5th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Evaluating kernel functions of support vector machines for supervised classification of land use classes

Sinan Bulut*[1]

[1]Çankırı Karatekin University, Faculty of Forestry, Department of Forest Management, Çankırı, Türkiye

Keywords
Support vector machine
Kernel functions
Land use
Sentinel-2
Supervised classification

Abstract

The aim of the study is to compare the accuracy of the kernel functions of the SVM method in terms of land use classification. The study was conducted in Abant Planning Unit within the north-west of Turkey. Supervised classification was performed using Sentinel-2 satellite image. Classification was made according to land use, and kernel functions of support vector machines method such as linear, polynomial, radial and sigmoid were used. According to the findings, the classification accuracies of the kernel functions were similar to each other. However, the sigmoid kernel function showed the highest classification success (kappa coefficient=0.775). When the confusion matrix was examined, the most accurately classified land classes were broadleaf forest, mixed forest, and other areas. Kernel functions were insufficient in classifying coniferous and degraded forests.

1. Introduction

Developing technology, satellite systems and new techniques that have been put into practice are developing and diversifying over time. Remote sensing has also been at the center of these developments and is used in many different study subjects. It has an intensive use area, especially in the field of forestry. Classification of satellite images and estimation of stand parameters are applications that have been extensively studied in forestry. There are many different remote sensing data that can be used in these studies and different techniques that can be used for these data. Especially in supervised and unsupervised classification, different satellite images and algorithms are frequently used. In studies on supervised classification, satellite images such as Landsat 8, Sentinel-2, IKONOS, Quickbird are used extensively. Supervised classification methods such as maximum likelihood, minimum distance, mahalanobis distance, support vector machine (SVM), neural net and decision trees are also used with these satellite images (Kavzoğlu and Çölkesen 2010; Otukei and Blaschke 2010; Srivastava et al. 2012; Taati et al. 2014; Üstüner et al. 2015; Kulkarni and Lowe 2016).

This study is comparative analysis between kernel functions of SVM technique. The aim of this study is to classify land use classes using different kernel functions of the SVM method based on Sentinel-2 satellite image and to determine the most appropriate approach.

2. Method

In this study, the Sentinel-2 satellite image and the stand map produced as a result of the forest management inventory were used as material. Kernel functions of the SVM method were used for land use classification.

2.1. Study area

The study was carried out at the Abant Planning Unit within the Bolu Forest Regional Directorate. The area is located in the north-west of Turkey (Fig. 1). The study area is totally 6319.1 ha. Approximately 82 percent of the study area is covered by forests (5179.9 ha). Areas of coniferous forest (CF), broadleaf forest (BF), mixed forest (MF), degraded forest (DF), and other areas (OA) are 981.5, 1011.2, 3090.3, 96.9 and 1139.2 ha, respectively (Fig. 2). Most of the study area consists of mixed stands (48.9%). The smallest areal distribution belongs to the stands with degraded structure (1.5%). The areal distribution of coniferous stands, broadleaf stands and other areas is close to each other.

* **Corresponding Author**

*(sbulut@karatekin.edu.tr) ORCID ID 0000-0001-6149-0910

Cite this study

Bulut S (2022). Evaluating kernel functions of support vector machines for supervised classification of land use classes. 5th Intercontinental Geoinformation Days (IGD), 37-40, Netra, India

Figure 1. Location of the study area

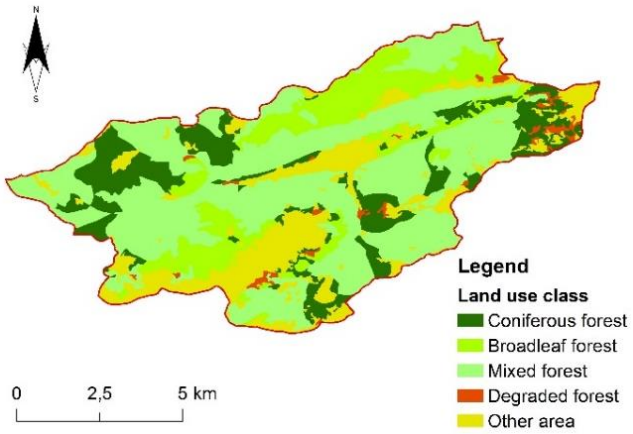

Figure 2. Land use map of the study area

2.2. Sentinel-2 satellite image

Sentinel-2 satellite image acquired 29 August 2015 was obtained free of charge from the USGS Earth Explorer data portal. A total of 10 bands from Sentinel-2 were used for classification. Bands 2, 3, 4 and 8 with a spatial resolution of 10 m, and bands 5, 6, 7, 8A, 11 and 12 with a spatial resolution of 20 m were combined and made ready for classification.

2.3. Support vector machine

The SVM, which is one of the supervised classification algorithms, is based on the structural risk minimization principle and statistical learning theory. SVM was designed for the classification of two-class linear data, then it was developed to solve the classification problem of multi-class and non-linear data. The aim in SVM is to obtain the optimal hyperplane that can separate the two classes. Different kernel functions such as linear, polynomial, radial and sigmoid can be used in a classification process using SVM.

2.4. Supervised classification of the Sentinel-2 satellite image

In this study, linear, polynomial, radial and sigmoid kernel functions of SVM were used to perform supervised classification. Stand map was used as reference data for the study area. The land uses of the study area were divided into five classes by using the stand map. These classes were CF, BF, MF, DF, and OA. Ten training fields were taken from Sentinel-2 satellite image for each land use classes. The same training fields were used in all classification processes to ensure equivalence in the comparison of applied linear, polynomial, radial and sigmoid kernel functions. ENVI 5.2 software was used for the processes carried out for the supervised classification.

3. Results

Overall accuracy and kappa coefficient were presented for linear, polynomial, radial and sigmoid kernel functions (Table 1). Overall accuracies were obtained between 82.82 and 83.80%. Kappa statistics were found from 0.765 to 0.775. Classification successes were similar among to kernel functions. But the highest kappa statistics was found for sigmoid kernel function (0.775).

Table 1. Classification performance of the kernel functions

Kernel function type	Overall accuracy	Kappa coefficient
Linear	83.46%	0.769
Polynomial	82.82%	0.768
Radial	83.25%	0.765
Sigmoid	83.80%	0.775

Confusion matrices for linear, polynomial, radial and sigmoid kernel functions were showed in Table 2-5. When the matrices were examined, it was seen that the kernel functions classified the CF class incorrectly. Kernel functions have classified the CF class mostly as MF class. The CF class was best classified by the polynomial function. BF class was distinguished by all kernel functions and classified with high accuracy. BF class was distinguished and classified with high accuracy. All kernel functions for the BF class showed similar success. The best classification for the MF class was made by the radial function, and the worst classification was by the polynomial function. Kernel functions were not successful in distinguishing the DF class. DF class was misclassified by the kernel functions and assigned to BF and MF classes. The classification success rate was high for the OA class. The most successful kernel function in classification for OA was linear. The most unsuccessful function was the sigmoid.

Table 2. Results of the confusion matrix for linear function

Class	CF	BF	MF	DF	OA	Total
CF	75	0	51	0	0	126
BF	3	1919	7	74	8	2011
MF	1051	5	2549	109	4	3718
DF	0	1	2	6	5	14
OA	2	3	1	11	2199	2216
Total	1131	1928	2610	200	2216	8085

Table 3. Results of the confusion matrix for polynomial function

Class	CF	BF	MF	DF	OA	Total
CF	580	0	601	30	0	1211
BF	3	1919	7	74	9	2012
MF	547	5	1998	81	8	2639
DF	0	1	2	3	3	9
OA	1	3	2	12	2196	2214
Total	1131	1928	2610	200	2216	8085

Table 4. Results of the confusion matrix for radial function

Class	CF	BF	MF	DF	OA	Total
CF	27	0	11	0	0	38
BF	3	1919	7	74	11	2014
MF	1100	5	2589	113	8	3815
DF	0	1	2	2	3	8
OA	1	3	1	11	2194	2210
Total	1131	1928	2610	200	2216	8085

Table 5. Results of the confusion matrix for sigmoid function

Class	CF	BF	MF	DF	OA	Total
CF	182	0	109	3	0	294
BF	4	1920	7	75	16	2022
MF	945	5	2491	114	16	3571
DF	0	2	2	2	4	10
OA	0	1	1	6	2180	2188
Total	1131	1928	2610	200	2216	8085

The classified land use maps obtained by the kernel functions were presented in the Fig. 3. When the maps were examined visually, it was seen that the DF class with limited area cannot be classified well. It has been seen that the CF class was assigned to more areas than it should be with the polynomial function. In linear and radial function maps, the CF class was classified independently and very dispersedly. The most stable map for the CF class was obtained with the sigmoid function. OA class was mapped similarly by all kernel functions.

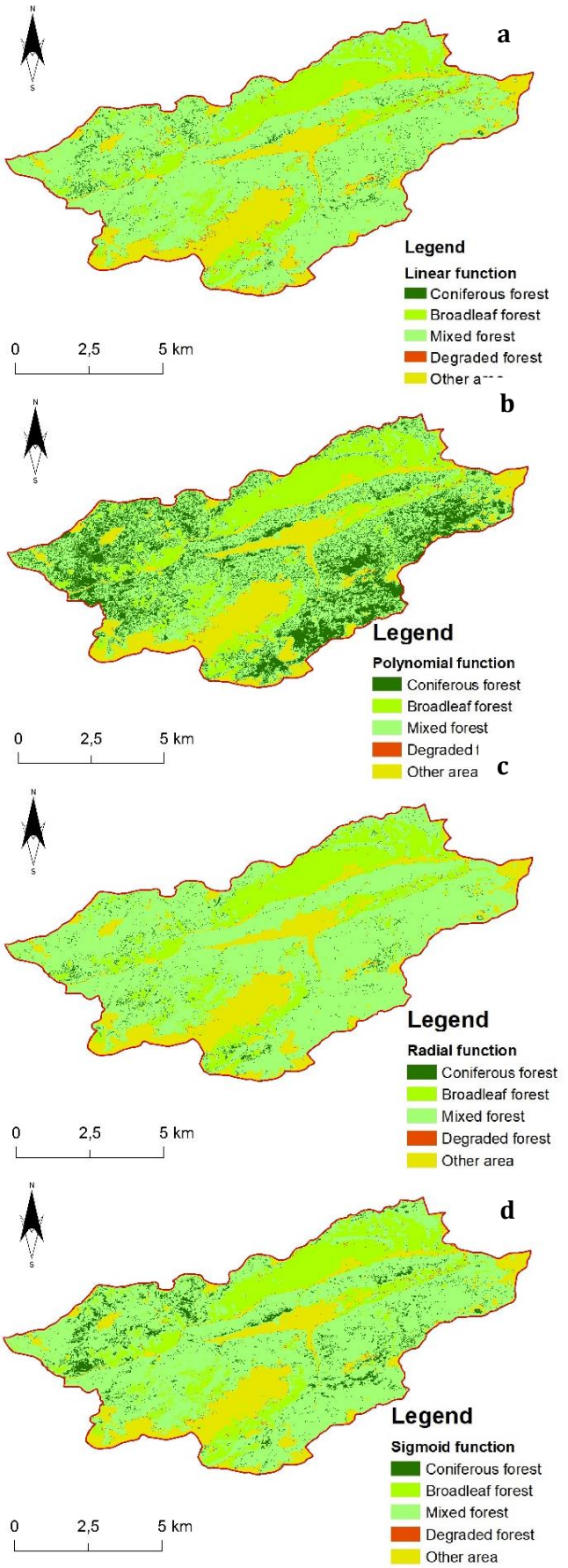

Figure 3. Classified land use maps by linear (a), polynomial (b), radial (c) and sigmoid (d) functions

4. Discussion

In the study, controlled classification of Sentinel-2 satellite image was made for land use classes. Linear, polynomial, radial and sigmoid kernel functions of SVM were used as classifier, and forest management stand map as referenced. Kappa statistics values ranged between 0.765 and 0.775. Highest success was achieved with the sigmoid function in terms of overall accuracy and kappa statistics. In this study, land use classification with success of 80 percent or higher could not be made. There are many studies in the literature on land use classification. Deilmai et al. (2014) generated land use and land cover maps using maximum likelihood (ML) and SVM method. They reported that the SVM method was better than ML method. Kappa statistics were 0.65 and 0.86 for ML and SVM method, respectively. Huang et al. (2002) classified to land cover classes with MODIS satellite imagery using SVM, maximum likelihood, artificial neural networks, and decision trees methods. The highest success was obtained with the SVM method, and the overall classification success was 75.62%. Kesikoğlu et al. (2019) used three different classifier method such as ML, ANN and SVM technique and land cover types were classified for 2005 and 2012. The highest success was found for ANN and SVM methods.

In the study, classification successes for CF and DF were quite low. Since the spatial resolution of the satellite image used was medium, coniferous areas in mixed stands can be distinguished. Therefore, the CF class was mostly classified as MF. The DF class had very limited area in the region. Because of this, kernel functions had a hard time separating these small areas from the others. The resolution characteristics of the satellite image such as spatial, radiometric, and spectral used affect the classification to a high degree. Especially when the spatial resolution of the satellite image is high, smaller details can be distinguished and classification can be made more sensitive. Lower resolution images can be used when the scale of the object to be classified is coarse. Therefore, it is important to select the appropriate satellite image according to the details of the classification in terms of classification success (Günlü 2012; Bulut and Günlü 2016; Abbas and Jaber 2020; Jamali and Karas 2022).

5. Conclusion

Kernel functions of the SVM method were used for land use classification. Classification success of all kernel functions was close to each other, and they were not especially successful in classifying forest types. Therefore, the dissemination of such studies, the use of satellite images with different resolutions and classification techniques will provide more accurate information for the selection of appropriate techniques and materials.

Acknowledgement

I would like to thank the Turkish General Directorate of Forestry for providing data.

References

Abbas, Z., & Jaber, H. S. (2020). Accuracy assessment of supervised classification methods for extraction land use maps using remote sensing and GIS techniques. In IOP Conference Series: Materials Science and Engineering (Vol. 745, No. 1, p. 012166). IOP Publishing.

Bulut, S., & Günlü, A. (2016). Arazi kullanım sınıfları için farklı kontrollü sınıflandırma algoritmalarının karşılaştırılması. Kastamonu University Journal of Forestry Faculty, 16(2), 528-535.

Deilmai, B. R., Ahmad, B. B., & Zabihi, H. (2014). Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia. In IOP conference series: Earth and environmental science (Vol. 20, No. 1, p. 012052). IOP Publishing.

Günlü, A. (2012). Landsat TM uydu görüntüsü yardımıyla bazı meşcere parametreleri (gelişim çağı ve kapalılık) ve arazi kullanım sınıflarının belirlenmesi. Kastamonu University Journal of Forestry Faculty, 12(1), 71-79.

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),725-749.

Jamali, A., & Karas, İ. R. (2022). Land Use Land Cover Mapping in Support of Land Degradation Mapping Using Tree-Based Classifiers. In Environmental Degradation in Asia (pp. 3-16). Springer, Cham.

Kavzoğlu, T., & Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. Harita Dergisi, Temmuz 2010, 144:73-82.

Kesikoglu, M. H., Atasever, U. H., Dadaser-Celik, F., & Ozkan, C. (2019). Performance of ANN, SVM and MLH techniques for land use/cover change detection at Sultan Marshes wetland, Turkey. Water Science and Technology, 80(3), 466-477.

Kulkarni, A. D., & Lowe, B. (2016). Random forest algorithm for land cover classification. Computer Science Faculty Publications and Presentations. Paper 1. http://hdl.handle.net/10950/341.

Otukei, J., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int J Appl Earth Obs Geoinf 12: 27–S31.

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M., & Islam, T. (2012). Selection of classification techniques for land use/land cover change investigation. Adv Space Res 50(9):1250–1265.

Taati, A., Sarmadian, F., Mousavi, A., Pour, C. T. H., & Shahir, A. H. E. (2014). Land use classification using support vector machine and maximum likelihood algorithms by Landsat 5 TM images. Walailak Journal of Science and Technology, 12(8): 681-687.

Üstüner, M., Şanlı, F. B., & Dixon, B. (2015). Application of support vector machines for landuse classification using high-resolution RapidEye images: a sensitivity analysis. European Journal of Remote Sensing, 48: 403-422.