TESTING PERFORMANCE OF HIGH SPATIAL RESOLUTION GOKTURK-2 SATELLITE IMAGE FOR URBAN AND RURAL ENVIRONMENT MAPPING

Gulcan Sarp[1] and Arzu Erener[2]
[1]Suleyman Demirel University, Department of Geography 32260 Isparta, Turkey,
Email: gulcansarp@sdu.edu.tr
[2]Kocaeli University, Department of Geomatics Engineering, 41380 Kocaeli, Turkey,
Email: arzu.erener@kocaeli.edu.tr

KEY WORDS: Göktürk-2, Support Vector Machines, Urban and Rural Environment Mapping

ABSTRACT: Advances in remote sensing in terms of increasing spatial/spectral resolutions have strengthened its ability of urban environmental analysis. Accurate estimation of urban and rural environment is necessary to arrange present and feature management and planning applications. In this study a new satellite Göktürk-2 images having 2.5 m. spatial resolution for panchromatic band and 5 m. for multispectral bands is tested for urban and rural environment mapping. In the study urban and rural areas are extracted from pan-sharpened Göktürk-2 satellite image using pixel based Support Vector Machines (SVMs) classification algorithm. The experimental results indicate a mean accuracy value around 80 % of the Göktürk-2 image for the urban and rural environment mapping which is very promising.

1. INTRODUCTION

Urban and Rural land cover information is one of the crucial data components for different field of studies especially in the urban planning and environmental applications. This kind of information practically derived from the aerial photos and satellite imageries. These data makes easy to accurately interpret small scale and inaccessible areas. The success of urban and rural area mapping mainly lies on the choice of appropriate image data and classification techniques. In the recent decade, urban researchers have advocated the use of high spatial resolution images (better than 5 m. spatial resolution), for different applications such as land-use/land-cover classification (Sugumaran et al., 2002; Wang et al., 2004; Lu and Weng, 2009; Lu et al., 2010). The land cover extraction object depends on physical properties and homogeneity of the objects. But, especially in urban areas, another important aspect to be considered is the local texture, spatial arrangement and geometric properties (size, shape, orientation etc.) of land cover objects (Zhan et al., 2002). High resolution imageries can provide more accurate distribution of the land surface objects. However, at high resolution imageries, an area that is formerly spectral uniform will be composed of pixels with a higher degree of spectral variation (Zhang and Feng, 2005).

Most researchers have used multi-class support vector machine (SVM) classification for land use detection of urban areas from high-resolution satellite images. SVMs are powerful tools for providing solutions to classification, regression and density estimation types of problems (Sagale and Kale, 2014). SVM has been selected as most promising classifier to be used within this study for the task of urban structure type pattern recognition within built-up areas. The classifier choice is based on a comprehensive study that was previously carried out to assess the performance of various learning machines to distinguish built-up from non-built-up areas (Wieland and Pittore 2014). Sarp et al. (2014) applied SVM classification to high resolution orthophotos to determine damaged building areas after Van-Ercis earthquake. Tuia et al. (2010) performed SVM classification using composite kernels for the classification of high-resolution urban images and concluded that a significant increase in the classification accuracy was achieved when the spatial information was used. Li et al. (2010) presented an object-oriented land cover classification method based on SVM. Their results indicate that fusion strategy and classification preprocessing increases classification accuracy.

In the study urban and rural areas are extracted from pan-sharpened Göktürk-2 satellite image using pixel based Support Vector Machines (SVMs) classification algorithm which successfully minimizes errors and maximizes the geometric characteristics of edge areas. The aim of this article is to evaluate the separability of urban and rural objects using Gokturk-2 image through SVM classification in the urban, suburban and rural areas of the Istanbul city (Turkey).

2. THE STUDY AREA AND DATA SETS

The study area is located in Istanbul province of Turkey. There test region selected for the study are a highly-dense built-up urban area, medium-dense built-up sub-urban area and low-dense built-up rural area which involves various urban objects such as buildings, different land uses like roads, bare soil, vegetation and shadow. In some places, buildings appear indistinguishable from roads, pavements and bare soil and may reflect fragmented characteristics

due to shading or they may be occluded by other buildings. Additionally, manmade structures are composed of different sizes and different roof materials such as concrete, brick, asphalt, metal, soil, etc.
The true and false color composite image of the urban, sub urban and rural areas test areas, covering a part of Istanbul of Turkey is given in the Fig.1 a, b, c and d, e, f respectively.

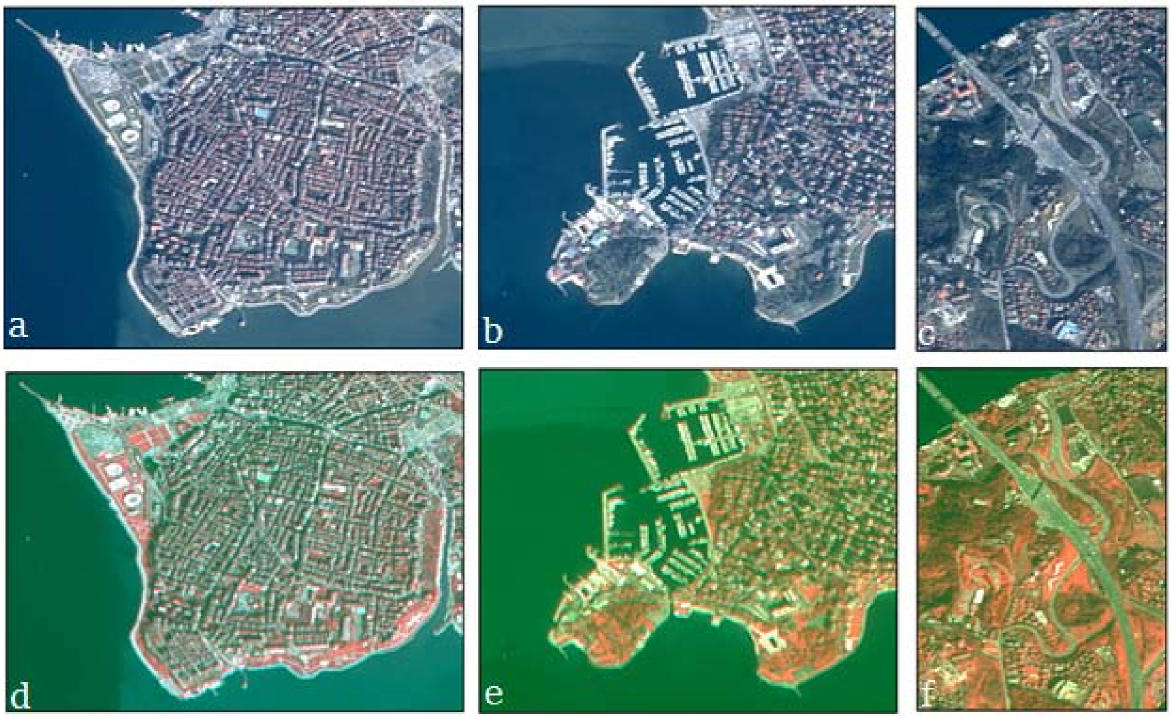

Figure 1. The subset of the test area (R: red band, G: Green band, B: Blue band) (a, b, c) and (R: NIR band, G: Green band, B: Blue band) (d, e, f)

Göktürk-2 is the second mini satellite of Turkey, which was built by TÜBİTAK UZAY and Turkish Aerospace Industries Inc. consortium. The Göktürk-2 spacecraft was launched on 18th December 2012. It has 2.5 m for panchromatic band and 5 m for multispectral bands which include NIR in addition to the usual RGB (Table 1). Göktürk-2 satellite is operated by Turkish Air Force (TUAF) and it supplies to national civilian and military imaging needs (Teke, 2016).

Table 1. Göktürk-2 specifications

Bands	Spatial Resolution	Revisit Time	Swath Width	Radiometric Resolution
Pan/ RGB-NIR	2.5 m/ 5 m	2.5 days (Avg.)	20 km	11-bit

A radiometric and geometrically corrected, pan-sharpened, multi-spectral Gokturk-2 sub-scene of 2.5 m pixel resolution is employed in the present study. This imagery is produced by fusing 11-bit of 2.5 m resolution panchromatic and 5 m resolution multi-spectral - blue, green, red and near infrared channels. The image data used in the study is provided by TUAF.

3. SUPPORT VECTOR MACHINE

SVM is a non-parametric classifier derived from statistical learning theory and originally developed by Vapnik (1995). This classifier uses kernel functions to project non-linearly separable classes into higher dimensional feature space, where non-linearly separable classes can be separated by a linear hyperplane (Figure, 2b). The optimal separating hyperplane between two classes is chosen by maximizing the margin between the separating hyperplane and the closest feature vectors. There will be an immeasurable number of hyperplanes and SVM will choose the hyperplane with maximum margin. The margin indicates the distance between the classifier and the training points (support vector) (Figure 2a). Therefore, only the closest training samples (support vectors) to the edge of the class distribution are used, which is why potentially SVM can deal well with small training data sets given that they are well selected

(Foody and Mathur, 2006). SVMs need training data that optimize the separation of the classes rather than describing the classes themselves (Foody and Mathur, 2006). Using a radial basis function (RBF), class distributions with non-linear boundaries can be mapped into a high dimensional space for linear separation (Huang et al., 2002). Training the SVM with a Gaussian radial basis function requires setting two parameters: C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the training error, while γ describes the kernel width. A small C-value tends to emphasize the margin while ignoring the outliers in the training data, while a large C-value may over fit the training data. A comprehensive description of SVMs can be found in Burges (1998) and Cristianini and Shawe-Taylor (2000).

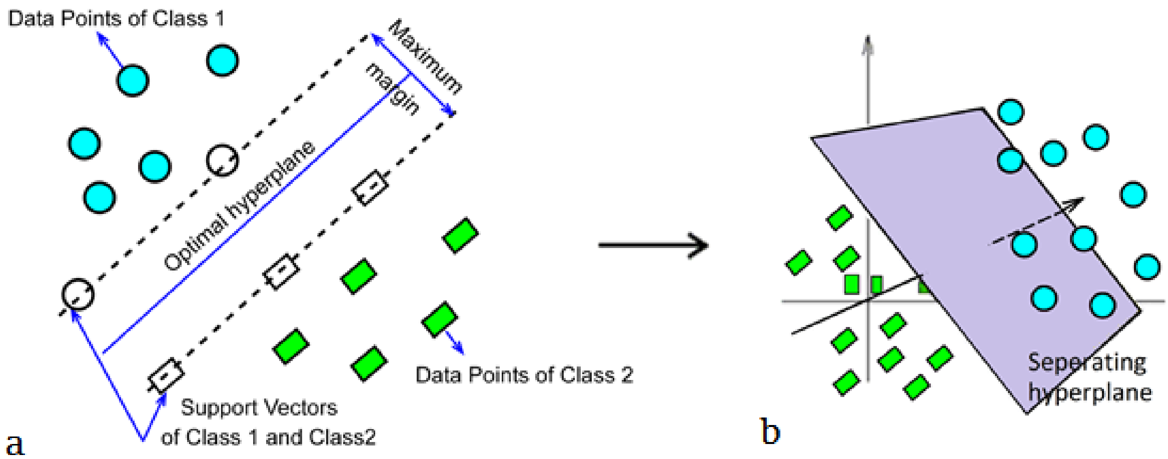

Fig. 2. Linear support vector machine example (modified from Burges (1998)).

SVM classifier offers: linear, polynomial, RBF, and sigmoid type kernels. The RBF kernel works well in most cases (ENVI Manual, 2004). The mathematical representation of each kernel is given in below equations [1-4],

$$Linear: K(x_i, x_j) = x_i T x_j \tag{1}$$

$$Polynomial: K(x_i, x_j) = (\gamma x_i T x_j + r)d, \gamma > 0 \tag{2}$$

$$RBF: K(x_i, x_j) = \exp(-\gamma \| x_i - x_j \| 2), \gamma > 0 \tag{3}$$

$$Sigmoid: K(x_i, x_j) = \tanh(\gamma x_i T x_j + \mathrm{r}) \tag{4}$$

where γ is the gamma term in the kernel function for all kernel types except linear, d is the polynomial degree term in the kernel function for the polynomial kernel, r is the bias term in the kernel function for the polynomial and sigmoid kernels, γ, d, and r are user depended parameters, as their correct definition significantly increases the accuracy of the SVM.

4. ACCURACY ANALYSIS

Error matrices and Cohen's kappa (*K*) are used for accuracy assessment. Kappa can be used as a measure of agreement between model predictions and reality (Congalton, 1991) or to determine if the values contained in an error matrix represent a result significantly better than random (Jensen, 1996). Kappa is computed using below equations;

$$K = \frac{\mathrm{N}\sum_{i=1}^{r} Xii - \sum_{i=1}^{r}(X_{i+} * + X_{+i})}{\mathrm{N}^2 - \sum_{i=1}^{r}(X_{i+} * + X_{+i})} \tag{5}$$

where N is the total number of sites in the matrix, r is the number of rows in the matrix, xii is the number in row i and column i, x+i is the total for row i, and xi+ is the total for column I (Jensen, 1996).

5. RESULTS & DISCUSSION

A comparison of spectral signatures among similar and different reflectance objects, illustrating the difficulty and simplicity in separating land-cover classes based on spectral signatures. 3D RGB surface plot and RGB profiles of the urban, sub-urban and rural areas indicates the complexity in the differentiation of the land cover objects. The case areas for the urban and sub-urban are located in the seaside for that reason the north-western part of the 3D plots and western part of the graphs represents water bodies. On the other hand, east and north-east part of the 3D plots which belongs to densely build up areas reveals of the difficulty in the differentiation of the land cover objects. Because in some places buildings are densely distributed especially in this areas buildings appear indistinguishable from roads, pavements and bare soil and may reflect fragmented characteristics due to shading or they may be occluded by other buildings. However, in the sub urban areas buildings are sparsely distributed as seen in its graphs this urban properties makes easy to differentiate sub-urban objects (Figure 3).

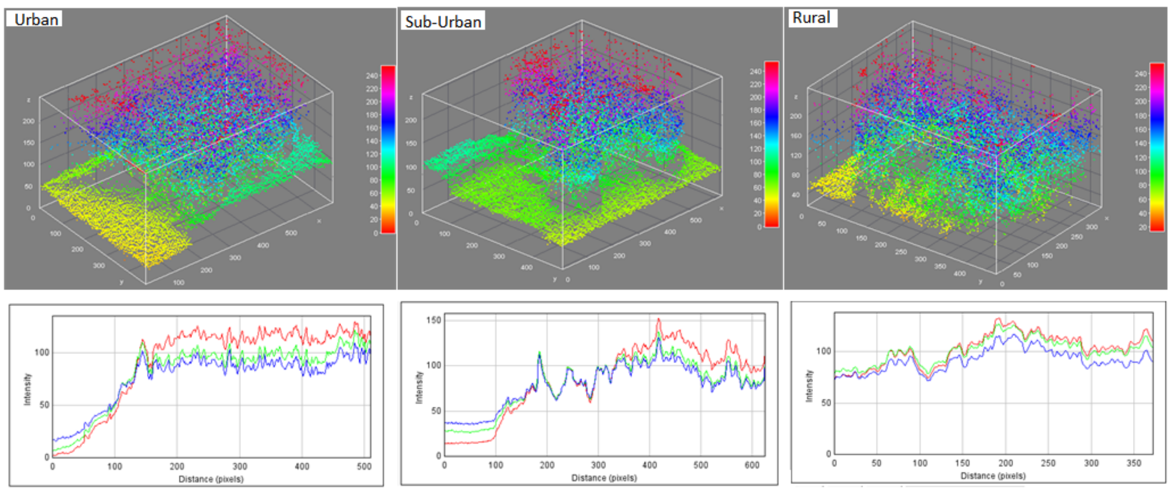

Figure 3. 3D RGB surface plot and RGB profiles of the urban, sub-urban and rural areas

For SVM classification training areas were created by choosing polygons that contain training pixels representing the land covers. This step is the most crucial part, since inaccurate training pixels can lead to serious misclassification. Although SVM can classify with only small training areas, in this study, medium to large training areas are used. This is due to the fact that such training areas tend to produce classification with a high accuracy. Visually, overall performance of SVM in land cover classification is good as it can classify all pixels effectively (Figure 4).

For accuracy assessment purposes, selection of ground truth pixels was done by random sampling. Accuracy analysis was carried by comparing the classified pixels with ground truth pixels using a confusion matrix. Table 2 shows confusion matrices for SVM classification in terms of number of pixels. The results were presented in terms of Kappa Coefficient and overall accuracy. According to results Kappa Coefficients of urban, sub-urban and rural areas are 0.79, 0.87 and 0.92, respectively. On the other hand overall accuracies of urban, sub-urban and rural areas are 88.17 %, 93.76 % and 95.96 % respectively. The results of the Kappa Coefficient and overall accuracies reveal that high spectral and spatial variation in the urban area, dense distribution of urban objects and shadow impacts of high rise buildings causes the decrease in the accuracies. The results of the study reveal that use of 2.5 m. Göktürk-2 image has important advances in urban studies.

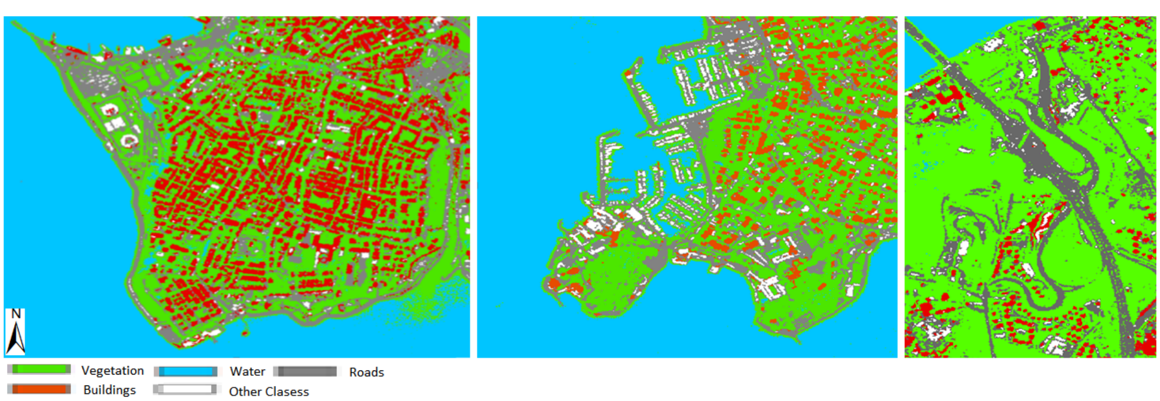

Figure 4. SVM classification results of urban, sub-urban and rural areas

Table 2. Error matrix of urban, sub-urban and rural SVM classification

URBAN AREA

	Ground Truth (Pixels)				
Class	Vegetation	Water	Road	Building	Other Classes
Vegetation	3639	2944	17	72	0
Water	90	16508	0	0	0
Road	29	0	1175	3	292
Building	5	0	8	3021	1
Other Classes	0	0	3	0	1487
Total	3763	19452	1203	3096	1780

Kappa Coefficient = 0.79 Overall Accuracy = (25830/29294) 88.17%

SUB-URBAN AREA

	Ground Truth (Pixels)				
Class	Vegetation	Water	Road	Building	Other Classes
Vegetation	5467	896	30	127	0
Water	743	24812	6	0	0
Road	140	25	1284	90	148
Building	96	0	3	3094	7
Other Classes	2	2	44	8	920
Total	6448	25735	1367	3319	1075

Kappa Coefficient = 0.87 Overall Accuracy = (35577/37944) 93.76%

RURAL AREA

	Ground Truth (Pixels)				
Class	Vegetation	Water	Road	Building	Other Classes
Vegetation	185173	0	62	2423	0
Water	1082	19012	1	0	0
Road	6882	0	58339	47	956
Building	506	0	83	17091	27
Other Classes	0	0	0	3	7193
Total	193643	19012	58485	19564	8176

Kappa Coefficient = 0.92 Overall Accuracy = (286808/298880) 95.96%

6. CONCLUSIONS

Overall, the SVM classification results of the Göktürk-2 were found very promising for urban and rural area classification. It has been shown that it can produce comparable or even better results than the other satellite imageries which have similar spectral, spatial and radiometric resolutions. Due to the quantity of details present at the 2.5 m resolution Gokturk-2 imagery, it enables analysis and mapping of the urban and rural area. The experimental results indicate a mean accuracy value around 80 % of the Göktürk-2 image for the urban and rural environment mapping which is very promising.

6.1 Acknowledgement

The study was supported by the Göktürk-2 satellite data which was acquired from Turkish Air Force (TUAF) under the theme study of "Determining Land Use and Industrial Zones of Marmara Region".

6.2. References

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2, pp. 121-167. Retrieved June 13, 2016, from http://dx.doi.org/10.1023/A:1009715923555.

Cristianini, N., Shawe-Taylor, J., 2000. An Introduction to Support Vector Machines and Other Kernelbased Learning Methods. Cambridge University Press. doi: http://dx.doi.org/10.1017/CBO9780511801389.

Congalton, R. G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37, pp.35-46.

ENVI Manual, 2004 . Retrieved June 13, 2016, from http://aviris.gl.fcen.uba.ar/Curso_SR/biblio_sr/ENVI_userguid.pdf.

Foody, G. M., Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sensing of Environment, 103 (2), pp. 179–189. doi: http://dx.doi.org/10.1016/j.rse.2006.04.001

Huang, C., Davis, L. S., Townshend, J. R. G., 2002. An Assessment of Support Vector Machines for Land Cover Classification. International Journal of Remote Sensing, 23 (4), pp. 725–749. doi: 10.1080/01431160110040323.

Jensen, J. R., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective (Second edition). Prentice Hall, Inc., Upper Saddle River, New Jersey, USA.

Joachims, T., 1998. Text categorization with support vector machines—learning with many relevant features. In Proceedings of the 10th European Conference on Machine Learning, Chemnitz, Germany. (Berlin: Springer), pp. 137–142.

Li Z., Gruen A., 2004. Automatic DSM Generation from Linear Array Imagery Data. XXth ISPRS congress, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Istanbul, Turkey, pp. 128-133. ISBN: 3-906467-55-4.

Lu, D., Hetrick, S., Moran, E., 2010. Land Cover Classification in a Complex Urban-Rural Landscape with QuickBird Imagery. Photogrammetric Engineering & Remote Sensing, October 2010, 76 (10), pp. 1159–1168.

Teke, M., 2016. Satellite Image Processing Workflow For Rasat And Göktürk-2. Journal Of Aeronautics And Space Technologies, January 2016, 9(1),pp.1-13.

Sagale, D., Kale, A., 2014. Combining Naive Bayesian and Support Vector Machine for Intrusion Detection System. International Journal of Computing and Technology, 1 (3), pp. 61-65.

Sugumaran, R., Zerr, D. and Prato, T, .2002. Improved urban landcover mapping using multitemporal IKONOS images for local government planning, Canadian Journal of Remote Sensing, 28, pp. 90–95.

Tuia, D., Ratle, F., Pozdnoukhov, A., Camps-Valls, G., 2010. Multisource composite kernels for urban-image classification. IEEE Geoscience and Remote Sensing Letters, 7 (1), pp. 88-92. doi: http://dx.doi.org/10.1109/LGRS.2009.2015341

Vapnik, V. N., 1995. The Nature of Statistical Learning Theory. (New York: Springer-Verlag).

Wang, L., Sousa, W.P., Gong, P. and Biging, G.S., 2004. Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of panama. Remote Sensing of Environment, 9, pp.432–440.

Zhan, Q., Molennar, M., Tempfli, K, 2002. Hierarchical object based structural analysis toward urban land use classification using high-resolution imagery and airbone LIDAR data. International Archives of Photogrammetry and Remote Sensing, 33(7), pp.1751-1758.

Zhang, X., Feng, X., 2005. Detecting urban vegetation from IKONOS data using an object oriented approach. In Proceedings of IGARSS 2005, Seoul, Korea, Jul 25-29, 2005.