Iraqi Journal of Science, 2018, Vol. 59, No.4C, pp: 2323-2335
DOI:10.24996/ijs.2018.59.4C.20

ISSN: 0067-2904

Outdoor Scene Classification Using Multiple SVM

Matheel E. Abdulmunem[1], Eman Hato[2*]
[1]Department of Computer Science, University of Technology-Iraq
[2]Department of Computer Science, Mustansiriyah University, Baghdad, Iraq

Abstract

This paper presents a hierarchical two-stage outdoor scene classification method using multi-classes of Support Vector Machine (SVM). In this proposed method, the gist feature of all the images in the database is extracted first to obtain the feature vectors. The image of database is classified into eight outdoor scenes classes, four manmade scenes and four natural scenes. Second, a hierarchical classification is applied, where the first stage classifies all manmade scene classes against all natural scene classes, while the second stage of a hierarchical classification classifies the outputs of first stage into either one of the four manmade scene classes or natural scene classes. Binary SVM and multi-classes SVMs are employed in the first and second stage of a hierarchical classification respectively. The proposed method is designed also to compare and find the most suitable multi-classes SVMs approach and the kernel function for classification task, where their performances are analyzed based on experimental results. The multi-classes SVMs used in this paper are One-versus-All (OvA) and One-versus-One (OvO), while the kernel functions used are linear kernel, Radius Basis Function (RBF) kernel and Polynomial kernel. Experimental results indicate that OvO classifier provides better performance than OvA classifier. The results, also show that the Polynomial kernel function is superior to others kernel function.

Keywords: Gist descriptor, OvA, OvO, RBF kernel, Polynomial kernel, SVM.

تصنيف المشاهد الخارجية بالإعتماد على المصنف المتعدد SVM

مثيل عماد الدين عبد المنعم[1]، إيمان هاتو هاشم[2*]
[1]قسم علوم الحاسوب، الجامعة التكنولوجية، بغداد، العراق.
[2]قسم علوم الحاسوب، الجامعة المستنصرية، بغداد، العراق.

الخلاصة

يقدم هذا البحث طريقة هرمية لتصنيف المشاهد الخارجية بالاعتماد على المصنف المتعدد SVM في هذه الطريقة المقترحة، يتم حساب ميزات النسيج المكاني لجميع الصور الموجودة في قاعدة البيانات للحصول على متجهات الميزات. قاعدة بيانات الصور مصنفة إلى ثمانية فئات من المشاهد الخارجية: أربعة فئات من المشاهد الصناعية وأربعة فئات من المشاهد الطبيعية. يتم بعدها تطبيق التصنيف الهرمي، حيث يتم في المرحلة الأولى تصنيف المشاهد الى صنفين وهما المشاهد الصناعية والمشاهد الطبيعية، أما في المرحلة الثانية من التصنيف الهرمي فيتم تصنيف مخرجات المرحلة الأولى إلى أربعة أصناف من المشاهد الصناعية وأربعة أصناف من المشاهد الطبيعية. يتم استخدام SVM الثنائي في المرحلة الأولى وSVM المتعدد في المرحلة الثانية من التصنيف الهرمي. يتم أيضاً في الطريقة المقترحة مقارنة وإيجاد طريقة SVM المتعدد والدالة الأكثر ملاءمة لمهمة التصنيف حيث يتم تحليل أدائها بناء على النتائج التجريبية. إن طرق SVM

*Email:emanhato@yahoo.com

المتعدد المستخدمة في هذه البحث هي One−versus−All(OvA) و One−versus−One(OvO) في حين إن أنواع الدوال المستخدمة هي Linear و Radius Basis Function (RBF) و Polynomial. أظهرت النتائج التجريبية أن المصنف OvO يوفر أداءً أفضل من المصنّف OvA. كما أظهرت النتائج إن دالة Polynomial يتفوق بالأداء على أنواع الدوال الأخرى.

1. Introduction

Digital images are the most important media in our daily life. With the advancement of modern telecommunication and multimedia technologies, a huge amount of digital images are traveled over shared networks and stored in various fields such as personal image collections, digital arts, medical imaging, aerial and satellites image.

The question that arises is how are images searched, accessed and stored in an efficient manner? One way to provide this effectiveness is to group images in one of predefined categories, this process is called image classification. The aim of image classification is to separate images using their visual content which is fully automatically extracted such as color, texture and shapes into two or more separate categories [1].

There are a lot of applications that can benefit from the image classification. Searching for images in a huge database, or in the internet is the most direct application of image classification. Personal photo organization is also one of the important applications of image classification, it provides the best dealing with a large library of photo memories according to visual topics. As well as other application such as image retrieval, medical applications, surveillance system and satellite imagery [2].

Although scenes classification is not a very difficult task for humans, but it difficult problem for computer programs due to their illumination variability, variations in scale ,ambiguity, and the gap between low level features extracted from image content and user perceptions of these features in the real world [3]. For example, the same mountain has different appearances depending on season of the year. It is snow covered in the winter, it has brown color in autumn and it is covered with green grass in the spring as shown in Figure - 1. A person can recognize the mountain in all these situations automatically, but recognizing these situations by computer still a challenge in computer vision.

Three mountain scenes with a different appearance

Three forest scenes with a different illumination

Figure 1- Scene classification with different conditions.

The prime factors of image classification include determination of suitable classifiers, and feature vectors. Several advanced classifiers such as Neural Networks, k Nearest Neighbor, SVM, Decision Tree Technique and Random Forest have been widely applied to image classification. Among the all classifiers SVM produced the highest performance and has a powerful learning ability [4, 5]. The vector of feature must be extracted carefully from an image to reduce huge and rich content of image and maintain the representation of the entire image. For example, the shape may be a good feature to distinguish between boats and cars but it is not good to distinguish between coast and forest [6].

Basic scene classification approaches based on low level features can be distinguished in the literature are: global and local representations. In global representations the feature vector is computed over the whole image that will be the input of classifier to give the final class of image. The local representations divide the image into several blocks, and then features vectors are extracted from each of these blocks; each feature vector is classified independently to obtain a class for each one. All categories of blocks are combined to obtain final image class. An example of a global and local representation is illustrated in the Figure - 2.

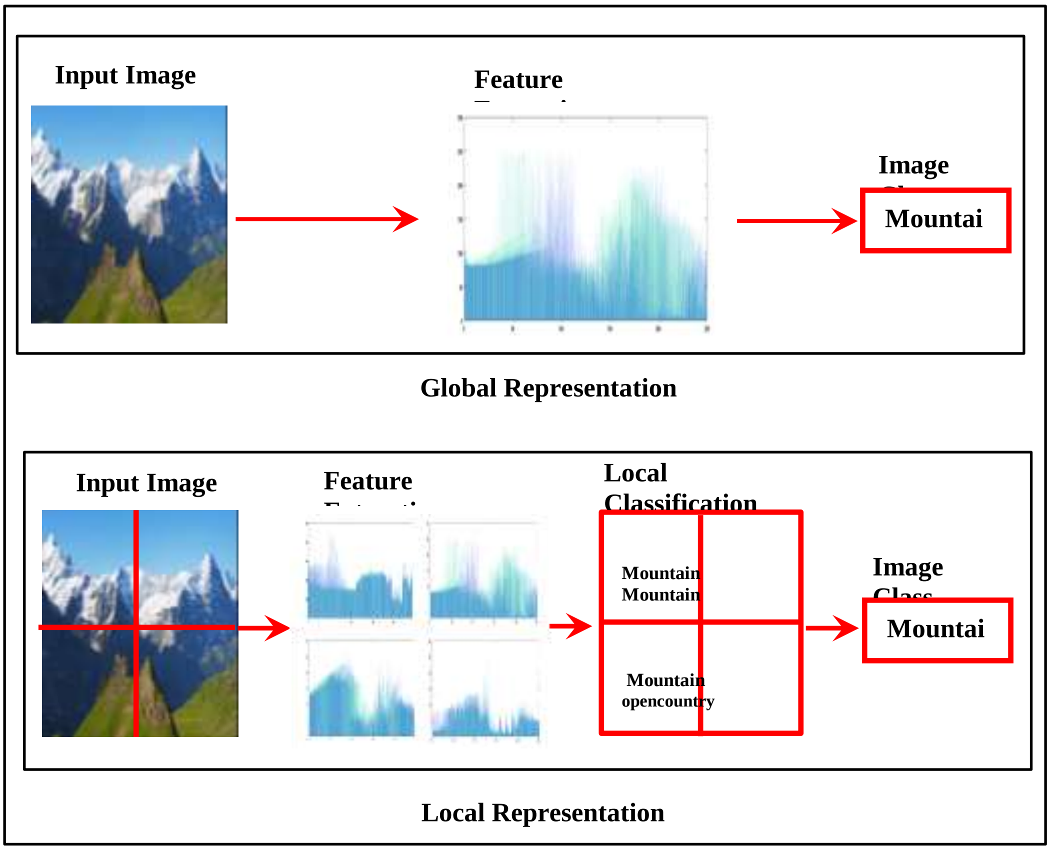

Figure 2- Global and local representation.

For example in [7] the neural network is used to perform dimensional reduction for multi visual features vectors including shape, texture, color layout, and color. It produces feature vector with low dimensional which is useful for effective classification. Three classifiers are employed K Nearest Neighbors (KNN), Gaussian Mixture Models (GMM) and SVM. The results indicate that the approach has robustness against visual distortions and noise. In [8] the Artificial Neural Network (ANN) and SVM are used to classify roman numerals image where the image is encoded as a shape matrix. The researchers in [9] divided each image into 36 blocks and then extracted 23 features from each block, which produce vector dimensions of 36 x 23 per image. The extracted features are edge directed histograms, entropy of wavelet coefficients and colour histogram. SVM is used as classifier with three kernels includes: linear, gaussian and polynomial kernel. The results show good performance and gaussian kernel outperform other kernels types.

The researchers in [10] used local representation where wavelet texture features and color features are extracted from sub blocks of image and classified separately using SVM as initial indoor outdoor prediction of different regions of the scene. The scene classes are then integrated by a Bayesian network designed to improved indoor-outdoor scene classification. In [11] images are segmented into regions to be matched. Edge direction histograms, colour histograms, colour moments and texture features are extracted of each region. The similarity between two images is then calculated as the cost of a pairwise matching of regions according to their related features. The method improves the performance of outdoor scenes classification. A new method are proposed in [12] where the color and texture feature are extracted from the block of region of interest (the middle area of image), and then each block is labeled by its relative semantic concept by using SVM classifier. Finally, the scene classifier gathers these semantic concepts to decide the final class of image. In order to improve performance of classification, local and global features are used together as proposed in [13].

Researchers in computer vision field are suggested the bag of words (BOW) technique recently to encode the contents of the image and then use it in the classification task. The first step in the BOW framework is image representation that includes features extraction and codebook construction. While, the second step is the image classification to specify the category to which the input image belongs to [14, 15].

The objective of this paper is to design a method to classify scene images using multi-classes SVM with low computation cost and preserving the efficiency and performance, as well as to analyze the effect of different multiclass SVMs approach and kernel functions on classification performance and accuracy. The remainder of this paper is structured as follows: In the next section description of gist features is presented. In section 3 SVM technique is discussed. Section 4 multi-classes SVM approaches are presented. The proposed scene classification method is explained in section 5. Section 6 discusses the experimental results obtained. Lastly, the conclusions are contained in section 7.

2. Gist Feature

The global gist descriptor is a low dimensional representation of the scene, which does not require any form of segmentation. A bank of Gabor filters is used in the frequency domain and tuned to different orientations and scales. The image is divided into a 4x4 grid where orientation histograms are calculated. The gist features produce a vector of dimension 512. The gist representation includes all levels of visual information ranging from low level features (contours, color and spatial frequencies) to intermediate features (texture, shapes) and high level features (objects, activation of semantic knowledge). Thus, gist can be represented as Perceptual gist (structural representation of a scene) and Conceptual gist (semantic information that is deduced while viewing a scene) [16]. The gist descriptor is scale invariable feature and shows good results for scene recognition and searching.

3. Support Vector machine (SVM)

The SVM is supervised classification technique that can be used as a binary classifier based on a linear discriminant function. SVM separates the data of two classes with a hyperplane. Let the labeled training examples are (x_i,y_i), where i=1,2,….n, $xi \in R_d$ is the i-th input vector, d is the dimensionality of the input vectors and $y_i \in \{+1,-1\}$ is the i-th output classes.

The basic idea of SVM is to find an optimal separating hyperplane that separates the training vectors. So, input vector lying on one side of the hyperplane are labeled as y=-1, and input vectors lying on the other side are labeled as y=+1. The training examples that locate closest to the hyperplane are called support vectors, and the distance that exists between the two support vectors is called margin [17] as shown in Figure - 3. If the set can be separated linearly, the SVM classifier defines the optimal separating hyperplane as a function:

$$f(x) = sign(w^t x + b) \quad (1)$$

Where w is weight vector, b is the bias, $w^t x$ is inner product defined as $w^t x = \sum_i w_i x_i$, and the sign of f(x) gives the label of the sample vector x.

The goal of the SVM is to maximize the margin between the optimal hyperplane and the support vector where the margin is given by: 2/||w||, where ||w|| is normal distance to the hyperplane. So instead of maximizing 2/||w|| it can equivalently minimize ||w||/2 and to do this Lagrange multiplier is employed to solve the problem [17, 18]:

$$w = \sum_{i=1}^{r} y_i \alpha_i x_i \quad (2)$$

$$f(x) = sign(\sum_{i=1}^{r} y_i \alpha_i x_i x + b) \quad (3)$$

Where α_i , i=1,2,…..,r is Lagrange multipliers.

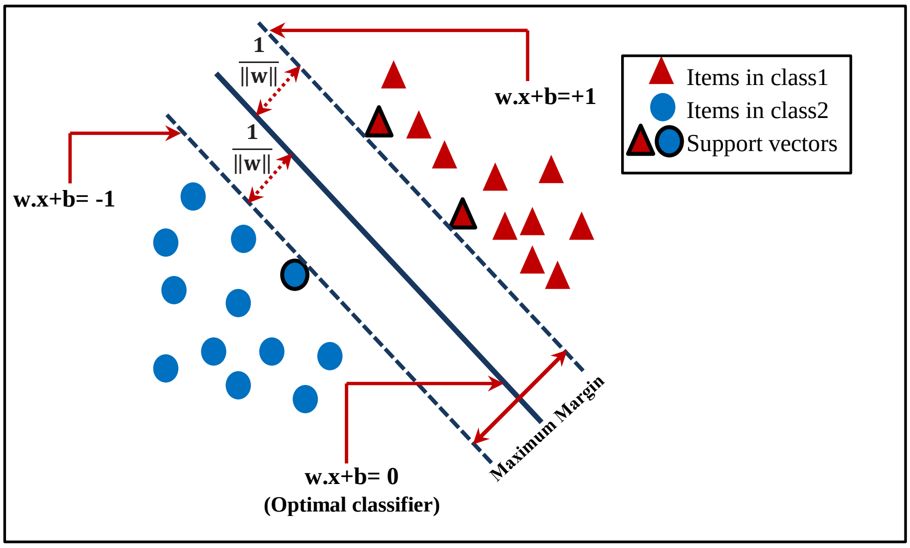

Figure 3- SVM linear classification.

The solution to this optimization problem can be easily obtained when the data are linearly separable. When the data is nonlinearly separable in the original space, classical SVMs fail to find an optimal linear classifier for separating classes. One solution is to transform the problem to a higher dimensional feature space, which can be separated linearly through a nonlinear transformation known as kernel function [19]. First the original inputs X is mapped into a high dimensional feature space F through a mapping function φ, then the optimal separating hyperplane can be defined in space F as follows:

$$f(x) = sign(\sum_{i=1}^{r} y_i \, \alpha_i \, \varphi(x_i) \, \varphi(x) + b) \tag{4}$$

Since a dot product is the only operation among the mapped examples, the kernel function is employed to avoid explicitly implement of the mapping function φ by:

$$K(x_i, x_j) = \varphi(x_i).\varphi(x) \tag{5}$$

The most popular kernel functions are the linear kernel, the Radius Basis Function (RBF) kernel and the polynomial kernel [19, 20] as defined in the following equationes:

Linear kernel:

$$K(x_i, x_j) = x_i^T x_j \tag{6}$$

RBF kernel:

$$K(x_i, x_j) = exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0 \tag{7}$$

Polynomial kernel:

$$K(x_i, x_j) = \left(\gamma x_i^T x_j + r\right)^d, \gamma > 0 \tag{8}$$

Although different SVM performance can be obtained by using different kernel functions, selecting the proper kernel function type for a specific problem will be difficult and important.

4. Multiclass SVM

Basically SVM separates only two classes with a maximum margin. The problems of real world need the recognition for more than two classes. In practice, the multi classes classification problems are partitioned into a predefined set of binary problems so that the binary SVM can be applied directly. The popular approach that solve the n-class classification problem by using n binary SVMs are One-versus-All (OvA) SVM, One-versus-One (OvO) SVM [21].

4.1One-versus-All (OvA) approach

OvA train one binary SVM for each class to recognize it from all the other classes. Hence classification k classes use separate k binary SVMs. The data of kth class is used as positive examples and the data of remaining k-1 classes is used as negative examples when training the kth binary SVM. One of the main problems of the OvA approach is the unbalanced training set [21, 22].

4.2 One-versus-One (OvO) approach

The other approach is OvO that trains k(k-1)/2 binary SVMs, each binary SVM recognizes between two of the k classes. A test example is examined by each classifier and one vote is given to the winning class. A test example is categorized to the class with the greatest number of votes. The classifiers size that is created by the OvO approach is much greater than that of the OvA approach. Moreover, compared with the OvA approach, the OvO approach is more symmetrical [21, 22].

5. Proposed Classification Method

A two stage hierarchical classification method is proposed to classify eight categories of outdoor scenes images (four manmade scenes and four natural scenes). Figure - 4 illustrates the general structure of the proposed method. It is composed of two parts: feature vectors generation and hierarchical classification.

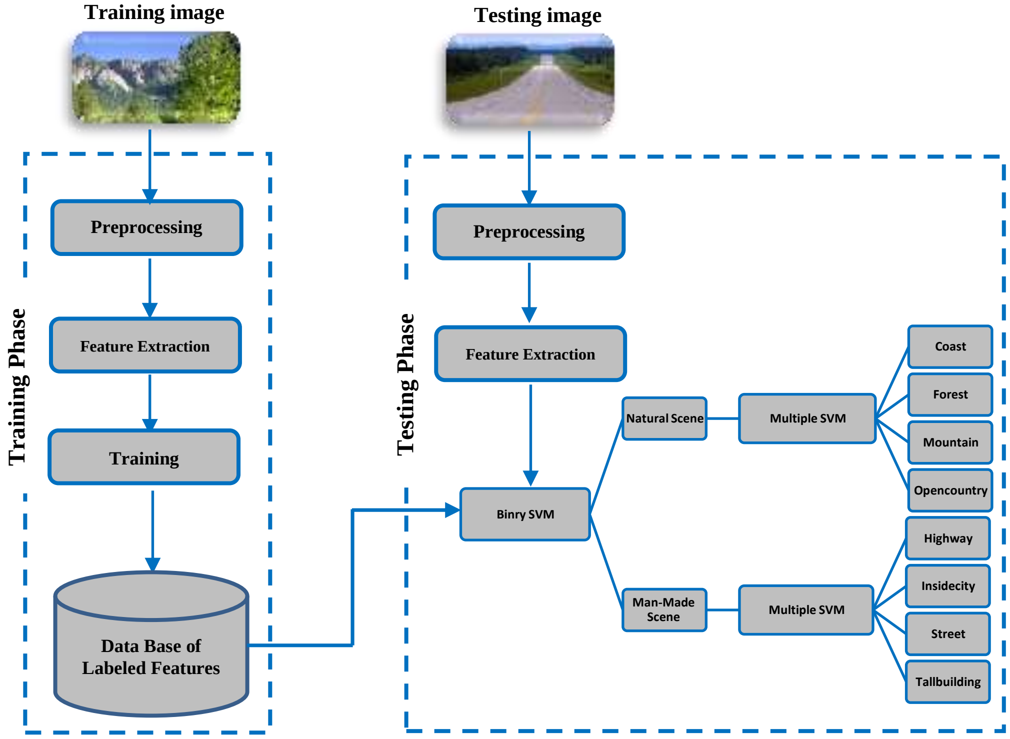

Figure 4 - Structure of proposed classification method.

5.1 Feature Vectors Generation

Firstly, all images are processed by resizing them to 256x256 and converting them to the YCbCr color space, where Y represents luminance and Cb, Cr represents color information (chrominance). After that, gist descriptor is computed for all images to obtain descriptor feature vectors. The gist features are computed by convoluting the Gabor filter with an image at different scales and orientations (intensity, colour and orientation channels). Gist is scale invariable feature because gist descriptor gives the scene of the image (it does not represent the details of an image), so changing the image size will not alter the gist features. The result of applying YCbCr color space and the gist features descriptor for a selected image are shown in Figure - 5.

5.2 Hierarchical Classification

A hierarchical classification is applied, where the first stage classifies all manmade scene classes versus all natural scene class by using binary SVM. In the second stage of a hierarchical classification multi-classes SVMs are employed to classify the outputs of first stage into either one of the four manmade scene classes or natural scene classes. The four manmade scene classes are highway, insidecity, street and tallbuilding while the natural scene classes include coast, forest, mountain, opencounty. OvA and OvO SVM are used in second stage of hierarchical classification and their performance is compared according to a different type of kernel functions. The types of kernel function used are Liner, RBF and Polynomial.

Original image YCbCr color space Gist descriptor

Figure 5 - YCbCr color space and gist descriptor applied for a selected image.

6.Experimental Results

The measure used to assess the proposed method performance is confusion matrix. The confusion matrix is used to analyze, visualize and measure the performance of most scene classification methods. The rows represent the actual classes, whereas the columns represent the predicted classes (classes in the classification result). The elements in the diagonal of matrix correspond to the number of images correctly classified for each class. In the confusion matrix, the class accuracy is measured as a fraction of correctly categorized images with reference to all images of that class. Total rating accuracy is calculated by average diagonal values of the confusion matrix. High score refers to high classification accuracy.

A database of 1440 outdoor scene images is employed for training and testing the SVMs classifiers. The dataset is randomly divided into training set of 1040 images and testing set of 400 images. So for each class there are 130 training images and 50 testing images. Figure - 6 shows samples of database images for each class.

Coast Forest Mountain Opencountry

Natural scene

Highway Insidecity Street Tallbuilding

Manmade scene

Figure 6- Some images of database.

The results of the experiments are presented in Table-1 to 6 as confusion matrixes. Each one of the confusion matrix details the classification accuracy of proposed method using OvO SVM and OvA SVM with three kernel function (linear, RBF and polynomial). It is clear from the Table - 1 to 6 that the effect of proposed method is good and it is able to correctly classify scene images.

Figure - 7 summarizes the overall accuracy of the proposed method in a graphical representation. It is evident that proposed method using OvO has achieved average accuracy higher than the proposed method using OvA.

Table 1-The result of proposed method using OvO with linear kernel function

Multi-classes SVM: One-versus-One / Kernel function: Linear kernel								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.90	0	0	0.1	0	0	0	0
Forest	0	0.90	0.08	0.02	0	0	0	0
Mountain	0	0.06	0.76	0.16	0	0.02	0	0
Opencountry	0.16	0.02	0.08	0.70	0.04	0	0	0
Highway	0.08	0	0	0	0.86	0.04	0.02	0
Insidecity	0	0.02	0	0	0.06	0.72	0.14	0.06
Street	0	0	0	0	0	0	0.98	0.02
Tallbuilding	0	0.02	0	0	0	0.12	0	0.86
Overall Accuracy	**0.835**							

Table 2-The result of proposed method using OvO with RBF kernel function

Multi-classes SVM: One-versus-One / Kernel function: RBF kernel								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.86	0	0	0.14	0	0	0	0
Forest	0	0.96	0.04	0	0	0	0	0
Mountain	0	0.04	0.82	0.12	0	0.02	0	0
Opencountry	0.12	0.08	0.16	0.60	0.02	0.02	0	0
Highway	0	0	0	0.08	0.82	0.02	0.06	0.02
Insidecity	0	0	0	0.02	0.02	0.70	0.16	0.1
Street	0	0	0	0	0.02	0	0.96	0.02
Tallbuilding	0	0	0.02	0	0	0	0	0.98
Overall Accuracy	**0.8375**							

Table 3-The result of proposed method using OvO with polynomial kernel function

Multi-classes SVM: One-versus-One / Kernel function: Polynomial kernel								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.86	0	0	0.14	0	0	0	0
Forest	0	0.94	0.02	0.04	0	0	0	0
Mountain	0.02	0.02	0.80	0.14	0	0.02	0	0
Opencountry	0.12	0.04	0.08	0.72	0	0	0.04	0
Highway	0.08	0	0	0	0.88	0	0.02	0.02
Insidecity	0	0	0	0.02	0.02	0.74	0.14	0.08
Street	0	0	0	0	0.02	0	0.96	0.02
Tallbuilding	0	0	0.02	0	0	0.06	0	0.92
Overall Accuracy	**0.8525**							

Table 4-The result of proposed method using OvA with linear kernel function

Multi-classes SVM: One-versus-All / Kernel function: Linear kernel								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.92	0	0.02	0.06	0	0	0	0
Forest	0	0.86	0.04	0.1	0	0	0	0
Mountain	0.06	0.04	0.76	0.12	0	0.02	0	0
Opencountry	0.26	0.08	0.22	0.40	0.04	0	0	0
Highway	0.08	0	0	0	0.84	0.04	0.04	0
Insidecity	0	0.02	0	0	0.1	0.72	0.1	0.06
Street	0	0	0	0	0.02	0	0.98	0
Tallbuilding	0	0.02	0	0	0	0.12	0	0.86
Overall Accuracy	**0.7925**							

Table 5-The result of proposed method using OvA with RBF kernel function

Multi-classes SVM: One-versus-All / Kernel function: RBF kernel								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.90	0	0	0.1	0	0	0	0
Forest	0	0.98	0	0.02	0	0	0	0
Mountain	0	0.02	0.84	0.12	0	0.02	0	0
Opencountry	0.16	0.06	0.14	0.60	0.02	0.02	0	0
Highway	0	0	0	0.08	0.86	0.02	0.04	0
Insidecity	0	0	0	0.02	0.04	0.74	0.14	0.06
Street	0	0	0	0	0.02	0	0.96	0.02
Tallbuilding	0	0	0.02	0	0	0.02	0	0.96
Overall Accuracy	**0.855**							

Table 6-The result of proposed method using OvA with polynomial kernel function

Multi-classes SVM: One-versus-All / Kernel function: Polynomial kernel								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.90	0	0	0.08	0	0	0	0
Forest	0	0.96	0.02	0.02	0	0	0	0
Mountain	0	0.02	0.82	0.14	0	0.02	0	0
Opencountry	0.18	0.04	0.16	0.58	0	0	0.04	0
Highway	0.08	0	0	0	0.86	0.02	0.04	0
Insidecity	0	0	0	0.02	0.04	0.74	0.14	0.06
Street	0	0	0	0	0.02	0	0.98	0
Tallbuilding	0	0	0.02	0	0	0.06	0	0.92
Overall Accuracy	**0.845**							

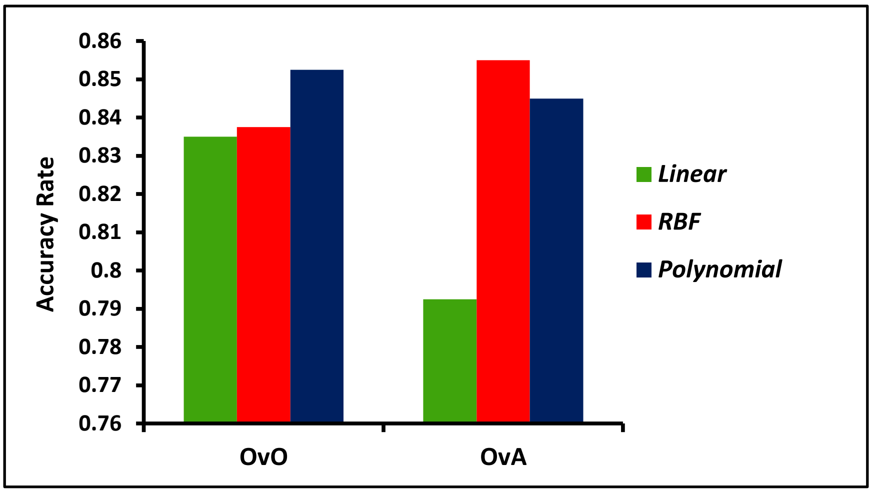

Figure 7- Accuracy of proposed method.

Different kernel functions vary the classification accuracy. One can observe the performance of the RBF kernel is more accurate than polynomial kernel when the proposed method using OvA. On the contrary, when the proposed method using OvO, polynomial kernel is achieved the highest accuracy. In both case the linear kernel is given the lowest performance. In general, the use of polynomial kernel has the highest performance comparative to other kernels whereas the use of linear kernel has the worst performance when considering the average rate of accuracy, where the average values of accuracy are 0.84875, 0.84625, 0.81375 for using polynomial, RBF, linear kernel respectively.

A high degree of ambiguity in various categories of scene images can greatly confuse image content and cause an error or misclassification. For example, the insidecity and the street classes are considered as two different scenes. However most of the insidecity images contain a street which will easily lead to classification errors as shown in Figure - 8.

Images are classified as street class where the correct class is insidecity

Images are classified as mountain class where the correct class is opencountry

Figure 8- Example of error classification.

The result of the proposed hierarchy classification method is compared to the result of applying the multi-classes SVM directly on the all classes. Table - 7 and 8 shows the confusion matrix of OvO and OvA respectively.

Table 7-The result of applying OvO

Multi-classes SVM: One-versus-One								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.82	0	0	0.14	0.02	0	0	0.02
Forest	0	0.90	0.06	0.04	0	0	0	0
Mountain	0	0.02	0.82	0.14	0	0	0.02	0
Opencountry	0.08	0.04	0.08	0.68	0.08	0	0.04	0
Highway	0.08	0	0	0.06	0.82	0.02	0.02	0
Insidecity	0.02	0	0	0	0	0.74	0.16	0.08
Street	0	0	0	0	0.02	0	0.98	0
Tallbuilding	0	0.02	0	0	0	0.06	0	0.92
Overall Accuracy	**0.835**							

Table 8-The result of applying OvA

Multi-classes SVM: One-versus-All								
Classes name	**Natural scene**				**Manmade scene**			
	Coast	Forest	Mountain	Opencountry	Highway	Insidecity	Street	Tallbuilding
Coast	0.82	0	0	0.14	0.04	0	0	0
Forest	0	0.94	0.02	0.02	0	0.02	0	0
Mountain	0	0.02	0.84	0.12	0	0	0.02	0
Opencountry	0.18	0.04	0.16	0.5	0.06	0	0.06	0
Highway	0.06	0	0	0.04	0.84	0.02	0.04	0
Insidecity	0	0	0.02	0	0.06	0.74	0.1	0.08
Street	0	0	0	0	0.02	0	0.98	0
Tallbuilding	0	0.02	0	0.02	0	0.04	0	0.92
Overall Accuracy	**0.8225**							

The comparison is made as shown in Figure-9 and it can be concluded from the accuracy rate that applying the multi-classes SVM directly has the lowest accuracy compared to the proposed method, which agrees with the expected results.

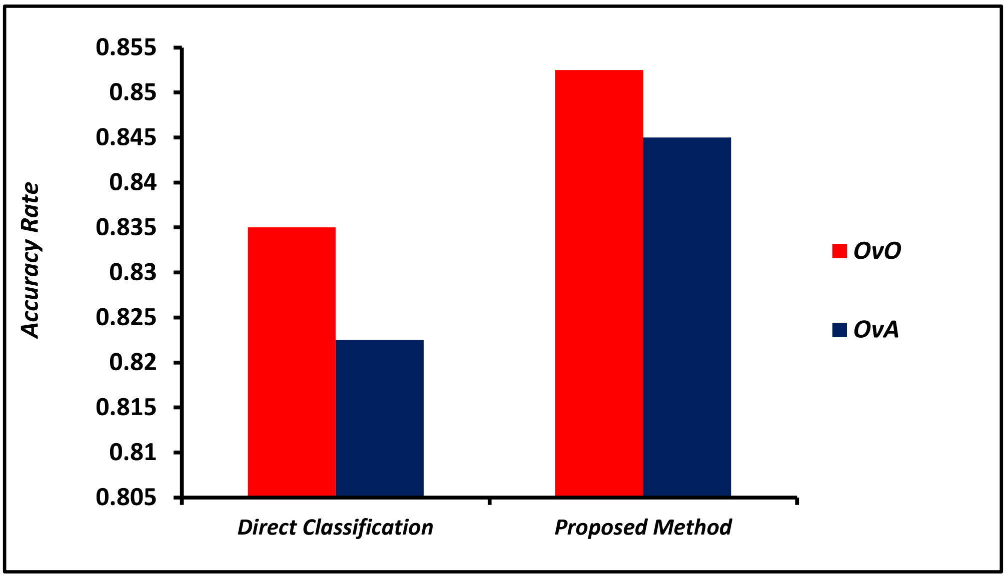

Figure 9- Accuracy comparison.

Computational complexity is an important aspect for efficient classification techniques. It is noticeable that the proposed method reduces the number of binary SVM that will be used in case of applying OvO directly. In OvO approach, the number of binary SVM required is k (k-1)/2 which mean 28 binary SVM in case of eight classes. While the total number of binary SVM classifier used in the proposed method when using OvO is only 13 (one binary SVM in first stage and 4(4-1)/2 for each multiple SVM in second stage). The proposed method also reduces the computation cost of the feature extractors because it uses only the gist features and it can still predict the outdoor scene categorization with satisfactory performance.

7. Conclusions

This paper presents an outdoor scene classification method using multi-classes SVM. The gist descriptor is employed in proposed method as generator of feature vectors and then two stages of hierarchical classification is applied. The first stag classifies manmade scene versus natural scene using binary SVM while, the second stage classifies the outputs of first stage into either one of the four manmade scene classes or natural scene classes using multi-classes SVM. The results indicate that the proposed method can achieve good level of performance for outdoor scene images classification and reduce the computational cost.

The performance of the OvO and OvA as an example of multi-classes of SVM is tested based on the accuracy rate of classification according to different kernel functions. Based on the results of the comparison performed, it is observed that the best average accuracy rate is obtained when using OvO in this method. The results also show that the best performance is achieved with polynomial kernel while the worst performance is achieved with linear kernel when considering the average accuracy rate.

8. References

1. Nath, S. S., Mishra, G., Kar, J., Chakraborty, S. and Dey, D. **2014**. A survey of image classification methods and techniques. In proceedings of IEEE International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), pp. 554-557.
2. Kamavisdar, P., Saluja, S. and Agrawal, S. **2013**. A survey on image classification approaches and techniques. *International Journal of Advanced Research in Computer and Communication Engineering*, **2**(1): 1005-1009.
3. Upadhyaya, N. and Dixit, M. **2016**. A Review: Relating Low Level Features to High Level Semantics in CBIR. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, **9**(3): 433-444.
4. Thanh Noi, P. and Kappas, M. **2017**. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors Journal*, **18**(1): 1:20.
5. Qian, Y., Zhou, W., Yan, J., Li, W. and Han, L. **2014**. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sensing*, **7**(1): 153-168.
6. Tian, D. P. **2013**. A review on image feature extraction and representation techniques. *International Journal of Multimedia and Ubiquitous Engineering*, **8**(4): 385-396.
7. Shen, J., Shepherd, J. and Ngu, A. H. **2005**. Semantic-sensitive classification for large image libraries. In proceedings of the 11th International Multimedia Modelling Conference, IEEE, pp. 340-345.
8. Thai, L. H., Hai, T. S. and Thuy, N. T. **2012**. Image classification using support vector machine and artificial neural network. *International Journal of Information Technology and Computer Science (IJITCS)*, **4**(5): 32-38.
9. Mandhala, V. N., Sujatha, V. and Devi, B. R. **2014**. Scene classification using support vector machines. In proceedings of the International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), IEEE, pp. 1807-1810.
10. Serrano, N., Savakis, A. E. and Luo, J. **2004**. Improved scene classification using efficient low-level features and semantic cues. *Pattern Recognition*, **37**(9): 1773-1784.
11. Kaick, O. V., and Mori, G. **2006**. Automatic classification of outdoor images by region matching. In proceedings of the 3rd Canadian Conference on Computer and Robot Vision, IEEE, pp. 9.

12. Lin, S., Chen, H. and Shih, S. **2013**. Real-time Image Scene Classification and Segmentation System. *International Journal on Computer Science and Engineering*, **5**(10): 864- 873.

13. Lisin, D. A., Mattar, M. A., Blaschko, M. B., Benfield, M. C., Learned-Miller, E. G. **2005**. Combining local and global image features for object class recognition. In proceedings of IEEE Computer society conference on Computer vision and pattern recognition-workshops, pp. 47-47.

14. Van Gemert, J. C., Snoek, C. G. M., Veenman, C. J., Smeulders, A. W. M. and Geusebroek, J. **2010**. Comparing compact codebooks for visual categorization. *Computer Vision and Image Understanding,* **114**(4): 450-462.

15. López-Sastre, R. J., Tuytelaars, T., Acevedo-Rodríguez, F. J. and Maldonado-Bascón, S. **2011**. Towards a more discriminative and semantic visual vocabulary. *Computer Vision and Image Understanding*, **115**(3): 415-425.

16. Douze, M., Jégou, H., Sandhawalia, H., Amsaleg, L. and Schmid, C. **2009**. Evaluation of gist descriptors for web-scale image search. In proceedings of the ACM International Conference on Image and Video Retrieval, p. 19.

17. Ye, F., Zhang, Z., Chakrabarty, K., Gu, X. **2017**. Diagnosis Using Support Vector Machines (SVM). In. Knowledge-Driven Board-Level Functional Fault Diagnosis. Springer, Cham, pp.23-42.

18. Murty, Maddipati Narasimha, and Rashmi Raghava. **2016**. *Support Vector Machines and Perceptrons: Learning, Optimization, Classification, and Application to Social Networks*. Springer.

19. Cho, G., Gantulga, N., Choi, Y. **2017**. A Comparative Study on Multi-class SVM & Kernel Function for Land Cover Classification in a KOMPSAT-2 Image. *KSCE Journal of Civil Engineering*, **21**(5): 1894–1904.

20. Yekkehkhany, B., Safari, A., Homayouni, S. and Hasanlou, M. **2014**. A comparison study of different kernel functions for SVM-based classification of multi-temporal polarimetry SAR data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, **40**(2): 281-285.

21. Wang, Z. and Xue, X. **2014**. Multi-Class Support Vector Machine. In. Y. Ma and G. Guo (eds.). Support Vector Machines Applications, Springer, Cham, pp. 23-48.

22. He, X., Wang, Z., Jin, C., Zheng, Y. and Xue, X. **2012**. A simplified multi-class support vector machine with reduced dual optimization. *Pattern Recognition Letters,* **33**(1): 71-82.