrankingscore.Whenthepositionsofthequerytag qare thesameforthetwoimages(i.e.,τ_i = τ_j), therankingscoreis decidedby$\delta_i$andδ_j ,namely,theimagethathasfewertagsis preferred.

B.Weak Bag Annotation Process

Inourframework,eachimageisconsideredasan"instance." Toconstruct"bags,"wepartitionthe relevant images intoclustersusingthe$k$-meansclustering methodbasedonvisualand textualfeatures. Afterthat, each cluster is consideredasa"bag." Tofacilitate(G)MIlearningmethodsinourframework,wehave toannotate positiveandnegativebagstotrainclassifiers.Note thatonlythebagsaretobeannotated, whilethelabelsofin-stancesineachbagarestillambiguous.Therefore, wereferto theannotationof abagas*weakbagannotation*. Specifically,foreach bag$\mathcal{B}_I$,itsbagrankingscore$S(\mathcal{B}_I)$is definedas theaverage instancerankingscore,i.e.,

$$S(B_I) = \frac{\sum_{x \in B_I} r(x)}{|B_I|} \quad (2)$$

where$|B_I|$standsfor the cardinalityof bag$\mathcal{B}_I$.

Inourautomaticbagannotationmethod,thetop-rankedbags with higher bagranking scores are used aspseudo positive bags, andthesamenumberofpseudonegativebagsisobtainedbyrandomlysampling afewirrelevantimages.Wewillshowinthe experimentsthatourGMIlearningmethodGMI-SVMwiththissimplebagannotation methodcanachievebetterretrievalper-formanceswhencomparedwiththosein[12]and[42].Notethat ourproposedautomaticweakbagannotationmethodis similar tothepseudo-RFalgorithmproposedin[37],whichcanannotateinstances, whereas ourapproach canannotatehigh-confidentbags.

C. GMI Learning

We denote the transpose of a vector/matrix by superscript'.We also define I as the identity matrix and 0 and 1$\in R^n$asthe zero vector and the vector of all 1's, respectively. Moreover,the element-wise product between matrices P and Q is representedas P ΘQ. Inequality u =$[u_1, u_2, \ldots, u_n]' \geq 0$ meansthat $u_i \geq 0$ for i = 1,…..,n. A positive or negative bag$\mathcal{B}_I$ isassociated with a bag label$Y_1 \in \{\pm 1\}$.We also denote the un-observed instance label of $x_i$as$y_i \in \{\pm 1\}$ .With this definitionof bags, we can define the GMI constraint on the instance labelsof positive and negative bags, respectively, as

$$\sum_{i:\mathbf{x}_i \in \mathcal{B}_I} \frac{y_i + 1}{2} \geq \mu\beta_I \text{for} Y_I = 1$$

$$\sum_{i:x \in \beta 1} \frac{y_i + 1}{2} \leq \gamma|\beta_I|, \text{for } Y_I = -1 \quad (3)$$

In other words, positive instances take up at leastportionμ of a positive bag, whereas positive instances occupy at mostportionγ of a negative bag. Note that traditional MI learning[1], [40] is actually a special case of GMI learning with$\mu = \frac{1}{\beta_I}$and$\gamma = 0$. In contrast to the restrictive MI assumptionin [1] and [40], the GMI constraint in (3) is more suitable to thisapplication.

We further denote y = $[y_1, \ldots\ldots, y_n]'$ as the vector of instance labels and y = $\{y/y_i \epsilon\{\div \pm 1\}$,and y satisfies (3)}as the domain of y. Then, the decision function of the GMIlearning can be learned by minimizing thefollowing structuralrisk functional:

$\min_{y\epsilon\gamma ,f}\ \Omega\,(\|f\|) + c\sum_{i=1}^{n} l(-y_i\, f(x_i))$ (4)

Where$\Omega(\|f\|)$ is the regularization term,$l(\cdot)$ is a loss functionfor each instance, and$C$ is the parameter that trades off the complexityand the fitness of the decision function$f$. Note that theconstraints in (3) are integer constraints; thus, the correspondingGMI problem (4) is usually formulated as a mixed integer programmingproblem.

D. GMI-SVMs

In this paper, we assume the decision function is in form of$f(x)$=$w'\varphi(x) + b$and the regularization term is (1/2)$\|w\|^2$.We adopt the formulation of the Lagrangian SVM, in which thesquare bias penaltyb^2 and the square hinge loss for each instanceare used in the objective function. The GMI optimizationproblem can be written as the following constrained optimizationproblem:

$\min_{y\in\gamma,w,b,\rho\,\xi_i}\frac{1}{2}(\|w\|^2 + b^2 + c\sum_{i=1}^{n}\xi_i^2) - \rho$ s.t.$y_i(w'\phi(x_i) + b) \geq \rho - \xi_i, i = 1, \ldots, n.$ (5)

Where $\xi_i$values are slack variables and $\rho/\|w\|$defines themargin separation. By introducing a dual variableα_i foreach inequality constraint in (5) and the kernel trick (i.e.,$k(x_i, x_j)$ =$\phi(x_i)'\phi(x_j)$), we arrive at the following minimaxsaddle-pointproblem:

$$\min_{y\epsilon y}.\max_{\alpha\epsilon A} -\frac{1}{2}\alpha'(\tilde{k}\odot yy' + \frac{1}{c}I)\alpha \quad (6)$$

Where $\alpha = [\alpha_1, \ldots\ldots, \alpha_n]'$ is the vector of the dual variablesand $\mathcal{A} = \{\alpha\alpha \geq 0, \alpha'1 = 1\}$is the domain of$\alpha$. Wealso define K=$[k(x_i, x_j)]$ as an n$\times$ n kernel matrix and K=K+11' as an n$\times$ $n$transformed kernel matrix forthe augmented feature mapping $\widetilde{\emptyset}(x) = [\emptyset(x)', 1]'$of kernel$\tilde{k}(x_i, x_j) = \widetilde{\emptyset}(x_i)'\widetilde{\emptyset}(x_j)$. Note that the instance labels$y_i$in (6) are also integer variables, and thus, (6) is a mixed integerprogramming problem, which is computationally intractable ingeneral.

Recently, Li *et al.* [18] proposed an efficient convex optimizationmethod to solve the mixed integer programming problemfor maximum margin clustering. In this paper, we extend theiralgorithm [18] to solve the mixed integer programming problemin (6). Our proposed method is then referred to as the GMISVM.

1) Convex Relaxation: First, let us consider interchangingthe order of$\min_{y\epsilon y}$. and $\max_{\alpha\epsilon A}$ and in (6). Then, we have

$\max\limits_{\alpha\epsilon A}\min_{y\epsilon y}. -\frac{1}{2}\alpha'(\tilde{k}\odot yy' + \frac{1}{c}I)\alpha$ (7)

According to the minimax theorem [16], the optimal objective of (6) is an upper bound of that of (7). By introducing θ, we can further rewrite (7) as follows:

$\max\limits_{\alpha\epsilon A}\{\ \max_\theta -\theta{:}\, \theta \geq . -\frac{1}{2}\alpha'(\tilde{k}\odot y^t y^{t'} + \frac{1}{c}I)\alpha\ , \forall y^t \in \gamma\}$ (8)

where$y^t$is any feasible solution in y . For the inner optimization subproblem of (8), we can obtain its Lagrangian L as follows by introducing a dual variable $d_t \geq 0$ for each constraint:

$L = -\theta + \sum_{t:y^t\in y} d_t\ (\theta - \frac{1}{2}\alpha'(\tilde{k}\odot y^t y^{t'}$
$+\frac{1}{c}I)\alpha\ )$ (9)

Setting the derivative of Lagrangian (9) with respect to θtozero, we have$\sum_{t:y^t\in y} d_t = 1$. Denote das a vector of $d_t$valuesand
$\mathcal{M} = \{d | d \geq 0, d'1 = 1\}$as the domain of d.We can thenarrive at its dual form as follows:

$\min_{d\in\mathcal{M}} - \frac{1}{2}\alpha'(\sum_{t:y^t\in y} d_t\tilde{k}\odot y^t y^{t'} + \frac{1}{c}I)\alpha$ (10)

Replacing the inner maximization subproblem in (8) with itsdual (10), we have the followingoptimization problem:

$$\max_{\alpha\epsilon A} \min_{d\in\mathcal{M}} -\frac{1}{2}\alpha'(\sum_{t:y^t\in y} d_t\tilde{k}\odot y^t y^{t'}+\frac{1}{c}I)\alpha = \min_{d\in\mathcal{M}} \max_{\alpha\epsilon A} -\frac{1}{2}\alpha'(\sum_{t:y^t\in y} d_t\tilde{k}\odot y^t y^{t'}+\frac{1}{c}I)\alpha \quad (11)$$

The equality holds as the objective function is concave in$\alpha$and linear in d, and thus, we can interchange the order of$\max_{\alpha\epsilon A}$and$\min_{d\in\mathcal{M}}$ in (11). Observe that (11) is analogousto the multiple kernel learning (MKL) problem [22], except thata label-kernel matrix, which is a convex combination of thebase label-kernel matrices$\tilde{k}\odot y^t y^{t'}$, is to be learned. Hence,(11) can be viewed as a multiple label-kernel learning (MLKL)problem.

2) Cutting-Plane Algorithm for GMI-SVM: Although is$y$finite and the MLKL problem (11) is a special case of MKL,there areO(2^n) candidates of the label vectory^t, and thus, the number of base label-kernel matrices$\tilde{k}\odot y^t y^{t'}$ is exponentialin size. Thus, it is not possible to directly apply recently proposedMKL techniques such as SimpleMKL [22] to our proposedGMI-SVM.

Algorithm 1: Cutting-plane algorithm for GMI-SVM.

1:Initialize$y_i = Y_I$ for$i \epsilon$ $B_I$as y^1, and set $\{y^1\}$;
2:Compute MKL to solve $\alpha$and din (11) based on C;
3:Useα to select the most violated $y^t$and set C=$y^t \cup C$;
4:Repeat lines 2 and 3 until convergence.

Fortunately, not all quadratic inequality constraints in (8) arenecessarily active at optimality, and only subset$C \subset \Upsilon$ ofthese constraints can usually lead to a very good approximationof the original optimization problem. Therefore, we canapply the cutting-plane method [15] to handle this exponential number of constraints. Moreover, the same strategy has beenalso applied in the recently proposed infinite kernel learning(IKL) [9], [10], in which the kernel is learned from an infiniteset of general kernel parameters, and thus, MLKL (with kernel$\sum_{t:y^t\in y} d_t\tilde{k}\odot y^t y^{t'}$) can be deemed as a variant of IKL. Asa result, our GMI-SVM enjoys the same convergence of IKL[9].The whole algorithm is summarized in Algorithm1. First,we set subset$C=\{y^1\}$, where the instance label vector y^1isinitialized according to the bag labels. Since $C$is no longer exponentialin size, one can apply MKL to learn the label kernel toobtain both $\alpha$and d. With a fixedα , the label vectory^t with aquadratic inequality constraint in (8), which is the most violatedone by the current solution, is then added to C. The process isrepeated until the convergence criterion (i.e., the relative changeof the objective values of (11) between two successive iterationsis less than 0.01) is met. After solving the MLKL problem, thedecision function can be obtained by
f(x)=$\sum_{i=1}^{n} \alpha_i \bar{y}_i\, \bar{k}(x,x_i)$

where$\bar{y}_i$=$\sum_{t:y^t\in C} d_t y_i^t$ and
$\bar{k}(x,x_i)= k(x,x_i)$+1

Algorithm 2: Finding the approximation of the most violate y^t.

1: Initialize $y_i = 1$for allx_i in positive bags $B_p$andy_i=-1 for allx_i in negative bagsB_N;
2:**for** each positive bagB_p **do**
3:Fix the labeling of instances in all the other bags, andfind the optimal $y_P$that maximizes the objective of (12) byenumerating the candidates of$y_i$inB_P;
4:**end for**
5:**for** each negative bagB_N **do**
6: Fix the labeling of instances in all the other bags, andfind the optimaly_N that maximizes the objective of (12) byenumerating the candidates of$y_i$inB_N ;
7:**end for**
8:Repeat lines 2–7 until convergence.

3) Finding the Approximation of the Most Violatedy^t: Similarto IKL, finding the most violated constraint (indexed by y^t)in MLKL is problem specific and is the most challenging part incutting-plane algorithms. Here, we discuss how to search for themost violated constraint to satisfy the GMI constraints in (3).
Referring to (8), to find the most violatedy^t , we have to solvethe following problem:
$\max_{y\epsilon\Upsilon} \quad \alpha'(\tilde{k}\odot yy')\alpha$(12)
Note that finding the most violated y^tthat maximizes (12) is acomputationally expensive problem when thebag size is large.

To accelerate our framework, we propose to use the instanceranking score defined in (1) to enforce the total number of instancesin each positive bag to be 15. Moreover, we can beforehand exclude a large numberof candidates ofy^t by checking our proposed GMI constraintin (3). In order to further speed up the process, we develop asimple but effective method. The basic idea is to enumerate thecandidates of $y_i$satisfying (3) for each bagB_I by fixing the labelingof other bags. Then, we iteratively choose the best$y_I$forB_I, which maximizes (12), where$y_I$is the vector of instance labelsinB_I. The procedure will be terminated when the relativechange of the objective values of (12) between two successiveiterations is less than 0.001. The detailed procedure is listed inAlgorithm 2.

III. Experimental Setup

In our experiments, for any given textual query (e.g., "fox"),the relevantWeb images that are associated with the word "fox"are firstly ranked using (1). We refer to this initial Web imagesearch method as *Init_Ranking*. We compare our bag-basedreranking framework and two existing methods, i.e., WEBSEIC[42] and information bottleneck (IB) reranking (IBRR) [12],for image reranking. It is worth noting that existing MI learningalgorithms can be readily adopted in our rerankingframework.we only employmi-SVM [1] and single-instance learning SVM (SIL-SVM)[2] in this paper as they are more suitable for predictingthe labels of bags rather than instances.

The assumption in ournewly proposed GMI-SVM is that positive instances compriseat least a certain portion of a positive bag, while a negative bagmay contain at most a few positive instances.81 images are employed as the database images, and all the 8 concept names are used astextual queries to perform the TBIR.Our frameworkcan achieve reasonable efficiency by using unoptimizedMATLAB code.we employ three types of global features-the grid color moment, the direction histogram feature and 128-D wavelet texture feature.We further concatenate all threetypes of visual features into lengthy feature vectors and normalizeeach feature dimension to zero mean and unit standarddeviation. To improve the speed and reduce the memory cost,principal component analysis is then applied for dimension reduction.

For the$i$ thimage, we further concatenate the visual feature$v_i$and the textual featurev_i together to form the lengthy featurevectorx_i, namely,$x_i = [\lambda v_i', t_i']'$, where the weight parameter$\lambda$is empirically fixed as 0.1 in the experiments. The databaseimages are grouped inton_B bags by using the k -means clusteringmethod with the distance metric defined as follows:

$$\boldsymbol{d}(x_i, x_j) = \sqrt{\lambda^2 \|v_i - v_j\|^2 + \|t_i - t_j\|^2} \quad (13)$$

wherev_i , $t_i$and v_j , t_j ,are the visual and textual features of the$i$th and jth images, respectively.We observethat it is computationally expensive to exploit the enumerationmethod for GMI-SVM if the number of instances in one bag is larger than 9. We therefore empirically set K=[(T/9)] inthe K-means clustering method, where T is the total number ofrelevant images.We throw away the clusters that have instances fewer than 9. For the remaining clusters, we only keep the top-ranked 9 instances with the highest instance ranking scoresto form one bag, and the remaining instances are discarded. Thebags are then ranked according to the average ranking score of the 9 instances in the bags. In the automatic bag annotation scheme, the top-ranked bags $n_B$are used as the positivebags, and we also randomly sample 9$n_B$irrelevant images toconstructn_B negative bags. Then_B positive and negative bagsare then used as the training data for GMI-SVM, mi-SVM, andSIL-SVM.ForGMI-SVM, we set proportion for positive bags andproportion for negative bags to fairly compare ourGMI-SVM and the other MI learning methods mi-SVM andSIL-SVM.

IV. Results of Retrieval Performances

GMI-SVM based on the convex relaxation in [18]can obtain a better optimal solution than other MI learningalgorithms for the bag-based reranking framework. The top-tenretrieved images of GMI-SVM, SIL-SVM, mi-SVM,WEBSEIC,IBRR, and Init_Ranking for the textual query "fox" are illustratedin Fig. 3. Again, we observe that GMI-SVM achieves thebest performance.

From Table I, we also observe that, for a fixedμ , GMI-SVMusing $\gamma = 0$generally achieves better performances comparedwith the results when setting $\gamma = 0.3$ and 0.5, which is consistentwith our observation that the negative bags generally donot contain positive instances.This observation demonstrates that, for thoseconcepts having more positive instances in the negative bags,

GMI-SVM can successfully cope with the ambiguities on theinstances in the negative bags and thus improve the retrieval performance.Considering that the MAP of GMI-SVM is the bestwhen setting$\gamma = 0$and$\mu = 0.5$, we fix$\gamma = 0$ and $\mu = 0.5$.We report the average central processing unit (CPU) time ofthe TBIR for different methods. For GMI-SVM, SIL-SVM, andmi-SVM, we still use one positive bag and one negative bag obtainedby using the automatic weak bag annotation process.

Fig. 3. Top-ten retrieved images of all methods for the textual query "fox." (Red boxes) Incorrect results.

	$\mu = 0.3$	$\mu = 0.5$	$\mu = 0.7$	$\mu = 0.9$
$\gamma = 0$	65.6%	**66.8%**	66.3%	66.2%
$\gamma = 0.3$	62.2%	66.7%	66.4%	65.9%
$\gamma = 0.5$	60.0%	66.4%	65.3%	65.7%

TABLE I

MAPS OVER 81 CONCEPTS OF GMI-SVM USING DIFFERENT POSITIVEPROPORTIONS (I.E., μANDγ) FOR POSITIVE AND NEGATIVE BAGS. EACHRESULT IN THE TABLE IS THE BEST AMONG THE RESULTS OBTAINED BYUSING DIFFERENT NUMBERS OF POSITIVE AND NEGATIVE TRAINING BAGS

V. CPU Time for Image Retrieval and Convergence Analysis

Allthe experiments are implemented with unoptimized MATLABcodes and performed on a workstation (3.33-GHz CPU with32-GB random access memory). The average CPU-time overall textual queries are shown in Table II.Ourproposed method GMI-SVM achieves reasonable efficiency forTBIR using unoptimized MATLAB codes. For GMI-SVM, onthe average, the iterative optimization algorithmtakes aboutsix iterations to converge for each concept. In Fig. 4, we take three concepts (i.e., "bus," "flower," and "horse") as examples toillustrate the convergence of GMI-SVM, in which the verticalaxis indicates the objective value of (11) and the horizontal axisgives the number of iterations. We have similar observationsfor other concepts.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a bag-based framework forlarge-scale TBIR. Given atextual query, relevant images are to be reranked after the initialtext-based search. Instead of directly reranking the relevantimages by using traditional image reranking methods, we havepartitioned the relevant images into clusters. By treating eachcluster as a "bag" and the images in a bag as "instances," wehave formulated this problem as a MI learning problem.To address the ambiguities on the instancelabels in both positive and negative bags, we have developed GMI-SVM to further enhance retrieval performance, inwhich the labels from the bag level have been propagated to the instance level.To facilitate (G)MI learning in our framework,we have propose an automatic bag annotation method to automaticallyfind positive and negative bags for training classifiers.Our framework using the automaticbag annotation method canachieve the best

	GMI-SVM	SIL-SVM	mi-SVM	WEBSEIC	IBRR	Init_Ranking
CPU	1.120	0.025	0.026	0.027	105.354	0.0005

TABLE II
AVERAGE CPU TIME (IN SECONDS) PER TEXTUAL QUERY FOR ALL METHODS

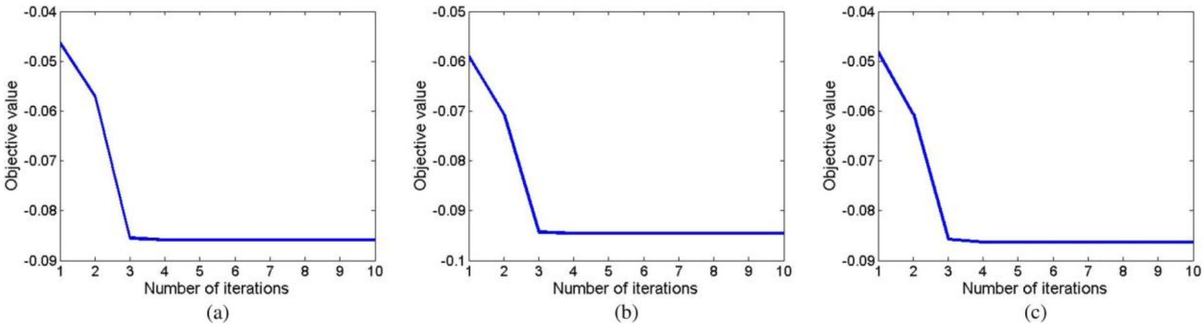


Fig. 4.Illustration of the convergence of GMI-SVM. (a) "Bus." (b) "Flower." (c) "Horse."

performance, as compared with other traditionalimage reranking methods on the NUS-WIDEdata set. Moreover, users are also allowed to manually annotate positive/negative bags during the RF process In order to decrease the CPU processing time principle component analysis(PCA) is employed and further future work is done in this direction.

REFERENCES:

[1] S. Andrews, I. Tsochantaridis, and T.Hofmann, "Support vector machines for multiple-instance learning," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2003, pp.561–568.

[2] R. C. Bunescu and R. J. Mooney, "Multiple instance learning for sparse positive bags," in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp.105–112.

[3] L. Chen, D. Xu, I. W. Tsang, and J. Luo, "Tag-based web photo retrieval improved by batch mode re-tagging," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.*, 2010, pp. 3440 3446.

[5] J. Cui, F. Wen, and X. Tang, "Real time google and live image search re-ranking," in *Proc. 16th ACM Int. Conf. Multimedia*, 2008, pp. 729–732.

[9] P. Gehler and S. Nowozin, "Infinite kernel learning" Max Planck Inst. Biol. Cybern., Tuebingen, Germany, Tech. Rep. TR-178, 2008.

[10] P. Gehler and S. Nowozin, "Let the kernel figure it out; principled learning of pre-processing for kernel classifiers," in *ProcIEEEComput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2009, pp. 2836–2843.

[12] W. H. Hsu, L. S. Kennedy, and S.-F.Chang, "Video search reranking via information bottleneck principle," in *Proc. 14th ACM Int. Conf.Multimedia*, 2006, pp. 35–44.

[14] Y. Jing and S. Baluja, "Pagerank for product image search," in *Proc.17th Int. Conf. World Wide Web*, 2008, pp. 307–316.

[15] J. E. Kelley, "The cutting plane method for solving convex programs," *SIAM J. Appl. Math.*, vol. 8, no. 4, pp. 703–712, Dec. 1960.

[16] S.-J. Kim and S. Boyd, "A minimax theorem with applications to machine learning, signal processing, and finance," *SIAM J. Optim.*, vol.19, no. 3, pp. 1344–1367, 2008.

[18] Y.-F. Li, I. W. Tsang, J. T. Kwok, and Z.-H. Zhou, "Tighter and convex maximum margin clustering," in *Proc. 22nd Int. Conf. Artif. Intell.Stat.*, 2009, pp. 344–351.

[22] A. Rakotomamonjy, F. R. Bach, and Y. Grandvalet, "SimpleMKL," *J.Mach. Learn. Res.*, vol. 9, pp. 2491–2521, 2008.

[31] X. Tian, D. Tao, X.-S.Hua, and X. Wu, "Active reranking for web image search," *IEEE Trans. Image Process.*, vol. 19, no. 3, pp. 805–820, Mar. 2010.

[32] X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, and X.-S.Hua, "Bayesian video search reranking," in *Proc. 16th ACM Int. Conf. Multimedia*, 2008, pp. 131–140.

[36] S.Wang, Q. Huang, S. Jiang, L. Qin, and Q. Tian, "Visual contextrank for web image re-ranking," in *Proc. 1st ACM Workshop Large-Scale Multimedia Retrieval Mining*, 2009, pp. 121–128.

[37] R.Yan,A.G.Hauptmann,andR.Jin,"Multimediasearchwithpseudo relevance feedback," in*Proc.ACMInt.Conf.ImageVideoRetrieval*, 2003,pp.238–247

[40] Q. Zhang and S. A. Goldman, "EM-DD: An improved multiple-instance learning technique," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2002, pp. 1073–1080.

[42] Z.-H. Zhou and H.-B.Dai, "Exploiting image contents in web search," in *Proc. 20th Int. Joint Conf. Artif.Intell.*, 2007, pp. 2928–2933.

AUTHOR BIOGRAPHIES:


R.SowmyareceivedherB.Tech,Degreein Electronics and communication engineering from Sri KottamTulasi Reddy Memorial College, Kurnool India,in 2009. Sheis pursuing her M.Tech in Digital Electronics and Communication Systems in Sri KottamTulasi Reddy Memorial College,Kondair India. Her area of interest is in the field of Content Based Image Retrieval(CBIR).


Mr.SharadkulkarnireceivedhisB.TechdegreeinElectronics and communication engineering from P.D.A college engineering, GulbargaIndia, in 1990.M.S in Electronics and Control from BITS pilani, Rajasthan India in 1995.He is pursuing Ph.D from Rayalaseema University Kurnool. He is currently working as Associate professor and Head of the Department E.C.E inSri KottamTulasi Reddy Memorial College,Kondair India.Hisresearch interestsincludeImage Processing and Content Based Image Retrieval(CBIR).