



## International Journal of Environment and Geoinformatics

### Research Article

### Open Access

# Classification of Urban Vegetation Utilizing Spectral Indices and DEM with Ensemble Machine Learning Methods



Şaziye Özge Atik<sup>1</sup>  

<sup>1</sup> İstanbul Technical University, Faculty of Civil Engineering, Department of Geomatic Engineering İstanbul Türkiye

### Abstract

Detection and monitoring urban vegetation are subjects in many sustainable development goal studies. Detection of green areas and their decreasing rates with increasing urbanization are followed with interest by municipalities and planners. Due to their high cost-effectiveness, unmanned aerial vehicles (UAVs) have been used extensively in agriculture and forest management. Overall, this study aimed to distinguish urban vegetation in a selected area with heterogeneous classes using multi-spectral UAV data and ensemble ML methods. It features groundwork for understanding the effects of different methods, the effectiveness of multi-spectral data, and the DEM effect for future studies on this subject. In this study, low vegetation, tree and non-vegetation classes were classified using ensemble machine learning methods on a university campus. After the pre-processing steps of the images obtained via UAV, datasets consisting of different bands and indices were created for classification, and the effect of the Digital Elevation Model (DEM) layer was also investigated. Four machine learning classifiers were implemented: XGBoost, LightGBM, Gradient Boosting and CatBoost. According to the results, the highest classification performances are achieved when vegetation indices and DEM are used together. The CatBoost algorithm obtained 90.2 % accuracy and 86.9% F1-score. It is understood that the classification of multispectral aerial images with machine learning has shown promising results in detecting vegetation in urban areas.


### Keywords

Multispectral UAV · Ensemble Machine Learning · Supervised Classification · Spectral Indices · DEM



Citation: Atik, Ş. Ö. (2025). Classification of urban vegetation utilizing spectral indices and DEM with ensemble machine learning methods. *International Journal of Environment and Geoinformatics*, 12(1), 43-53. <https://doi.org/10.26650/ijegeo.1640878>

 This work is licensed under Creative Commons Attribution-NonCommercial 4.0 International License. 

 2025. Atik, Ş. Ö.

 Corresponding author: Şaziye Özge Atik [donmezsaz@itu.edu.tr](mailto:donmezsaz@itu.edu.tr)



## Introduction

Urban vegetation is essential for ensuring the well-being of human communities, ecological balance, and reducing floods. It also can create a pollution barrier. Studies on rapid and effective detection of changes in urban areas have recently increased (Gazioğlu et al., 2017; Atik, 2023). Some studies in this field have demonstrated the advantages of using unmanned aerial vehicles (UAVs). Thanks to the use of machine learning (ML) algorithms for UAV data, real-time studies are possible for urban monitoring. UAVs provide many advantages with their lightweight, easy portability, low cost, low altitude flight, and pilotless flight features. By using UAV, it is possible to obtain spectral images with much higher resolution than satellite images and at times and places where satellite images cannot be obtained. According to user requirements, users can acquire data with extremely high spatial (centimetric spatial resolution) and temporal resolutions using UAVs that are lightweight, high-quality, cost-effective and small vehicles (Bayırhan & Gazioğlu, 2020; López-García et al., 2022). With the help of ML algorithms, very high-resolution UAV images can be processed faster than the operations that a manual operator would perform. Moreover, despite the increase in image processing speed thanks to hyperparameter optimization in ML algorithms, it will be possible to maintain compelling classification features by applying ensemble methods. The goal of ensemble ML is to combine several classifiers to produce a higher-performing classifier. Ensemble ML models offer some benefits over single models when resolving complicated issues. Ensemble ML models balance the trade-off between bias and variance, lowering the chance of both overfitting and underfitting (Atik & Atik, 2024). Multispectral cameras are imaging equipment that collects data from more than three spectral bands. They are often used in agricultural and environmental monitoring of plant health, soil conditions, and vegetation types. Due to the visible spectrum (red, green, and blue), red edge, and near-infrared bands, multispectral cameras use several analyses (Sefercik et al., 2024) in addition to health monitoring, cellular structure of plants, water absorption, and chlorophyll. Recently, ML algorithms have been commonly used for processing UAV-derived multi-spectral data. The effects of the vegetation indices on land cover classification were investigated using ML algorithms. The effects of vegetation indices on plant species identification in ML-based supervised classification were examined (Öztürk and Çölkese, 2021). Shu et al. investigated the effects of various image bands obtained from a multispectral UAV camera on the phenotypic trait estimation performance of ensemble classification algorithms (2022). Abdollahnejad & Panagiotidis proposed a machine

learning-based methodology for tree species detection and health status assessment from multispectral UAV images (2020). SVM classifier and nine VIs were used in the study. Wang et al. have a study on using multi-spectral and visible VI conducted various ML algorithms and calculated root mean square errors (RMSE) for the regression models (2022). They also use NIR and RE bands and visible ones to predict desert vegetation cover. Lan et al. used UAV multi-spectral data for citrus greening detection using VIs in a large-scale area (2020). The study includes pre-processing data, feature extraction using principal component analysis (PCA), conducting ML algorithms and enhancing accuracy steps.

In this study, urban vegetation classification is determined to be used by supervised ensemble ML methods such as Gradient Boosting, Categorical Boosting (CatBoost), Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM). Orthomosaic and Digital Elevation Model (DEM) data generated with multispectral UAV images were used for classification. This study sought answers to the following questions:

1. Does using multispectral UAV data increase efficiency in urban vegetation detection?
2. Do popular vegetation indices calculated with near-infrared, red edge, and visible bands in urban vegetation detection improve ML classification?
3. Does the use of DEM data positively affect the separation of vegetation from other classes? Or does it help distinguish low and high vegetation?
4. Does the classification accuracy increase when different supervised ensemble ML algorithms are optimized and compared with accuracy metrics?

## Materials and Method

### Study area

Istanbul Technical University Ayazaga Campus was selected as the study area. A UAV flight was carried out in a region where settlement and vegetation are located together within the campus. The produced orthomosaic and DEM cover an area of 207 x 267 meters. The study area is presented in [Figure 1](#).

### UAV-camera system and data collection

The research was carried out with the DJI Mavic 3M RTK UAV. The technical specifications of the DJI Mavic 3M RTK UAV are presented in [Table 1](#). The UAV collected images at an altitude of 60 m above the take-off point. 80% longitudinal overlap ratio and 70% side lap ratio were selected in flight planning. Moreover, ground sampling distance (GSD) is approximately



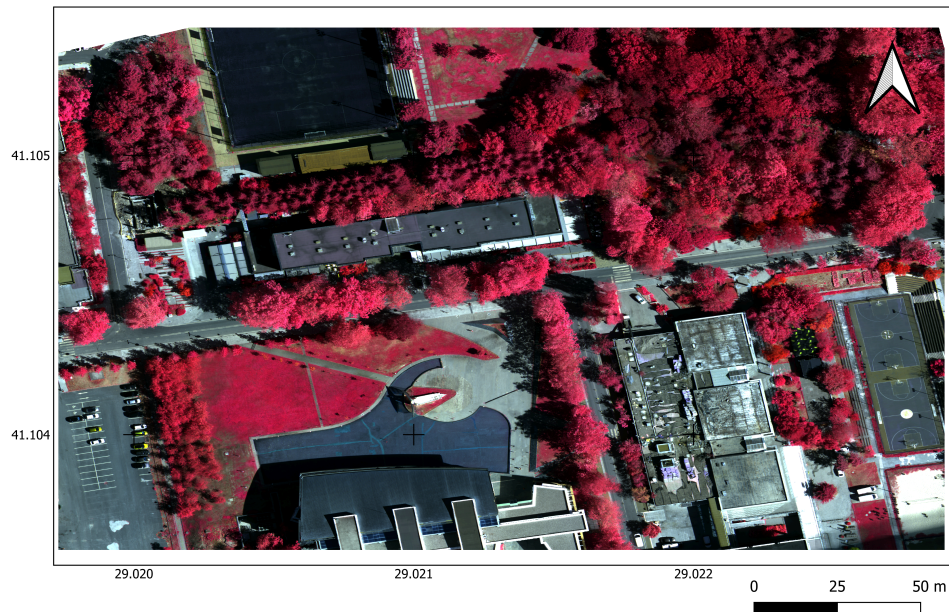


Figure 1. Study area

Table 1. Specification of the camera on DJI Mavic 3 multispectral UAV (Specs—DJI Mavic 3 Enterprise, 2025)

Equipment	Category	Feature
UAV	Model	DJI Mavic 3M
	Camera Sensor	20 Megapixels
	Max Flight Time (without wind)	43 minutes
	Max Takeoff Weight	1050 g
	Max Flight Speed (at sea level, no wind)	15 m/s (Normal Mode)
	GNSS	G+ R <sup>1</sup> +E+C
	RTK Positioning Accuracy	RTK Fix: H: 1 cm + 1 ppm V: 1.5 cm + 1 ppm
	Operating Temperature	10° to 40° C (14° to 104° F)

1.6 cm/pixel. A total of 179 multispectral aerial images were obtained from nadir angles.

Each image consists of four bands: red (R), green (G), near-infrared (NIR), and red-edge (RE) and center wavelength of the band are shown in Table 2. Experts generally use different colors in the visible part of the electromagnetic spectrum and the amount of infrared reflection reflected from the plant to decide the density of green cover on the land. Five vegetation indices were calculated with the arithmetic of these four bands. Normalized Difference Vegetation Index (NDVI) (Peñuelas et al., 1993), Leaf Chlorophyll Index (LCI) (Pu et al., 2008), Green Normalized Difference Vegetation Index (GNDVI) (Gitelson et al., 1996), Optimized Soil-Adjusted Vegetation Index (OSAVI) (Rondeaux et al., 1996) and Normalized Ratio Drought Index (NDRE) (Barnes et al., 2000) were used in this study. The flowchart of the study is shown in Figure 2.

Table 2. Center wavelength values of the bands of the UAV

Band	Center Wavelength
Red (R)	560 ± 16 nm
Green (G)	650 ± 16 nm
Red Edge (RE)	730 ± 16 nm
NIR	860 ± 26 nm

## Data processing

The RTK system achieved vertical and horizontal position errors of less than 0.1 m when georeferencing the raw photos. Using DJI Terra software, raw pictures were acquired for the creation of georeferenced spectral reflectance and VI mosaic calibrated production. Aerial triangulation, which established a distinct direction for every stereo model of the photogrammetric block, was the initial stage of the UAV photogrammetry procedure. Bundle adjustment was performed by the DJI Terra using the multiview stereo (MVS) and structure from motion (SfM) methods. The simultaneous

calculation of the internal and external orientations of all images is called aerial triangulation. In aerial triangulation, the interior and exterior orientation parameters are calculated to estimate the pose of each camera.

### Structure-from-Motion (SfM)

SfM is a computer vision method that generates 3D models using 2D images taken from stereo images. In this algorithm, prior to scene reconstruction, the SfM approach does not require knowledge of the 3D location of a set of control points and external parameters. The algorithm automatically detects similar features in the corresponding photos and reconstructs the camera pose and scene geometry (Schonberger & Frahm, 2016).

The first step of SfM is extraction and matching of the corresponding features using feature extraction methods such as SIFT (Lowe, 2004). The intrinsic projective geometry between two views, known as the epipolar geometry, is used to evaluate matching features in two images. The epipolar

geometry is expressed using a  $3 \times 3$  matrix, which is known as the fundamental matrix ( $F$ ) (Atik and Arkali, 2025).

$$x'F x = 0 \quad (1)$$

where  $x'$  and  $x$  refer to the projected points of the 3D point onto the camera image plane (Zhao et al., 2021).

The normalized eight-point approach is used to match features in order to compute the fundamental matrix in a preliminary approach. Using the combined set of camera and scene parameters, Bundle Adjustment attempts to reduce the distance between the projected and observed points progressively:

$$g(C, X) = \sum_{i=1}^n \sum_{j=1}^m \omega_{ij} \| x_{ij} - P(C_i, X_j) \|^2 \quad (2)$$

where  $\omega_{ij}$  is an indicator variable:  $\omega_{ij} = 1$  if camera  $i$  observes point  $j$ ; otherwise,  $\omega_{ij} = 0$ ;  $C$  represents the collection of camera parameters for a single camera with  $C_i = \{M, R, t, k1, k2\}$ ;  $x_{ij}$  is the observed image point; and  $P(C_i, X_j)$  is the projected image point (Zhao et al., 2021).

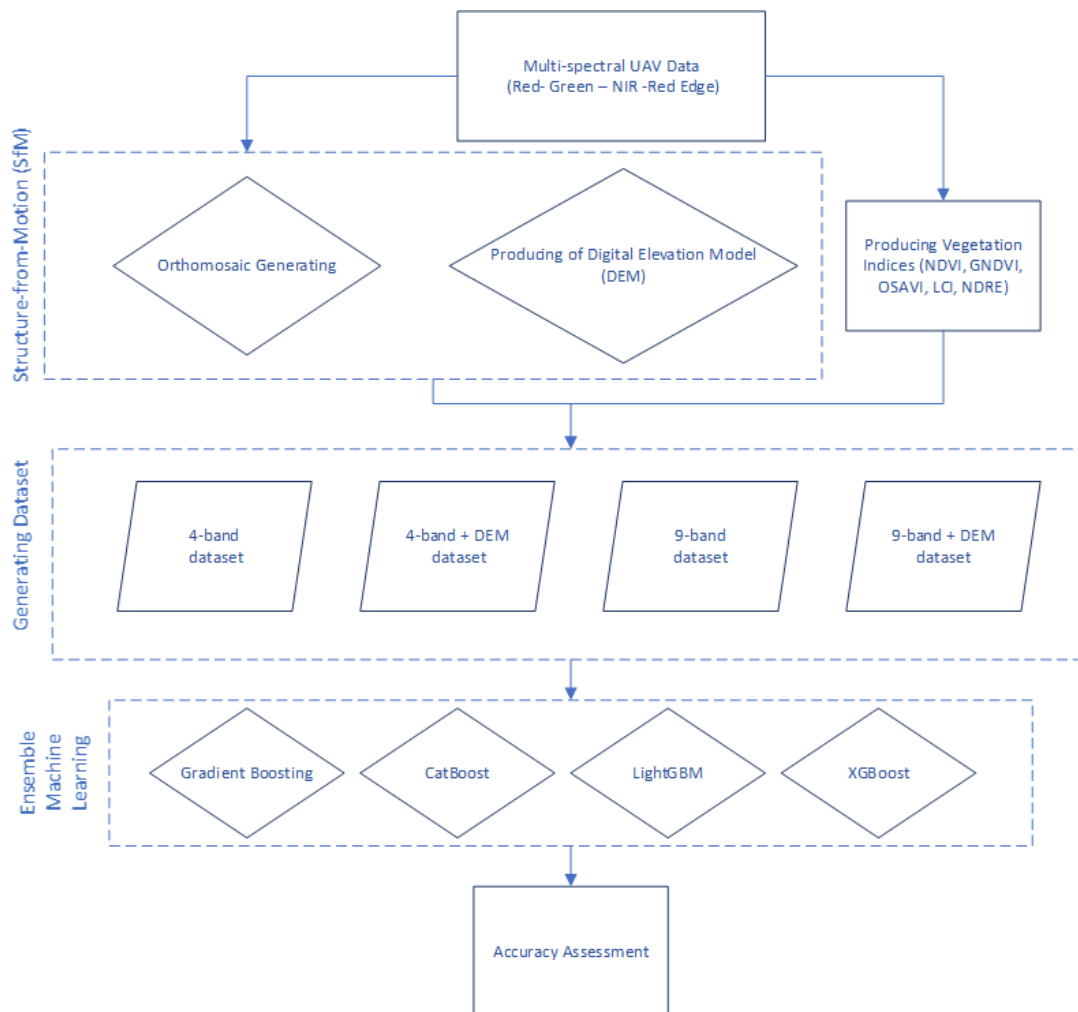


Figure 2. Flowchart of the study



## Spectral Indices

Spectral indices used in vegetation analysis are widely used in agricultural management, environmental and ecosystem monitoring, and drought and climate change. Agricultural management includes yield estimation, irrigation amount planning, and fertilization management studies. For environmental and ecosystem monitoring, studies are carried out to monitor plant health, estimate biomass, and monitor changes in the ecosystem. The spectral indices created are presented in Figure 3.

NDVI (Equation 3) is a vegetation index commonly employed to assess plant photosynthetic activity. It is determined by the normalized ratio of NIR and Red bands and utilized to enhance the understanding of plant activity dynamics. This

is due to the high absorption in the visible range of 0.4–0.7  $\mu\text{m}$  for the color pigment chlorophyll in green leaves. As the number of leaves on a plant increases, so does the reflection of wavelengths of light.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3)$$

LCI (Leaf Chlorophyll Index) (Equation 4) monitors the chlorophyll content in plant leaves. Chlorophyll is an important indicator of plant health. By using LCI, chlorophyll content is measured precisely and is frequently used in agricultural productivity analyses (Gallardo-Salazar & Pompa-García, 2020).

$$LCI = \frac{R_{850} - R_{710}}{R_{850} + R_{680}} \quad (4)$$

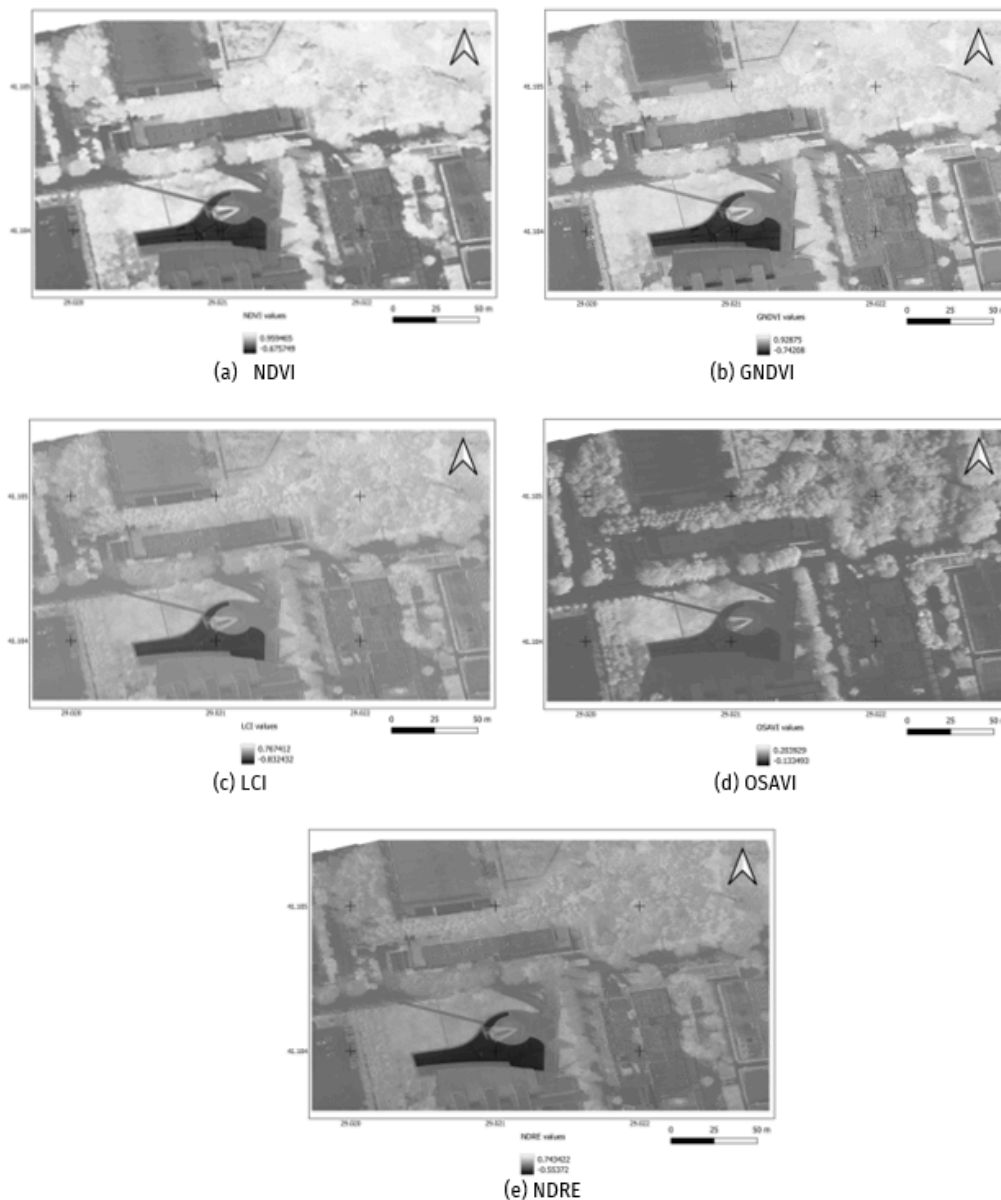


Figure 3. Produced vegetation indices for the study area

GNDVI (Equation 5) is similar to NDVI. However, it measures the green spectrum between 540 and 570 nm instead of the red spectrum. Compared to NDVI, this index is more sensitive to chlorophyll concentration (Gitelson et al., 1996).

$$GNDVI = \frac{NIR - GREEN}{NIR + GREEN} \quad (5)$$

The NDVI and the GNDVI were also calculated for the study area, which is very sensitive to chlorophyll concentration. GNDVI is an index used to assess plant nitrogen content and general vegetation health. OSAVI (Equation 6), LCI, and NRDE are other indices widely used to distinguish vegetational classes.

OSAVI is used in areas with low vegetation cover with a high ratio soil ratio and in estimating vegetation density.

$$OSAVI = \frac{NIR - RED}{NIR + RED + 0.16} \quad (6)$$

NRDE is a vegetation index based on a combination of red and red-edge bands. To understand sugar content and chlorophyll, NRDE (Equation 7) may be used for plants (Gallardo-Salazar & Pompa-García, 2020).

$$NRDE = \frac{RE1 - RE2}{RE1 + RE2} \quad (7)$$

## Ensemble Machine Learning Algorithms for Classification

Feature extraction is a leading application in remote sensing on four basics of the image. It uses a composite combination of color (spectral), pattern, shape (geometry information), and context (Momm and Easson, 2011). Spectral and spatial functions are used in feature extraction. ML methods, which have been widely used in recent years, provide significant advantages and have been demonstrated in many studies to improve classification accuracy.

### Extreme Gradient Boosting (XGBoost)

XGBoost is an efficient method and is used for both classification and regression. It is a scalable approach based on a gradient-boosting tree (Chen and Guestrin, 2016). A tree ensemble model that uses functions of  $K$ .  $K$  is an additive used to estimate the dataset  $D = \{(x_i, \hat{y}_i)\}$  with  $n$  samples and  $m$  features. The following formula is used for the algorithm,

$$\hat{y}_i = \Phi(x_i) = \sum_{k=1}^K f_k(x_i), \quad f_k \in \varphi \quad (8)$$

$$\varphi = \{f(x) = \omega_{q(x)}\} (q: R^m \rightarrow T, \omega \in R^T) \quad (9)$$

$x_i$  refers to an observation,  $f_k(x_i)$  is the projected score for the given observation, and  $y_i$  is the model's forecast Eq. (8). The collection of regression trees with the independent tree structure  $q$  is denoted by  $\varphi$  [Eq. (9)]. The weights of the leaves

is represented by  $\omega$ , and  $q$ .  $T$  is the number of leaves on the tree (Atik et al., 2024).

The standard loss function and the model's complexity construct XGBoost's goal function Eq. (10) - (11). XGBoost aims to minimize the regularized objective function,

$$L(\Phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (10)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (11)$$

### Light Gradient Boosting Machine (LightGBM)

LightGBM is an effective ML classification algorithm that includes gradient-boosting decision tree (GBDT), a method developed by Microsoft in 2017 (Ke et al., 2017). This algorithm aims to obtain efficient prediction result using high-dimensional data. LightGBM is a developing version of XGBoost. The algorithm includes gradient-based one-sided sampling (GOSS) and exclusive feature packing (EFB). GOSS is a state-of-the-art sampling approach for GBDT that preserves accuracy while reducing the number of data samples. EFB is an innovative technique for efficiently reducing the number of features.

### Categorical Boosting Algorithm (CatBoost)

Prokhorenkova et al. (2018) developed CatBoost algorithm. It is produced as an alternative to XGBoost, LightGBM and gradient-boosting machine algorithms. It is particularly appropriate for use in categorical data. It has been observed that it shows higher performance on large data sets. It is implemented in classification and regression applications. The basic formulation in Equation 12 of CatBoost to optimize the loss function is similar to classical Gradient Boosting:

$$\hat{y} = F(X) = \sum_{m=1}^M \gamma_m h_m(X) \quad (12)$$

$X$  is input feature vector,  $M$  is the number of trees in the ensemble,  $h_m$  refers to the  $m$ th decision tree and  $\gamma_m$  is the learning rate is also weight for tree  $m$ . The CatBoost algorithm can produce results with good performance without much hyperparameter tuning. It has a fast model training feature without encountering overfitting problems. It is a faster algorithm compared to GB, especially on categorical data.

### Gradient Boosting

The Gradient Boosting algorithm, first proposed by Friedman (2002), has later been developed into versions such as XGBoost, LightGBM and CatBoost. Gradient Boosting is an ensemble ML algorithm that minimizes errors step by step using decision trees. It is one of the supervised learning methods and is used in classification and regression applications. It follows an iterative optimization step to



minimize the propagation function in its formula in Equation 13.

$$F_m(X) = F_{m-1}(X) + \gamma_m h_m(X) \quad (13)$$

$F_m(X)$  is prediction at iteration  $m$ ,  $F_{m-1}(X)$  refers to the prediction from the previous iteration,  $h_m(X)$  is the new weak learner and  $\gamma$  is the learning rate and steps size.

Gradient Boosting combines multiple weak learners to create a stronger model. Since it uses decision trees, it is considered a tree-based approach. Its loss functions can be customized.

## Experimental Details

Orthomosaic and DEM were produced with DJI Terra software from multispectral aerial images obtained with appropriate flight planning. The generated ortho mosaic and DEM have an image size of 5207 x 4063 pixels. Five plant indices were calculated with appropriate band calculations: NDVI, OSAVI, NDRE, LCI, and GNDVI. Four data sets were obtained using orthomosaic, DEM and plant indices.

- 4-band: R, G, NIR, REDEGE
- 4-band + DEM: R, G, NIR, REDEGE, DEM
- 9-band: R, G, NIR, REDEGE, NDVI, OSAVI, NDRE, LCI, GNDVI
- 9-band +DEM: R, G, NIR, REDEGE, NDVI, OSAVI, NDRE, LCI, GNDVI, DEM

Training and test samples were selected from the image to apply ML classifiers. It selected 194860 samples for training and 2446940 samples for testing. The distribution of train and test samples are shown in Figure 4. XGBoost, LightGBM, Gradient Boosting and CatBoost methods were used to evaluate these data. The most appropriate parameter selection for each classifier was determined by GridSearch method. Accordingly, training and testing were performed with the combinations of values determined for each parameter in the parameter group of a method. The parameter combination with the highest accuracy was selected. All classification

was performed in Python software language and using the scikit learn library. For the experiments, i7-11800H, 2.30 GHz processor, GTX 3070 graphics card, and 64 GB RAM hardware was used.

The performance of the classifiers was measured by precision, recall, F1-score and accuracy metrics in Equation 14-17.

$$Precision = \frac{TP}{TP + FP} \quad (14)$$

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

$$F1 - Score = \frac{2 \times OA \times Recall}{OA + Recall} \quad (16)$$

$$Overall Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

The number of properly detected pixels when the real label and the forecast are both positive is known as the real Positive (TP). The number of properly classified pixels where the real label and the forecast are both negative is known as the real Negative (TN). False Positive (FP) is the term used to describe the quantity of pixels that are mistakenly positively detected when the genuine label is negative but the forecast is positive. Pixels with an improperly negative classification—where the genuine label is positive but the forecast is negative are known as False Negatives (FN) (Alpaydin, 2020).

## Results and Discussion

The results in Table 3 compare the performance of different band combinations (4 bands, 4 bands + DEM, 9 bands, 9 band + DEM) and different ensemble algorithms (XGBoost, LightGBM, Gradient Boosting, CatBoost) using Accuracy and F1-score metrics. When 4-Band dataset is used, LightGBM algorithm achieves the highest value (85.3% and 81.3%) in terms of Accuracy and F1-score. CatBoost algorithm has a slightly lower performance in this combination compared to the other three algorithms.

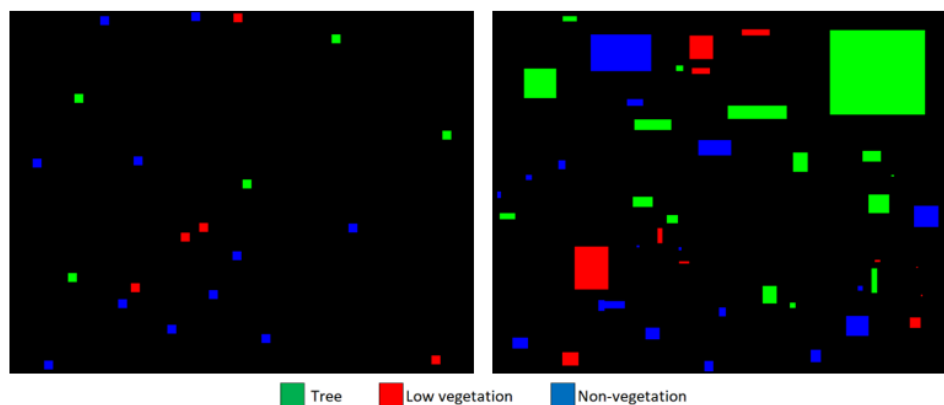


Figure 4. Distribution of train (left) and test (right) samples

**Table 3.** Results of the experiments according to accuracy and F1-scores of the algorithms. All values are in %

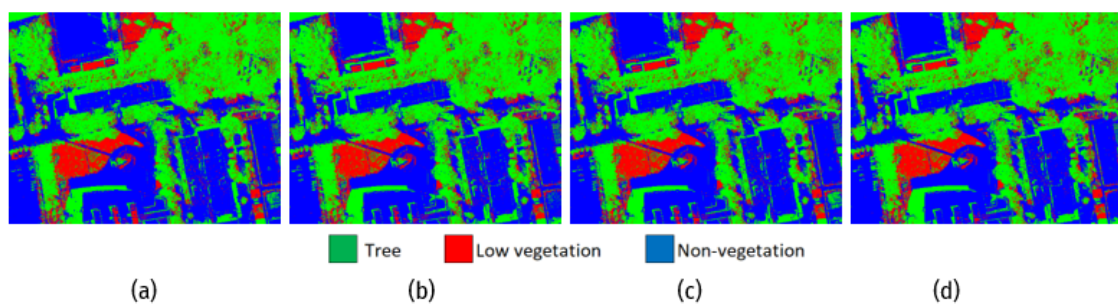
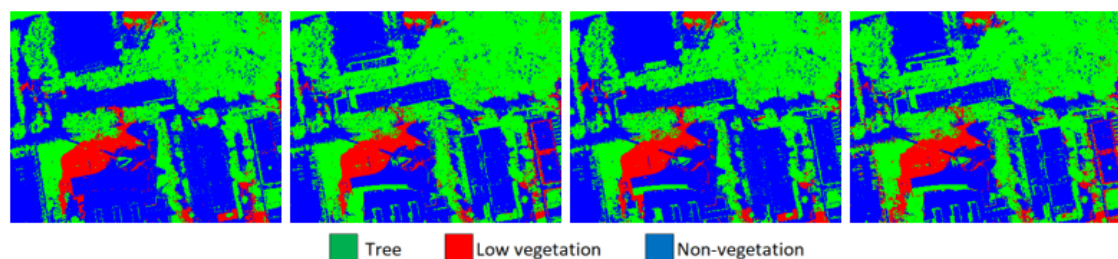
Classifier	4-BAND		4-BAND + DEM		9-BAND		9-BAND + DEM	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
XGBoost	85.1	81.1	84.1	75.4	85.4	81.6	88.7	84.9
LightGBM	<b>85.3</b>	<b>81.3</b>	<b>85.8</b>	77.6	<b>85.5</b>	<b>82.0</b>	<b>90.2</b>	86.5
Gradient Boosting	85.0	81.1	84.0	75.3	85.2	81.2	87.8	83.8
CatBoost	84.5	80.7	85.5	<b>78.2</b>	<b>85.5</b>	<b>82.0</b>	<b>90.2</b>	<b>86.9</b>

Adding DEM data to the dataset provided an increase in accuracy for LightGBM and CatBoost algorithms compared to 4-band data (in Figure 5 - Figure 6). In this dataset, LightGBM had the highest Accuracy (85.8%) while CatBoost achieved the highest F1-score (78.2%). Adding DEM for XGBoost and Gradient Boosting gave slightly lower or similar results (especially in XGBoost, Accuracy dropped from 85.1% to 84.1%).

When 9-Band data is used (Figure 7), all algorithms show similar and relatively high performance (approximately 85 Accuracy, approximately 81–82 F1-score). It has been observed that the results of the LightGBM and CatBoost algorithms are at the same level. The combination of 9 bands + DEM gives the highest results for all classifiers. LightGBM and CatBoost algorithms have neck and neck performance on Accuracy. In the F1 score, CatBoost classifier (86.9%) produced slightly better results than LightGBM (86.5%). It has been observed that XGBoost and Gradient Boosting show significant improvements compared to their previous combinations with 9-Band + DEM.

It has been observed that increasing the number of bands from 4 bands to 9 bands and adding DEM information generally increases the classification performance. The highest performance was achieved with the combination of 9-Band + DEM. Among the algorithms, LightGBM and CatBoost were mostly seen to give the best results. Especially CatBoost (general accuracy = 90.2 % , F1-score = 86.9%) and LightGBM (general accuracy = 90.2%, F1-score = 86.5%) have the best performance with 9-Band + DEM. These results show that increasing the use of spectral bands for classification (9 bands) and using additional elevation information (DEM) increased the discrimination power of the models. In addition, it was seen that CatBoost and LightGBM were partly more stable and provided higher success among different ensemble algorithms.

When we evaluate the results on a class basis, the precision for the tree class in all models is quite high, usually in the range of 92–98%. Recall increases with the number of bands and the addition of DEM; in the 4-band dataset, accuracies reach 90% with 9-band + DEM. In particular, CatBoost and

**Figure 5.** Qualitative results of the methods for 4-Band. (a) XGBoost; (b) LightGBM; (c) Gradient Boosting; (d) Catboost**Figure 6.** Qualitative results of the methods for 4-Band + DEM. (a) XGBoost; (b) LightGBM; (c) Gradient Boosting; (d) Catboost



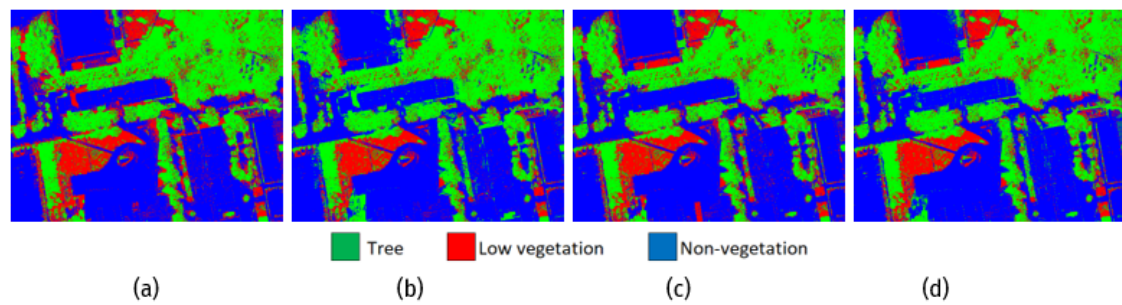


Figure 7. Qualitative results of the methods for 9-Band. (a) XGBoost; (b) LightGBM; (c) Gradient Boosting; (d) Catboost

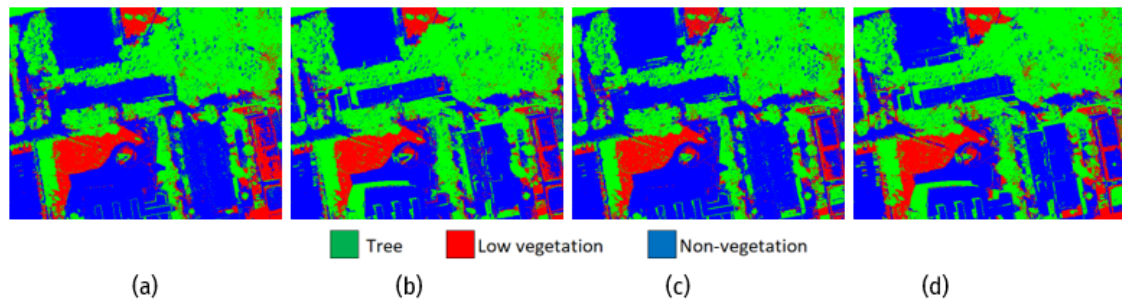


Figure 8. Qualitative results of the methods for 9-Band+DEM/ (a) XGBoost; (b) LightGBM; (c) Gradient Boosting; (d) Catboost

LightGBM algorithms reach approximately 97–98% precision and 91–92% recall classification accuracy for the tree class (Figure 8) when the 9-Band + DEM dataset is used (Table 4). The low vegetation class was observed to have the lowest accuracy of the models. In the 4-Band dataset, precision was obtained between 57–61% and recall values were obtained between 73–75%. It was observed that when 4-Band + DEM was added, the values in precision metrics increased significantly in Figure 6. In addition, recall values were observed to decrease significantly. This indicates that the model labeled

the low vegetation class more selectively, thus decreasing false positives but increasing false negatives.

Although 4-band data (R, G, B and NIR) provides reasonable performance on its own, better results were observed especially in low vegetation class and partially in tree classes when 9-band dataset was used. Adding DEM band increased the recall value of tree class and it was found that low vegetation class also increased its precision value significantly. However, adding DEM to 4-band dataset caused a significant decrease in low vegetation recall metric. 9-band + DEM generally gave the highest accuracy and best F1

Table 4. Results of the experiments according to classes of the algorithms. All values are in %

Classifier	Class	4-BAND		4-BAND + DEM		9-BAND		9-BAND + DEM	
		Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
XGBoost	Tree	97.5	84.5	92.5	90.9	97.6	84.0	97.8	90.8
	Low veg.	60.2	74.1	92.2	38.8	61.2	73.2	84.1	72.1
	Non-veg.	79.4	94.9	66.9	99.4	79.2	98.3	73.2	95.6
LightGBM	Tree	97.3	84.7	94.1	92.7	97.8	83.5	97.4	92.5
	Low veg.	61.1	73.6	88.7	43.2	62.8	74.7	89.2	70.5
	Non-veg.	79.2	95.7	69.1	98.8	78.6	99.4	76.2	98.8
Gradient Boosting	Tree	97.5	84.4	92.8	91.1	97.8	84.6	96.1	90.1
	Low veg.	59.3	73.9	89.9	39.8	60.4	74.1	88.7	66.2
	Non-veg.	80.0	95.2	66.2	97.3	79.0	95.3	71.5	97.6
CatBoost	Tree	97.4	83.2	93.8	92.1	98.1	83.8	98.0	91.4
	Low veg.	57.8	74.6	88.8	47.0	60.2	76.5	86.9	74.1
	Non-veg.	80.3	95.3	68.5	96.3	80.7	97.1	76.4	99.2



scores, and when looking at the table, it was seen that it produced more balanced results for all three classes. When the differences on class basis were considered, CatBoost and LightGBM achieved higher precision or more balanced precision-recall especially in low vegetation class.

According to the findings obtained in the study, it was found that using more spectral bands (9 bands) and adding DEM information generally increased the classification quality. The performance improvement is especially noticeable in a class that is harder to distinguish, such as low vegetation. While the tree class already has a high precision value in all datasets; the recall value also increased with the increase in the number of DEM and bands. The very high recall value in the non-vegetation class showed that the models tended not to miss pixels belonging to this class. However, the precision metric was slightly lower during this time. CatBoost and LightGBM algorithms gave the highest results in both precision and recall metrics in the Low vegetation class in the 9-band + DEM dataset. Depending on the purpose of the application, if the low vegetation class is intended to be classified with the highest accuracy, in other words, if high recall is required, optimization can be made with parameter settings or class weights. On the other hand, decision-making thresholds or model hyperparameters can be optimized to reduce false positives, i.e. to increase precision. Consequently, CatBoost or LightGBM models using 9-band + DEM data; it provides the best balance in both general accuracy and F1-score metrics and in class-based evaluation, especially in the low vegetation class. Changing hyperparameter settings or decision thresholds according to the class priorities required by the application, such as requiring a higher recall for the low vegetation class, can provide even more optimized results.

## Conclusion

In this study, ensemble ML-based supervised classification of urban vegetation was performed using multispectral orthomosaic, DEM, and vegetation indices generated from UAV images. As a result, it was observed that CatBoost or LightGBM models using 9-band + DEM data provided the best balance in both overall accuracy and F1-score metrics and especially in evaluating the low-vegetation class. The outcomes demonstrate that multispectral UAVs offer high-resolution, economical, and time-efficient data for monitoring in areas with diverse environmental conditions and especially for vegetation and tree classes.

The results of this study are promising, but more research is needed on the data. The difference in the density of the vegetation and the changes in different seasons increase the complex structure of the data used. Therefore, the algorithm

to be used in such applications needs to be improved. Like Kaya and Dervisoglu's urban areas classification study (2023), the application can be enhanced with new studies to which many spectral indices will be added.

Overall, this study aimed to distinguish urban vegetation in a selected area with heterogeneous classes using multi-spectral UAV data and ensemble ML methods. It features groundwork for understanding the effects of different methods, the effectiveness of multi-spectral data, and the DEM effect for future studies on this subject. The study can be more interpretable and transparent in future applications by adding explainable artificial intelligence techniques to ML methods.



Peer Review	Externally peer-reviewed.
Conflict of Interest	The author has no conflict of interest to declare.
Grant Support	The author declared that this study has received no financial support.

## Author Details

### Şaziye Özge Atik

<sup>1</sup> Istanbul Technical University, Faculty of Civil Engineering, Department of Geomatic Engineering İstanbul Türkiye

ORCID: 0000-0003-2876-040X Email: donmezszaz@itu.edu.tr

## REFERENCES

- Abdollahnejad, A., & Panagiotidis, D. (2020). Tree species classification and health status assessment for a mixed broadleaf-conifer forest with UAS multispectral imaging. *Remote Sensing*, 12(22), 3722.
- Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.
- Atik, M. E., & Arkali, M. (2025). Comparative Assessment of the Effect of Positioning Techniques and Ground Control Point Distribution Models on the Accuracy of UAV-Based Photogrammetric Production. *Drones*, 9(1), 15. <https://doi.org/10.3390/drones9010015>
- Atik, M. E., Duran, Z., & Seker, D. Z. (2024). Explainable Artificial Intelligence for Machine Learning-Based Photogrammetric Point Cloud Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
- Atik, Ş. Ö. (2023). Çok Yüksek Çözünürlüklü Uydur Görüntülerinden Bina Çıkarımında Derin Öğrenme ve Çoklu-Çözünürlüklü Bölütleme Kullanılarak Nesne-Tabanlı Entegrasyon. *Türkiye Uzaktan Algılama Dergisi*, 5(2), 67-77.
- Atik, S. O., & Atik, M. E. (2024). Optimal band selection using explainable artificial intelligence for machine learning-based hyperspectral image classification. *Journal of Applied Remote Sensing*, 18(4), 042604-042604.
- Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., ... & Moran, M. S. (2000, July). Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In *Proceedings of the fifth international conference on precision agriculture*, Bloomington, MN, USA (Vol. 1619, No. 6).
- Bayirhan, İ., & Gazioglu, C. (2020). Use of Unmanned Aerial Vehicles (UAV) and Marine Environment Simulator in Oil Pollution Investigations. *Baltic Journal of Modern Computing*, 8(2).
- Cao, Q., Li, M., Yang, G., Tao, Q., Luo, Y., Wang, R., & Chen, P. (2024). Urban Vegetation Classification for Unmanned Aerial Vehicle Remote Sensing Combining Feature Engineering and Improved DeepLabV3+. *Forests*, 15(2), 382. <https://doi.org/10.3390/f15020382>



- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4), 367-378.
- Gallardo-Salazar, J. L., & Pompa-García, M. (2020). Detecting individual tree attributes and multispectral indices using unmanned aerial vehicles: Applications in a pine clonal orchard. *Remote Sensing*, 12(24), 4144.
- Gazioğlu, C., Varol, Ö. E., Şeker, D. Z., & Çağlar, N. (2017, December). Determination of the Environmental Impacts of
- Gitelson, A.A., Kaufman, Y.J., & Merzlyak, M.N. (1996). Use of a green channel in remote sensing of global vegetation from EOSMODIS. *Remote Sens Environ*, 58, 289-298. *Inequity*.
- Kaya, Z., & Dervisoglu, A. (2023). Determination of urban areas using Google Earth Engine and spectral indices; Esenyurt case study. *International Journal of Environment and Geoinformatics*, 10(1), 1-8.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.
- Lan, Y., Huang, Z., Deng, X., Zhu, Z., Huang, H., Zheng, Z., ... & Tong, Z. (2020). Comparison of machine learning methods for citrus greening detection on UAV multispectral images. *Computers and electronics in agriculture*, 171, 105234.
- López-García, P., Intrigliolo, D., Moreno, M. A., Martínez-Moreno, A., Ortega, J. F., Pérez-Álvarez, E. P., & Ballesteros, R. (2022). Machine Learning-Based Processing of Multispectral and RGB UAV Imagery for the Multitemporal Monitoring of Vineyard Water Status. *Agronomy*, 12(9), 2122. <https://doi.org/10.3390/agronomy12092122>
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60, 91-110.
- Marine Accidents Using UAV and RS Technologies. In *19th MESAEP Symposium on Environmental and Health*
- Momm, H., & Easson, G. (2011, April). Feature extraction from high-resolution remotely sensed imagery using evolutionary computation. In *Evolutionary Algorithms*. IntechOpen.
- Munyati, C. (2000). Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International journal of remote sensing*, 21(9), 1787-1806.
- 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.
- Peñuelas, J., Filella, I., Biel, C., Serrano, L., & Save, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *International journal of remote sensing*, 14(10), 1887-1905.
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31.
- Pu, R., Gong, P., & Yu, Q. (2008). Comparative analysis of EO-1 ALI and Hyperion, and Landsat ETM+ data for mapping forest crown closure and leaf area index. *Sensors*, 8(6), 3744-3766.
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote sensing of environment*, 55(2), 95-107.
- Schonberger, J. L., & Frahm, J. M. (2016). Structure-from-motion revisited. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4104-4113).
- Sefercik, U. G., Nazar, M., Aydin, I., Büyüksalih, G., Gazioglu, C., & Bayirhan, I. (2024). Comparative analyses for determining shallow water bathymetry potential of multispectral UAVs: case study in Tavşan Island, Sea of Marmara. *Frontiers in Marine Science*, 11, 1388704.
- Shu, M., Fei, S., Zhang, B., Yang, X., Guo, Y., Li, B., & Ma, Y. (2022). Application of UAV multisensor data and ensemble approach for high-throughput estimation of maize phenotyping traits. *Plant Phenomics*.
- Specs—DJI Mavic 3 Enterprise. Retrieved 15 February 2025 from <https://ag.dji.com/mavic-3-m>
- Wang, N., Guo, Y., Wei, X., Zhou, M., Wang, H., & Bai, Y. (2022). UAV-based remote sensing using visible and multispectral indices for the estimation of vegetation cover in an oasis of a desert. *Ecological Indicators*, 141, 109155.
- Zhao, S., Kang, F., Li, J., & Ma, C. (2021). Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction. *Automation in Construction*, 130, 103832.
- Zheng, Z., Yuan, J., Yao, W., Kwan, P., Yao, H., Liu, Q., & Guo, L. (2024). Fusion of UAV-Acquired Visible Images and Multispectral Data by Applying Machine-Learning Methods in Crop Classification. *Agronomy*, 14(11), 2670. <https://doi.org/10.3390/agronomy14112670>

