



## Semantic segmentation of very-high spatial resolution satellite images: A comparative analysis of 3D-CNN and traditional machine learning algorithms for automatic vineyard detection

Özlem Akar<sup>\*1</sup>, Ekrem Saralioğlu<sup>2</sup>, Oğuz Güngör<sup>3</sup>, Halim Ferit Bayata<sup>4</sup>

<sup>1</sup>Erzincan Binali Yıldırım University, Department of Architecture and City Planning, Türkiye, ozlemerden@gmail.com

<sup>2</sup>Artvin Coruh University, Department of Geomatics Engineering, Türkiye, ekremsaralioglu@artvin.edu.tr

<sup>3</sup>Ankara University, Department of Real Estate Development and Management, Türkiye, ogungor@ankara.edu.tr

<sup>4</sup>Erzincan Binali Yıldırım University, Department of Civil Engineering, Türkiye, hfbayata@erzincan.edu.tr

Cite this study:

Akar, Ö, Saralioğlu, E., Güngör, O., & Bayata, H. F. (2024). Semantic segmentation of very-high spatial resolution satellite images: A comparative analysis of 3D-CNN and traditional machine learning algorithms for automatic vineyard detection. *International Journal of Engineering and Geosciences*, 9(1), 12-24

<https://doi.org/10.26833/ijeg.1252298>

### Keywords

Machine Learning  
Deep Learning  
Random Forest  
Gabor  
Image Classification

### Research Article

Received:17.02.2023

Revised: 09.05.2023

Accepted:26.06.2023

Published:02.01.2024



### Abstract

The Erzincan (Cimin) grape, which is an endemic product, plays a significant role in the economy of both the region it is cultivated in and the overall country. Therefore, it is crucial to closely monitor and promote this product. The objective of this study was to analyze the spatial distribution of vineyards by utilizing advanced machine learning and deep learning algorithms to classify high-resolution satellite images. A deep learning model based on a 3D Convolutional Neural Network (CNN) was developed for vineyard classification. The proposed model was compared with traditional machine learning algorithms, specifically Support Vector Machine (SVM), Random Forest (RF), and Rotation Forest (ROTF). The accuracy of the classifications was assessed through error matrices, kappa analysis, and McNemar tests. The best overall classification accuracies and kappa values were achieved by the 3D CNN and RF methods, with scores of 86.47% (0.8308) and 70.53% (0.6279) respectively. Notably, when Gabor texture features were incorporated, the accuracy of the RF method increased to 75.94% (0.6364). Nevertheless, the 3D CNN classifier outperformed all others, yielding the highest classification accuracy with an 11% advantage (86.47%). The statistical analysis using McNemar's test confirmed that the  $\chi^2$  values for all classification outcomes exceeded 3.84 at the 95% confidence interval, indicating a significant enhancement in classification accuracy provided by the 3D CNN classifier. Additionally, the 3D CNN method demonstrated successful classification performance, as evidenced by the minimum-maximum F1-score (0.79-0.97), specificity (0.95-0.99), and accuracy (0.91-0.99) values.

### 1. Introduction

Most grape varieties are known to be derived from *vitis vinifera*, the ancient grape often mentioned in the Bible. It was originated in the southern parts of the region between the Caspian and the Black Sea and has been carried all over the world by civilized people [1]. Today this perennial species is cultivated worldwide, including subtropical regions, since it is tolerant of most climates and soil types, and easy to grow. Grapes can be cultivated in the temperate climate zone, generally between 30° and 50° latitudes in both hemispheres. Depending on the latitude, it can be grown up to 1600-

1800 m altitudes [2,3]. According to the Food and Agriculture Organization of the United Nations (FAO), the grape was cultivated on 7.7 million hectares of land worldwide in 2019 and Spain had a 22.7-percent share followed by France, China, Italy, and Türkiye. In 2019, the global export of fresh grapes reached a total of 10 million tons, with Türkiye accounting for 1.7 million tons. South Africa and Spain ranked second and third in fresh grape exports, with 1.3 million tons and 1.2 million tons respectively, following Türkiye. In the same year, global raisin export was 3 million tons and 73 percent (2,1 million tons) was exported by Türkiye alone, followed by the USA and Iran. Fresh grape yield increased by 2.7

percent in 2020 compared to the previous year, reaching 4.2 million tons [4]. According to the statistics of the Erzincan Directorate of Provincial Agriculture and Forestry [5], 5402 tons of fresh grapes were produced in 2021 from the 949.5-hectare vineyard in the province of Erzincan, located in north-eastern Anatolia.

In Erzincan, the "Cimin grape" dominates the grape varieties grown, with its cultivation mainly concentrated in the Üzümlü district, named after its reputation for abundant grape production. After the application of Üzümlü Municipality, the Erzincan Cimin grape was patented by the Turkish Patent Institute in 2001. Since the Cimin grape is the most important endemic agricultural product for Erzincan's economy, conservation, monitoring, and dissemination of this variety is increasingly becoming vital for the region. This can be accomplished with modern agricultural practices powered by remote sensing and geographic information systems, which have become increasingly popular in recent years.

Due to increasing radiometric, spatial, spectral, and temporal resolutions, remote sensing data is used by different disciplines in many different applications including land cover/use mapping, urban and environmental change analysis, object extraction, crop monitoring, disease detection, yield estimation, etc. Furthermore, extracting information classes and determining their spatial distributions in a scene using different multispectral image classification algorithms provide valuable information for various applications needing geo-spatial data.

Multispectral image classification has been used in monitoring tropical forests, which are an important and rich source in terms of biological diversity [6], in determining the spatial distribution of vineyards [7-9], in monitoring coastal changes [10], in monitoring urban development [11], in agriculture [12], designing rangeland information systems [13], in object extraction [14], in land cover classification [15-20], burning area mapping [21], and in classification of different product types [22].

The process of labelling pixels according to their intensity values to transform them into meaningful land cover data is also called image classification or information extraction [23]. Image classification is the process of categorizing pixels in an image using logical decision rules in the spatial domain or statistical decision rules in the spectral domain. The spectral values of the remotely sensed data are used to classify images in the spectral domain. The geometric size, form, texture, and pattern of pixels or objects are at the foreground of spatial domain decision criteria [23]. Various machine learning-based methods such as random forest, artificial neural networks, and support vector machines have been used in recent years to obtain reliable and most accurate information from satellite images efficiently by image classification. There are many studies in which the RF classifier stands out in terms of classification accuracy when it comes to the use of machine learning methods in agriculture [24-29]. In the realm of precision agriculture applications, deep learning algorithms, which are a subset of machine learning techniques, have gained

significant popularity in recent years due to their ability to provide more precise and reliable detection of agricultural products on images [30]. Grinblat et al. were able to detect plant species with high accuracy by using deep learning algorithms to identify plants from their vascular structures [31]. Deep learning algorithms were used by Ferentinos et al. to distinguish diseased plants [32]. Among 25 different plant species, they detected diseased ones with 99.53 percent accuracy. Chlingaryan et al. identified plant species by classifying images with 99.58 percent accuracy using deep learning algorithms [33], and thus made a crop yield estimation [34]. Zhao et al. used deep learning models to produce crop type mapping with sufficient accuracy [35]. Zhong et al. also used RF, SVM, and 1-D Convolutional Neural Network Model (Conv1D) to classify agricultural products and highlighted that the Conv1D yielded satisfactory results [36].

The objective of this study was to identify the cultivation areas of Cimin, an indigenous fruit cultivated in the Üzümlü region, utilizing a deep learning method based on CNN and commonly used machine learning algorithms. The findings demonstrate that by leveraging a pre-trained CNN, vineyards throughout the region can be automatically detected without requiring additional supervised learning. This ability stands as a significant advantage of deep learning architectures over traditional machine learning methods.

## 2. Study area and data set

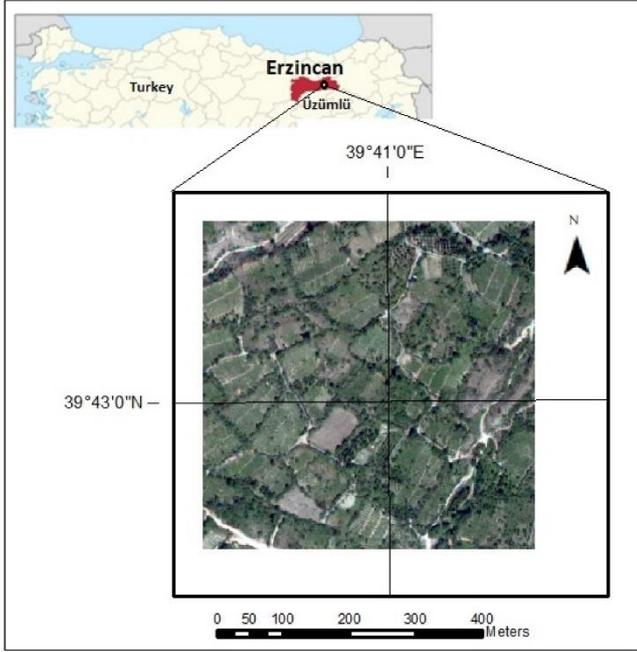
In the Üzümlü Town, where the Erzincan grape is cultivated, a pilot study area measuring 25 hectares was designated (Figure 1). This study area is situated at coordinates 39° 41' 00" East and 39° 43' 00" N, within the Upper Euphrates section of Eastern Anatolia, in the province of Erzincan. Around 80% of the administrative boundary of the Üzümlü district is located in the Esence Mountains region to the north of the Erzincan basin, while the remaining 20% lies within the Erzincan plain. With an area of 410 km<sup>2</sup>, Üzümlü is the second smallest district of Erzincan [37].

For this study, the satellite image used was the Worldview-2 (WV-2) image, which covered a significant portion of Cimin vineyards in the Üzümlü region. The WV-2 satellite image consists of 8 multispectral (MS) bands with a spatial resolution of 2 meters, along with a panchromatic band offering a higher spatial resolution of 0.5 meters. The 8 MS bands encompass the following spectral ranges: Coastal, Blue, Green, Yellow, Red, Red Edge, Near-Infrared 1, and Near-Infrared 2. It is worth noting that the satellite image had undergone prior atmospheric, radiometric, and geometric corrections, enabling its direct utilization for classification purposes without requiring any pre-processing steps.

## 3. Method

This study had two purposes: (a) to determine the locations and distributions of vineyards accurately and robustly by using widely used machine learning algorithms such as SVM, RF, ROTF and CNN-based deep

learning technique (b) to investigate the performance of CNN-based deep learning technique by comparing it aforementioned machine learning algorithms. To better distinguish the vineyards by taking advantage of the sub-meter spatial resolution of PAN image, the WV-2 MS and WV-2 PAN bands were fused with the Hyper Spherical Color Space pan-sharpening method (HCS). The literature contains numerous studies showcasing the efficacy of Padwick's HCS image fusion method [38] in maintaining the spectral and spatial characteristics of both multispectral and panchromatic input images during the fusion process, particularly when applied to Worldview-2 images [38-41].



**Figure 1.** Study area.

The pan-sharpened image (1000x1000 pixels) was categorized into five land cover classes, namely vineyard, forest, soil, road, and shadow. ENVI software was utilized to select a total of 70505 pixels, employing a random feature selection approach in MATLAB, to generate the training and test data. Subsequently, the fused image was subjected to classification using 3D CNN, SVM, RF, and ROTF algorithms. To determine the optimal parameters for classification, a trial-and-error strategy was employed. The optimum parameters for this study were determined as  $m=3$ ,  $N=350$  for RF,  $K=3$ ,  $L=3$  for ROTF, and  $C=100$  for SVM. The 3D CNN model employed in this study utilized both spectral features and texture features, as detailed in [42], to classify the image. Moreover, the

Gabor filter was utilized to extract texture data, which was then integrated into the RF classifier as explained in [26]. Among the machine learning methods employed in the study, the RF classifier exhibited the highest classification accuracy. Consequently, a comparison was conducted to evaluate the performances of these two methods. To identify the optimal parameter values and filter sizes for texture extraction using the Gabor filter, a trial-and-error method was employed to determine the values that most accurately represented the vineyards. Subsequently, the image was classified using these identified parameter values (Figure 2).

### 3.1. 3D Convolutional neural network model

Deep learning, which is usually defined by neural networks with more than two hidden layers, has been named one of the top ten breakthrough technologies of 2013 [43]. Deep learning model used in this study is created on the structure of CNN. Three-dimensional (3D) convolution is naturally suitable for spatial-temporal studies. Recently, some studies have been conducted on learning spatial-temporal features from video [44, 45], LIDAR point clouds [46], temporal images [47] and hyperspectral images [48]. In general, 3D CNN is not as widely used as 2D CNN because the temporal dimension is typically ignored in machine learning and computer vision applications. Remote sensing images, on the other hand, frequently provide dynamic or temporal information from which more information can be extracted. CNNs, which are widely used in image processing, are also useful for classifying satellite images [49].

High spatial resolution multispectral images with more than three spectral bands contain a lot of spectral information. To extract both spatial and spectral information from a multidimensional image, 3D convolution is preferred. By utilizing 3D convolution, the interaction between various spectral bands can be effectively modeled, encompassing both spatial and spectral information. Unlike 2D convolution, which focuses solely on spatial details, 3D convolution takes into consideration both spatial and spectral aspects, as highlighted in reference [50]. The integration of spatial and spectral information is crucial for improving the accuracy of satellite image classification. In this particular study, 3D convolution layers were employed to capture the spectral relationships among the 8 bands present in the WV-2 image. Consequently, these 8 bands were utilized as input data for the 3DCNN model.

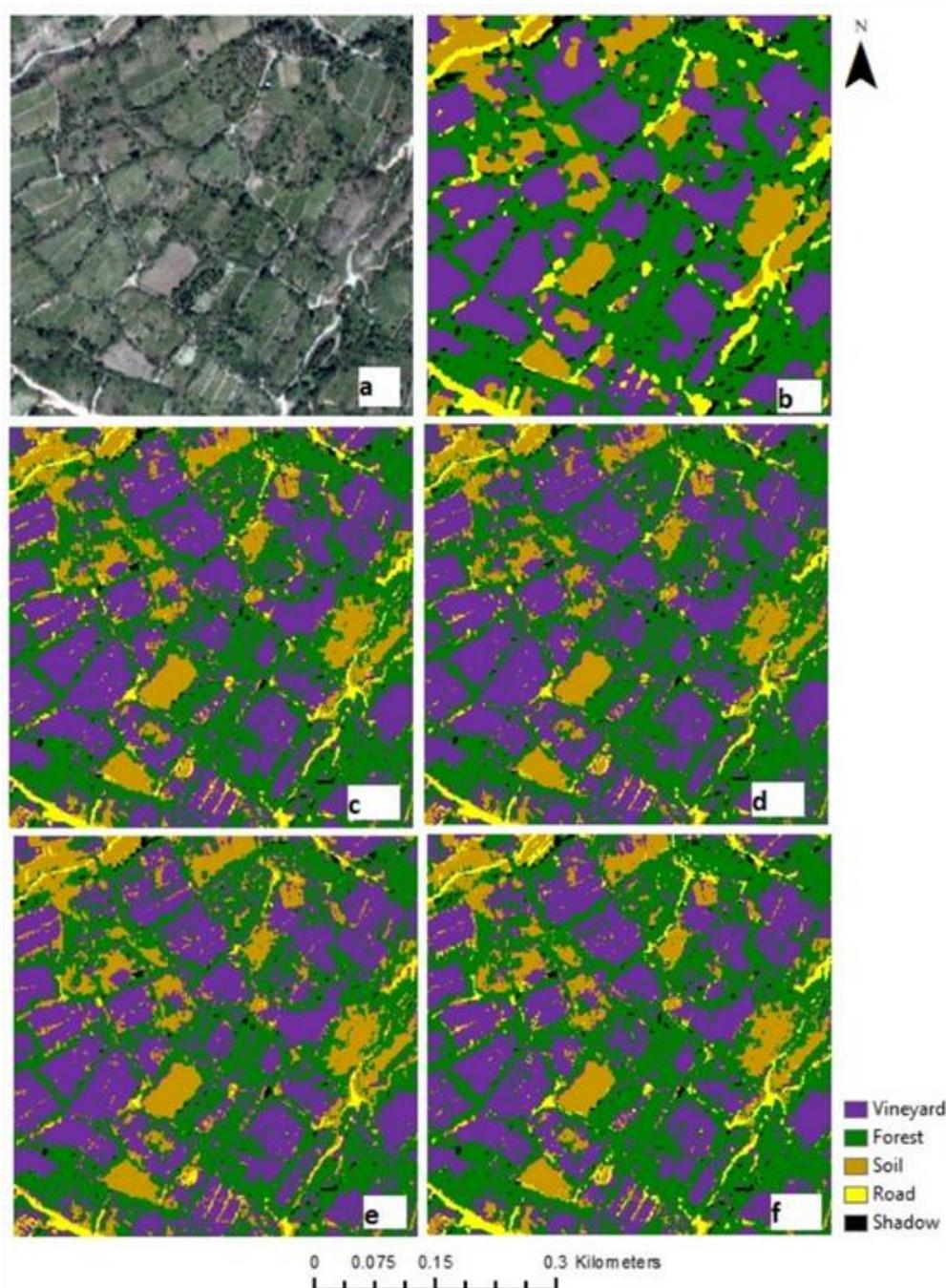
$$v_{ij}^{xyz} = f(b_{ij} + \sum_{p=0}^{P_i-1} \sum_{s=0}^{S_{i-1}} \sum_{q=0}^{Q_{i-1}} \sum_{r=0}^{R_{i-1}} w_{ijp}^{qrs} v_{(i-1)p}^{(x+q)(y+r)(z+s)}) \quad (1)$$

In Equation 1,  $v$  represents the output of feature maps,  $S, Q, R$  defines spectral and spatial kernel dimensions where  $(s, q, r)$  are kernel, and  $(x, y, z)$  are feature map indices. While  $w$  specifies the kernel parameters,  $i, j, p$  represent the input layer, output layer, and feature map indices, respectively.  $P$  is the number of

feature maps.  $P_i$  represents the feature maps in the  $i^{th}$  layer. While the bias term is denoted by  $b$ ,  $f$  represents the activation function of PReLU used in the model. Python programming language on Jupyter notebook and TensorFlow and Keras library in the background were used to create the model and classify the image with this

model. Jupyter Notebooks is an open-source and browser-based tool that combines interpreted languages, libraries, and visualization tools [51]. A Jupyter Notebook can run locally or in the cloud. Typical outputs include text, tables, charts, and graphics. The computer on which the study was conducted has 24 GB of RAM, NDVI GTX 1650 GPU, and an i7 9750h processor. For this study, a model was created using four 3D convolution layers. Each layer had a filter size of 3x3. The first, second, third, and fourth layers were configured with 128, 64, 32, and 16 filters, respectively. Two fully connected layers were used after the convolution layers. The first layer is a dense layer that performs a rough classification of the extracted features from the convolutional layer. The second layer is the model's final layer, and it is used to extract class scores with a Softmax

classifier. Softmax is a technique employed in multi-class classification tasks. It computes probability values for each class in order to classify a given input. The probability value associated with each class falls within the range of 0 to 1, and the sum of all probability values across all classes is equal to 1. Consequently, the Softmax classifier determines the probability values for each class in multi-class classification, ultimately selecting the class with the highest probability as the predicted class. The activation function was the Parametric Rectified Linear Unit (PReLU), the optimization method was Adam, and the subtraction function was categorical cross-entropy. The deep learning model created has 3118405 parameters in total. Figure 3 depicts the 3D CNN model that was used.



**Figure 2.** a) Fused image and classified images obtained by classification with b) 3D CNN, c) RF, d) ROTF, e) SVM, f) RF\_Gabor.

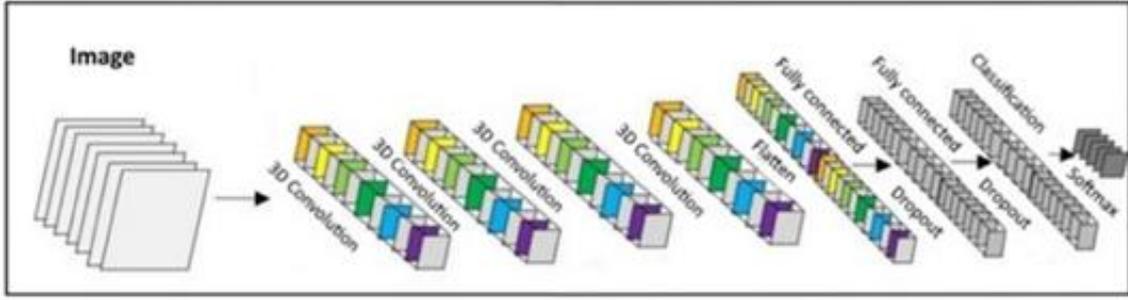


Figure 3. Architecture of 3D CNN model.

### 3.2. Random Forest

Random Forest is a classifier that can classify multiple variables and classes without the use of complex models or parameters [52]. The RF classifier outperforms many tree-based algorithms [53-56]. RF aims to develop a tree by dividing each node according to the GINI index, basically. Randomly selected variables are used instead of using all variables at each node [26]. Decision trees (DTs) are trained in RF using random bootstrap samples with the replacement of an original dataset [52]. The user defines two parameters:  $m$  and  $N$ .  $m$  (The number of variables) is used to determine the best split at each node.  $N$  is the number of DTs to be developed. Bootstrap samples are generated by randomly selecting two-thirds of the training Dataset. Trees are then built from these boot samples without employing pruning. The remaining one-third of the training dataset is reserved for use as the test dataset. The RF algorithm builds a large number of trees to determine the class of each pixel. The labeling of a pixel can vary across different trees, indicating that the total number of trees determines the frequency of the pixel's class assignment. The final class of the pixel is determined based on the class that has the highest occurrence among the candidate pixel's labels.

### 3.3. Rotation Forest

Rodriguez et al. [57] proposed the Rotation Forest algorithm as an ensemble method for encouraging both individual accuracy and classifier member [58]. The ROTF is a linear transformation method that creates a new performing space within another space [59]. Theoretically, the ROTF algorithm and the RF algorithm share similarities. Both methods aim to grow more than one tree in classification. However, ROTF creates the dataset in the Principle Component Analysis (PCA) feature space. It generates a large number of DTs from training datasets defined within a different feature space. The training dataset is divided into subsets, and feature extraction is performed using the feature space chosen from each subset. The ROTF algorithm has two user-specified parameters,  $K$  and  $L$ , which are required to determine the dataset used to grow each DT. This approach enables classification by training all classifiers in parallel [59].

### 3.4. Support Vector Machine

The Support Vector Machine classifier can distinguish linearly and nonlinearly separable data by finding the

best hyperplane for separating the classes [60]. If classes are linearly separable, it finds the planes separating them and uses these planes to construct a linear discriminant function. If classes cannot be separated linearly, the data is transformed to a higher-dimensional space in which the classes can be linearly separated by using a positive  $C$  parameter and a kernel function that minimizes classification error while maximizing the distance between planes [51-64]. The Radial basis function is the most commonly used kernel function because it performs well [65, 66]. The radial basis function is widely recognized for its exceptional performance in classification accuracy, making it the preferred choice as the most commonly used kernel function [65, 66].

### 3.5. Accuracy assessment

Congalton and Green [67] proposed the multinomial distribution to calculate the minimum number of samples needed to statistically calculate the classification accuracy. The minimum number of samples is calculated with the Equation 2 and 3 using the multinomial distribution approach;

$$n = \frac{B\Pi_i(1 - \Pi_i)}{b_i^2} \quad (2)$$

$$B = \left(\frac{\alpha}{k}\right) \times 100th \quad (3)$$

where  $n$  represents the number of reference pixels,  $\alpha$  represents the confidence interval,  $k$  represents the number of classes,  $\Pi_i$  is the ratio of the area of the  $i$ th class to the total area, and  $b_i$  represents the required accuracy. If no prior information about  $\Pi_i$  is available, the sample number is calculated using Equation 4 [68].

$$n = \frac{B}{4b^2} \quad (4)$$

In the study, the analysis of classification accuracy required determining the minimum number of reference points. For this purpose, a 95% confidence interval was utilized, with the number of classes set at 5. Therefore, in the calculation of the  $B$  value,  $\alpha/k = 0.05/5 = 0.01$  is used to find the corresponding value at 1 degree of freedom in the  $\chi^2$  distribution table as  $\chi^2_{(1,0.01)}=6.635$ . Accordingly, the minimum number of reference points was calculated as 664 as follows, yet 665 was used instead in the accuracy analysis (Equation 5).

$$n = \frac{B}{4b^2} = \frac{6.635}{4(0.05^2)} = 664 \quad (5)$$

The appropriate number of reference points required to create error matrices that evaluate the accuracy of each classification outcome was determined. The stratified method was utilized to distribute these points evenly across the image. Kappa ( $\kappa$ ) analysis is another approach that performs accuracy analysis by determining whether one error matrix is statistically significantly different from another. The  $\kappa$  value, calculated within the range of 0 to 1, provides a statistical assessment of the agreement among the utilized categories or classes. This value serves as a measure of fit and classification accuracy, with higher values indicating better alignment and accuracy (approaching 1) and lower values suggesting poorer fit and lower classification accuracy (approaching 0). Kappa, as pointed out by Pontius and Millones [69], has been criticized for attempting to compare accuracy with a baseline of randomness. As an alternative, they proposed the utilization of allocation and quantity disagreements, which leverage the distinctions between a reference map and a comparison map. The quantity disagreement focuses on disparities in the category proportions between the reference and comparison maps, while the allocation disagreement addresses differences in the spatial distribution of categories between the reference and comparison maps [70]. In light of this suggestion, allocation and quantification values were calculated in addition to the kappa coefficient for each image's post-classification accuracy assessment. Subsequently, the McNemar test was employed to assess the presence of statistically significant differences. The classification results obtained from all algorithms utilized in the study were compared pairwise using the McNemar test. The nonparametric McNemar test can be computed using the Equation 6 [71].

$$\chi^2 = \frac{(|f_{12} - f_{21}| - 1)^2}{f_{12} + f_{21}} \quad (6)$$

Where  $f_{12}$  denotes the number of pixels incorrectly classified by the second method but correctly classified by the first one, and  $f_{21}$  denotes the number of pixels

incorrectly classified by the first method but correctly classified by the second one.

Furthermore, the performance evaluation of both 3DCNN and the other methods employed in this study was conducted using metrics in Equation 7-11 including F1-Score, specificity, and accuracy.

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \quad (7)$$

$$F1 - Score = \frac{2xPrecisionxRecall}{Precision + Recall} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (11)$$

Where TP and TN represent true positive and true negative observations in the confusion matrix, respectively. FP and FN are false positive and false negative observations [72]. For each class in the confusion matrix of each method, the metrics were computed.

#### 4. Results and Discussion

By analyzing the error matrices, the classification accuracies of the classified images were examined for the SVM, RF, ROTF algorithms, and the 3D CNN model. The overall classification accuracies obtained from the error matrices were 86.47%, 70.53%, 66.92%, and 62.41% for the 3D CNN, RF, ROTF, and SVM, respectively (Table 1). In addition, the Mean IoU value for 3D-CNN, which includes the average of all classes, was found to be 84.93%. These results indicate that the 3D CNN method exhibited superior performance compared to RF by a margin of 16%, ROTF by 20%, and SVM by 24%. This analysis is further supported by the Kappa coefficients presented in Table 1.

**Table 1.** Overall classification accuracies and kappa analyses.

	3D CNN	RF	ROTF	SVM
Overall Accuracy	86.47	70.53	66.92	62.41
Kallocation	87.35	86.64	81.81	78.86
Kquantity	92.35	57.90	56.58	50.18
Khisto	95.11	72.47	71.09	66.54
Kcongaltion	83.08	62.79	58.16	52.47
Kcohen's	85.91	69.07	65.67	60.10

Upon examining the  $\kappa$  allocation values, it was evident that 3D CNN showcased superiority over RF by 1%, ROTF by 6%, and SVM by 8%. In terms of  $\kappa$  amount values, 3D CNN surpassed RF by 34%, ROTF by 36%, and SVM by 42%. Moreover, the  $\kappa$  histo values indicated that 3D CNN demonstrated a 24% improvement compared to RF, a 24% improvement compared to ROTF, and a 29%

improvement compared to SVM. The  $\kappa$  congaltion and  $\kappa$  cohen values presented in Table 1 further confirm the superiority of 3D CNN over other methods.

In order to enhance the accuracy of RF and assess its performance relative to 3D CNN, the texture features extracted using the Gabor filter were incorporated. The integration of RF texture features resulted in a

classification accuracy of 75.94%, representing a 5% improvement compared to the original RF performance (Table 2).

After comparing the Kappa values, it becomes clear that RF\_Gabor exhibits a 3% improvement over RF.

Moreover, Table 3 presents the computed Quantity and Allocation disagreement/agreement values, which were proposed by Pontius and Millones [69] for accuracy assessment.

**Table 2.** Contribution of texture to classification accuracy for RF.

	Vineyard	Forest	Soil	Road	Shadow	ΣRow	PA (%)	UA(%)	
Vineyard	136	20	5	15	1	177	92.52%	76.84%	
Forest	10	119	3	14	86	232	82.07%	51.29%	
Soil	1	6	120	24	8	159	91.60%	75.47%	
Road	0	0	3	59	0	62	52.68%	95.16%	
Shadow	0	0	0	0	35	35	26.92%	100.00%	
ΣColumn	147	145	131	112	130	665			
Overall accuracy	70.53%								
Khisto	72.47 %								

**Table 3.** Contribution of texture to classification accuracy for RF\_Gabor.

	Vineyard	Forest	Soil	Road	Shadow	ΣRow	PA (%)	UA(%)	
Vineyard	142	8	2	12	0	164	96.60%	86.59%	
Forest	4	135	4	19	88	250	93.10%	54.00%	
Soil	1	2	121	11	5	140	92.37%	86.43%	
Road	0	0	4	70	0	74	62.50%	94.59%	
Shadow	0	0	0	0	37	37	28.46%	100.00%	
ΣColumn	147	145	131	112	130	665			
Overall accuracy	75.94%								
Khisto	75.14%								

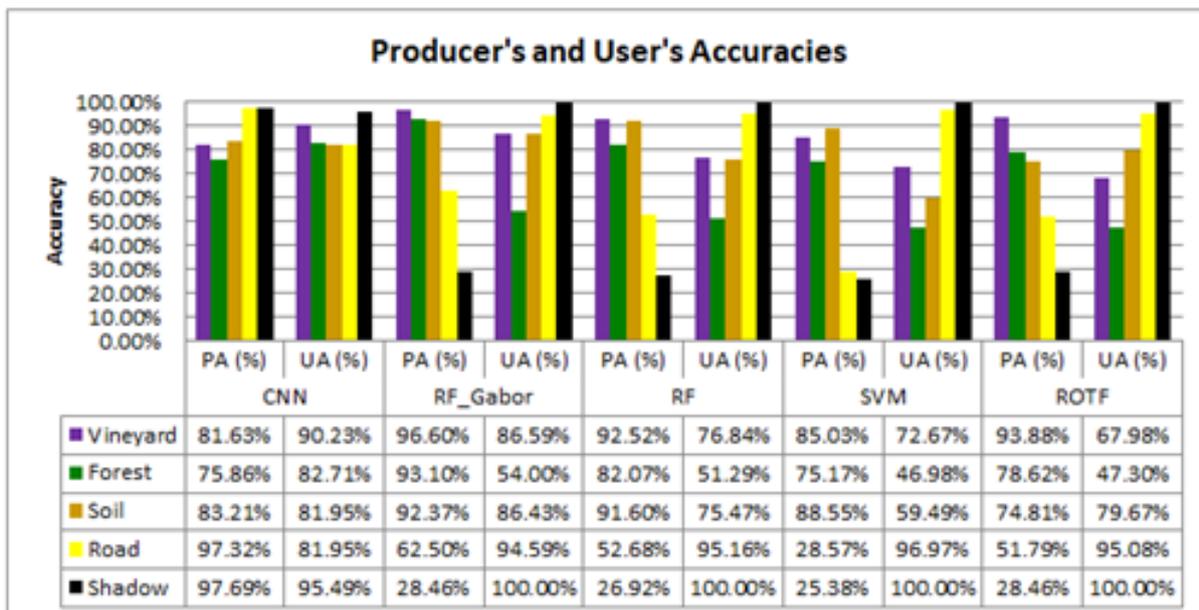
**Table 4.** The agreement and disagreement values.

Agreement/Disagreement (%)	3D CNN	RF_Gabor	RF	ROTF	SVM
Chance agreement	10	10	10	10	10
Quantity agreement	10	11	11	11	11
Allocation agreement	66	55	50	46	41
Allocation disagreement	10	4	8	10	11
Quantity disagreement	4	20	22	23	26

The overall disagreement, sum of both allocation and quantity disagreements, is lowest for 3D CNN at 14%, followed by RF\_Gabor at 24%, RF at 30%, ROTF at 33%, and SVM at 37%. This indicates that 3D CNN effectively reduces disagreement and enhances classification

accuracy. Additionally, the incorporation of texture features decreases the disagreement of RF by 6%.

Producer (PA) and User (UA) accuracies that can be obtained from error matrices are also examined (Figure 4).



**Figure 4.** Producer's (PA) and User's (UA) in the error matrices of the classified images obtained with 3DCNN, RF, ROTF, SVM, and RF\_Gabor.

Upon comparing the PA (Producer's Accuracy) values, it is evident that RF, ROTF, SVM, and RF\_Gabor achieved PA values of 11%, 12%, 3%, and 15% respectively. These results indicate that the other methods demonstrated more accurate classification of the vineyard class compared to 3D CNN. For soil class, RF\_Gabor, RF and SVM performed 9%, 8% and 5% better than 3D CNN, respectively, yet 3D CNN performed 8% more successful classification than ROTF. Also, 3D CNN classified road and shadow classes better than other methods. According to UA values (Figure 4), 3D CNN outperforms RF\_Gabor by 4%, RF by 13%, SVM by 18%, and ROTF by 22% for the vineyard class. When it comes to the forest class, 3D CNN surpassed RF\_Gabor by 29%, RF by 31%, SVM by 36%, and ROTF by 35%. Regarding the soil class, 3D CNN outcompeted RF by 7%, SVM by 23%, and ROTF by 2%, while RF\_Gabor outperformed 3D CNN by 5%. Based on the UA results, 3D CNN especially discriminated vineyard, forest and soil classes better than RF, ROTF, and SVM, but was not succeeded as the other methods in road and shadow classes. The high

similarity in spectral properties between the forest and vineyard classes has led to a significant amount of confusion between these classes. Similarly, the spectral properties of very dark pixels in earth roads, soil, and forest classes closely resemble those of the shadow class, resulting in confusion among these classes as well. However, the incorporation of texture information has contributed to improved identification and extraction of vineyards.

In order to evaluate the significance of performance differences between 3D CNN and other classifiers using the McNemar test,  $\chi^2$  values were computed for RF\_Gabor, RF, ROTF, and SVM as 19.512, 44.100, 63.515, and 90.289, respectively (Table 5).

The  $\chi^2$  values exceeding the reference value of 3.84 indicate that the 3D CNN classifier yields a significant improvement in accuracy at the 95% confidence interval. Lastly, Figure 5 presents the land-use map generated from the thematic classified image using the 3D CNN classifier, which demonstrates the highest classification accuracy.

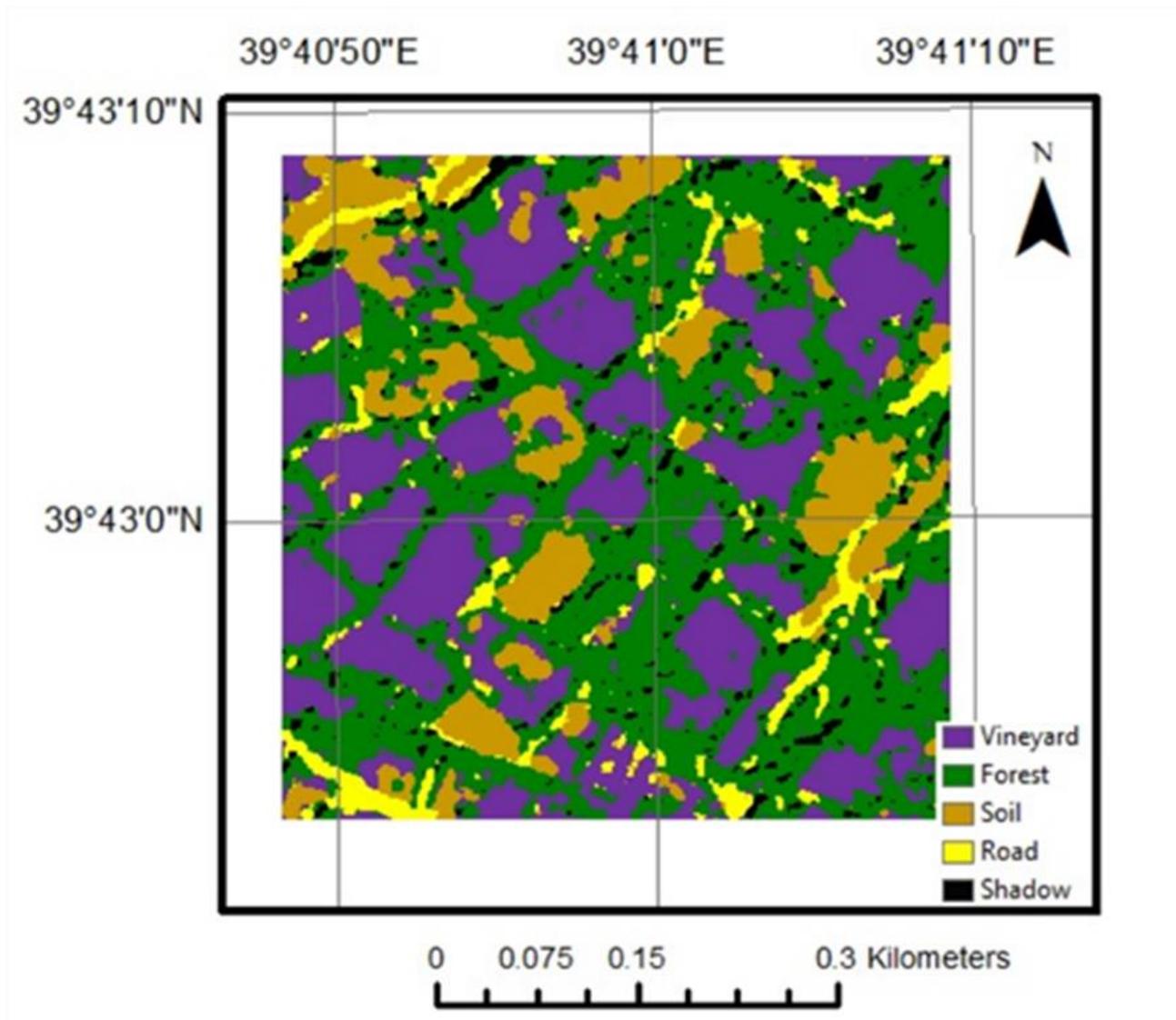


Figure 5. Land-use map was produced from the image classified with the 3D CNN classifier.

**Table 5.** McNemar test results for 3D CNN and other classifiers.

	$f_{11}$	$f_{12}$	$f_{21}$	$f_{22}$	Total	$\chi^2$
3D CNN-RF_Gabor	418	157	87	3	665	19.512
3D CNN-RF	397	178	72	18	665	44.100
3D CNN-RTF	379	196	66	24	665	63.515
3D CNN-SVM	355	220	60	30	665	90.289

$f_{11}$  is the number of correctly classified pixels in both cases.  $f_{22}$  represents the number of misclassified pixels in both cases.  $f_{12}$  represents number of correctly classified pixels with first classifier but misclassified pixels with second classifier.  $f_{21}$  is the number of misclassified pixels with second classifier but correctly classified pixels with first classifier

It has been determined that the 3D CNN model, incorporating 3D convolution layers, outperforms RF, ROTF, and SVM, which are commonly utilized in vineyard detection in existing literature. An important aspect to note is that the developed 3D CNN model achieves high accuracy directly from the image data, eliminating the requirement for additional data. For instance, while texture features were added to the WV-2 image using Gabor to enhance the classification accuracy of the RF algorithm, the addition of supplementary textures proved to be less effective compared to the 3D CNN model, which can automatically extract textures from images. However, one drawback of the 3D CNN model, which achieves accurate vineyard detection in WV-2 images without necessitating user intervention, is the significantly longer training time (20 hours) when compared to other machine learning models.

In order to assess the classification performance of the 3DCNN method and other approaches, the metrics presented in Table 5 were analyzed for each method and

class. Upon examining the results in Table 6, it is observed that the 3DCNN method demonstrates successful classification, supported by the minimum and maximum values of F1-score (0.79-0.97), specificity (0.95-0.99), and accuracy (0.91-0.99). The RF\_Gabor method exhibits classification performance that closely resembles that of 3DCNN, as indicated by the F1-score (0.44-0.91), specificity (0.76-1.00), and accuracy (0.80-0.95) (Table 7). When averaging the metric values across the classes, the 3DCNN method achieves an F1-score of 0.866, specificity of 0.963, and accuracy of 0.942. For the RF\_Gabor method, the corresponding values are an F1-score of 0.737, specificity of 0.930, and accuracy of 0.892. ROTF yields an F1-score of 0.653, specificity of 0.895, and accuracy of 0.839 (Table 8). SVM demonstrates an F1-score of 0.584, specificity of 0.872, and accuracy of 0.809 (Table 9). These results align with the overall classification accuracies obtained from the respective methods.

**Table 6.** Evaluation the performance of the methods for 3DCNN.

	PRECISION	RECALL	F1-SCORE	SPECIFICITY	ACCURACY
VINEYARD	0.902	0.816	0.857	0.972	0.935
FOREST	0.827	0.759	0.791	0.953	0.908
SOIL	0.820	0.832	0.826	0.951	0.926
ROAD	0.820	0.973	0.890	0.951	0.955
SHADOW	0.955	0.977	0.966	0.987	0.985

**Table 7.** Evaluation the performance of the methods for RF\_Gabor.

	PRECISION	RECALL	F1-SCORE	SPECIFICITY	ACCURACY
VINEYARD	0.866	0.966	0.913	0.943	0.949
FOREST	0.540	0.931	0.684	0.763	0.802
SOIL	0.864	0.924	0.893	0.953	0.946
ROAD	0.946	0.625	0.753	0.991	0.917
SHADOW	1.000	0.285	0.443	1.000	0.844

**Table 8.** Evaluation the performance of the methods for ROTF.

	PRECISION	RECALL	F1-SCORE	SPECIFICITY	ACCURACY
VINEYARD	0.680	0.939	0.789	0.825	0.857
FOREST	0.473	0.786	0.591	0.723	0.738
SOIL	0.797	0.748	0.772	0.933	0.885
ROAD	0.951	0.518	0.671	0.992	0.886
SHADOW	1.000	0.285	0.443	1.000	0.827

**Table 9.** Evaluation the performance of the methods for SVM.

	PRECISION	RECALL	F1-SCORE	SPECIFICITY	ACCURACY
VINEYARD	0.727	0.850	0.784	0.861	0.857
FOREST	0.470	0.752	0.578	0.713	0.723
SOIL	0.595	0.885	0.712	0.791	0.815
ROAD	0.970	0.286	0.441	0.997	0.837
SHADOW	1.000	0.254	0.405	1.000	0.811

## 5. Conclusion

Accurately identifying vineyards is crucial for contributing to the national economy, effectively managing viticulture processes, monitoring crops, and implementing site-specific automated crop management. With this objective in mind, this study aimed to determine the spatial distribution of Cimin grape using high-resolution satellite images. The performance of the developed 3D CNN model was compared to that of the RF, ROTF, and SVM algorithms. The classification accuracies obtained were 86.47%, 75.94%, 70.53%, 66.92%, and 62.41% for the 3D CNN, RF\_Gabor, RF, ROTF, and SVM methods, respectively. The 3D CNN method outperformed the RF\_Gabor (second-ranked) by 11% and the RF (third-ranked) by 16% in terms of classification performance, resulting in more accurate vineyard classification. The evaluation of metrics such as F1-Score, specificity, accuracy, Kappa analyses, and  $\chi^2$  values obtained from the McNemar test further confirm the success of the 3D CNN method. Preliminary results indicate that the proposed 3D CNN-based deep learning model can effectively classify Cimin vineyards and determine their spatial distributions. Future work will focus on evaluating the performance of different CNN-based architectures for the same problem.

## Acknowledgement

This work was supported by Erzincan Binali Yıldırım University Scientific Research Project [Grant Number: 636].

## Author contributions

**Özlem Akar:** Processed and classified the images using machine learning methods, and analyzed the results.  
**Ekrem Saraloğlu:** Classified the study area by developing a 3D convolutional neural network model.  
**Oğuz Güngör:** Wrote and reviewed the manuscript.  
**Halim Ferit Bayata:** Wrote and reviewed the manuscript.

## Conflicts of interest

The authors declare no conflicts of interest.

## References

- Weaver, R. J. (1976). Grape growing. John Wiley & Sons.
- Akpınar, E., & Çelikoğlu, Ş. (2016). Karaerik (Cimin) üzümünün Erzincan ekonomisine ve tanıtımına katkıları. Uluslararası Erzincan Sempozyumu, 2, 15-23.
- Bulut, İ. (2006). Genel tarım bilgileri ve tarımın coğrafi esasları (Ziraat Coğrafyası). Gündüz Eğitim ve Yayıncılık, Ankara, 255.
- Republic of Turkey Ministry of Agriculture and Forestry. (2021). 2021-January Agricultural Products Markets Report: GRAPE, <https://arastirma.tarimorman.gov.tr/tepge/Menu/27/Tarim-Urunleri-Piyasaları>
- Erzincan Directorate of Provincial Agriculture and Forestry (2022). <https://erzincan.tarimorman.gov.tr/Menu/66/Tarimsal-Veriler>
- Christian, B., & Krishnayya, N. S. R. (2009). Classification of tropical trees growing in a sanctuary using Hyperion (EO-1) and SAM algorithm. Current Science, 96(12), 1601-1607.
- Prins, A. J., & Van Niekerk, A. (2020). Regional Mapping of Vineyards Using Machine Learning and LiDAR Data. International Journal of Applied Geospatial Research (IJAGR), 11(4), 1-22. <https://doi.org/10.4018/IJAGR.2020100101>
- Darra, N., Psomiadis, E., Kasimati, A., Anastasiou, A., Anastasiou, E., & Fountas, S. (2021). Remote and proximal sensing-derived spectral indices and biophysical variables for spatial variation determination in vineyards. Agronomy, 11(4), 741. <https://doi.org/10.3390/agronomy11040741>
- Vélez, S., Ariza-Sentís, M., & Valente, J. (2023). Mapping the spatial variability of Botrytis bunch rot risk in vineyards using UAV multispectral imagery. European Journal of Agronomy, 142, 126691. <https://doi.org/10.1016/j.eja.2022.126691>
- Gungor, O., Boz, Y., Gokalp, E., Comert, C., & Akar, A. (2010). Fusion of low and high resolution satellite images to monitor changes on coastal zones. Scientific Research and Essays, 5(7), 654-662.
- Chi, M. V., Thi, L. P., & Si, S. T. (2009, October). Monitoring urban space expansion using Remote sensing data in Ha Long city, Quang Ninh province in Vietnam. In 7th FIG Regional Conference Spatial Data Serving People: Land Governance and the Environment-Building the Capacity Hanoi, Vietnam, 19-22.
- Kaya, Y., & Polat, N. (2023). A linear approach for wheat yield prediction by using different spectral vegetation indices. International Journal of Engineering and Geosciences, 8(1), 52-62. <https://doi.org/10.26833/ijeg.1035037>
- Akar, A., & Gokalp, E. (2018). Designing a sustainable rangeland information system for Turkey. International Journal of Engineering and Geosciences, 3(3), 87-97. <https://doi.org/10.26833/ijeg.412222>
- Zhang, W., Xue, X., Sun, Z., Guo, Y. F., Chi, M., & Lu, H. (2007). Efficient feature extraction for image classification. IEEE 11th International Conference on Computer Vision, 1-8. <https://doi.org/10.1109/ICCV.2007.4409058>
- Huang, Y., Fipps, G., Lacey, R. E., & Thomson, S. J. (2011). Landsat satellite multi-spectral image classification of land cover and land use changes for GIS-based urbanization analysis in irrigation districts of Lower Rio Grande Valley of Texas. Journal of Applied Remote Sensing, 2(1), 27-36.
- Akar, Ö., & Tunç Görmüş, E. (2019). Göktürk-2 ve Hyperion EO-1 uydu görüntülerinden rastgele orman sınıflandırıcısı ve destek vektör makineleri ile arazi kullanım haritalarının üretilmesi. Geomatik, 4(1), 68-81. <https://doi.org/10.29128/geomatik.476668>

17. Ahady, A. B., & Kaplan, G. (2022). Classification comparison of Landsat-8 and Sentinel-2 data in Google Earth Engine, study case of the city of Kabul. *International Journal of Engineering and Geosciences*, 7(1), 24-31. <https://doi.org/10.26833/ijeg.860077>
18. Sefercik, U. G., Kavzoğlu, T., Çölkesen, I., Nazar, M., Öztürk, M. Y., Adalı, S., & Dinç, S. (2023). 3D positioning accuracy and land cover classification performance of multispectral RTK UAVs. *International Journal of Engineering and Geosciences*, 8(2), 119-128. <https://doi.org/10.26833/ijeg.1074791>
19. Cengiz, A. V. C. I., Budak, M., Yağmur, N., & Balçık, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences*, 8(1), 1-10. <https://doi.org/10.26833/ijeg.987605>
20. Tirmanoğlu, B., İsmailoğlu, I., Kokal, A. T., & Musaoğlu, N. (2023). Yeni nesil multispektral ve hiperspektral uydu görüntülerinin arazi örtüsü/arazi kullanımı sınıflandırma performanslarının karşılaştırılması: Sentinel-2 ve PRISMA Uydusu. *Geomatik*, 8(1), 79-90. <https://doi.org/10.29128/geomatik.1126685>
21. Çömert, R., Matci, D. K., & Avdan, U. (2019). Object based burned area mapping with random forest algorithm. *International Journal of Engineering and Geosciences*, 4(2), 78-87. <https://doi.org/10.26833/ijeg.455595>
22. Sun, Z., Di, L., Fang, H., & Burgess, A. (2020). Deep learning classification for crop types in north dakota. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 2200-2213. <https://doi.org/10.1109/JSTARS.2020.2990104>
23. Gao, J. (2009). *Digital analysis of remotely sensed imagery*. McGraw-Hill Education, New York. ISBN: 9780071604659
24. Jay, S., Lawrence, R., Repasky, K., & Keith, C. (2009). Invasive species mapping using low-cost hyperspectral imagery. In *ASPRS Annual Conference*.
25. Ok, A. O., Akar, O., & Gungor, O. (2012). Evaluation of random forest method for agricultural crop classification. *European Journal of Remote Sensing*, 45(1), 421-432. <https://doi.org/10.5721/EuJRS20124535>
26. Akar, Ö., & Güngör, O. (2015). Integrating multiple texture methods and NDVI to the Random Forest classification algorithm to detect tea and hazelnut plantation areas in northeast Turkey. *International Journal of Remote Sensing*, 36(2), 442-464. <https://doi.org/10.1080/01431161.2014.995276>
27. Ntouros, K. D., Gitas, I. Z., & Silleos, G. N. (2009, August). Mapping agricultural crops with EO-1 Hyperion data. In *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 1-4. <https://doi.org/10.1109/WHISPERS.2009.5289057>
28. Kpienbaareh, D., Sun, X., Wang, J., Luginaah, I., Bezner Kerr, R., Lupafya, E., & Dakishoni, L. (2021). Crop type and land cover mapping in northern Malawi using the integration of sentinel-1, sentinel-2, and planetscope satellite data. *Remote Sensing*, 13(4), 700. <https://doi.org/10.3390/rs13040700>
29. Wang, S., Azzari, G., & Lobell, D. B. (2019). Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote sensing of environment*, 222, 303-317. <https://doi.org/10.1016/j.rse.2018.12.026>
30. Akar, Ö., Saralioğlu, E., Güngör, O., & Bayata, H. F. (2021). Determination of vineyards with support vector machine and deep learning-based Image classification. *Intercontinental Geoinformation Days*, 3, 26-29.
31. Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). Deep learning for plant identification using vein morphological patterns. *Computers and Electronics in Agriculture*, 127, 418-424. <https://doi.org/10.1016/j.compag.2016.07.003>
32. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, 145, 311-318. <https://doi.org/10.1016/j.compag.2018.01.009>
33. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, 151, 61-69. <https://doi.org/10.1016/j.compag.2018.05.012>
34. Abdullahi, H. S., Sheriff, R., & Mahieddine, F. (2017). Convolution neural network in precision agriculture for plant image recognition and classification. *Seventh International Conference on Innovative Computing Technology (INTECH)*, 10, 256-272.
35. Zhao, H., Duan, S., Liu, J., Sun, L., & Reymondin, L. (2021). Evaluation of five deep learning models for crop type mapping using sentinel-2 time series images with missing information. *Remote Sensing*, 13(14), 2790. <https://doi.org/10.3390/rs13142790>
36. Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. *Remote Sensing of Environment*, 221, 430-443. <https://doi.org/10.1016/j.rse.2018.11.032>
37. TR Erzincan Governorate. (2021). <http://www.erkincan.gov.tr/erkincan-uzumu>
38. Padwick, C., Deskevich, M., Pacifici, F., & Smallwood, S. (2010). WorldView-2 pan-sharpening. In *Proceedings of the ASPRS 2010 Annual Conference*, San Diego, CA, USA, 2630, 1-14.
39. Akar, Ö. (2019). Göktürk-2 ve Worldview-2 Uydu Görüntüleri için Görüntü Keskinleştirme Yöntemlerinin Değerlendirilmesi. *Erzincan University Journal of Science and Technology*, 12(2), 874-885.
40. Li, H., Jing, L., & Tang, Y. (2017). Assessment of pansharpening methods applied to WorldView-2 imagery fusion. *Sensors*, 17(1), 89. <https://doi.org/10.3390/s17010089>
41. Anshu, S. K., Pande, H., Tiwari, P. S., & Shukla, S. (2017). Evaluation of Fusion Techniques for High Resolution Data-A Worldview-2 Imagery.

- International Journal of Applied Remote Sensing and GIS, 4, 10-22.
42. Fu, L., Ma, J., Chen, Y., Larsson, R., & Zhao, J. (2019). Automatic detection of lung nodules using 3D deep convolutional neural networks. *Journal of Shanghai Jiaotong University (Science)*, 24, 517-523. <https://doi.org/10.1007/s12204-019-2084-4>
  43. Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *Geoscience and Remote Sensing Magazine*, 5(4), 8-36. <https://doi.org/10.1109/MGRS.2017.2762307>
  44. Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3D convolutional neural networks for human action recognition. *Transactions on Pattern Analysis and Machine Intelligence*, 35(1), 221-231. <https://doi.org/10.1109/TPAMI.2012.59>
  45. Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, 4489-4497.
  46. Xu, Z., Guan, K., Casler, N., Peng, B., & Wang, S. (2018). A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery. *ISPRS journal of photogrammetry and remote sensing*, 144, 423-434. <https://doi.org/10.1016/j.isprsjprs.2018.08.005>
  47. Ji, S., Zhang, C., Xu, A., Shi, Y., & Duan, Y. (2018). 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sensing*, 10(1), 75. <https://doi.org/10.3390/rs10010075>
  48. Mei, S., Yuan, X., Ji, J., Zhang, Y., Wan, S., & Du, Q. (2017). Hyperspectral image spatial super-resolution via 3D full convolutional neural network. *Remote Sensing*, 9(11), 1139. <https://doi.org/10.3390/rs9111139>
  49. Saralioglu, E., & Gungor, O. (2022). Semantic segmentation of land cover from high resolution multispectral satellite images by spectral-spatial convolutional neural network. *Geocarto International*, 37(2), 657-677. <https://doi.org/10.1080/10106049.2020.1734871>
  50. Li, Y., Zhang, H., & Shen, Q. (2017). Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, 9(1), 67. <https://doi.org/10.3390/rs9010067>
  51. Pérez, F., & Granger, B. E. (2007). IPython: a system for interactive scientific computing. *Computing in Science & Engineering*, 9(3), 21-29. <https://doi.org/10.1109/MCSE.2007.53>
  52. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
  53. Watts, J. D., & Lawrence, R. L. (2008). Merging random forest classification with an object-oriented approach for analysis of agricultural lands. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7), 579-582
  54. Waske, B., Heinzl, V., Braun, M., & Menz, G. (2007). Random forests for classifying multi-temporal sar data. *Envisat Symposium*, 2007, 23-27.
  55. Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2004). Random forest classification of multisource remote sensing and geographic data. In *IGARSS 2004. International Geoscience and Remote Sensing Symposium*, 2, 1049-1052. <https://doi.org/10.1109/IGARSS.2004.1368591>
  56. Pal, M. (2003, July). Random forests for land cover classification. In *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)* 6, 3510-3512. <https://doi.org/10.1109/IGARSS.2003.1294837>
  57. Rodriguez, J. J., Kuncheva, L. I., & Alonso, C. J. (2006). Rotation forest: A new classifier ensemble method. *IEEE transactions on pattern analysis and machine intelligence*, 28(10), 1619-1630. <https://doi.org/10.1109/TPAMI.2006.211>
  58. Xia, J., Du, P., He, X., & Chanussot, J. (2013). Hyperspectral remote sensing image classification based on rotation forest. *IEEE Geoscience and Remote Sensing Letters*, 11(1), 239-243. <https://doi.org/10.1109/LGRS.2013.2254108>
  59. Liu, K. H., & Huang, D. S. (2008). Cancer classification using rotation forest. *Computers in biology and medicine*, 38(5), 601-610. <https://doi.org/10.1016/j.compbiomed.2008.02.007>
  60. Vapnik, V. (1999). *The nature of statistical learning theory*. Springer science & business media.
  61. Özkan, Y. (2008). *Veri Madenciliği Yöntemleri*, Papatya Yayıncılık, İstanbul.
  62. Mather, P., & Tso, B. (2016). *Classification methods for remotely sensed data*. CRC press.
  63. Stephens, D., & Diesing, M. (2014). A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *PloS one*, 9(4), e93950. <https://doi.org/10.1371/journal.pone.0093950>
  64. Çölkesen, İ., & Yomraloğlu, T. (2014). Arazi örtüsü ve kullanımının haritalanmasında WorldView-2 uydu görüntüsü ve yardımcı verilerin kullanımı. *Harita Dergisi*, 152(2), 12-24.
  65. Thanh Noi, P., & Kappas, M. (2018). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors*, 18(1), 18. <https://doi.org/10.3390/s18010018>
  66. Kavzoglu, T., & Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11(5), 352-359. <https://doi.org/10.1016/j.jag.2009.06.002>
  67. Congalton, R. G., & Green, K. (2019). *Assessing the accuracy of remotely sensed data: principles and practices*. CRC Press.
  68. Congalton, R. G., & Green, K. (2019). *Assessing the accuracy of remotely sensed data: principles and practices*. CRC Press.
  69. Pontius Jr, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation

- disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15), 4407-4429. <https://doi.org/10.1080/01431161.2011.552923>
70. Akar, A. (2022). Improving the accuracy of random forest-based land-use classification using fused images and digital surface models produced via different interpolation methods. *Concurrency and Computation: Practice and Experience*, 34(6), e6787. <https://doi.org/10.1002/cpe.6787>
71. Foody, G. M. (2004). Thematic map comparison. *Photogrammetric Engineering & Remote Sensing*, 70(5), 627-633. <https://doi.org/10.14358/PERS.70.5.627>
72. Amini, S., Saber, M., Rabiei-Dastjerdi, H., & Homayouni, S. (2022). Urban land use and land cover change analysis using random forest classification of landsat time series. *Remote Sensing*, 14(11), 2654. <https://doi.org/10.3390/rs14112654>



© Author(s) 2024. This work is distributed under <https://creativecommons.org/licenses/by-sa/4.0/>