Mersin Photogrammetry Journal

https://dergipark.org.tr/en/pub/mephoj

e-ISSN 2687-654X

The impacts of vegetation indices from UAV-based RGB imagery on land cover classification using ensemble learning

Muhammed Yusuf Öztürk [*1], İsmail Çölkesen [1]

[1] *Gebze Technical University, Engineering Faculty, Department of Geomatics Engineering, Kocaeli, Turkey*

Keywords
Ensemble learning
UAV
LULC
LightGBM
XGBoost

ABSTRACT

The production of land use and land cover (LULC) maps using UAV images obtained by RGB cameras offering very high spatial resolution has recently increased. Vegetation indices (VIs) have been widely used as an important ancillary data to increase the limited spectral information of the UAV image in pixel-based classification. The main goal of this study is to analyze the effect of frequently used RGB-based VIs including green leaf index (GLI), red-green-blue vegetation index (RGBVI) and triangular greenness index (TGI) on the classification of UAV images. For this purpose, five different dataset combinations comprising of RGB bands and VIs were formed. In order to evaluate their effects on thematic map accuracy, four ensemble learning methods, namely RF, XGBoost, LightGBM and CatBoost were utilized in classification process. Classification results showed that the use of RGB UAV image with VIs increased the overall accuracy (OA) values in all cases. On the other hand, the highest OA values were calculated with the use of Dataset-5 (i.e. RGB bands and all VIs considered). Additionally, the classification result of Dataset-4 (i.e. RGB bands and TGI) showed superior performance compared to Dataset-2 (i.e. RGB bands and GLI) and Dataset-3 (i.e. RGB bands and RGBVI). All in all, the TGI was found to be useful for improving classification accuracy of UAV image having limited spectral information compared to GLI and RGBVI. The improvement in overall accuracy reached to 2% with the use of RGB bands and TGI index. Furthermore, within the ensemble algorithms, CatBoost produced the highest overall accuracy (92.24%) with the dataset consist of RBG bands and all VIs considered.

1. Introduction

Gathering accurate and reliable land use and land cover (LULC) information about the Earth's surface is a prerequisite for the success of a wide range of applications carried out at local, regional and global scales (Colkesen and Ertekin, 2020). Recent developments in the field of unmanned aerial vehicle (UAV) technologies and imaging sensor systems have led to a renewed interest in extracting required information about surface objects from high spatial resolution UAV images (Yao and Qin, 2019).

Supervised pixel-based image classification that one of the popular classification techniques to produce LULC maps in the literature (Huth et al., 2012; Tehrany et al., 2014; Goldblatt et al., 2018). Pixel-based image classification is generally based on the assignment of the image pixels into pre-defined LULC classes using their digital numbers. The RGB-UAV-based platform is an alternative and low-cost aerial platform technology ensuring the capturing surface images at very high spatial and temporal resolutions. Although the RGB cameras are able to provide high spatial information about the surface, their spectral resolutions are limited for distinguishing spectrally similar pixels (Jang et al., 2020). In order to overcome this limitation, the ancillary data such as vegetation indices, texture features and principal components have been widely used in image classification process. Combinations of various vegetation indices (VIs) and RGB bands have been frequently used in the literature to improve the classification performance of RGB-UAV images (Sumesh et al., 2021). Many vegetation indices based on different sensor specifications have been developed since the launch of the first remote sensing satellite, Landsat. They are widely used for quantitative and qualitative evaluations of vegetation information (Xue and Su, 2017).

* **Corresponding Author**

*(m.ozturk2020@gtu.edu.tr) ORCID ID 0000 – 0001 – 6459 – 9356
(icolkesen@gtu.edu.tr) ORCID ID 0000 – 0001 – 9670 – 3023

Research Article / DOI: 10.53093/mephoj.943347

Cite this article

Öztürk M Y & Çölkesen İ (2021). The impacts of vegetation indices from UAV-based RGB imagery on land cover classification using ensemble learning. Mersin Photogrammetry Journal, 3(2), 41-47

Received: 26/05/2021; Accepted: 28/08/2021

Choosing the most appropriate classification algorithm for classifying UAV images having limited spectral information is also one of the important factors effecting the thematic map accuracy. In recent years, there has been great interest in classifying digital images using ensemble learning methods in the literature due to their robust, effective and fast performance (Zhiwei et al., 2016). The main idea behind the ensemble learning is to combine predictions of multiple learners (e.g. decision trees) to final decision on a given unknown sample (Tonbul et al., 2020). Previous studies confirmed that decision tree based ensemble learning algorithms such as bagging, boosting, RF, XGBoost, LightGBM and CatBoost perform better than utilize of single decision tree classifier (Sagi and Rokach, 2018; Shi et al., 2021).

The main purpose of this study is to analyse the effect of the use of RGB based vegetation indices on the LULC classification accuracy. For this purpose, three widely preferred vegetation indices, namely green leaf index (GLI), red-green-blue vegetation index (RGBVI) and triangular greenness index (TGI) were formed as ancillary dataset. RF, XGBoost, LightGBM and CatBoost ensemble learning algorithms were utilized to perform classification process. Classification results were evaluated using overall accuracy (OA), Kappa coefficient and F-score measures.

2. Study area and dataset

The study area covers the north-eastern part of Gebze Technical University located in Gebze district of Kocaeli province. As shown in Figure 1, within the boundaries of the study area, faculty buildings, other man-made structures, green vegetation and bare soil areas exist. Study area consists of six main LULC classes: concrete including gray stone floor, road, gray and white roofs, forest class including deciduous trees, coniferous trees and grass, parkour including bicycle road, basketball and tennis court, shadow, soil and tile roof.

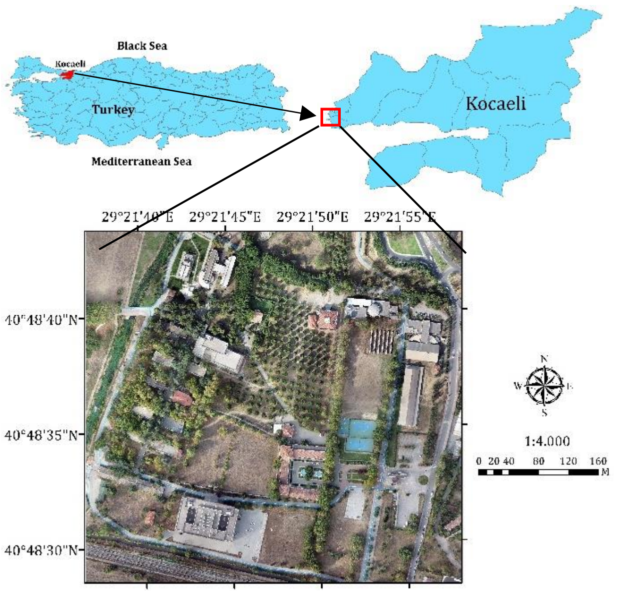

Figure 1. Study area

In this study, UAV-based high-resolution remote images were acquired on 24 September 2020 by Phantom IV Pro V2.0 drone equipped with a 20 MP RGB camera (Table 1). "Pix4Dmapper" application was preferred for flight planning. The images collected from 80 m flight altitude with 80% forward overlap and %70 side overlap, resulting in ground sampling distance 2.3 cm. Agisoft PhotoScan software was used to process the obtained images and as a result, an 8-bit ortho-mosaic with a spatial resolution of 5 cm was produced.

Table 1. Specifications of Phantom IV Pro V2.0

Technical Specifications	Value
Sensor	1 inch 20MP
Weight	1375 gr
Max Flight Time	30 min
Max Speed	45 mph (72 km/h)
Max Ascent Speed	6 m/s
Max Descent Speed	4 m/s
GNSS Module	GPS/GLONASS
Hover Accuracy Range	Vertical: ±0.1 m Horizontal: ±0.3 m

3. Methodology

In this study, the effect of the use of RGB based GLI, RGBVI and TGI vegetation indices on the accuracy of thematic maps produced from UAV image were analyzed. For this purpose, training and validation pixels for each LULC classes were determined on UAV images. Four robust ensemble learning algorithms, namely, RF, XGBoost, LightGBM and CatBoost, were utilized to construct classification model using training samples. Then, the datasets consisting of the combinations of UAV images and vegetation indices were classified with the constructed classification models and thematic maps were produced. In order to conduct the accuracy assessment, OA, kappa coefficient and F-scores were utilized, and derived results analyzed.

3.1. Vegetation indices (VIs)

VIs are obtained from the mathematical equations applied to two or more spectral bands to emphasize the vegetation characteristics. Various VIs based on RGB bands have been developed. In this study, GLI, RGBVI and TGI indices, frequently used in various studies in the literature, were evaluated.

GLI was developed by the Louhaichi et al. (2001) for determination of wheat planted areas using 8-bit RGB camera. GLI values take values between -1 and +1. Negative values correspond to soil and lifeless features, whereas positive values correspond to green vegetation.

$$GLI = \frac{2 \times Green - Red - Blue}{2 \times Green + Red + Blue} \tag{1}$$

RGBVI was developed for biomass estimation by Bendig et al. (2015). It can be described as the normalized difference of the squared green spectral band and the product of blue×red bands.

$$RGBVI = \frac{Green^2 - Blue \times Red}{Green^2 + Blue \times Red} \quad (2)$$

TGI, based on red, green and blue spectral bands, is sensitive to chlorophyll content at leaf and canopy (Hunt et al. 2011). Since this indice uses the bands in the visible region, chlorophyll content can be estimated with TGI on images acquired from UAVs equipped with an RGB camera.

$$TGI = Green - 0.39 \times Red - 0.61 \times Blue \quad (3)$$

3.2. Ensemble learning methods

3.2.1. Random forest (RF)

The RF algorithm, proposed by Breiman (2001), is one of the most popular decision tree based ensemble learning algorithms used for performing pixel-based image classification procedure due to its robust and efficient performance (Nitze et al., 2015; Fu et al., 2017). RF utilizes multiple decision trees in that each tree trained using bootstrapped samples of input dataset for constructing classification model. Majority voting rule is applied to make the final prediction and simple majority rule is applied for final prediction (Colkesen and Kavzoglu, 2017). Based on bootstrapping strategy, decision trees are trained using two thirds of input dataset and the remaining one-third of input dataset is utilized to evaluate the classification error (Tonbul et al., 2020). The results of each tree are aggregated, and final model output is composed. RF requires two main parameters to employ RF, such as the number of sample trees (*ntree*) and the number of variables suitable for splitting (*mtry*).

3.2.2. Extreme gradient boosting (XGBoost)

XGBoost, one of the advanced and effective tree-based algorithm presented by Chen and Guestrin (2016), has been used in various remote sensing applications due to its effective and fast performance (Zou et al., 2019; Abdi, 2020). It works based on essential of gradient boosting that construct multiple decision trees iteratively and transform weak learners to strong learners in each iteration (Sahin, 2020). The main difference of XGBoost than other tree-based algorithms, it uses the loss function to correct the error of weak learners of previous model in each iteration and employs regularization parameter to prevent overfitting to produce accurate classification model (Hamedianfar et al., 2020; Ustuner et al., 2020). XGBoost consists of several tuning parameters that should be defined by user-side. Seven parameters including eta, gamma, min_child_weight, subsample, colsample_bytree, max_depth, nround were optimized for XGBoost ensemble model in this study.

3.2.3. Light gradient boosting machines (LightGBM)

LightGBM, developed by Microsoft (2017), is one of the most preferred open-source and gradient boosting based method for regression and classification problems (Ma et al., 2018; Sun et al., 2020). It uses a histogram-based model that speeds up the training process and enables a more accurate model to be constructed (Al Daoud, 2019). The main difference between LightGBM and other gradient-based methods are that it utilizes gradient-based one-side sampling (GOSS) algorithms that divide training samples into smaller subsamples and leaf-wise growth strategy (Chen et al., 2019). LightGBM has several parameters and "boosting", "learning_rate", "num_leaves", "min_data", "sub_features", "feature_fraction", "bagging_fraction" and "max_depth" were tuned to construct classification model (Ke et al., 2017).

3.2.4. Categorical boosting (CatBoost)

CatBoost, novel gradient boosting method, was developed by Yandex (2018) for handle different datasets such as categorical features using random permutation technique and minimize overfitting problems (Pham et al., 2020). CatBoost consists of two main training steps. In the first step, training data is randomly divided into subsets and labels are transformed into integer. Categorical features are converted into numerical in the second step. The maximum depth of trees (depth), the control of training time (learning_rate), the number of trees in model (iteration), coefficient at the L2 regularization (l2_leaf_reg), the percentage of variables to utilize at each split selection (rms) and the controlling number of splits for numerical variables (border_count) were utilized for implementation of CatBoost ensemble classification model.

4. Results

In this study, the effect of VIs on pixel based LULC classification of RGB image acquired by UAV was investigated. To achieve this purpose, three VIs (i.e., GLI, RGBVI and TGI) were calculated using equation given in section 3.2 and stretched to 0-255 pixel values. In order to construct classification model and to evaluate accuracies of thematic maps, totally 30,000 pixels (i.e. 5,000 pixels for each LULC class) were selected as training and 6,000 pixels (i.e. 1,000 pixels for each LULC class) were selected as validation. Five datasets were created using RGB bands and different combination of VIs to evaluate classification results: Dataset-1 includes only RGB band, Dataset-2 consists of RGB bands and GLI, Dataset-3 consists of RGB bands and RGBVI, Dataset-4 consists of RGB bands and TGI and Dataset-5 corresponds to combination of RGB bands and all VIs considered. On the hand, parameters of each classifier should be determined by user side to obtain optimal classification models. Tuning parameters required for each ensemble methods were determined by grid search algorithm and estimated values were given in Table 2. Note that all classification processes and accuracy assessments were performed in R software. Additionally, the "randomForest", "xgboost", "lightgbm" and "catboost" packages were utilized for implementation of RF, XGBoost, LightGBM and CatBoost ensemble methods, respectively and "caret" package was utilized to calculate accuracy assessment measures.

Table 2. Optimal parameters of ensemble methods

Method	Parameter	Value
RF	ntree	380
	mtry	2
XGBoost	eta	0.3
	gamma	0
	min_child_weight	0.6
	subsample	0.8
	colsample_bytree	1
	max_depth	4
	nround	400
LightGBM	boosting	goss
	learning_rate	0.3
	num_leaves	20
	min_data	80
	sub_features	0.8
	feature_fraction	1
	bagging_fraction	1
	max_depth	4
CatBoost	depth	4
	learning_rate	0.3
	iteration	400
	l2_leaf_reg	0.7
	rms	0.95
	border_count	128

Accuracy assessment results of each Dataset were given in Table 3. As could be seen from the table, the highest OA values estimated for Dataset-5 as 91.2% (Kappa value of 0.89), 92.2% (Kappa value of 0.90), 92.0% (Kappa value of 0.90) and 92.4% (Kappa value of 0.91) by the use of RF, XGBoost, LightGBM and CatBoost, respectively. On the other hand, the lowest OA values were observed for Dataset-1 with all ensemble learning algorithms. Furthermore, the OA values of Dataset-4 calculated from all classifiers were very close to the OA values obtained with Dataset-5 (lower about 0.1%). Additionally, it was observed that calculated OA values increased by about 0.3% with the use of Dataset-2 compared to use of Dataset-3. When the classification results based on OA accuracy were evaluated it was seen that while the CatBoost showed quite similar performance to XGBoost and LightGBM, the highest OA value was obtained as 92.24% with the CatBoost algorithm. On the other hand, it has been observed that the classification performance of the RF algorithm is lower than that of the others for all datasets except for Dataset-1.

In order to evaluate and compare class-based accuracy performances, F-score values were also calculated, and derived statistics were given in Table 3. It can be seen that concrete, forest and parkour classes were classified with over 91% classification accuracy by all ensemble algorithms. Furthermore, the highest F-score values were estimated for parkour class, whereas the worst class-based accuracy calculated for soil class. The reason why the class-level accuracy of the soil class was lowest may be that various substances that were mixed into the soil and have similar spectral properties with other LULC classes could be easily distinguished in the images obtained with the UAV. Moreover, with the classification of the RGB bands with all calculated indices (Dataset-5), the F-score value of soil class produced by all classifiers increased up to 4% compared to results of Dataset-1. Furthermore, the accuracies of concrete class estimated by XGBoost, LightGBM and CatBoost for Dataset-4 were significantly higher than result of RF (about 3%).

Table 3. Classification results of each dataset

Method	LULC Class	F-scores				
		D-1	D-2	D-3	D-4	D-5
RF	Concrete	91.3	91.8	92.4	92.4	94.6
	Forest	97.5	97.0	96.9	97.2	96.7
	Parkour	98.5	99.4	99.5	99.7	99.2
	Shadow	89.2	90.4	90.0	90.5	89.9
	Soil	79.5	81.2	79.5	80.0	80.0
	Tile roof	84.5	85.8	85.9	87.2	87.3
	OA	90.0	90.8	90.6	91.1	91.2
	Kappa	0.88	0.89	0.89	0.89	0.89
XGBoost	Concrete	91.6	92.8	93.4	95.5	95.8
	Forest	97.4	96.3	96.1	97.9	97.8
	Parkour	98.0	98.7	98.4	98.7	99.4
	Shadow	88.7	91.1	91.1	89.8	91.1
	Soil	79.7	82.2	80.2	81.3	81.2
	Tile roof	84.4	86.6	86.7	88.5	88.2
	OA	89.9	91.2	90.9	92.0	92.2
	Kappa	0.88	0.89	0.89	0.90	0.90
LightGBM	Concrete	92.9	92.5	93.8	95.3	95.7
	Forest	97.1	96.6	96.6	97.6	97.1
	Parkour	97.1	98.8	98.8	98.6	98.8
	Shadow	88.4	90.8	91.4	90.2	91.2
	Soil	79.2	81.4	81.2	81.9	81.2
	Tile roof	85.1	86.2	86.7	88.3	88.2
	OA	89.8	91.0	91.3	91.9	92.0
	Kappa	0.88	0.90	0.90	0.90	0.90
CatBoost	Concrete	92.2	92.6	95.5	95.3	95.8
	Forest	97.1	97.7	96.5	97.8	97.7
	Parkour	97.9	99.7	99.1	99.1	99.6
	Shadow	89.5	90.7	90.4	90.7	91.4
	Soil	80.1	82.1	81.3	82.5	81.2
	Tile roof	85.6	86.8	88.4	88.9	88.3
	OA	90.4	91.5	91.8	92.3	92.4
	Kappa	0.89	0.90	0.90	0.91	0.91

To visual comparison of datasets, thematic maps were produced by RF and CatBoost classifiers which yield the lowest and highest OA values and presented in Figure 2. Determined misclassification error on the thematic maps are highlighted with a dashed white circle. According to visual analysis, thematic maps produced by CatBoost were smoother than those of RF. On the other hand, main classification errors were occurred among concrete, soil, shadow and tile roof. It was observed that both ensemble methods were insufficient in distinguishing soil, shadow and concrete classes. This visual result is consistent with the F-score values of the soil class. As can be seen from marked areas, CatBoost outperformed to RF in assigning the pixels corresponding to the concrete class to the correct land cover class. Additionally, the noise generated in the concrete class was reduced in the thematic map produced by RF using Dataset-5 compared to other thematic maps by RF.

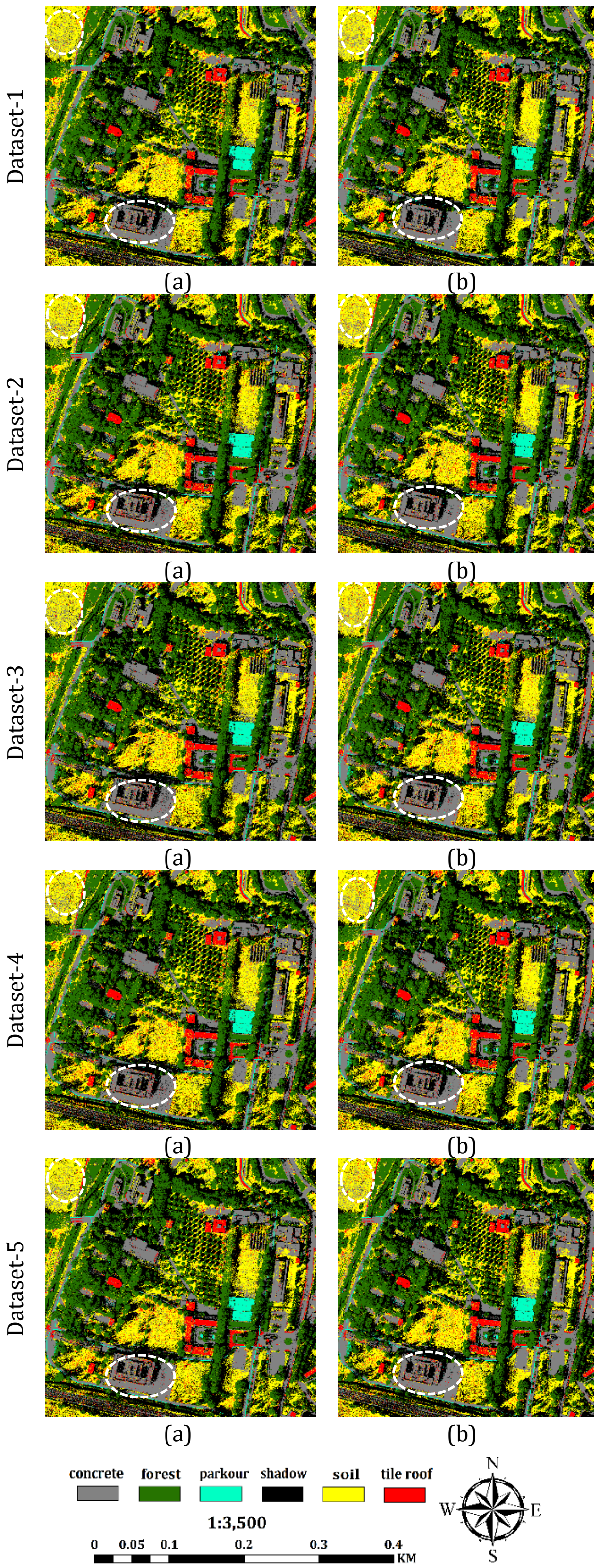

Figure 2. Thematic maps of each dataset produced by (a) RF, (b) CatBoost.

4. Discussion and conclusions

In recent years, there has been an increasing interest in the production of thematic maps with UAV images using ensemble learning algorithms due to their effective classification performance. VIs used to highlight the features of the Earth's provide a great advantage in increasing the spectral information in the pixel-based classification of RGB images. In this context, the effects of widely used three VIs (i.e. GLI, RGBVI and TGI) on the classification of UAV image having three spectral bands were investigated. For this purpose, five datasets containing different combination of VIs and UAV image were created and each dataset was classified with RF, XGBoost, LightGBM and CatBoost.

The following conclusions can be made by analyzing the classification results obtained. OA values obtained by the classification of UAV image with VIs (Dataset-5) increased by 1% with the use of RF and by 2% with other remaining methods compared to results of visible spectral bands only (Dataset-1). This could be probably result of the increase in spectral information by means of vegetation indices usage. Kerkech et al. (2018), Wan et al. (2018) and Lu et al. (2021) analyzed the performance of several VIs derived from RGB drone images. They reported that while the highest classification and regression accuracies were computed in the processing of the RGB image with all evaluated vegetation indices, the accuracy decreased as the spectral information decreased. On the other hand, the OA values of Dataset-4 (combination of RGB bands and TGI) generated by all the algorithms are about 0.1% less than the classification results of Dataset-5. Additionally, all ensemble learning methods yielded higher classification results in the use of UAV image with TGI compared to datasets consisting of aggregation of RGB band with other indices (Dataset-2 and Dataset-3). These results clearly showed that TGI was found to be the most useful indices to identification of LULC classes in classifying three-band UAV images for considered dataset used in this study. Fuentes-Peailillo et al. (2018) analyzed the various RGB-based vegetation indexes for distinguishing soil and vegetation areas and they found that TGI index showed superior performance than other VIs. Starý et al. (2020) conducted comparative study using seven RGB-based VIs including GLI, RGBVI and TGI indexes for estimating hops plants in hop gardens and TGI outperformed to others in their study. Hindersah et al. (2018) also found similar results with our study. Moreover, CatBoost, relatively new ensemble learning algorithm, showed superior classification performance in separation of LULC classes compared to other algorithms. The implementation of CatBoost for classification and regression problems of remote sensing problems in the literature is very limited. In addition, pixel-based classification of UAV images has not been made with this algorithm until now. However, it was verified by Samat et al. (2020), Ha et al. (2021) and Pham et al. (2021) that CatBoost have effective and superior classification and regression performance compared to other bagging, boosting and other classifiers (i.e., RF, XGBoost, SVM). Our findings also contribute their studies. Different vegetation indices can be evaluated with various robust classifiers (e.g., support vector machines, canonical corelation forest, rotation forest) in order to better analyse the effect of VIs on the classification of the RGB image. In addition, further studies are required to evaluate the effectiveness of the use of RGB derived vegetation indices on object-based classification accuracy.

Acknowledgement

This article is presented in "2nd Intercontinental Geoinformation Days" 2021 and selected for publication in Mersin Photogrammetry Journal.

Author contributions

Muhammed Yusuf Öztürk: Literature review, Field study, Modelling, classification, Writing **İsmail Çölkesen:** Modelling, classification, Writing-editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

Abdi A M (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, *57*(1), 1–20. https://doi.org/10.1080/15481603.2019.1650447

Al Daoud E (2019). Comparison between XGBoost, LightGBM and CatBoost Using a Home Credit Dataset. *International Journal of Computer and Information Engineering*, *13*(1), 6–10.

Breiman L (2001). Random Forests. In *Machine Learning* (pp. 5–32). Chapman and Hall/CRC. https://doi.org/10.1023/A:1010933404324

Chen Tianqi & Guestrin C (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-Augu*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen Tingting, Xu J, Ying H, Chen X, Feng R, Fang X, Gao H & Wu J (2019). Prediction of Extubation Failure for Intensive Care Unit Patients Using Light Gradient Boosting Machine. *IEEE Access*, *7*, 150960–150968. https://doi.org/10.1109/ACCESS.2019.2946980

Colkesen I & Ertekin O H (2020). Performance Analysis of Advanced Decision Forest Algorithms in Hyperspectral Image Classification. *Photogrammetric Engineering & Remote Sensing*, *86*(9), 571–580. https://doi.org/10.14358/PERS.86.9.571

Colkesen I & Kavzoglu T (2017). The use of logistic model tree (LMT) for pixel- and object-based classifications using high-resolution WorldView-2 imagery. *Geocarto International*, *32*(1), 71–86. https://doi.org/10.1080/10106049.2015.1128486

Fu B, Wang Y, Campbell A, Li Y, Zhang B, Yin S, Xing Z & Jin X (2017). Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. *Ecological Indicators*, *73*, 105–117. https://doi.org/10.1016/j.ecolind.2016.09.029

Fuentes-Peailillo F, Ortega-Farias S, Rivera M, Bardeen M & Moreno M (2018). Comparison of vegetation indices acquired from RGB and Multispectral sensors placed on UAV. *2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA)*, 1–6. https://doi.org/10.1109/ICA-ACCA.2018.8609861

Goldblatt R, Stuhlmacher M F, Tellman B, Clinton N, Hanson G, Georgescu M, Wang C, Serrano-Candela F, Khandelwal A K, Cheng W H & Balling R C (2018). Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. *Remote Sensing of Environment*, *205*(December 2017), 253–275. https://doi.org/10.1016/j.rse.2017.11.026

Ha N T, Manley-Harris M, Pham T D & Hawes I (2021). Detecting multi-decadal changes in seagrass cover in tauranga harbour, new zealand, using landsat imagery and boosting ensemble classification techniques. *ISPRS International Journal of Geo-Information*, *10*(6). https://doi.org/10.3390/ijgi10060371

Hamedianfar A, Gibril M B A, Hosseinpoor M & Pellikka P K E (2020). Synergistic use of particle swarm optimization, artificial neural network, and extreme gradient boosting algorithms for urban LULC mapping from WorldView-3 images. *Geocarto International*, *0*(0), 1–19. https://doi.org/10.1080/10106049.2020.1737974

Hindersah R, Handyman Z, Indriani F N, Suryatmana P & Nurlaeny N (2018). JOURNAL OF DEGRADED AND MINING LANDS MANAGEMENT Azotobacter population, soil nitrogen and groundnut growth in mercury-contaminated tailing inoculated with Azotobacter. *J. Degrade. Min. Land Manage*, *5*(53), 2502–2458. https://doi.org/10.15243/jdmlm

Hunt E R, Daughtry C S T, Eitel J U H & Long D S (2011). Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index. *Agronomy Journal*, *103*(4), 1090–1099. https://doi.org/10.2134/agronj2010.0395

Huth J, Kuenzer C, Wehrmann T, Gebhardt S, Tuan V Q & Dech S (2012). Land cover and land use classification with TWOPAC: Towards automated processing for pixel- and object-based image classification. *Remote Sensing*, *4*(9), 2530–2553. https://doi.org/10.3390/rs4092530

Jang G, Kim J, Yu J, Kim H, Kim Y, Kim D, Kim K, Lee C W & Chung Y S (2020). Remote sensing Review : Cost-E ff ective Unmanned Aerial Vehicle (UAV) Platform for Field Plant Breeding Application. *Remote Sensing*, *12*(6), 998.

Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q & Liu T Y (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *2017-Decem*(Nips), 3147–3155.

Kerkech M, Hafiane A & Canals R (2018). Deep leaning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Computers and Electronics in Agriculture*, *155*(October), 237–243. https://doi.org/10.1016/j.compag.2018.10.006

Lu J, Cheng D, Geng C, Zhang Z, Xiang Y & Hu T (2021). Combining plant height, canopy coverage and vegetation index from UAV-based RGB images to estimate leaf nitrogen concentration of summer maize. *Biosystems Engineering*, *202*, 42–54.

https://doi.org/10.1016/j.biosystemseng.2020.11.010

Ma X, Sha J, Wang D, Yu Y, Yang Q & Niu X (2018). Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGboost algorithms according to different high dimensional data cleaning. *Electronic Commerce Research and Applications*, *31*, 24–39. https://doi.org/10.1016/j.elerap.2018.08.002

Nitze I, Barrett B & Cawkwell F (2015). Temporal optimisation of image acquisition for land cover classification with Random Forest and MODIS time-series. *International Journal of Applied Earth Observation and Geoinformation*, *34*(1), 136–146. https://doi.org/10.1016/j.jag.2014.08.001

Pham T D, Yokoya N, Nguyen T T T, Le N N, Ha N T, Xia J, Takeuchi W & Pham T D (2020). Improvement of Mangrove Soil Carbon Stocks Estimation in North Vietnam Using Sentinel-2 Data and Machine Learning Approach. *GIScience & Remote Sensing*, 1–20. https://doi.org/10.1080/15481603.2020.1857623

Pham T D, Yokoya N, Nguyen T T T, Le N N, Ha N T, Xia J, Takeuchi W & Pham T D (2021). Improvement of Mangrove Soil Carbon Stocks Estimation in North Vietnam Using Sentinel-2 Data and Machine Learning Approach. *GIScience & Remote Sensing*, *58*(1), 68–87. https://doi.org/10.1080/15481603.2020.1857623

Sagi O & Rokach L (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), 1–18. https://doi.org/10.1002/widm.1249

Sahin E K (2020). Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto International*, *0*(0), 1–25. https://doi.org/10.1080/10106049.2020.1831623

Samat A, Li E, Du P, Liu S, Miao Z & Zhang W (2020). CatBoost for RS Image Classification With Pseudo Label Support From Neighbor Patches-Based Clustering. *IEEE Geoscience and Remote Sensing Letters*, 1–5. https://doi.org/10.1109/LGRS.2020.3038771

Shi J, Shao T, Liu X, Zhang X, Zhang Z & Lei Y (2021). Evolutionary Multitask Ensemble Learning Model for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 936–950. https://doi.org/10.1109/JSTARS.2020.3037353

Starý K, Jelínek Z, Kumhálova J, Chyba J & Balážová K (2020). Comparing RGB-based vegetation indices from uav imageries to estimate hops canopy area. *Agronomy Research*, *18*(4), 2592–2601. https://doi.org/10.15159/AR.20.169

Sumesh K C, Ninsawat S & Som-ard J (2021). Integration of RGB-based vegetation index, crop surface model and object-based image analysis approach for sugarcane yield estimation using unmanned aerial vehicle. *Computers and Electronics in Agriculture*, *180*(July 2020), 105903. https://doi.org/10.1016/j.compag.2020.105903

Sun X, Liu M & Sima Z (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. *Finance Research Letters*, *32*(November 2018), 101084. https://doi.org/10.1016/j.frl.2018.12.032

Tehrany M S, Pradhan B & Jebuv M N (2014). A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping using SPOT 5 imagery. *Geocarto International*, *29*(4), 351–369. https://doi.org/10.1080/10106049.2013.768300

Tonbul H, Colkesen I & Kavzoglu T (2020). Classification of poplar trees with object-based ensemble learning algorithms using Sentinel-2A imagery. *Journal of Geodetic Science*, *10*(1), 14–22. https://doi.org/10.1515/jogs-2020-0003

Ustuner M, Abdikan S, Bilgin G & Balik Sanli F (2020). Hafif Gradyan Artırma Makineleri ile Tarımsal Ürünlerin Sınıflandırılması. *Türk Uzaktan Algılama ve CBS Dergisi*, *1*(2), 97–105.

Wan L, Li Y, Cen H, Zhu J, Yin W, Wu W, Zhu H, Sun D, Zhou W & He Y (2018). Combining UAV-Based Vegetation Indices and Image Classification to Estimate Flower Number in Oilseed Rape. *Remote Sensing*, *10*(9), 1484. https://doi.org/10.3390/rs10091484

Xue J & Su B (2017). Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors*, *2017*, 1–17. https://doi.org/10.1155/2017/1353691

Yao H & Qin R (2019). *Unmanned Aerial Vehicle for Remote Sensing Applications — A Review*. *11*(12), 1443.

Zhiwei Y, Juan Y, Xu Z & Zhengbing H (2016). Remote Sensing Textual Image Classification based on Ensemble Learning. *International Journal of Image, Graphics and Signal Processing*, *8*(12), 21–29. https://doi.org/10.5815/ijigsp.2016.12.03

Zou X, Liang A, Wu B, Su J, Zheng R & Li J (2019). UAV-based high-throughput approach for fast growing Cunninghamia lanceolata (Lamb.) cultivar screening by machine learning. *Forests*, *10*(9). https://doi.org/10.3390/f10090815