16| Оригинални научни рад
Research paper
doi 10.7251/STP1813247J ISSN 2566-4484

KARTIRANJE KORIŠĆENJA ZAMLJIŠTA/ZEMLJIŠNOG POKRIVAČA UPOTREBOM SENTINEL-2 I ALGORITAMA ZA MAŠINSKO UČENJE

Gordana Jakovljević, *gordana.jakovljevic@aggf.unibl.org*, University of Banja Luka, Faculty of Architecture, Civil Engineering and Geodesy
Dragoljub Sekulović, *sekulovicdr@yahoo.co.uk*, University of Defense
Miro Govedarica, *miro@uns.ac.rs*, University of Novi Sad, Faculty of Technical Science

Abstract: Korišćenje zemljišta/zemljišni pokrivač (LULC) ima značajan uticaj na degradaciju zemljišta, eroziju i dostupnost vode, stoga kartiranje prostorne raspodjele LULC je od suštinskog značaja za upravljanje zemljištem. Precizno kartiranje LULC klasa pomoću upotrebom daljinske detekcije zahtijeva robusne metode klasifikacije. Posljednjih se godina koriste razni algoritmi klasifikacije i satelitske slike. Za ovu studiju korišćeni su Sentinel-2 satelitski snimci umjerene prostorne rezolucije. Kako bi se ocijenio potencijal ulaznih snimaka i izradila karta korišćenja zemljišta u složenom urbanom području Banja Luke, Republika Srpska s najvećom mogućom preciznošću, primijenjeni su algoritmi za mašinsko učenje: Supported Vector Machine (SVM) i Random Forst (RF). Ukupna tačnost klasifikacije od 90,82% s kappa vrijednošću 0,87 i 88,29 s kappa vrijednšću od 0,84 postignuta je pprimjenm SVM i RF. Rezultati, prezentovani u ovom radu, su pokazali da algoritmi za mašinsko učenje i Sentinel-2 satelitski snimic mogu efiksno koristiti za izradu karata korišćenja zemljišta. Rezultati, prezentovani u ovom radu, su pokazali da algoritmi za mašinsko učenje i Sentinel-2 satelitski snimic mogu se efiksno koristiti za izradu karata o korišćenju zemljišta.

Keywords: korišćenje zemljišta/zemljišni pokrivač, Sentinel 2, SVM, RF

LAND USE / LAND COVER MAPPING FROM SENTINEL 2 DATA USING MACHINE LEARNING ALGORITHMS.)

Abstract: Land cover/land use (LULC) have an important impact on land degradation, erosion and water availability therefore mapping of patterns and spatial distribution of LULC is essential for land management. Accurate mapping of complex land cover and land use classes using remotely sensed data requires robust classification methods. Various classification algorithms and satellite images have been used in recent years. For this study, moderate resolution Sentinel-2 image was used. In order to evaluate the potential of the input image and derive land cover map in complex urban area of Banja Luka, Republic of Srpska with highest possible precision, two machine learning algorithms where applied: Supported Vector Machines (SVM) and Random Forst (RF). An overall classification accuracy of 90,82% with kappa value of 0,87 and 88,29 with kappa value of 0,84 was achieved using SVM and RF. The study showed that of machine learning algorithms on Sentinel-2 imagery can results in accurate land cover maps.

Keywords: land cover/ land use, Sentinel 2, SVM, RF

1. INDRODUCTION

Land cover refers to the physical characteristics of earth's surface, captured I distribution of vegetation, water, settlements while land use refers to the way in which land has been used by humans and their habitat [1]. Understanding landscape patterns, changes and interaction between human activities and natural phenomena are essential for proper land management, decision improvement and understanding different changes such as degradation level of forest and wetlands, rate of urbanization, intensity of agricultural activities, landslide, erosion e.g. Land use/ land cover classification is time-consuming and expensive process. Remotely sensed data from various earth observation satellites can provide accurate and timely geospatial information of urban and per-urban areas at diverse spectral, spatial and temporal scales [2]. In recent decades, remotely sensed dataset are become an attractive alternative to ground based survey and mapping methods for documentation, characterization and quantification of the LULC. The accuracy of the produced maps is affected by spatial and spectral resolution of remotely sensed imagery, cloud cover, quality of training data and image classification techniques. RS data classification is based on a unique relationship between a land cover class and its reflected radiation at certain wavelength (reflectance) contained in a spectral band of an image, e.g., a one-pixel–one-class relationship [3]. The classification of LULC in urban areas is challenging dual heterogeneous landscapes and multiple object within pixel. Mixed pixel is a common confounding factor in classification using moderate resolution datasets due to the large pixel size. To resolve these issues investigators have utilize different supervised classification techniques such as maximum likehood, artificial neutral network, decision tree, supported vector machine, random forest etc. Table 1. Shows application of different classification techniques for LULC class delineation.

Table 1. Classification techniques for LULC class delination

Authors	Platform	Classification technique	Accuracy	
			Kappa	Overall
[4]	Rapideye	SVM RF	0.92 0.92	91.80% 93.07%
[5]	AwiFS	Decision tree		91.81%
[6]	WorldView 2	RF ORF-SVM 1A11	0.82 0.83	83.70% 85.20%
[7]	Landsat	RF	0.73	82,00%
[8]	Sentinel 1	SVM RF K-NN	0.75 0.73 0.73	80.80% 79.60% 79.48%
[9]	Landsat 8	SVM MLC	0.89 0.88	91.3% 90.4%
[10]	Landsat	ANN SVM		97.15% 96.25%
[11]	Landsat	STARFM		74.50%
[20]	Landsat TM	RF ML	0.92 0.83	

Whereas the application of classification algorithm as random forest and supported vector machine has been explored in LULC classification using multispectral imagery, there is a paucity of knowledge on the performance od those algorithm on medium resolution Sentinel-2 imagery. This study compared the performance of RF and SVM classifiers on the Sentinel 2 image in a heterogeneous urban landscape of Banja Luka, Republic of Srpska. The main objectives of the study are to assess (i) accuracy of RF and SVM algorithm for LULC classification and (ii) potential of medium resolution open source Sentinel-2 image for LULC mapping.

2. MATERIALS AND METHODS

2.1. STUDY AREA

The City of Banja Luka is the study area chosen for this paper. It is located in the South-Western part of the Republic of Srpska, i.e. in the Western part of Bosnia and Herzegovina. The City of Banja Luka is the biggest political and territorial unit occupying 1239 km2 with the population of 250000. Situated in a basin 164 m above sea level, where the Dinaric Alps from the south descend into the Pannonian Basin in the north, Banja Luka has temperate continental climate with the prevailing influences from the Pannonian plain [27]. The biggest part of study area is covered by forest and agricultural land.

1 oblique random forest using support vector machines and the one against one combination approach;

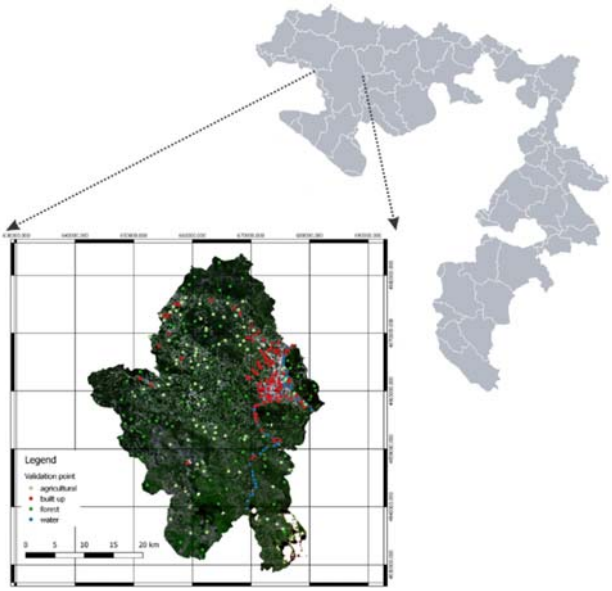


Figure 1. Study area

2.2. Sentinel-2

SENTINEL-2 is a European wide-swath, high-resolution, multi-spectral imaging mission. The full mission specification of the twin satellites flying in the same orbit but phased at 180°, is designed to give a high revisit frequency of 5 days at the Equator. Sentinel-2 multispectral images are used in this study. Sentinel-2 images consists of 13 spectral bands: four bands at 10 m (B2, B3, B4 visible and B8 Near infrared specter)), six bands at 20 m (B5, B6, B7, B8a Near Infrared and B11, B12 Shortwave Infrared) and three bands at 60 m spatial resolution [28]. The atmospherically and terrain corrected Bottom-of-atmosphere BoA (surface) reflectance Sentinel 2 Level 2A image from 18.05.2017. used in classification are provided from [29]. The input data comprise reflectance values of the tree 10 m bands (Band 2, 3 and 4) resampled at 20 m and the reflectance value of the 20 m bands (Band 8a, 11 and 12).

2.3. Software environment R

SVM and RF classification are performed in R. R is a language and environment for statistical computing, graphics and data manipulation and is available as free software under the terms of the Free Software Foundation's GNU General Public License in source code form. R can be extended via packages. Package contains code, data, documentation, tests. Linking code to a package makes it easy to share with other users who can easily capture, install and learn how to use it.

In this paper following packages are used: sp (provide plotting spatial data as maps, spatial selection, summary, print), raster (provide reading, writing, manipulating, analyzing and modeling of gridded spatial data), rasterVis (provide methods for enhanced visualization and interaction with raster data), caret (set of function for creating predictive models), snow (Support for simple parallel computing in R), rgdal (provides bindings to GDAL and

access to projection/transformation operations from the PROJ.4), randomForest and e1071 (Functions for latent class analysis, short time Fourier transform, fuzzy clustering, support vector machines, shortest path computation, bagged clustering, naive Bayes classifier…).

2.4. Classification

A pixel-based image analysis was carried out in order to identify four classes (water, forest, built up and agricultural), using a Support Vector Machine and Random Forest learning algorithm. The ground truth samples (training data) were located according to Google Earth. The same training dataset was used for both classification approaches. The 214 training polygons was created.

2.4.1. SUPPORTED VECTOR MACHINE

Supported Vector Machine is a supervised machine learning algorithm, proposed by Vapnik [12], which can be used for both classification and regression. SVM is suitable to distinguish the patterns and objects and it can be used for pixel-based and object-based classification. Satellite image classification with SVM require a training data. Training data are represented by $\{x_i, y_i\}, i = 1,\ldots..r, y_i = \{1,-1\}$ where r is a number of training samples and Training vector consists of two classes $y_i = 1$ for class α_1 and $y_i = -1$ for class α_2. If classes are linearly separable it is possible to define at least one hyperplane defined by vector w with bias b which can separate the classes properly (training error is 0) (1):

$$w \cdot x + b = 0 \tag{1}$$

To find such hyperplane w and b are estimated in the way that $y_i(w \cdot x_i + b) \geq 1$ for $y_i = 1$ (class α_1) and $y_i(w \cdot x_i + b) \geq -1$ for $y_i = -1$ (class α_2). These two can be combined to provide (2):

$$y_i(w \cdot x_i + b) - 1 \geq 0 \tag{2}$$

There are many hyperplanes which could be fitted to separate two classes but there is only one n dimensional optimal hyperplane. Optimal hyperplane between two classes is founded by maximizing the gap between the classes closest point (see Figure 2). The training points who are closest to the optimal hyperplane and lying on the two boundaries, given with $w \cdot x_i + b = \pm 1$, are called support vectors and the middle of the margin is optimal separating hyperplane.

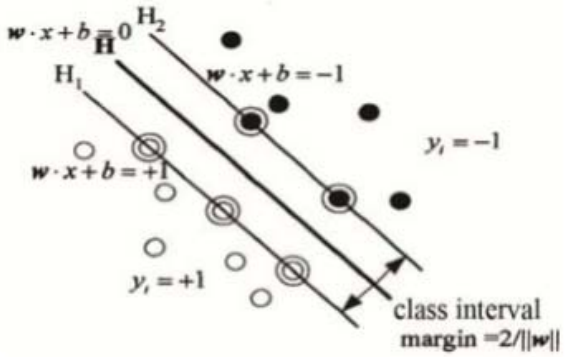


Figure 2. Optimal hyperplane [13]

Mathematically, this means that we want maximize the distance between supported vectors. This distance is equal to $\frac{2}{\|w\|}$. This is expressed as:

$$\min \frac{1}{2}\|w\| \tag{3}$$

Subject to following constraints:

$y_i(w \cdot x_i + b) \geq 1$

Where |(|w|)| is the norm of the hyperplane. Using the Lagrangian multiplier, the cost function can be defined as (4):

$$\frac{1}{2}\|w\|^2 - \sum_{i=1}^{r} a_i \left(y_i\left((w \cdot x_i) + b\right) - 1\right) \tag{4}$$

Where a_i is the Lagrangian multiplier.

For the non-linearly separable classes, and the constrain of equation 2 cannot be satisfied. To deal with such cases using only linear separate boundaries set of new variables that the distance the case is from the optimal hyperplane and so the amount of coloration of the consistence may be introduces [14].

The Equation (1) becomes (5):

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{r} \xi_i \tag{5}$$

Under the constraints of $y_i(w \cdot x_i + b) \geq 1 - \xi_i$ i=1,…, r. Where C controls the magnitude of the penalty associated with training samples that lie on the wrong side of the decision boundary. To generalize the above method to non-linear discriminant functions, the Support Vector Machine maps the input vector x by non-linear mapping $\phi(x)$ into a high-dimensional feature space and then constructs the optimal separating hyperplane in that space [15]. According to the Marcer's theorem the inner product of the vectors in the mapping space can be expressed as a function of the inner products of the corresponding vector in the original space [16]. The inner product operation (6):

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

Where is $K(x_i, x_j)$called kernel function. Radial basis function defined with [14], [16], [17] (7):

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

is one of the most powerful kernels. If we choose $K(x_i, x_j)$ the non-linear SVM is reduced to its linear version. The classical linearly constrained optimization problem can be translated (using a Lagrangian formulation) into following dual problem (8):

Maximize: $$\sum_{i=1}^{r} a_i - \frac{1}{2}\sum_{i=1}^{r}\sum_{i=1}^{r} a_i a_j y_i y_j K(x_i, x_j) \tag{8}$$

Subject to $\sum_{i=1}^{r} y_i a_i = 0$ $0 \le a_i \le C$, i=1,…, r. a_i is a Lagrange multiplier corresponding with each constraint in original problem.

Using appropriate kernel function in optimal classification surface can achieve linear classification after nonlinear transformation, while computational complexity does not increase. The final result is the discriminant function f(x) conveniently expressed as a function of the data in original (lower) dimensional feature space (9):

$$f(x) = \sum_{i \in S} a_i y_i K(x_i, x_j) + b \quad (9)$$

For this research, SVM was run in R software. The parameters which had to be defined in order to apply the algorithm were: input data (all bands from the Sentinel-2 image), SVM-type (C-classification), SVM-kernel (Radial Basis Function), cost (C = 100) and gamma (γ=0.5). The determination of parameters was solved by cross validation and grid search on the training data set.

2.4.2. RANDOM FOREST

The random forests algorithm is a machine learning technique proposed by Breiman [18], Consists of a collection of tree-structured classifiers $\{h(x, \Theta_k), k = 1,...\}$ where the $\{\Theta_k\}$ are independent identically distributed random vectors and each tree casts a unit vote for most frequent class to the input vector (x). A RF uses a random subset of input features or predictive variables in the division of every node, instead of using the best variables, which reduces the generalization error. Additionally, to increase the diversity of the trees, a RF uses bagging or bootstrap aggregating to make the trees grow from different training data subsets [20]. In bagging a each randomly selected subset (without replacement) of certain proportion of the training dataset is used to grow each tree [19]. The samples which are not used in the training subset are included as part of another subset called out-of-bag (oob). OOB elements can be classified by the tree to evaluate performance. RF use the Gini Index as a measure for the best split selection, which measures the impurity of a given element with respect to the rest of the classes [18]. For a given training dataset T, the Gini Index can be expressed as:

$$\sum \sum_{j \neq i} (f(C_i, T)/|T|)(f(C_j, T)/|T|) \quad (10)$$

Where $f(C_i, T)/|T|$ is the probability that a selected case belongs to class C_i. Thus, by using a given combination of features, a decision tree is made to grow up to its maximum depth (with no pruning). Hence, RF, as it grows without pruning, presents an added advantage. This RF also provides an assessment of the relative importance of the different features or variables during the classification process [21]. To assess the importance of each feature (e.g. satellite image band), the RF switches one of the input random variables while keeping the rest constant, and it measures the decrease in accuracy which has taken place by means of the oob error estimation and of Gini Index decrease [18].

The RF algorithm was implemented using carter and raster package within R. Two parameters are required to construct an RF framework: the number of decision tree (k) in the ensemble and the number of input predictors (m) randomly selected at each node.

2.5. Accuracy assessment

Confusion (error) matrix is frequently used for standard pixel-based accuracy assessment. Confusion matrix is simple cross tabulation of the predicted class label against the reference data for a sample of cases at the specific locations, it provides a foundation on which both classification accuracy and characterize errors can be define [22].

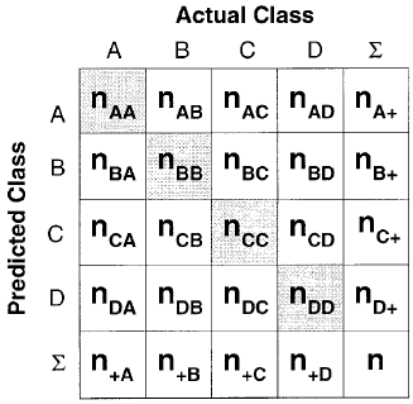


Figure 3. Confusion matrix. Diagonal of the matrix contain the number of pixel correctly classified for each class, whereas the off-diagonal elements represent pixel where is disagreement in the predicted and actual class [23].

Many measures of classification accuracy can be derived from a confusion matrix: kappa coefficient, overall, commission and omission error.

Overall accuracy describes the proportion of the total number of correct classified for all class and total number of pixel in confusion matrix (sum of diagonal members of matrix divided by total sum of pixels) Eq (11). Commission error (CE) and Omission error (OE) describe the errors related to individual classes. CE represents pixels that belong to another class but are labeled as belonging to the target class (i.e. the percentage of pixels classified as water but which do actually not belong to that class). OE represents the pixels that belong a class but fail to be classified into that class (i.e. percentage of pixels which are water but which were not classified as such)

$$Overallaccuracy = \frac{\sum_{i=1}^{k} n_{ii}}{n} \cdot 100 \qquad (11)$$

Overall accuracy have often been criticized because in the some cases may have been allocated to the correct class purely by chance [23]. For highlights the difference between correctly classified pixels and the chance agreement presented by sum of the rows and columns Cohen [24] introduced Kappa statistics Eq(12). Kappa statistics is used as a measure of classification accuracy reduced for accidentally correct class agreement.

$$Kappacoefficient = \frac{n\sum_{i=1}^{k} n_{ii} - \sum_{i=1}^{k} n_{i+}n_{+i}}{n^2 - \sum_{i=1}^{k} n_{i+}n_{+i}} \quad (12)$$

Registered value reflects the overall classification accuracy and consistency between the image and the reference grid with a random distribution of pixels in the classes. Interpretation of Kappa coefficient proposed by Landis and Koch [25] is showed in Table 3.

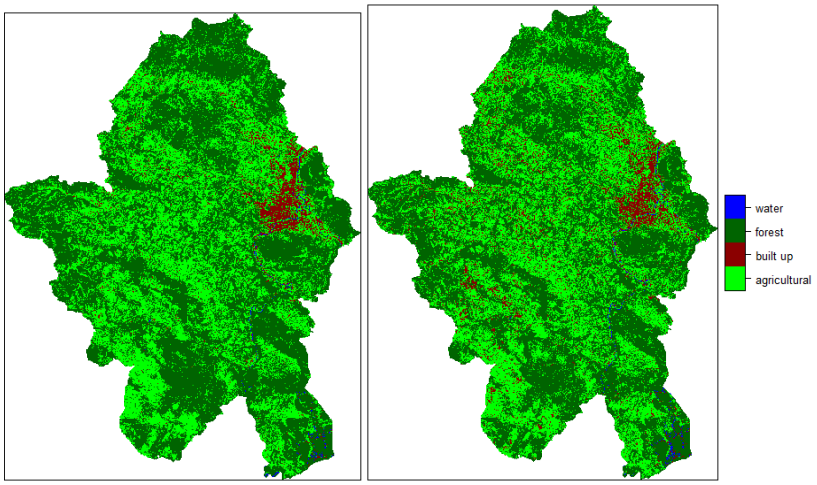
Table 2. Kappa coeficient interpretation [25]

Kappa coefficient value	Interpretation
$0.81 \leq \widehat{K} \leq 1$	Perfect agreement
$0.61 \leq \widehat{K} \leq 0.8$	substantial agreement
$0.41 \leq \widehat{K} \leq 0.6$	moderate agreement
$0.21 \leq \widehat{K} \leq 0.4$	Fair agreement
$0.0 \leq \widehat{K} \leq 0.2$	Poor agreement

It is a well-known phenomenon in binary classification that a training set consisting of different numbers of representatives from either class may result in a classifier that is biased towards the more frequent class. The balanced accuracy can be defined as the average accuracy obtained on either class which avoids inflated performance estimates on imbalanced datasets [26].

3. RESULTS AND DISCUSION

The results of classification are presented in Figure 3 .



(a) (b)

Figure 4. Results of Sentinel 2 classification based on (a) SVM (b) RF

The results of the image classification was validated by pixel-by-pixel accuracy assessment based on validation points. The validation points were selected over different locations representing different land cover/land use classes. A 316 validation points where created from which 89, 104, 75 and 48 respectively for agricultural, built up, forest and water class. Same points are used for the RF and SVM classifier. Table 4 provide overall and per class classification accuracies achieved by random forest and SVM classifier.

Table 3. Accuracy assessment of Sentinel 2 classification based on SVM and RF

Algorithm	Class	Kappa	Overall accuracy	Omission	Commission	Balanced Accuracy
SVM	Agricultural	0.87 (0.78-0.88)	90.82 (83.89-91.44)	14.61 (6.78-22.44)	21.84 (14.33-29.35)	0.90
	Built up			10.58 (4.21-16.95)	6.06 (0.83-11.29)	0.93
	Forest			5.33 (-0.48-11.14)	13.41 (5.49-21.33)	0.95
	Water			2.08 (-3.43-7.59)	2.08 (-3.43-7.59)	0.99
RF	Agricultural	0.84 (0.81-0.90)	88.29 (84.22-91.62)	17.98 (9.52-26.43)	21.35 (12.90-29.80)	0.87
	Built up			13.46 (6.46-20.46)	12.62 (5.76-19.48)	0.90
	Forest			2.67 (-1.82-7.16)	3.95 (-1.32-9.22)	0.97
	Water			0	2.08 (-3.42-7.59)	0.99

Both classifier produce high overall accuracy and perfect agreement with reality. Although, SVM classifier had a higher accuracy than the RF there were no significant difference ($p<0.05$) in the agreement with reality between the classification obtained using RF and SVM classifier. In general water class achieved lowest value (higher accuracy) of commission and omission error for both classifier while lowest accuracy was produced for agricultural class. RF classifier produce higher commission error for built up class (12.62%) then the SVM (6.06%) while SVM produce higher omission and commission error for forest class (5.33;13.41) then RF algorithm (2.67; 3.95). The high omission error of the agricultural class was caused by the confusion with built up since some crop surface have similar spectral signature as built up. Results of accuracy assessment were validated through computation of area of delineated LULC classes. Results are shown in Table 4.

Table 4. Estimated area of LULC classes [km2]

Class	RF	SVM	RF-SVM
Agricultural	507.78	516.76	-8.98
Built up	72.16	65.91	6.25
Forest	654.27	652.25	2.02
Water	4.58	3.88	0.70

4. CONCLUSION

This research evaluated the potential of space borne multispectral sensor Sentinel-2, with advanced classification techniques RF and SVM to delineate land cover/land use classes in a heterogeneous urban landscape. Results in this study showed that RF and SVM classifier can be successfully used for land use/land cover mapping from Sentinel 2 imagery achieving a kappa coefficient of 0.84 and 0.87 respectively. Comparing the two classifiers, results demonstrate a slightly better overall performance of SVM. Classification results illustrated that water class could be successfully identified using both RF and SVM while highest omission and commission error is detected for agricultural class.

REFERENCE

[1] Rawat, J.S., Kumar, M. (2015). Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India, Ehe Egyptian Journal of Remote Sensing and Space Sciences, 18, 77-84. http://dx.doi.org/10.1016/j.ejrs.2015.02.002

[2] Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. Remote Sensing of Environment, 117, 162e176.

[3] Paul, M. Mather, (2004). Computer Processing of Remotely-Sensed Images-An Introduction, Wiley Publications, 3rd edition.

[4] Adam, E., Mutanga, O., Odindi, J., Abdel-Rahman, E.M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers, International Journal of Remote Sensing, 35, 10, 3440-3458. http://dx.doi.org/10.1080/01431161.2014.903435.

[5] Punia, M., Joshi, P.K., Porwal, M.C. (2010). Decision tree classification of land use land cover for Delhi, India using IRS-P6 AWiFS data, Expert Systems with Applications, doi:10.1016/j.eswa.2010.10.078.

[6] Zaakirah, B., Urmilla, B., Zoltan, S., Riyad, I. (2016). Land cover and land use mapping of the iSimangaliso Wetland Park, South Africa: comparison of oblique and orthogonal random forest algorithms, J. Appl. Remote Sens. 10(1), doi: 10.1117/1.JRS.10.015017.

[7] Ghosh, A., Sharma, R., Joshi, P.K. (2014). Random forest classification of urban landscape using Landsat archive and ancillary data: Combining seasonal maps with decision level fusion, Applied Geography, 48, 31-41. http://dx.doi.org/10.1016/j.apgeog.2014.01.003

[8] Ustuner, M., Bilgin, G., Abdikan, S. (2017). Land Use and Cover Classification of Sentinel-1A SAR Imagery: A Case Study of Istanbul, Signal Processing and Communications Applications Conference (SIU). DOI: 10.1109/SIU.2017.7960373.

[9] Jia, K., Wei, X., Gu, X., Yao, Y., Xie, X., Li, B. (2014). Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China, Geocarto International, 29, 8, 941-951. http://dx.doi.org/10.1080/10106049.2014.894586

[10] Kolios, S., Stylios, C.D. (2013). Identification of land cover/land use changes in the greater area of the Preveza peninsula in Greece using Landsat satellite data, Applied Geography, 40, 150-160. http://dx.doi.org/10.1016/j.apgeog.2013.02.005

[11] Senf, C., Leitao, P., Pflugmacher, D., Hostert, P. (2015). Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery, Remote Sensing of Environment, 156, 527-536. DOI: 10.1016/j.rse.2014.10.018.

[12] Vapnik, V.N. (1995). The nature of statistical learning theory [M]. New York: Springer-Verlag.

[13] Naiyang, D., Yingjie, T. (2004). New Data Mining Method - Support Vector Machines [M] Beijing: Science Press.

[14] Foody, G. M. and Mathur, A. (2004). A relative evaluation of multi-class image classification by support vector machines, IEEE Transactions on Geoscience and Remote Sensing, Vol. 42, pp. 1335-1343.

[15] Tzotsos, A. (2006). A Support Vector Machine approach for object based image analysis. In Proceedings of 1st International Conference on Object-based Image Analysis, OBIA 2006, Salzburg, Austria, July 4-5. ПРОЧИТАЈТЕ ИСПОД??

[16] Vapnik, V.N. (1998). Statistical Learning Theory. John-Wiley and Sons, Inc.

[17] Melgani, F., Bruzzone, L. (2004). Classification of Hyperspectral Remote Sensing Images with Support Vector Machines, IEEE transactions on geoscience and remote sensing, Vol. 42, No. 8, 10.1109/TGRS.2004.831865.

[18] Breiman, L. (2001). Random Forest, Machine learning, 45, 1.

[19] Breiman, L. (1996). Bagging Predictors, Machine Learning, 26, 2, 123-140.

[20] Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M., Jeganathan, C. (2012). Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture, Remote Sensing of Environment, 121, 93-107. doi:10.1016/j.rse.2011.12.003.

[21] Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification, ISPRS Journal of Photogrammetry and Remote Sensing, 67, 93–104. doi:10.1016/j.isprsjprs.2011.11.002.

[22] Campbell, J. B. (1996). Introduction to remote sensing (2nd ed.). London: Taylor and Francis.

[23] Foody, G. (2002). Status of land cover classification accuracy assessment, Remote Sensing of Environment, Vol. 80, pp. 185-201.

[24] Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. Educ. Psychol. Meas. Vol. 20, pp. 37-46, doi:10.1177/001316446002000104.

[25] Landis, J.R., Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, Vol. 33, pp.159-174.

[26] Brodernes, K. H., Ong, C. S., Stephan, K. E., Buhmann, J.M. (2010). The balanced accuracy and its posterior distribution, International Conference on Pattern Recognition, DOI:10.1109/ICPR.2010.764.
[27] ERD: http://www.banjaluka.rs.ba/front/category/139/ (accessed 06.03.2018).
[28] ERD: https://sentinel.esa.int/web/sentinel/missions/sentinel-2/overview (accesed 06.11.2017).
[29] ERD: https://scihub.copernicus.eu/dhus/(accesed 05.12. 2017).