Web Based Image Retrieval Incorporating Generalized Multi-Instance (GMI) Learning And Bag-Based Reranking

R.SOWMYA*, SHARAD KULKARNI**
(*Department of Electronics and communication engineering, Sri Kottam Tulasi Reddy Memorial College, JNTU-H university, Kondair. AP)
(**Department of Electronics and communication engineering, Sri Kottam Tulasi Reddy Memorial College, JNTU-H university, Kondair. AP)

ABSTRACT—This paper incorporates two newschemes called Generalized Multi-Instance (GMI) Learning and Bag-Based Reranking for 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 each cluster 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, we compare the performance of the proposed method with other existing ones like SIL-SVM, mi-SVM. The experimental results 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 ambiguous as shown in Fig 1.

To address this issue, many image reranking methods have been developed [5], [12]–[14], [31], [32], [36], [42] to rerank the initially 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 images as a “bag” and the images inside the cluster as “instances.”

Fig. 1. Web images with noisy tags.
Web Based Image Retrieval Incorporating generalized Multi-Instance (GMI) Learning And

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 is very 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 MI learning, 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 at most 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 to as GMI-SVM, which uses the recently proposed "Label Generation" strategy [18] and maximum margin criterion to effectively rerank the relevant images by propagating the labels from the bag level to the instance level.

In our setting, each bag (cluster) can have a rough estimate of the proportion of positive instances (images). For example, the positive bags consist of at least 10% positive instances, whereas the negative bags have at most 2% positive instances. Note that our new assumption is different from the conventional MI assumption in two aspects: 1) it removes the strict assertion of the negative bags and 2) it provides more information for positive bags. To address the ambiguities on the instance label in both positive and negative bags, we generalize the MI learning problem under the new setting and develop a GMI-SVM algorithm for label prediction on instances (images) to enhance the retrieval performance.

To facilitate (G)MI learning in our framework, we conduct so-called weak bag annotation process to automatically find positive and negative bags for training classifiers. First, we introduce an instance ranking score defined by the similarity between the textual query and each relevant image. Then, we obtain a bag ranking score for each bag by averaging the instance ranking scores of the instances in this bag. Finally, we rank all bags with the bag ranking score. In our automatic bag annotation method, the top ranked bags are used as the pseudopositive bags, and pseudonegative bags are obtained by randomly sampling a few irrelevant images that are not associated with the textual query. After that, these bags are used to train a classifier that is then used to rerank the database images.

Fig. 2 shows the overall flowchart of our proposed bag-based framework for the TBIR. We will show in the experiments that our framework with the automatic bag annotation method performs much better than the existing image reranking methods [12], [42]. Moreover, users are also allowed to manually annotate positive/negative bags during the RF process, and our experiments show that the retrieval performance of GMI-SVM can be further improved by using the manually labeled training bags.

II. BAG-BASED WEB IMAGE RERANKING FRAMEWORK

Here, we present our proposed bag-based reranking framework for large-scale TBIR. Our goal is to improve the Web image retrieval in Internet image databases, such as Flickr. These Web images are usually accompanied by textual descriptions. For the Web image, the low-level visual feature

![Fig: 2. Bag-based image reranking framework for large-scale TBIR.](image-url)
A. Initial Reranking

After the user provides a textual query, our system exploits the inverted-file method[19] to automatically find relevant Web images whose surrounding text contains the textual query tag q, as well as irrelevant Web images whose surrounding text does not contain q. For each retrieved relevant image x, a instanceranking score can be defined as follows[3]:

\[
\tau(x) = \tau + \frac{1}{\delta}(1)
\]

where \(\delta\) isthetotalnum bersoftagsinimage x and \(\tau\) istherankpositionofthequery tag q in the taglist of image x. If \(\tau_i < \tau_j\) and \(i \neq j\), then we have \(\tau_i > \tau_j\). In other words, when nonerelevant imagecontains the textual query tag q at the position of the query tag q, the same for the Web images (i.e., \(\tau_i = \tau_j\)), the ranking score is decided by \(\delta\) and \(\tau\), namely, the image that has a lower weight is preferred.

B. Weak Bag Annotation Process

In our framework, each image is considered as an “instance.” To construct “bags,” we partition the relevant images into clusters using the -means clustering method based on visual and textual features. After that, each cluster is considered as a “bag.” To facilitate GMI learning methods in our framework, we have to annotate positive and negative bags to train classifiers. Note that only the bags are to be annotated, while the label information of instances in each bag is still ambiguous. Therefore, we refer to the annotation of abagasweak bag annotation.

Specifically, for each bag \(B_i\), its bag ranking score \(S(B_i)\) is defined as the average instance ranking score, i.e.,

\[
S(B_i) = \frac{\sum_{x \in B_i} \tau(x)}{|B_i|} \quad (2)
\]

where \(|B_i|\) stands for the cardinality of bag \(B_i\).

In our automatic bag annotation method, the top-ranked bags with higher bag ranking scores are used as asp eudo positive bags, and the same number of pseudonegative bags is obtained by randomly sampling a few irrelevant images. We will show in the experiments that our GMI learning method GMI-SVM with this simple bag annotation method can achieve better retrieval performance when compared with those in [12] and [42]. Note that our proposed automatic weak bag annotation method is similar to the pseudo-RF algorithm proposed in [37], which can annotate instances, whereas our approach can annotate high-confidence bags.

C. GMI Learning

We denote the transpose of a vector/matrix by superscript’. We also define I as the identity matrix and 0 and \(1 \in \mathbb{R}^n\) asthe zero vector and the vector of all 1’s, respectively. Moreover, the element-wise product between matrices P and Q is represented as P IQ. Inequality \(u \equiv [u_1, u_2, \ldots, u_n] \geq 0\) means that \(u_i \geq 0\) for \(i = 1, \ldots, n\). A positive or negative bag \(B_i\) is associated with a bag label \(Y_i \in \{\pm 1\}\). We also denote the unobserved instance label of \(x_i\) as \(y_i \in \{\pm 1\}\). With this definition of bags, we can define the GMI constraint on the instance label of positive and negative bags, respectively, as

\[
\sum_{i : x_i \in B_i} \frac{y_i + 1}{2} \geq \mu \beta_i \text{ for } Y_i = 1
\]

\[
\sum_{i : x_i \in B_i} \frac{y_i + 1}{2} \leq \gamma |\beta_i| \text{ for } Y_i = -1 \quad (3)
\]

In other words, positive instances take up at least portion \(\mu\) of a positive bag, whereas positive instances occupy at most portion \(\gamma\) of a negative bag. Note that traditional MI learning [11], [40] is actually a special case of GMI learning with \(\mu = \frac{1}{\beta_i}\) and \(\gamma = 0\). In contrast to the restrictive MI assumption in [11] and [40], the GMI constraint in (3) is more suitable to this application.

We further denote \(y = [y_1, \ldots, y_n]'\) as the vector of instance labels and \(y = [y/y_i \in \{\pm 1\}]\), and y satisfies (3) as the domain of y. Then, the decision function of the GMI learning can be learned by minimizing the following structural risk functional:

\[
\frac{1}{n} \sum_{i = 1}^{n} \log(1 + \exp(-y_i f(x_i))) + \frac{\lambda}{2} \sum_{i = 1}^{n} \left( \frac{1}{\beta_i} \|w_i\|_2^2 \right)
\]

www.irjes.com 24 | Page
min_{\gamma \epsilon \mathcal{F}} \Omega(\|f\|) + c \sum_{i=1}^{n} l(-y_i, f(x_i)) \quad (4)

Where \( \Omega(\|f\|) \) is the regularization term, \( l(\cdot) \) is a loss function for each instance, and \( c \) is the parameter that trades off the complexity and the fitness of the decision function \( f \). Note that the constraints in (3) are integer constraints; thus, the corresponding GMI problem (4) is usually formulated as a mixed integer programming problem.

D. GMI-SVMs

In this paper, we assume the decision function is in form \( f(x) = w' \phi(x) + b \) and the regularization term \( \|w\|^2 \). We adopt the formulation of the Lagrangian SVM, in which the square bias penalty \( b^2 \) and the constraint on the margin are slack variables.

Replacing the inner maximization subproblem in (8) with its dual (10), we have the following optimization problem:

\[
\min_{y \epsilon \mathcal{Y}, w, b, \rho, \xi} \left( \frac{1}{2} \|w\|^2 + b^2 + c \sum_{i=1}^{n} \xi_i^2 \right) - \sum_{i=1}^{n} y_i \left( w' \phi(x_i) + b \right) \geq \rho - \xi_i, \quad i = 1, \ldots, n. \quad (5)
\]

Where \( \xi_i \) values are slack variables and \( \rho / \|w\| \) defines the margin separation. By introducing a dual variable \( \alpha_i \) for each inequality constraint in (5) and the kernel trick \( \langle k(x_i, x_j), \phi(x_i), \phi(x_j) \rangle \), we arrive at the following minimax saddle-point problem:

\[
\min_{y \epsilon \mathcal{Y}, \max_{\alpha \in \mathcal{A}}} -\frac{1}{2} \alpha' (\bar{K} \bar{y} \bar{y} + \frac{1}{c} I) \alpha \quad (6)
\]

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

Recently, Li et al. [18] proposed an efficient convex optimization method to solve the mixed integer programming problem for maximum margin clustering. In this paper, we extend their algorithm [18] to solve the mixed integer programming problem (6). Our proposed method is then referred to as the GMISVM.

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

\[
\max_{\alpha \in \mathcal{A}} \min_{y \epsilon \mathcal{Y},} -\frac{1}{2} \alpha' (\bar{K} \bar{y} \bar{y} + \frac{1}{c} I) \alpha \quad (7)
\]

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

\[
\max_{\alpha \in \mathcal{A}} \left\{ \max_{\gamma} -\theta; \theta \geq -\frac{1}{2} \alpha' (\bar{K} \bar{y} \bar{y} + \frac{1}{c} I) \alpha, \forall y \epsilon \gamma \right\} \quad (8)
\]

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

\[
L = -\theta + \sum_{i \epsilon \gamma} \gamma_i \left( \theta - \frac{1}{2} \alpha' (\bar{K} \bar{y} \bar{y} + \frac{1}{c} I) \alpha \right) \quad (9)
\]

Setting the derivative of Lagrangian (9) with respect to \( \theta \) to zero, we have \( \sum_{i \epsilon \gamma} \gamma_i = 1 \). Denote \( d \) as a vector of \( d_i \) values and \( \mathcal{M} = \{d | d \geq 0, d 1 = 1\} \) as the domain of \( d \). We can then arrive at its dual form as follows:

\[
\min_{d \epsilon \mathcal{M}} -\frac{1}{2} \alpha' (\bar{K} \bar{y} \bar{y} + \frac{1}{c} I) \alpha \quad (10)
\]

Replacing the inner maximization subproblem in (8) with its dual (10), we have the following optimization problem:
Web Based Image Retrieval Incorporating generalized Multi-Instance (GMI) Learning And

\[
\max_{\alpha \in \mathcal{A}} \min_{d \in \mathcal{D}} -\frac{1}{2} \alpha^t (\sum_{i \in y^t} d_i \bar{k}y^ty^t + \frac{1}{c} I)\alpha = \min_{\alpha \in \mathcal{A}} \max_{d \in \mathcal{D}} -\frac{1}{2} \alpha^t (\sum_{i \in y^t} d_i \bar{k}y^ty^t + \frac{1}{c} I)\alpha
\]  

(11)

The equality holds as the objective function is concave in \( \alpha \) and linear in \( d \), and thus, we can interchange the order \( \max \) and \( \min \) in (11). Observe that (11) is analogous to the multiple kernel learning (MKL) problem [22], except that the label-kernel matrix, which is a convex combination of the base label-kernel matrices \( \bar{k}y^ty^t \), 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 isyfinite and the MLKL problem (11) is a special case of MKL, there are \( O(2^n) \) candidates of the label vector \( y^t \), and thus, the number of base label-kernel matrices \( \bar{k}y^ty^t \) is exponential in size. Thus, it is not possible to directly apply recently proposed MKL techniques such as SimpleMKL [22] to our proposed GMI-SVM.

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

1: Initialize \( y_t = Y_t \) for \( e B_j \) as \( y^t \), and set \( \{ y^t \} \);
2: Compute MLKL to solve \( \alpha \) and \( d \) in (11) based on \( C \);
3: Use \( \alpha \) 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) are necessarily active at optimality, and only subset \( C \subset Y \) of these constraints can usually lead to a very good approximation of the original optimization problem. Therefore, we can apply the cutting-plane method [15] to handle this exponential number of constraints. Moreover, the same strategy has been also applied in the recently proposed infinite kernel learning (IKL) [9], [10], in which the kernel is learned from an infinite set of general kernel parameters, and thus, MLKL (with kernel \( \sum_{i \in y^t} d_i \bar{k}y^ty^t \)) can be deemed as a variant of IKL. As a result, our GMI-SVM enjoys the same convergence of IKL [9]. The whole algorithm is summarized in Algorithm 1. First, we set subset \( C = \{ y^t \} \), where the instance label vector \( y^t \) is initialized according to the bag labels. Since \( C \) is no longer exponential in size, one can apply MKL to learn the kernel that obtains both \( r \) and \( d \). With a fixed \( d \), the label vector \( y^t \) with quadratic inequality constraint in (8), which is the most violated one by the current solution, is then added to \( C \). The process is repeated until the convergence criterion (i.e., the relative change of the objective values of (11) between two successive iterations less than 0.01) is met. After solving the MLKL problem, the decision function can be obtained by

\[ f(x) = \sum_{i=1}^{p} \alpha_i \bar{y}_i \bar{k}(x, x_i) \]

where \( \bar{y}_i = \sum_{i \in y^t} \bar{d}_i y^t \) and \( \bar{k}(x, x_i) = k(x, x_i) + 1 \)

**Algorithm 2:** Finding the approximation of the most violate \( y^t \).

1: Initialize \( y_t = Y_t \) for all \( x_i \) in positive bags \( B_p \) and \( y_t = -1 \) for all \( x_i \) in negative bags \( B_{neg} \);
2: for each positive bag \( B_p \) do
3: Fix the labeling of instances in all the other bags, and find the optimal \( y_p \) that maximizes the objective of (12) by enumerating the candidates of \( y_p \) in \( B_p \);
4: end for
5: for each negative bag \( B_{neg} \) do
6: Fix the labeling of instances in all the other bags, and find the optimal \( y_{neg} \) that maximizes the objective of (12) by enumerating the candidates of \( y_{neg} \);
7: end for
8: Repeat lines 2−7 until convergence.

3) Finding the Approximation of the Most Violated \( y^t \): Similar to IKL, finding the most violated constraint (indexed by \( y^t \)) in MLKL is problem specific and is the most challenging part in cutting-plane algorithms. Here, we discuss how to search for the most violated constraint to satisfy the GMI constraints in (3). Referring to (8), to find the most violated \( y^t \), we have to solve the following problem:

\[
\max_{y^t} \alpha (\bar{k}y^t y^t) \alpha (12)
\]

Note that finding the most violated \( y^t \) that maximizes (12) is computationally expensive problem when the bag size is large.
To accelerate our framework, we propose to use the instanceranking score defined in (1) to enforce the total number of instances in each positive bag to be 15. Moreover, we can beforehand exclude a large number of candidates of $y_i$ by checking our proposed GMI constraint (3). In order to further speed up the process, we develop a simple but effective method. The basic idea is to enumerate the candidates of $y_i$ satisfying (3) for each bag $B_i$ by fixing the labeling of other bags. Then, we iteratively choose the best $y_i$ for $B_i$, which maximizes (12), where $\gamma$ is the vector of instance labels in $B_i$. The procedure will be terminated when the relative change of the objective values of (12) between two successive iterations is less than 0.001. The detailed procedure is listed in Algorithm 2.

### III. Experimental Setup

In our experiments, for any given textual query (e.g., "fox"); the relevant Web 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-based reranking 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 learning algorithms can be readily adopted in our reranking framework. We only employ mi-SVM [1] and single-instance learning SVM (SIL-SVM)[2] in this paper as they are more suitable for predicting the labels of bags rather than instances.

The assumption in our newly proposed GMI-SVM is that positive instances comprise at least a certain portion of a positive bag, while a negative bag may contain at most a few positive instances. 81 images are employed as the database images, and all the 8 concept names are used as textual queries to perform the TBIR. Our framework can achieve reasonable efficiency by using unoptimized MATLAB 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 three types of visual features into lengthy feature vectors and normalize each feature dimension to zero mean and unit standard deviation. To improve the speed and reduce the memory cost, principal component analysis is then applied for dimension reduction.

For the $ith$ image, we further concatenate the visual feature $v_i$ and the textual feature $y_i$ together to form the lengthy feature vector $x_i$, namely $x_i = [\lambda v_i, t_i]^T$, where the weight parameter $\lambda$ is empirically fixed as 0.1 in the experiments. The database images are grouped into $n_B$ bags by using the $k$-means clustering method with the distance metric defined as follows:

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

where $v_i, t_i, v_j, t_j$ are the visual and textual features of the $ith$ and $jth$ images, respectively. We observe that it is computationally expensive to exploit the enumeration method for GMI-SVM if the number of instances in one bag is larger than 9. We therefore empirically set $K = \lceil T/9 \rceil$ in the $k$-means clustering method, where $T$ is the total number of relevant 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 score for each bag, and the remaining instances are discarded. The bags 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 positive bags, and we also randomly sample $9n_B$ irrelevant images to reconstruct $n_B$ negative bags. Then, positive and negative bags are then used as the training data for GMI-SVM, mi-SVM, and SIL-SVM. For GMI-SVM, we set proportion for positive bags and proportion for negative bags to fairly compare our GMI-SVM and the other MI learning methods mi-SVM and SIL-SVM.

### IV. Results of Retrieval Performances

GMI-SVM based on the convex relaxation in [18] can obtain a better optimal solution than other MI learning algorithms for the bag-based reranking framework. The top-ten-retrieved images of GMI-SVM, SI-L-SVM, mi-SVM, WEBSEIC, IBRR, and Init_Ranking for the textual query “fox” are illustrated in Fig. 3. Again, we observe that GMI-SVM achieves the best performance. From Table I, we also observe that, for a fixed $\mu$, GMI-SVM using $\gamma = 0$ generally achieves better performances compared with the results when setting $\gamma = 0.3$ and 0.5, which is consistent with our observation that the negative bags generally do not contain positive instances. This observation demonstrates that, for those concepts having more positive instances in the negative bags, GMI-SVM can successfully cope with the ambiguities in the instances in the negative bags and thus improve the retrieval performance. Considering that the MAP of GMI-SVM is the best when setting $\gamma = 0$ and $\mu = 0.5$, we fix $\gamma = 0$ and $\mu = 0.5$. We report the average central processing unit (CPU) time of the TBIR for different
Web Based Image Retrieval Incorporating Generalized Multi-Instance (GMI) Learning And Methods. For GMI-SVM, SIL-SVM, andmi-SVM, we still use one positive bag and one negative bag obtained by using the automatic weak bag annotation process.

![Image]

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

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\mu = 0.3$</th>
<th>$\mu = 0.5$</th>
<th>$\mu = 0.7$</th>
<th>$\mu = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0$</td>
<td>65.6%</td>
<td><strong>66.8%</strong></td>
<td>66.3%</td>
<td>66.2%</td>
</tr>
<tr>
<td>$\gamma = 0.3$</td>
<td>62.2%</td>
<td>66.7%</td>
<td>66.4%</td>
<td>65.9%</td>
</tr>
<tr>
<td>$\gamma = 0.5$</td>
<td>60.0%</td>
<td>66.4%</td>
<td>65.3%</td>
<td>65.7%</td>
</tr>
</tbody>
</table>

**TABLE I**

Maps over 81 concepts of GMI-SVM using different positive proportions (i.e., $\mu$ and $\gamma$) for positive and negative bags. Each result in the table is the best among the results obtained by using different numbers of positive and negative training bags.

V. CPU Time for Image Retrieval and Convergence Analysis

All the experiments are implemented with unoptimized MATLAB codes and performed on a workstation (3.33-GHz CPU with 32-GB random access memory). The average CPU-time overall textual queries are shown in Table II. Our proposed method GMI-SVM achieves reasonable efficiency for TBIR using unoptimized MATLAB codes. For GMI-SVM, on the average, the iterative optimization algorithm takes about six iterations to converge for each concept. In Fig. 4, we take three concepts (i.e., “bus,” “flower,” and “horse”) as examples to illustrate the convergence of GMI-SVM, in which the vertical axis indicates the objective value of (11) and the horizontal axis gives the number of iterations. We have similar observations for other concepts.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a bag-based framework for large-scale TBIR. Given a textual query, relevant images are to be reranked after the initial text-based search. Instead of directly reranking the relevant images by using traditional image ranking methods, we have partitioned the relevant images into clusters. By treating each cluster as a “bag” and the images in a bag as “instances,” we have formulated this problem as a MI learning problem. To address the ambiguities on the instance labels in both positive and negative bags, we have developed GMI-SVM to further enhance retrieval performance, in which the labels from the bag level have been propagated to the instance level. 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. Our framework using the automatic bag annotation method can achieve the best...
Web Based Image Retrieval Incorporating generalized Multi-Instance (GMI) Learning And

<table>
<thead>
<tr>
<th></th>
<th>GMI-SVM</th>
<th>SIL-SVM</th>
<th>mi-SVM</th>
<th>WEBSEIC</th>
<th>IBRR</th>
<th>Init_Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>1.120</td>
<td>0.025</td>
<td>0.026</td>
<td>0.027</td>
<td>105.354</td>
<td>0.0005</td>
</tr>
</tbody>
</table>

**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.”](image)

performance, as compared with other traditional image reranking methods on the NUS-WIDE data 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:**


AUTOR BIOGRAPHIES:

R. Sowmya received her B.Tech. degree in Electronics and Communication Engineering from Sri Kottam Tulasi Reddy Memorial College, Kurnool India, in 2009. She is pursuing her M.Tech in Digital Electronics and Communication Systems in Sri Kottam Tulasi Reddy Memorial College, Kondair India. Her area of interest is in the field of Content Based Image Retrieval (CBIR).

Mr. Sharad Kulkarni received his B.Tech. degree in Electronics and Communication Engineering from P.D.A College Engineering, Gulbarga India, 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 in Sri Kottam Tulasi Reddy Memorial College, Kondair India. His research interests include Image Processing and Content Based Image Retrieval (CBIR).