Abstract
This paper deals with the semantic-based web mining for image retrieval by means of enhanced Support Vector Machine (SVM). Generally, conventional Content-Based Image Retrieval (CBIR) systems are unsuccessful to satisfy users’ requirement because of the 'semantic gap' among the derived features and the user’s query. A large amount of existing approaches shows certain predetermined semantic category and allocate the images to suitable categories through certain learning processes. In contrast, these approaches constantly require human involvement and depend on content-based features. Here, semantic-based web mining for image retrieval by means of enhanced Support Vector Machine (SVM) is introduced. The outcome of this text mining process includes two maps which expose the semantic associations among images and keywords, respectively. These maps are employed to carry out image retrieval processes. The experimental results reveal the effectiveness of the enhanced SVM and high retrieval accuracy and relevance of retrieved images.

Keywords: Semantic-Based Image Retrieval, Text Mining, Self-Organizing Map, SVM, PCA.

Introduction
Due to the arrival of Internet and its vast applications, huge numbers of images are now available freely in online [1]. The quick development in the amount of such images can formulate the process of finding and accessing image of interest more difficult, since there are huge numbers of users. As a result, some supplementary processing is desired with the aim of making these collections searchable in a constructive way. The existing web image retrieval search engines, together with Google image search, Lycos and AltaVista photo finder, make use of text for the purpose of searching images, without taking the image content into account [2, 4]. Also, there is dissimilarity among what image features can differentiate and what people recognizes from the image [3]. CBIR systems make use of 'contents' for the purpose of retrieving images, Semantic-Based Image Retrieval (SBIR) systems makes an attempt to find out the factual semantic implication of an image and utilize it to retrieve appropriate images. On the other hand, understanding and discovering the semantics of a portion of information are high-level cognitive processes, and as a result complicated to automate.

Numerous efforts have been made to deal with this complication. Here, semantic-based web mining for the purpose of effective image retrieval by means of Enhanced Support Vector Machine is applied. Initially explicit knowledge is given into the framework which is integrated by representing the images with their surrounding texts in the web pages. In addition, these representations solve the complicatedness of semantic demonstration. The process of semantic extraction is accomplished with the help of a text mining process on these texts. The semantic relevance measure is intended for the purpose of matching the user’s query and images in the complete collection. This idea approaches from the recognition that it is extremely complicated to directly mine semantics from images. As a result direct access is not provided for the image contents, which is commonly time-consuming and inaccurate. As an alternative, image semantics is acquired from its environmental texts, which are usually contextually appropriate to the images. At last, the image retrieval for the keyword based query is acquired with the help of Enhanced SVM classifier. The rest of the paper is organised as follows: Section 2 describes the text mining process and its application on semantic image retrieval. Section 3 discusses the experiments and results. Finally Section 4 concludes the paper.

Image Mining for Semantic Image Retrieval
This section clearly gives the overview of the enhanced SVM based image retrieval method. The functional constituents and the processing flow are portrayed in Fig. 1. During the pre-processing and feature extraction phases, appropriate web pages are collected and converted into feature vectors, then subsequently which are indexed and accumulated in a database. The image feature vector of every image has been gathered by means of Principal Component Analysis (PCA). Subsequently the text mining process is executed on this set of feature.
Feature Extraction using Principal Components Analysis (PCA)

As the collection of high-dimensional vectors into a two-dimensional image feature space is performed as explained in feature extraction. The eigenvectors introduce a linear transformation from the original image feature space to a new space in which attributes are uncorrelated. The best n eigenvectors those one with largest eigenvalues (m < n) are selected as new features, while the rest are discarded. Mean is determined using the following equation:

\[ \mu = \frac{1}{N} \sum_{i=1}^{N} y_t - \]  

(1)

The covariance matrix of \( y_t \) is given as,

\[ c = \frac{1}{N} \sum_{i=1}^{N} (y_t - \mu)(y_t - \mu)^T \]  

(2)

\[ cv_t = \lambda_i v_t \]  

(3)

where \( \lambda_i \) (i = 1, 2, ..., n) represent the eigenvalues and \( v_i \) (i = 1, 2, ..., n) indicate the equivalent eigenvectors. In order to characterize document records with low-dimensional vectors, it is essential to compute the m eigenvectors in proportion to those \( m \) largest eigenvalues (m < n). Consider,

\[ \phi = [v_1, \ldots, v_m], \Lambda = \text{diag} \{\lambda_1, \ldots, \lambda_m\}, c\phi = \phi \Lambda - \]  

(4)

Extracting environmental texts:

The remaining segment of the web page is further processed in order to extract essential texts. These texts are known as Environmental Texts (ETs). For instance, HTML-format web page close tags, captions, alternate texts and filenames are recognized. Here, two categories of ETs are obtained, namely ET_Caption and ET_Normal. The ET_Caption contains texts that are synchronized by HTML tags, like image URL and ALT text. Conversely, ET_Normal contains normal texts which are placed outside any HTML tags. ET_Caption is simple to be extracted by means of the <img> tags. If both the ET_Normal and ET_Caption are within each image have been obtained, then a word extractor is employed for the purpose of segmenting these ETs into a list of terms. Subsequently, the resulting list of terms is employed to characterize the image. These terms are known as Environmental Keywords (EKs) of their related image. Fig. 2 demonstrates the ETs and EKs of an image obtained from an example web page.

<table>
<thead>
<tr>
<th>Image</th>
<th>Environmental text</th>
<th>Environmental keyword</th>
</tr>
</thead>
<tbody>
<tr>
<td>ET-Normal: “The winner of the Fleurieu Art Prize 2013, Fiona Lowry, Alone With You”</td>
<td>Fleurieu, Art, Prize, Australian, Landscape, Fiona, Lowry</td>
<td>Alone</td>
</tr>
<tr>
<td>ET-Caption: “./files/fionalowry0847_600 max.jpg”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: The ETs and EKs of the pointed out image in the web page with URL: http://www.artprize.com.au/finalists-2013.html. ET_Normals only consists of the paragraph where the image comes.

Discovering Image Semantics by Text Mining

Document clustering using self-organizing maps:

This subsection deals with describing the organization of images and EKs into clusters by their similarities due to co-occurrence. First, the encoding of images in the collection is performed as explained in feature extraction. Additionally, similar clusters should be ‘near’ in some way or the other. The unsupervised learning algorithm of SOM networks [5] is employed. The SOM algorithm arranges a collection of high-dimensional vectors into a two-dimensional map of neurons appropriate to the similarities between the vectors. Similar vectors, i.e. vectors spaced by small distances will get their mapping to the same or neighboring neurons after the training process. For the determination of the cluster in which an image or an EK is present, a labeling process is employed for the images and the EKs, correspondingly. Once the labeling process is over, an association is made between each image and a neuron in the map. The recording of the association yields the image cluster map. In the same way, each EK is labeled for the map and the keyword cluster map is created. Then these two maps are utilized for image semantics...
discovery. Some denotations are defined and described in the training process. Let \( V_{i}=\{V_{i}(j,n)\in[0,1] \mid 1\leq n\leq N, 1\leq j\leq M \} \), denote the encoded vector of the ith image within the corpus, in which N and M are the corresponding overall amount of EKs and images within the corpus. These vectors are taken as the inputs for training to the SOM network. The network comprises of a uniform grid of neurons which each includes N synapses. Consider \( X_{j}=\{X_{j}(j,n)\mid 1\leq n\leq N, 1\leq j\leq J \} \), represent the synaptic weight vector of the jth neuron within the network, where J indicates the quantity of neurons inside the network. The closeness is sometimes specified based on the geometrical distance between every pair of neurons. Training process halts after time t, which is large enough such that the selection of each vector. Both the training gain and neighbourhood size decrease with the increase in t.

**Mining Image and Keyword Associations**

When the image clustering process is completed, a labeling process is conducted on the trained network for establishing the association between every image and one neuron. The labeling process is explained as below. Every image’s feature vector \( V_{i}, 1 \leq i \leq M \) is compared with every neuron’s weight vector in the map. Once the labeling process is over, every image is labeled to a neuron. The results from labeling are recorded and the ICM is obtained. In the ICM, every neuron is labeled with a collection of images which are regarded to be similar and lie in the same cluster. This means that the SOM algorithm must do the clustering of images. After this the KCM is obtained by labeling each of the neuron in the trained network with specific EKs is developed. The labeling method may not entirely label each of the EK in the corpus. These EKs are called as unlabeled keywords. Unlabeled keywords happen in two situations. One condition is when many neurons contend for a particular EK at some stage in the training process. The consequence of the competition is frequent imperfect convergence of weights, as a result of which few EKs possibly will not be learned adequately, to be precise; their associated components may not have values close to 1 in any of the neuron’s weight vector. The other situation is when an EK is not standard in that neuron’s labeled images. This means, the nth EK is labeled to the jth neuron if, 

\[
X_{jn} = min_{1\leq ks\neq j} |label_k| \neq 0. \tag{5}
\]

Where label\( _j \) is the set of images that are labeled to neuron j. For instance, an EK labeled to a neuron j when its overall related component value over a set of adjacent neurons gets over the threshold. In this work, the 0-order process is applied.

**Image Retrieval by Semantics using Enhanced SVM**

After obtaining the image clusters and keyword clusters, use them for semantic image retrieval. The semantic retrieval by keywords using enhanced SVM classifier discussed given below.

**Enhance Support Vector Machine (SVM)**

Image retrieval is based on SVM Classification. The classification is performed by SVM is by generating a hyper plane. The optimal hyper plane separates the image in two categories. Generally support vector machine can be defined as the training method if polynomial radial basis function where the weights are calculated by solving QP problem. The method of selecting the appropriate representation is called feature selection. The followings will describe the overall principle of enhance SVM in linear separable case.

A set of linear separable training samples \( (x_{i}, y_{i})_{1\leq i\leq m} \), \( x_{i} \in R^d, y_{i} \in (-1,1) \) is called as the class label. The common form of linear classification function is \( g(x) = w \cdot x + b \), corresponding to a separating hyper plane \( w \cdot x + b = 0 \). Then normalize \( g(x) \) to meet \( |g(x)| \geq 1 \) for all \( x_i \), such that the distance from the nearest point to the hyperplane is \( 1/|w| \).

Between the isolating hyper planes, the one for which the space to the nearest point is maximum is referred to as Optimal Separating Hyper plane (OSH). As the distance to the point closest is \( 1/||w|| \), finding the OSH relates to the reducing \( ||w|| \) and the objective function is:

\[
min \phi(w) = \frac{1}{2}||w||^2. \tag{6}
\]

The solution w has an expansion \( w = \sum a_{ij}y_i x_i \) in terms of a subset of training patterns, known as support vectors, which settle on the margin. Classification function can as a result be given as

\[
f(x) = sign(\sum a_{ij}y_i (x_i \cdot x) + b) \tag{7}
\]

When the data is not separable linearly, one way, SVM incorporates slack variables and a penalty factor so that the modification of the objective function can be given as

\[
\phi(w) = \frac{1}{2}||w||^2 + c(\sum \xi_i) \tag{8}
\]

In another way, input data is mapped by means of some nonlinear mapping into a high-dimensional feature space, in which the optimal separating hyper plane is built. Thus the dot product can be denoted by \( k(x,y) := \phi(x) \cdot \phi(y) \) when the kernel k meets Mercer’s condition \[6\] function is obtained as

\[
f(x) = sign(\sum a_{ij}y_i k(x_i,x) + b) \tag{9}
\]

Conducting the sorting images based on their distance to the hyper plane, helps in obtaining a better result in classification.

**Semantic image retrieval by keywords:**

When a user presents keywords as a query, the images which have semantic relevance to this query can be got back by utilizing the KCM. For this purpose, the query is translated to a query vector in the similar way as the image vector. Here let \( q = \{q_i \in \{0,1\} \mid 1 \leq i \leq N \} \) denote the query vector. And also perform the transformation of every keyword cluster in KCM to a vector as given below: let \( k_j = \{k_{jn} \in \{0,1\} \mid 1 \leq i \leq N \} \) be the encoded vector for the keyword cluster corresponding to neuron j, where \( k_{jn} = 1 \) when the ith EK is labeled to neuron j. The similarity among the query vector \( q \) and an image vector \( x \) is computed with an extended cosine measurement in the vector space model:

\[
S_{q,x} = A \frac{|q \cdot k_j|}{||q|| ||k_j||} + \frac{|q \cdot x|}{||q|| ||x||} \tag{10}
\]

Let \( x \) denote the encoded vector of some image associated with neuron j. The primary term of the Right Hand Side (RHS) in Eq. (10) determines the similarity among the query vector and the cluster vector. Second term gives the measurement of the similarity among the query vector and the given image vector that is connected with neuron j. A indicates a scaling parameter that is sufficiently big to distinguish the contributions from cluster and individual image. Also allow this neuron to be labeled by means of images 1, 7, 20 and 34. The primary term of the RHS in Eq. (10) possibly will find that neuron 20 is the
nearest to the query. Because a well-trained map should have a nearly even distribution of images on the neurons, an acceptable value of $A$ is $\frac{B-M}{J}$ where $B$ indicates a positive number that is bigger than 1, $M$ represents the number of images in the corpus, and $J$ indicates the number of neurons in the SOM. Through the parameter $A = 0$ in Eq. (10), they influence the computation of the similarity among the query vector and each image vector. The similarity between two images is computed as follows:

$$S'_{q,x} = \mathcal{B}(\|G(q) - G(x)\| + \frac{\|q.x\|}{\|q\|\|x\|}) - (11)$$

where function $G: \mathbb{R}^N \rightarrow \mathbb{R}^2$ provides the two-dimensional grid coordinates of its argument in the ICM. $G$ argument indicates an encoded image vector or the query vector.

**Semantic image retrieval by example web page images:**

If an image is taken as a query, as a first step its ETs from its associated web page are found. In order to get rid of the requirement of presenting a complete page as query, which is diverse from the normal practice, subsequently connect each image with a representative web page or a collection of web pages and extract ETs from this set of web pages. To speed up the matching process, the similarity computation divided into two stages. In the first stage, only the first terms in Eqs. (10) and (11) are computed. In the second stage, the images in accordance with the second terms are ordered. This process takes benefit of the fact that the number of neurons, $J$, is normally much smaller than the number of images, $M$. The second phase ten or twenty more computations of the second term. In view of the fact that the user is normally interested only in the top-rank images, this two-stage scheme could get satisfactory results.

**Experimental Results**

In this section, the explored method is tested with a set of web pages that were collected by human effort in accordance with the Yahoo! Or Google web site directory. The Yahoo! directory hierarchy was best due to the fact that it has been a universal test bed for categorization and semantics development of web pages. All categories are categorized in a hierarchical approach in the Yahoo! hierarchy, through 14 top-level categories in the Yahoo! directory. In this technical work, the “Arts and Humanities” category is utilized as source of web pages. The WebPages are considered as table 1 and table 2.

<table>
<thead>
<tr>
<th>Table 1: The list of level-1 categories in ‘Art’ hierarchy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art history</td>
</tr>
<tr>
<td>Art Weblogs@</td>
</tr>
<tr>
<td>Awards</td>
</tr>
<tr>
<td>Booksellers@</td>
</tr>
<tr>
<td>Censorship</td>
</tr>
<tr>
<td>Crafts</td>
</tr>
</tbody>
</table>

To train the image vectors a self-organizing map is constructed which consists of 900 neurons in a $30 \times 30$ grid form. The preliminary gain is predetermined to 0.4 and the maximum training time is predetermined to 500 in the training process. A portion of the labelling results is done by using the following links, http://www.webcorps.com, http://www.wildnatureimage.com, http://www.ovis.net, http://www.northlight-images.co.uk

**Figure 3: Precision Recall Curve**

The recall is the percentage of the correct documents which are retrieved over all the correct documents that correspond to a query, i.e. recall $= \frac{|Ra|}{R}$, where $Ra$ indicates the set of correct documents being retrieved and $R$ is the collection of correct documents for a query $q$. The precision is defined as the percentage of correct documents in the documents that are retrieved in response to $q$, i.e. precision $= \frac{Ra}{A}$ where $A$ stands for the set of documents retrieved for $q$. All images are utilized as queries and the average precision vs recall curve is illustrated as shown in Fig. 3. First the precision versus recall curve approach [6] is explained.

The image retrieval using SVM approach produces better accuracy rate shown in Fig.4 which is much greater accuracy results than existing image retrieval methods such as ontology based image retrieval and picSOM based image retrieval methods. When the number of images increases, the accuracy of the result is also increases. This approach produces high accuracy rate when compared to existing system. Fig.6 shows the number of selected images for different feature vector dimensions and classification of images retrieval (IR) is done using methods such as SVM based IR, ontology based IR, and picSOM based IR. The best accuracy rate is achieved using the linear and non-linear feature vector.
The used image retrieval using SVM produces high F-measure shown in Fig.6 is higher than the existing image retrieval methods. When the number of images increases the F-measure of the result is increases. This approach produces effective F-measure rate when compared to existing systems.

**Figure 4: Accuracy Comparison**

**Figure 5: F-Measure Comparison**

**Figure 6: Mean Absolute Error Comparison**

**Conclusion**

This paper illustrates a semantic-based web mining for image retrieval using enhanced support vector machine approach is suggested to discover semantically related images. This approach depends on the accurate segmentation of the environmental texts. The ETs of each image are further clustered in accordance with their semantic similarities through the SOM algorithm, which is a fraction of the text mining procedure. The image retrieval is done by the enhanced SVM classification and it achieves high accurate result. However, collections of skeleton images like medical images or photo albums would not suit for this method except when some form of annotations were applied to these images either manually or automatically. In fact, automatic annotation of images is a kind of text mining process application, and it will be incorporated into this work for getting over this skeleton image problem in the future.

**References**


Authors Details:

Mrs. P. Sumathi M.Sc., M.Phil., working as an Assistant Professor with eleven years of experience in teaching. She has published four papers in National Conference, one paper in International Conference, four papers in International Journals. She also has participated in seminars, conference, workshops and refresher courses at National Level. Now she is pursuing Ph.D Computer Science in Dr. Mahalingam Center for Research and Development at NGM College, Pollachi.

Dr. Manickachezian Rajakannu, Associate Professor of Computer Science born in the year 1965 in Tamil Nadu, India, is born with the thirst to acquire knowledge and share it to the world at large. He had his Master of Science in Applied Science at PSG College of Technology, Coimbatore and his thirst to update himself and sharpen his skills more made him to pursue M.S. in Software Systems at Birla Institute of Technology and Science, Pilani, Rajasthan and also his doctorate in Computer Science at School of Computer Science and Engineering, Bharathiar University, Coimbatore, India.

His motto being, knowledge earned is to be dissipated, he took up Teaching as his vocation. Many a Rural student has been benefited and risen to higher positions in his rich teaching experience for nearly 3 decades (2 years at PSG College of Technology, Coimbatore and 27 years at NGM College, Pollachi, Coimbatore). His rich experience for over a period of 18 years in Research and his constant updation and expertise has been beneficial for UG / PG students, M.Phil and Ph.D scholars. To his credit, he has completed 300 UG/PG Projects, 22 M.Phil, 4 Ph. D’s and 8 Ph.D Scholars are currently pursuing under his guidance and genius. His scholarliness and research spirit enabled him to receive the “Best Research Supervisor Award”. He is also a recipient of many awards like Desha Mithra Award, Best Paper Award twice in International Conferences and Life Time Achievement Award for his meritorious 30 years service in the field of Computer Science. He received National Talented Students’ Scholarship during his schooling.

He has to his credit 85 publications in International Journals and 55 publications in International Conferences / Seminars. With his learned expertise and experience he has served as a Board of Studies Member / Board of Examiners in various Universities and Colleges. He has acted as a Program Committee Member / Editor of Conference Proceedings and Abstracts / Paper Reviewer more than 20 times in International conferences conducted in Australia, Indonesia, Malaysia and India. He is also an Editor of Journal of Applied Sciences Research published in Jordan. Currently, serves as an Associate Editor / Editorial Board Member in a couple of International Journals. He authored a book on the topic “Internet and Java Programming” for PGDCA students.

He has chaired in more than 10 National / International Seminars. Has acted as a Convenor / Member of Inspection Commissions to start new courses including Ph. D in numerous colleges under Bharathiar University, India. He served as a Subject Expert / Advisor in various Doctoral Committee Meetings and for recruiting Assistant Professors in various colleges in Tamil Nadu, India. Has also associated himself with the various Professional Bodies like ISCA, CSI, IAENG and CSTA. His research interests include but not limited to Data Mining, Network Databases, Mobile Networks, Bioinformatics, Cloud Computing and Big Data Analytics.