COMBINING CONTENT AND CONTEXT INFORMATION FOR SEMANTIC IMAGE ANALYSIS AND CLASSIFICATION

G. Th. Papadopoulos[1,2], *V. Mezaris*[1], *I. Kompatsiaris*[1] *and M. G. Strintzis*[1,2]

[1] Informatics and Telematics Institute
1st Km Thermi-Panorama Road
Thessaloniki, GR-57001 Greece
e-mail: {papad, bmezaris, ikom}@iti.gr

[2] Information Processing Laboratory
Electrical and Computer Engineering Department
Aristotle University of Thessaloniki
Thessaloniki, GR-54124, Greece
e-mail: strintzi@eng.auth.gr

ABSTRACT

In this paper, a learning approach to semantic image analysis and classification is proposed that combines global and local information, with explicitly defined knowledge in the form of an ontology. The ontology specifies the selected domain, its sub-domains, the concepts related to each sub-domain as well as contextual information. Support Vector Machines (SVMs) are employed in order to provide image classification to one of the defined sub-domains based on global image descriptions and, after a segmentation algorithm is applied, to perform an initial mapping between region low-level visual features and the concepts in the ontology. Then, a decision function, that receives as input the region to concepts associations together with contextual information, realizes image classification based on local-level information. The contextual information used is in the form of frequency of appearance of each concept in every particular sub-domain. A fusion mechanism combines the intermediate classification results, provided by the local- and global-level information processing, and decides on the final classification. A Genetic Algorithm (GA) is employed for optimizing the fusion process. Experiments with images from the personal collection domain demonstrate the performance of the proposed approach.

1. INTRODUCTION

During the recent years, important advances in the hardware technology have led to the development of devices for capturing high-quality images quickly and conveniently, devices with great storage capabilities and the appropriate mediums for fast and efficient distribution of the acquired content. As a consequence, literally vast multimedia databases have been created and the problem of effective and efficient manipulation of the available content has emerged [1]. This has triggered intense research efforts towards the development of appropriate systems and algorithms for solving this challenging issue. Most emerging approaches adopt the fundamental principle of shifting image manipulation techniques towards the process of the visual content at a semantic level, thus attempting to bridge the so called *semantic gap* [2]. To this end, research efforts have concentrated on the semantic analysis and classification of images, often adopting techniques that exploit *a priori* domain specific knowledge [3], so as to result in a high-level representation of them [1].

Image classification is an important component of image manipulation attempts. Several approaches have been proposed in the relevant literature regarding the task of the categorization of images in a number of predefined classes based on global image descriptions [4][5]. However, image manipulation based solely on global low-level features does not always lead to the best results. Coming one step closer to treating images similarly as human does, image analysis tasks (including classification) shifted to treating images at a finer level of granularity, i.e. at the region or local level, taking advantage of segmentation techniques applied to the image [6][7]. Furthermore, incorporating knowledge into classification techniques emerges as a promising approach for improving classification efficiency. Such an approach provides a coherent semantic domain model to support "visual" inference in the specified context [8].

In this paper, a semantic image analysis and classification approach is proposed that combines global and local information with explicitly defined knowledge in the form of an ontology. The ontology specifies the selected domain, its sub-domains, the concepts related to each sub-domain as well as contextual information. SVMs are employed for image classification in one of the defined sub-domains based on global image descriptions, by generating an image to sub-domain association hypothesis set. Additionally, after a segmentation algorithm is applied and the image is divided into regions, SVMs are again employed, this time for performing an initial mapping between region low-level visual features and the concepts in the ontology (i.e. generating region to concept association hypothesis set for every region). Then, a *decision function*, that receives as input the regions to concepts association hypothesis sets together with contextual information, realizes image classification based on local-level information. Contextual information consists of the frequency of appearance of each concept in every particular sub-domain. A fusion mechanism combines the global and local features-based classification information and decides on the final image classification. A GA is introduced for optimizing the performed information fusion step. Finally, a refined region to concept association is performed, while considering only the concepts that are related to the sub-domain to which the image is classified to.

The paper is organized as follows: Section 2 presents the overall system architecture. Sections 3 and 4 describe the low-level information extraction and the employed high-level knowledge, respectively. Sections 5 and 6 detail the individual system components. Experimental results are presented in Section 7 and conclusions are drawn in Section 8.

2. SYSTEM OVERVIEW

The first step in the development of the proposed semantic image analysis and classification architecture is the definition of an appropriate knowledge infrastructure. This is defined in the form of an ontology suitable for describing the semantics of the selected domain. The proposed ontology comprises of a set of sub-domains to which images of the domain can be classified to and a set of concepts, each associated with at least one of the afore-mentioned sub-domains, which represent objects of interest that may be depicted in the images. In addition to the above, the ontology also includes contextual information in the form of the frequency of appearance of each concept in the images of each sub-domain.

Then, a set of images, $\mathcal{Q}$, belonging to the aforementioned domain was assembled. Each image was manually annotated (i.e. manually generated image classification and -after segmentation is applied- region-concept associations) according to the ontology def-

The work presented in this paper was supported by the European Commission under contracts FP6-001765 aceMedia, FP6-027685 MESH and FP6-027026 K-Space, and by the GSRT under project DELTIO.

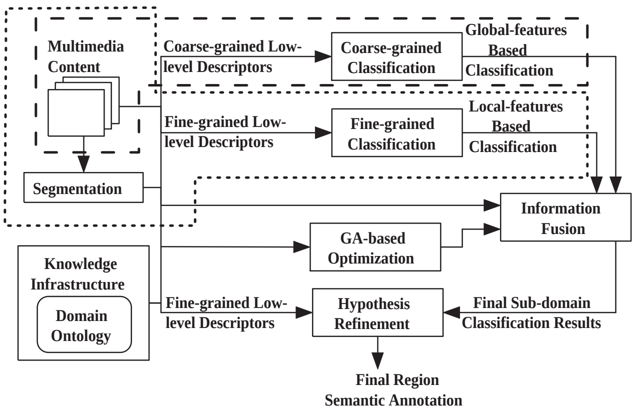

Figure 1: System architecture

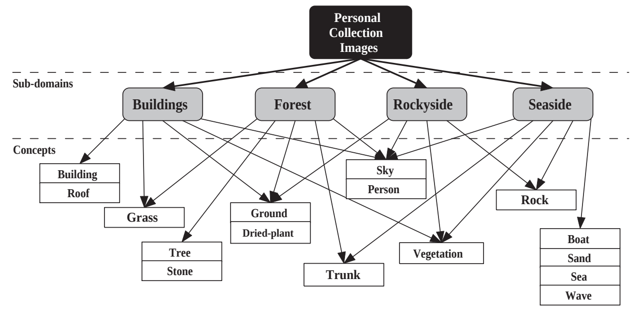

Figure 2: Sub-domains and concepts of the ontology developed for the personal collection domain

initions. This set was divided into two equal in terms of amount sub-sets, $\mathcal{Q}_{tr}$ and $\mathcal{Q}_{te}$. $\mathcal{Q}_{tr}$ was used for training purposes, while $\mathcal{Q}_{te}$ served as a test set for the evaluation of the proposed system performance, as will be described in the sequel.

At the signal level, low-level global image descriptors are extracted for every image and form an *image feature vector*. This is utilized for performing image classification based on global descriptions to one of the sub-domains, supplied as input to a set of SVMs, each trained to detect images that belong to a certain sub-domain. Every SVM returns a numerical value which denotes the degree of confidence to which the corresponding image is assigned to the sub-domain associated with the particular SVM; the maximum of the degrees of confidence over all sub-domains for an image indicates its classification based on global features.

In parallel to this process, a segmentation algorithm is applied to the image and low-level descriptions are estimated for every resulting segment. These are employed for generating initial hypotheses regarding the region's association to an ontology concept. This is realized by evaluating the low-level *region feature vector*, using a second set of SVMs, where each SVM is trained to identify instances of a single concept defined in the ontology. SVMs were selected for the aforementioned tasks due to their reported generalization ability [9]. The computed hypothesis sets are subsequently introduced to a *decision function* which realizes image classification based on local-level and the ontology provided contextual information.

Then, a fusion mechanism is introduced, which implements the fusion of the computed global- and local-features based classification information, in order to make a final image classification decision. A genetic algorithm is employed for optimizing the parameters of the fusion mechanism. The choice of a GA for this task is based on its extensive use in a wide variety of global optimization problems [10], where they have been shown to outperform other traditional methods.

Once the image sub-domain is selected, a second region-concept association hypothesis generation process is performed. The procedure is similar to the one described at the previous stage, the difference being that now only the SVMs that correspond to concepts of the selected sub-domain are employed. Thus, sub-domain-specific hypothesis sets are computed and final region-concepts association is accomplished. The overall architecture of the proposed system for semantic image analysis and classification is illustrated in Fig. 1.

3. LOW-LEVEL VISUAL INFORMATION PROCESSING

The global image classification procedure, as will be described in detail in the sequel, requires that appropriate low-level descriptions are extracted at the image level for every examined image and form an *image feature vector*. The image feature vector employed in this work comprises of three different descriptors of the MPEG-7 standard, namely the *Scalable Color*, *Homogeneous Texture* and *Edge Histogram* descriptors. Following their extraction, the image feature vector is produced by stacking all extracted descriptors in a single vector. This vector constitutes the input to the SVMs structure which realizes the image classification using global features, as described in Section 5.1.

Moreover, in order to perform the region-concept association procedure, the examined image has to be segmented into regions and suitable low-level descriptions have to be extracted for every resulting segment. In the current implementation, an extension of the Recursive Shortest Spanning Tree (RSST) algorithm has been used for segmenting the image [11]. Output of this segmentation algorithm is a segmentation mask S, $S = \{s_i\ ,\ i = 1,..., N\}$, where s_i, $i = 1,...N$ are the created spatial regions. For every generated image segment, the following MPEG-7 descriptors are extracted: *Scalable Color*, *Homogeneous Texture*, *Region Shape* and *Edge Histogram*. The above descriptors are then combined to form a single *region feature vector*. This vector constitutes the input to the SVMs structure which computes the region-concept association hypothesis sets for every segment, as described in Section 5.2.

4. KNOWLEDGE INFRASTRUCTURE

Among the possible domain knowledge representations, ontologies [12] present a number of advantages, the most important being that they provide a formal framework for supporting explicit, machine-processable semantics definition and they enable the derivation of new knowledge through automated inference. Thus, ontologies are suitable for expressing multimedia content semantics so that automatic semantic analysis and further processing of the extracted semantic descriptions is allowed [8]. Following these considerations, an ontology was developed for representing the knowledge components that need to be explicitly defined under the proposed approach. More specifically, the images of concern belong to the personal collection domain. Thus, in the developed ontology, a number of sub-domains, related to the broader domain of interest, are defined and denoted by D_l, $l = 1,...L$. For every sub-domain, the particular semantic concepts of interest are also defined and denoted by c_j, where $C = \{c_j,\ j = 1,...J\}$ being the set of all defined concepts.

Contextual information in the form of frequency of appearance of each concept in every particular sub-domain, is also included in the ontology and is acquired according to a certain ontology population procedure. For that purpose, a set of segmented images, $\mathcal{Q}_{tr}$, with ground truth classification and annotations, which serves as a training set, is assembled as described in Section 2. Then, the reported frequency of appearance of each concept c_j with respect to the sub-domain D_l, $freq(c_j, D_l)$, is defined as the percentage of the images belonging to sub-domain D_l where concept c_j appears. The computed values are stored in the developed domain ontology. The sub-domains and concepts of the ontology employed in this work are presented in Fig. 2.

5. IMAGE CLASSIFICATION AND REGION-CONCEPT ASSOCIATION

5.1 Image Classification Using Global Features

In order to perform the classification of the examined images to one of the sub-domains defined in the ontology using global image descriptions, an *image feature vector* is initially formed, as described in Section 3. Then, a SVM structure is utilized to compute the class to which every image belongs. This comprises L SVMs, one for each defined sub-domain D_l, each trained under the *'one-against-all'* approach. For the purpose of training the SVMs, the sub-domain membership of the images belonging to the training set, $\mathscr{Q}_{tr}$, assembled in Section 2 is employed. The aforementioned image feature vector constitutes the input to each SVM, which at the evaluation stage returns for each image of unknown sub-domain membership a numerical value in the range $[0,1]$ denoting the degree of confidence to which the corresponding image is assigned to the sub-domain associated with the particular SVM. The metric adopted is defined as follows: For every input feature vector the distance z_l from the corresponding SVM's separating hyperplane is initially calculated. This distance is positive in case of correct classification and negative otherwise. Then, a sigmoid function is employed to compute the respective degree of confidence [13], h_l^D, as follows:

$$h_l^D = \frac{1}{1+e^{-t \cdot z_l}} \quad , \tag{1}$$

where the slope parameter t is experimentally set. For each image, the maximum of the L calculated degrees of membership indicates its classification, whereas all degrees of confidence, h_l^D, constitute its sub-domain hypotheses set H^D, where $H^D = \{h_l^D,\ l = 1,...L\}$. The SVM structure employed for image classification based on global features, as well as for the region-concept association tasks described in the following sections, was realized using the SVM software libraries of [14].

5.2 Image Classification Using Local Features and Initial Region-Concept Association

As already described in Section 2, the SVM structure used in the previous section for global image classification is also utilized to compute an initial concept-region association for every image segment. Similarly to the global case, an individual SVM is introduced for every concept c_j of the employed ontology, to detect the corresponding association. For that purpose, a training process identical to the one performed for global image classification is followed. The differences are that now the region feature vector, as defined in Section 3, is utilized and that each SVM returns a numerical value in the range $[0,1]$ which in this case denotes the degree of confidence to which the corresponding segment is assigned to the concept associated with the particular SVM. The respective metric adopted for expressing this degree is defined as follows: Let $h_{ij}^C = I_M(g_{ij})$ denote the degree to which the visual descriptors extracted for segment s_i match the ones of concept c_j, where g_{ij} represents the particular assignment of c_j to s_i. Then, $I_M(g_{ij})$ is defined as

$$I_M(g_{ij}) = \frac{1}{1+e^{-t \cdot z_{ij}}} \quad , \tag{2}$$

where z_{ij} is the distance from the corresponding SVM's separating hyperplane for the input feature vector used for evaluating the g_{ij} assignment. The pairs of all supported concepts and their respective degree of confidence h_{ij}^C computed for segment s_i comprise the segment's concept hypotheses set H_i^C, where $H_i^C = \{h_{ij}^C,\ j = 1,...J\}$.

After the concept hypotheses sets, H_i^C, are generated for every image region s_i, a decision function is introduced for realizing image classification based on local features, i.e. estimating the sub-domain membership of the image on the basis of the concept hypotheses sets of its constituent regions and the contextual information:

$$g(D_l) = \sum_{s_i,\ where\ c_j \in D_l} I_M(g_{ij}) \cdot \mathscr{F}(s_i,\ c_j,\ a_l,\ D_l) \tag{3}$$

$$\mathscr{F}(s_i,\ c_j,\ a_l,\ D_l) = a_l \cdot freq(c_j, D_l) + (1-a_l) \cdot area(s_i) \tag{4}$$

where $freq(c_j, D_l)$ is the concept frequency defined in Section 4 and $area(s_i)$ is the percentage of the image area captured by region s_i. Parameters a_l are introduced for adjusting the importance of the aforementioned frequencies against the regions' areas for every defined sub-domain. Their values are estimated according to the procedure described in Section 6.

5.3 Information Fusion for Image Classification and Final Region-Concept Association

After image classification has been performed using global, h_l^D, and local, $g(D_l)$, information, a fusion mechanism is introduced for deciding upon the final image classification. This has the form of a weighted summation, based on the following equation:

$$G(D_l) = \mu_l \cdot g(D_l) + (1-\mu_l) \cdot h_l^D \tag{5}$$

where μ_l, $l = 1,...,L$ are sub-domain-specific normalization parameters, which adjust the magnitude of the global features against the local ones upon the final outcome and their values are estimated according to the procedure described in Section 6. The domain with the highest $G(D_l)$ value constitutes the final image classification. Since the final image classification decision is made, a refined region-concept association procedure is performed, where now only concepts associated with the estimated sub-domain can be detected. Thus, final region-concept association is accomplished.

6. OPTIMIZING INFORMATION FUSION

In Sections 5.2 and 5.3, variables a_l and μ_l are introduced for adjusting the importance of the frequency of appearance against the region's area and the global against the local information on the final image classification decision, respectively. A genetic algorithm is employed for estimating their values, as outlined in Section 2.

Initially, the image set $\mathscr{Q}_{tr}$, that was assembled as described in Section 2, is divided into two subsets, namely a sub-training $\mathscr{Q}_{tr}^2$ and a validation $\mathscr{Q}_v^2$ set. $\mathscr{Q}_{tr}^2$ is used for training the employed SVMs framework and $\mathscr{Q}_v^2$ for validating the overall system classification performance.

Subject to the problem of concern is to compute the values of parameters a_l and μ_l that lead to the highest correct image classification rate. For that purpose, *Classification Accuracy*, *CA*, is used as a quantitative performance measure and is defined as the fraction of the number of the correctly classified images to the total number of images to be classified.

Under the proposed approach, each chromosome C represents a possible solution, i.e. a candidate set of values for parameters a_l and μ_l. In the current implementation, the number of genes of each chromosome is predefined and set equal to $2 \cdot l \cdot 2 = 4 \cdot l$. The genes represent the decimal coded values of parameters a_l and μ_l assigned to the respective chromosome, according to the following equation:

$$C \equiv [\, c_1\ c_2\ ...c_{4 \cdot l}\,] = [\mu_1^1\ \mu_1^2 ... \mu_l^1\ \mu_l^2 a_1^1\ a_1^2 ... a_l^1\ a_l^2] \tag{6}$$

where $c_i\ \varepsilon\ \{0,1,...9\}$ represents the value of gene i and μ_r^t, a_r^t represent the t^{th} decimal digit of variable μ_r, a_r, respectively. The genetic algorithm is provided with an appropriate *fitness function*, which denotes the suitability of each solution. More specifically, the fitness function $f(C)$ is defined as equal to the CA metric already defined, $f(C) \equiv CA(C)$, where $CA(C)$ is calculated over all images that comprise the validation set $\mathscr{D}_v^2$, after applying the fusion mechanism (Section 5.3) with parameter values for a_l and μ_l denoted by the genes of chromosome C.

Regarding the GA's implementation details, an initial population of 100 randomly generated chromosomes is employed. New generations are iteratively produced until the optimal solution is reached. Each generation results from the current one through the application of the following operators:

- Selection: a pair of chromosomes from the current generation are selected to serve as parents for the next generation. In the proposed framework, the Tournament Selection Operator [10], with replacement, is used.
- Crossover: two selected chromosomes serve as parents for the computation of two new offsprings. Uniform crossover with probability of 0.2 is used.
- Mutation: every gene of the processed offspring chromosome is likely to be mutated with probability of 0.4.

To ensure that chromosomes with high fitness will contribute to the next generation, the overlapping populations approach was adopted. More specifically, assuming a population of m chromosomes, m_s chromosomes are selected according to the employed selection method, and by application of the crossover and mutation operators, m_s new chromosomes are produced. Upon the resulting $m + m_s$ chromosomes, the selection operator is applied once again in order to select the m chromosomes that will comprise the new generation. After experimentation, it was shown that choosing $m_s = 0.4m$ resulted in higher performance and faster convergence. The above iterative procedure continues until the diversity of the current generation is equal to/less than 0.001 or the number of generations exceeds 30. The final outcome of this optimization procedure are the optimal values of parameters a_l and μ_l, used in Eq. 3 and 5. The above GA-based optimization procedure was realized using the GA software libraries of [15].

7. EXPERIMENTAL RESULTS

In this section, experimental results of the application of the proposed approach to images belonging to the personal collection domain are presented. Initially, an appropriate ontology was developed for representing the knowledge components that need to be explicitly defined, as described in detail in Section 4. Then, a set of 800 randomly selected images, $\mathscr{D}$, belonging to the aforementioned domain was assembled, as described in detail in Section 2.

According to the SVMs training process (Section 3), a polynomial function was used as a kernel function by each SVM for both global image classification and region-concept association cases. The respective low-level *image feature vector* and *region feature vector* are composed of 398 and 433 values respectively, normalized in the interval $[-1,1]$.

Based on the trained SVMs structure, global image classification is performed as described in Section 5.1. Then, initial concept hypotheses are generated for each image segment and a decision function realizes image classification based on local features (Section 5.2). Afterwards, the approach described in Section 5.3 is employed for implementing the fusion of the global and the local classification information and computing the final image classification, where the values of its parameters are estimated according to a GA-based optimizer (Section 6). Using the computed image classification decision, the final region-concept association procedure is performed.

In Fig. 3 and 4 indicative classification results are presented, showing the input image, the image classification using only global (row 2) and only local (row 3) information and the final classification after the implementation of the fusion mechanism. In Table 1, quantitative performance measures of the image classification algorithms are given in terms of accuracy for each sub-domain and overall. Accuracy is defined as the percentage of the images, belonging

Input image		
Global Image Classification	Buildings :0.44 Rockyside:**0.58** Forest :0.56 Seaside :0.30	Buildings :0.21 Rockyside:0.34 Forest :**0.61** Seaside :0.20
Local (i.e. region-based) Image Classification	Buildings :**0.29** Rockyside:0.17 Forest :0.11 Seaside :0.08	Buildings :0.11 Rockyside:**0.27** Forest :0.22 Seaside :0.18
Final Image Classification Using Information Fusion	**Buildings**	**Forest**

Figure 3: Indicative image-sub-domain association results

to a particular sub-domain, that are correctly classified. The results presented in Table 1 show that the global classification method generally leads to better results than the local one. Additionally, it must be noted that the performance of both algorithms is sub-domain dependent, i.e. some sub-domains are more suitable for classification based on global features (e.g. *Rockyside* and *Forest*), whereas for other sub-domains the application of a region-based image classification approach is advantageous (e.g. *Buildings*). For example, in the *Rockyside* sub-domain the presented color distribution and texture characteristics are very similar among the corresponding images. Thus, image classification based on global features performs better than the local-level case. On the other hand, for sub-domains like *Buildings*, where the color distribution and the texture characteristics of the depicted real-world objects may vary significantly (i.e. buildings are likely to have many different colors and shapes), the image classification based on local-level information presents increased classification rate. Furthermore, it can be verified that the proposed global and local classification information fusion approach leads to a significant performance improvement. Regarding the exploitation of the contextual information in the form of frequency of concept appearance in every sub-domain, extensive experiments were conducted, while using different image sets. This has resulted into negligible variations of the local-features based as well as the overall image classification performance.

The respective representative examples and performance measures for the concept detection case along the sequential steps of the proposed approach are illustrated in Fig. 5 and Table 2, respectively. It can be shown that the overall as well as all the sub-domain specific concept detection accuracies are improved after the implementation of the proposed classification algorithm, compared to the performance which corresponds to the initial region-concept association. This increase in performance justifies the assumption that the reduction of the total number of concepts to be detected, after image classification is performed, leads to better concept detection results. It must be noted that for the numerical evaluation, any concept present in the examined image test set that was not included in the ontology definitions, e.g. umbrella in the seaside sub-domain, was not taken into account.

8. CONCLUSIONS

In this paper, an approach to semantic image analysis and classification that combines global, local and contextual information with explicitly defined knowledge in the form of an ontology was presented and produced encouraging results in a relatively broad domain. The proposed framework can easily be extended by including additional sub-domains and concepts, provided that the employed knowledge representation is appropriately extended and that the employed training set is enriched with suitable training samples.

Input image		
Global Image Classification	Buildings :0.12 Rockyside:0.23 Forest :0.42 Seaside :**0.84**	Buildings :0.14 Rockyside:**0.52** Forest :0.33 Seaside :0.09
Local (i.e. region-based) Image Classification	Buildings :0.14 Rockyside:0.22 Forest :0.31 Seaside :**0.79**	Buildings :0.23 Rockyside:0.36 Forest :**0.38** Seaside :0.14
Final Image Classification Using Information Fusion	**Seaside**	**Rockyside**

Figure 4: Indicative image-sub-domain association results

Table 1: Sub-domain detection accuracy

Accuracy	Method		
	Global Image Classification	Local (Region-based) Image Classification	Final Image Classification Using Information Fusion
Buildings	44.64%	76.92%	**85.72**%
Rockyside	**73.24**%	42.86%	72.73%
Forest	**85.72**%	50.00%	63.38%
Seaside	78.18%	81.40%	**90.48**%
Overall	71.02%	63.75%	**77.55**%

Input image | Initial region-concept association | Final region-concept association

Figure 5: Indicative region-concept association results

Table 2: Concept detection accuracy

Accuracy	Algorithm Stage	
	Initial Region-Concept Association	Final Region-Concept Association
Buildings	48.18%	**51.30**%
Rockyside	56.76%	**58.30**%
Forest	45.09%	**48.01**%
Seaside	59.14%	**61.48**%
Overall	51.22%	**53.80**%

REFERENCES

[1] W. Al-Khatib, Y.F. Day, A. Ghafoor, and P.B. Berra, "Semantic modeling and knowledge representation in multimedia databases," *IEEE Trans. on Knowledge and Data Engineering*, vol. 11, pp. 64–80, Jan.-Feb. 1999.

[2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 11349–1380, 2000.

[3] N. Zlatoff, B. Tellez, and A. Baskurt, "Image understanding and scene models: a generic framework integrating domain knowledge and Gestalt theory, " in *International Conference on Image Processing, ICIP '04*, Vol. 4, pp. 2355-2358.

[4] J. Tang, C. Zhang, and B. Luo, "A graph and PNN-based approach to image classification," in *Proceedings of International Conference on Machine Learning and Cybernetics, ICMLC 2005*, Vol. 8, pp. 5122-5126.

[5] H. Tanaka, H. Sakano, and S. Ohtsuka, "Retrieval method for multi-category images," in *Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004*, Vol. 2, pp. 965-968.

[6] V. S. Tseng, M. Wang, and J. Su, "A New Method for Image Classification by Using Multilevel Association Rules," in *21st International Conference on Data Engineering, ICDE 2005*, pp. 1180-1180.

[7] Z. Wang, D. Feng, and Z. Chi, "Region-based binary tree representation for image classification," in *Proceedings of the International Conference on Neural Networks and Signal Processing, ICNNSP 2003*, Vol. 1, pp. 232-235.

[8] L. Hollink, S. Little, and J. Hunter, "Evaluating the Application of Semantic Inferencing Rules to Image Annotation," in *3rd International Conference on Knowledge Capture, K-CAP 2005*, Banff, Canada.

[9] K. I. Kim, K. Jung, S. H. Park, and H. J. Kim, " Support vector machines for texture classification," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1542–1550, Nov. 2002.

[10] M. Mitchell, *An introduction to genetic algorithms*. Address: MIT Press, 1995.

[11] T. Adamek, N. O'Connor, and N. Murphy, "Region-based Segmentation of Images Using Syntactic Visual Features," in *Proc. of Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2005*, Switzerland.

[12] S. Staab and R. Studer, *Handbook on ontologies*. Address: Springer-Verlag, 2004.

[13] J. Platt, *Probabilities for SV Machines*, In: Advances in Large Margin Classifiers, MIT Press, 2000.

[14] C.-C. Chang and C.-J. Lin., *"LIBSVM: A library for support vector machines"*, http://www.csie.ntu.edu.tw/∼cjlin/libsvm, 2001.

[15] M. Wall, *GAlib: A C++ Genetic Algorithm Library, MIT,* http://lancet.mit.edu/ga, 2005.