

Vineyard classification using OBIA on UAV-based RGB and multispectral data: A case study in different wine regions

Luís Pádua ^{a,b}, Alessandro Matese ^a, Salvatore Filippo Di Gennaro ^{a,*}, Raul Morais ^{b,c}, Emanuel Peres ^{b,c}, Joaquim J. Sousa ^{c,d}

^a Institute of BioEconomy (IBE), National Research Council (CNR), Via Caproni 8, 50145 Florence, Italy

^b Centre for the Research and Technology of Agro-Environmental and Biological Sciences, University of Trás-os-Montes e Alto Douro, 5000-801 Vila Real, Portugal

^c School of Science and Technology, University of Trás-os-Montes e Alto Douro, 5000-801 Vila Real, Portugal

^d Centre for Robotics in Industry and Intelligent Systems (CRIIS), INESC Technology and Science (INESC-TEC), 4200-465 Porto, Portugal



ARTICLE INFO

Keywords:

Precision viticulture
Aerial imagery
Object-based image analysis
Artificial neural network
Random forest
Support vector machine

ABSTRACT

Vineyard classification is an important process within viticulture-related decision-support systems. Indeed, it improves grapevine vegetation detection, enabling both the assessment of vineyard vegetative properties and the optimization of in-field management tasks. Aerial data acquired by sensors coupled to unmanned aerial vehicles (UAVs) may be used to achieve it. Flight campaigns were conducted to acquire both RGB and multispectral data from three vineyards located in Portugal and in Italy. Red, green, blue and near infrared orthorectified mosaics resulted from the photogrammetric processing of the acquired data. They were then used to calculate RGB and multispectral vegetation indices, as well as a crop surface model (CSM). Three different supervised machine learning (ML) approaches—support vector machine (SVM), random forest (RF) and artificial neural network (ANN)—were trained to classify elements present within each vineyard into one of four classes: grapevine, shadow, soil and other vegetation. The trained models were then used to classify vineyards objects, generated from an object-based image analysis (OBIA) approach, into the four classes. Classification outcomes were compared with an automatic point-cloud classification approach and threshold-based approaches. Results shown that ANN provided a better overall classification performance, regardless of the type of features used. Features based on RGB data showed better performance than the ones based only on multispectral data. However, a higher performance was achieved when using features from both sensors. The methods presented in this study that resort to data acquired from different sensors are suitable to be used in the vineyard classification process. Furthermore, they also may be applied in other land use classification scenarios.

1. Introduction

It is well-established that spatial variability within a vineyard can be duly monitored by resorting to precision viticulture (PV) approaches (Bramley, 2001). As spatial variability weighs in grapevines' (*vitis vinifera* L.) biophysical parameters, their analysis can be used not only to identify this variability, but also to have improved management approaches for defined areas within the vineyard (Campos et al., 2019).

Remote sensing is widely used in PV. Indeed, aerial imagery acquired by satellites, manned aircrafts, and unmanned aerial vehicles (UAVs), when coupled to different types of sensors, provide must-needed heterogeneous data for decision support systems. As a matter of fact, this data enables several diverse outcomes. Some examples are:

- orthophoto mosaics, either RGB or colour infrared (CIR) composites, that can be used to visually assess the vineyard;
- digital elevation models (DEM), which may include objects' altitude values above the surface level (digital surface model, DSM) or at terrain level only (digital terrain model, DTM) (Panagiotidis et al., 2017);
- point clouds, obtained either through light detection and ranging (LiDAR) sensors or through cost-effective photogrammetric approaches (Guimarães et al., 2020);
- spectral signatures, when dealing with hyperspectral imagery. It provides a greater number of narrow bands, covering the most part of the visible and near-infrared within the electromagnetic spectrum (Adão et al., 2017); and

* Corresponding author.

E-mail address: salvatorefilippo.digennaro@cnr.it (S.F. Di Gennaro).

- vegetation indices, resulting from arithmetic operations among the different available spectral bands (Salamí et al., 2014).

Satellites and manned aircrafts are remote sensing platforms that have the distinct advantage of addressing areas at regional or even national scales (Pádua et al., 2017). However, this does not come without some drawbacks. Indeed, spatial resolution, cost and data availability (low revisiting time) are some of the issues that may influence the spatio-temporal vineyard variability mapping process (Matese et al., 2015). Regarding UAVs, they can acquire high-resolution data meaning more true grapevine pixels present when and where it is required, using different coupled sensors. Moreover, their operation is a relatively cheap, when compared with the other two remote sensing platforms (Khaliq et al., 2019), which makes them just right to map vineyards' variability (Sozzi et al., 2020).

Grapevines' segmentation and vineyard classification are key to optimize the management of vineyard plots. Resorting to high-resolution imagery acquired by UAVs coupled with different sensors provides several approaches to achieve this. Some of the most meaningful examples are: image segmentation techniques (Comba et al., 2015; Nolan et al., 2015; Pádua et al., 2018), grapevines' height filtering based in DEMs (Burgos et al., 2015; Di Gennaro and Matese, 2020; Matese et al., 2016), and point clouds (Comba et al., 2018; Di Gennaro and Matese, 2020; Jurado et al., 2020); vegetation indices thresholding (Campos et al., 2019); object-based image analysis (OBIA) (de Castro et al., 2018); unsupervised clustering techniques (Cinat et al., 2019; Fuentes-Peñailllo et al., 2018; Poblete-Echeverría et al., 2017); supervised machine learning (Poblete-Echeverría et al., 2017) or deep learning techniques (Kerkech et al., 2020a, 2020b). However, most of the studies found in the literature refer to well-managed commercial vineyards, with good contrast between grapevines and soil. In those contexts, very few or no missing grapevines exist (Nolan et al., 2015; Poblete-Echeverría et al., 2017). Furthermore, the use of high-resolution RGB or CIR data is not possible in some cases, either due to logistic or to economic reasons. As such, available data is acquired solely by RGB and/or multispectral sensors with a lower resolution (approximately of one megapixel). This strongly affects spatial resolution and therefore the resulting orthorectified products (Cinat et al., 2019). Sometimes, the use

of manual data extraction procedures using geographic information systems (GIS) is required to obtain grapevines' parameters.

Given that vineyards are composed of elements other than grapevines' vegetation, the main purpose of this study is to classify them into four different classes: bare soil, shadowed areas, grapevine, and other vegetation. Furthermore, this must be achieved using UAV-acquired aerial imagery (RGB and/or multispectral) of vineyards that do not follow the typical structure of those already analysed in most published studies: commercial vineyards, with almost no slope, few to none missing grapevine plants and almost no inter-row vegetation. Hence, the path presented in this study is (i) to assess different supervised machine learning (ML) approaches, commonly used in remote sensing (Belgiu and Drăguț, 2016; Mahmon and Ya'acov, 2014; Mountrakis et al., 2011) in vineyards' classification, namely support vector machine (SVM), random forest (RF) and artificial neural network (ANN); (ii) evaluate the suitability of different features for vineyards' classification estimated from RGB and/or multispectral data, along with or without height features; and (iii) use an OBIA approach to predict the four different classes distribution and to compare grapevine' estimation with an automatic point cloud classification method. To have data acquired from diverse geographical, agronomic and cultural contexts, the methods presented in this study were applied in vineyards with distinct characteristics, located in different wine regions, with specific grapevine rows structure, and training systems.

2. Materials and methods

2.1. Study areas characterization

Three vineyards from distinct wine regions were chosen as study areas. Two are located in Portugal: one in the north-east, within the Douro Demarcated Region (DDR), and the other in the north-west, within the Demarcated Vinho Verde Region (DVVR). The third vineyard is located in Italy, near Castellina in Chianti (Siena), within the Chianti Classico domain (CC). Besides their geographic location, these vineyards were also selected due to their different planting styles and topography. DDR vineyard (Fig. 1a) is located in a steep slope terrain and its rows follow the terrain hillside. It is also worth mentioning that

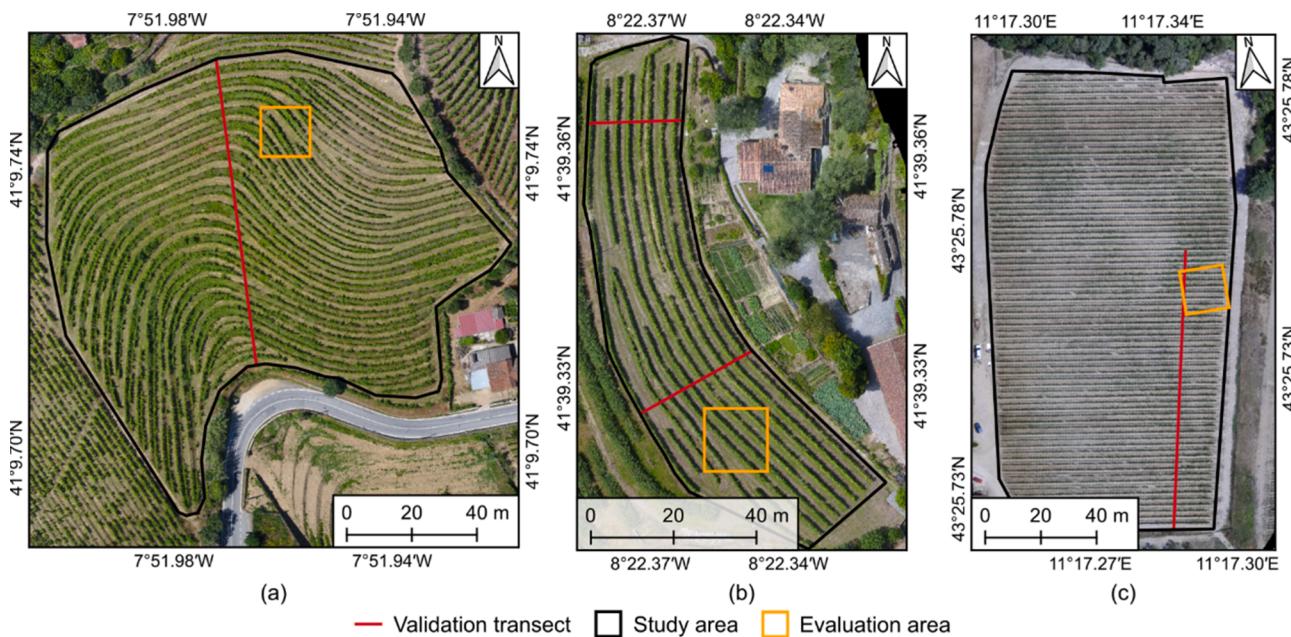


Fig. 1. General overview of the selected vineyards: (a) in the Douro Demarcated Region, Portugal; (b) Demarcated Vinho Verde Region, Portugal; and (c) in the Chianti Classico domain, Italy.

Table 1

Unmanned aerial systems and sensors used in each flight campaign, along with some characteristics (MSP: multispectral).

Vineyard	Flight date	UAV	Sensors (resolution)		Flight height (m)	Spatial Resolution (m)
			RGB	MSP		
A	19 June 2018	DJI Phantom 4	DJI FC330 (12 MP)	Parrot sequoia (1.2 MP)	40	RGB: 0.03 MSP: 0.05
B	13 June 2018				50	RGB: 0.02 MSP: 0.05
C	6 August 2019	HiSystems Mikrokopter	Thermal Capture Daylight (1.92 MP)	Tetracam ADC SNAP (1.3 MP)	50	RGB: 0.03 MSP: 0.02

the aerial imagery was acquired before the removal of excess shoots, meaning that grapevines present some excess vegetative volume. Grapevines are trained in a double guyot vertical shoot positioned (VSP) trellis system, with a height of approximately 1.5 m. DVVR vineyard (Fig. 1b) is positioned in a flat terrace, with grapevines in the flowering stage. Its rows present some curvature with a single-cordon VSP trellis system. Plants can reach a maximum height ranging from 2.5 to 2.7 m, following DVVR traditional viticulture practices. CC site (Fig. 1c) is a 1.4 ha vineyard (355 m above sea level), planted in 2008 with NW–SE rows orientation. It is located in a steep southern slope terrain and its rows follow the terrain's slope. Sangiovese cv grapevines were in an early stage of veraison. They are trained at single cordon, with a VSP trellis system.

2.2. Unmanned aerial vehicles and sensors

Vineyards' aerial imagery was acquired by using different multi-rotor UAVs and two types of sensors (Table 1). Both RGB and multispectral data was captured in all flight campaigns, being acquired close to solar noon and with optimal light conditions. While the same setup was used to acquire data in the Portuguese vineyards, a different UAV and sensors were used to survey the Italian vineyard. In both cases, green, red, blue and near-infrared (NIR) data is available. The spectral properties of the two multispectral sensors used are: for the Portuguese vineyards, green (550 nm, 40 nm bandwidth), red (660 nm, 40 nm bandwidth) and NIR (790 nm, 40 nm bandwidth); and for the Italian vineyard, green (560 nm, 80 nm bandwidth), red (660 nm, 60 nm bandwidth) and NIR (840 nm, 160 nm bandwidth).

Regardless of having all flights done in similar conditions, spatial resolution varies from 0.02 to 0.05 m/pixel between campaigns. This is explained by the resolution of each sensor, as well as differences in flight height. Indeed, while that in campaigns carried in the Portuguese sites the flight height is related to the UAV take-off position, in the Italian campaigns, the UAV followed the terrain's altitude, thus maintaining the same height throughout the whole flight. Imagery overlap was of 80% longitudinal and 70% lateral for the Portuguese vineyards, with a camera angle of approximately 70° for the RGB sensor. The multispectral sensor was pointed in a nadiral position. As for the Italian vineyard, the overlap was of 75% in both sides and both sensors were pointed in a nadiral position.

2.3. Aerial imagery processing

The photogrammetric processing of the acquired UAV data was done with Pix4Dmapper Pro (Pix4D SA, Lausanne, Switzerland). First off, a sparse point-cloud is generated based on key points found among the images. Then, several terrain elements are identified in the RGB and multispectral images: they are used to align data from both sensors. The sparse point cloud is then re-optimized. With the data properly aligned, a high-density point cloud is generated and automatically classified. An automatic point cloud classification function (Becker et al., 2018) available in Pix4Dmapper Pro was used to classify the generated dense point clouds. This approach is based on ML techniques trained in rural,

construction and vegetation areas. As features, both the geometry and the colour information of each point are used. Then, every point is assigned to a given class, namely: ground, road surface, high-vegetation, building, and human made object.

Dense point clouds are then interpolated to generate several raster outcomes: an orthophoto mosaic (regarding RGB data), a DSM and a DTM. A Crop Surface Model (CSM) is also generated by subtracting the DTM altitude values to the DSM, creating a raster product with the height of all objects above soil level. Moreover, each band was

Table 2

Band normalization equations for RGB and multispectral data. R, G, B and N are the pixel values of the Red, Green, Blue and near-infrared bands, respectively.

Sensor	Equation
RGB	$r_{RGB} = \frac{R}{R + G + B}$
	$g_{RGB} = \frac{G}{R + G + B}$
	$b_{RGB} = \frac{B}{R + G + B}$
Multispectral	$n_{MSP} = \frac{N}{N + G + R}$
	$g_{MSP} = \frac{G}{N + G + R}$
	$r_{MSP} = \frac{R}{N + G + R}$

Table 3

Computed vegetation indices found in the literature and their respective equations. R, G, B and N are the pixel values of the Red, Green, Blue and near-infrared bands, respectively. L = 0.5.

Sensor	Name	Equation	Reference
RGB	Green-Red Vegetation Index	$GRVI = \frac{G - R}{G + R}$	(Tucker, 1979)
	Green-Blue Vegetation Index	$GBVI = \frac{G - B}{G + B}$	(Kawashima and Nakatani, 1998)
	Red Green Blue Vegetation Index	$RGBVI = \frac{G^2 - R \times B}{G^2 + R \times B}$	(Bendig et al., 2015)
	Excess Green	$ExG = 2g_{RGB} - r_{RGB} - b_{RGB}$	(Woebbecke et al., 1995)
	Visible Atmospherically Resistant Index	$VARI = \frac{G - R}{G + R - B}$	(Gitelson et al., 2002)
	Normalized Difference Vegetation Index	$NDVI = \frac{N - R}{N + R}$	(Rouse et al., 1974)
	Green Normalized Difference Vegetation Index	$GNDVI = \frac{N - G}{N + G}$	(Gitelson et al., 1996)
	Soil Adjusted Vegetation Index	$SAVI = \frac{N - R}{N + R + L} \times 1 + L$	(Huete, 1988)
	Green Soil Adjusted Vegetation Index	$GSAVI = \frac{N - G}{N + G + L} \times 1 + L$	(Sripada et al., 2006)
	Ratio Vegetation Index	$RVI = \frac{R}{N}$	(Baret and Guyot, 1991)

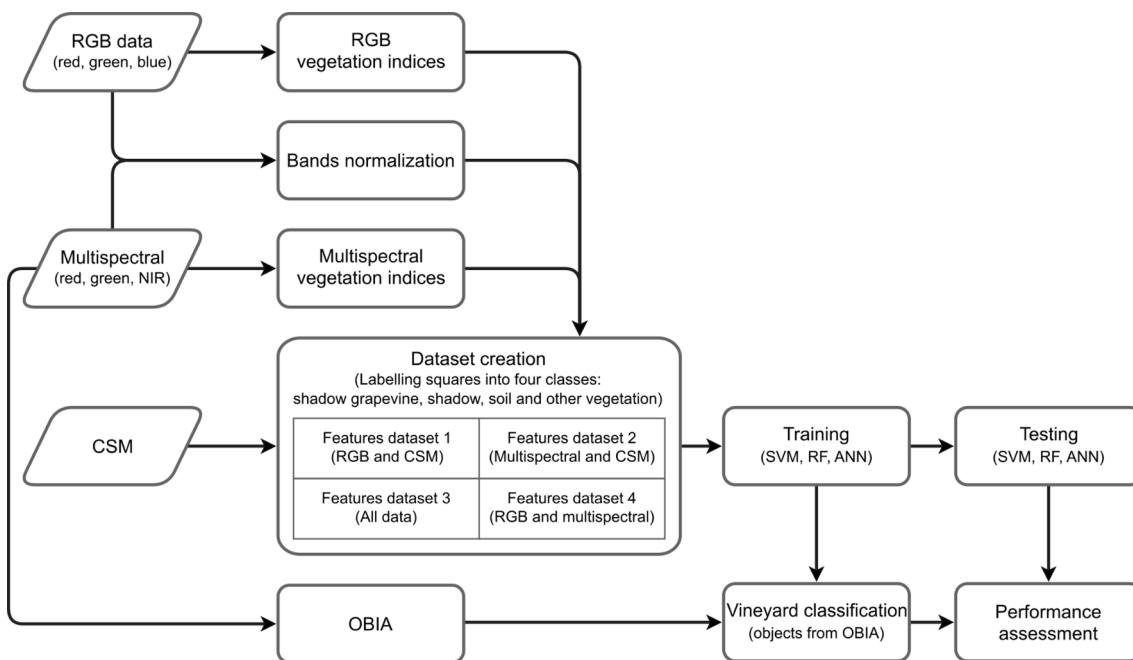


Fig. 2. General overview of the methodology used in the vineyard classification process.

normalized using only the bands from the same sensor: red, green, and blue in the RGB imagery and NIR, green and red in the multispectral imagery (equations in [Table 2](#)). This procedure renders the same range of pixel values (0 to 1), regardless the source sensor.

Alongside with the normalized bands, ten vegetation indices were calculated: five based on RGB data and another five using multispectral data, as presented in [Table 3](#). These vegetation indices are known to present good results to attain different purposes, such as the estimation of biophysical crop parameters and vegetation detection.

2.4. Vineyard classification using OBIA and machine learning

To achieve vineyard classification, the raster outcomes computed as described in [Section 2.3](#) are used as features in creating different datasets for each vineyard, as described in [Fig. 2](#). These datasets, combining different features, supported both the training and performance evaluation of ML classifiers. Trained models are then used to classify the whole vineyard into four different classes: (1) soil, areas without active vegetation mainly composed by bare soil, small rocks and dry grass; (2) shadow, caused by grapevines, vineyard posts and surrounding trees; (3) other vegetation, composed by vegetation other than grapevines, as inter-row vegetation and weeds; and (4) grapevine, addressing exclusively vegetation belonging to grapevines. Please note that regardless of dense point clouds classification, it was not considered in the labelling process.

2.4.1. Dataset creation and labelling

Both training and testing samples are required for training and assessing selected ML classifiers. As such, a set of 320 squares (0.2×0.2 m) were spread throughout each vineyard, evenly distributed between the four defined classes. A total of 80 samples per class was achieved. The number and size of the squares was chosen to keep class balance, cover a higher diversity of cases and to avoid the presence of multiple

classes in a single square. From the 320 samples, 200 (50 per class) were used exclusively for training purposes. The remainder 120 squares were exclusively used for testing the classifiers (30 per class). Transects—perpendicular to the grapevine rows—covering all classes were also considered (location in [Fig. 1](#)). Their length is of approximately 95 m in DDR, 100 m in CC vineyard, and in the DVVR they measure 23 and 31 m (54 m in total). Those were buffered and then split to create 0.2×0.2 m squares, labelled to one of the analysed classes. The result was a total of 477, 269 and 515 for DDR, DVVR and CC vineyards, respectively. Thus, the trained models were tested using squares generated from the transects, and the 120 squares not used at the training stage.

To classify vineyard elements each study area was submitted to an OBIA process. For this purpose a large-scale segmentation based on the mean-shift algorithm ([Michel et al., 2014](#)) was applied. By providing an input raster, a set of polygon objects are generated in a vector format, according to their spectral similarity. This process was carried out in QGIS software, using the Orfeo ToolBox (OTB) ([Inglada and Christophe, 2009](#)). NIR, green and red bands from the multispectral imagery were used as input and the spatial and range radius were set to 10 and 25, respectively. These values were chosen to prevent objects to simultaneously represent multiple classes. Minimum segment size was defined to 100, ensuring that regions smaller than this parameter are merged into similar neighbouring regions.

Following the creation and labelling of the training and testing samples, and the creation of objects by the OBIA procedure, different features were estimated for each square and object. These rely on the mean values of the vegetation indices ([Table 3](#)), the mean value of the normalized bands ([Table 2](#)) and the mean height values from the CSM.

2.4.2. Machine learning classifiers

For classification purposes, three ML algorithms were selected and evaluated: SVM, RF and ANN. In [Pádua et al. \(2020\)](#) these models were previously evaluated using only data derived from a RGB sensor in a

single vineyard plot. Trained models obtained from each ML classifier were used to classify objects identified using the mean-shift algorithm.

The SVM model, based in the LIBSVM (Chang and Lin, 2011), used a linear kernel, with a cost parameter C set to one. As for ANN and RF classifiers, both are based in the OpenCV ML library. The RF model was configured as: five maximum tree depth, 10 as minimum number of samples in each node, 100 as maximum number of trees in the forest and an out-of-bag (OOB) error of 0.01. With regard to the ANN classifier, it was trained with a resilient back-propagation algorithm using six neurons in each intermediate layer, a symmetrical sigmoid function for neuron activation with alpha and beta parameters set to one. The training process ended when one of two conditions was met: (i) a maximum number of 1000 iterations; or (ii) an epsilon value of 0.01.

2.4.3. Performance evaluation

To assess the performance and importance of each sensor in this study, the datasets were divided as follows: (1) RGB data only, its normalized bands (Table 2), the CSM height values and the RGB-based vegetation indices (Table 3); (2) multispectral data only, its normalized bands (Table 2), the CSM height values and the vegetation indices using multispectral data (Table 3); (3) all features from both sensors, excluding the height information from the CSM; and (4) using all extracted features.

Models resulting from the training with the multiple datasets and classifiers combinations had their performance assessed by resorting to confusion matrices. Indeed, several metrics were extracted to assess performance, namely: precision, recall and F1 score. Both precision and recall consider the number of true positives (TP). However, while the first tallies false positives (TP/TP+FP), the second considers false negatives (TP/TP+FN). F1 score is the harmonic mean of precision and recall. Furthermore, overall accuracy (OA) and kappa coefficient were used to evaluate the models' general performance. Lastly, the overall model predictions were visually compared.

Prediction outcomes from the automatic point cloud classification function available in Pix4Dmapper Pro were visually compared to assess grapevine discrimination capabilities. Points assigned to the "high-vegetation" class were compared since in a typical vineyard, high-vegetation usually means grapevines. Therefore, the generated RGB dense point clouds were exported to CloudCompare (version 2.11.3, GPL software, <https://www.cloudcompare.org/>) and all points belonging to the "high vegetation" class were rasterized. Moreover, thresholding techniques commonly found in studies present in the state-of-the-art, were also used for comparison: NDVI was used with a threshold value based on the Otsu's method (Otsu, 1979) and also by applying a height value of 0.5 m to the CSMs.

3. Results

3.1. Data characterization for each vineyard

Before training and testing the different ML classifiers and datasets, the behaviour of features selected to be used in this study was analysed with regard to each defined class. By doing so, it is possible to assess differences (if any) among vineyards and sensors. As such, the mean value of the features in each sample square spread throughout the vineyards and evenly distributed by the defined classes was used: 80 samples per class and a total of 320 samples per vineyard. Fig. 3 presents the overall mean value for the three case-study areas.

With regard to the normalized red, green and blue bands obtained from the RGB imagery, it shows values between 0.3 and 0.4 for both soil and shadow classes. It is noteworthy that values are higher for the red

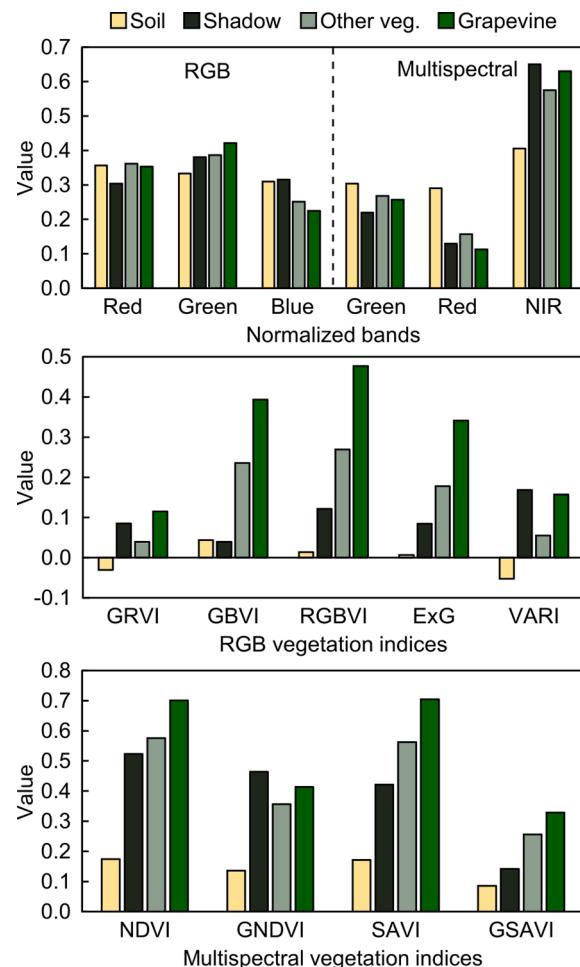


Fig. 3. Overall mean value of the square samples in the studied vineyards.

band in the soil class, and for the green band in the shadow class. As for the vegetation classes (grapevine and other vegetation) the green band presents higher mean values, followed by the red band. As for RGB-based vegetation indices—excluding the VARI that presents the higher values in the shadow class—they showed higher mean values in the grapevine class. The other vegetation class was the second with the highest values for GBVI, RGBVI and ExG, while in the GRVI, the shadow class presented higher values than other vegetation. The soil class had the lowest values, being close to or lower than zero. With GBVI, the shadow class presented the lowest values.

As for the multispectral data normalized bands, NIR presents the higher mean values in all classes, while the red band presents the lower ones. Most vegetation indices computed presented the higher values in the grapevine class, with GNDVI as the only exception (the shadow class presents higher mean values). The soil class presents the lowest values in all vegetation indices. In RVI, the higher values were achieved by the grapevine class (7.45), followed by the shadow class (5.28) and other vegetation (4.44), while soil class presents the lowest mean values (1.53).

Lastly, the mean height values extracted from the CSM present a value of 1.65 m in the grapevine class, 0.53 m in the shadow class, 0.33 m in the other vegetation class and 0.23 m in the soil class. Mean values observed for each studied vineyard are provided in Fig. A1.

3.2. Performance of each classifier per vineyard and dataset

The number of samples used to assess ML classifiers' performance in each vineyard, considering the transects (location in Fig. 1) and 120 squares (30 per class), was a total of 597 in DDR vineyard, 389 in DVVR and 635 in CC vineyard. DDR, DVVR and CC vineyards are hereinafter referred as vineyards A, B, and C, respectively. As for the distribution of the number of validation samples per class, the shadow class was the least represented with 11%, 23% and 12% of the total samples for A, B and C vineyards (65, 88 and 75 samples, respectively). In vineyard A, the soil class represents 27% of the total samples, while grapevine and other vegetation classes represent each 31%. In vineyard B, the grapevine class had 24% of the samples, the soil class 26%, and the other vegetation class 27%. Lastly, vineyard C had the soil class representing 33% of the total number of samples, while the other vegetation class represents 29% and the grapevine class 26%.

3.2.1. Overall performance of each classifier

The performance of the different classifiers when dealing with the four selected classes showed disparate results depending on the used features. The OA, F1 score and kappa coefficient values per classifier, dataset and vineyard are presented in Fig. A2, while precision and recall values are available in Fig. A3. When analysing the OA of each dataset, higher rates are achieved when using all features. The classifier with the best performance was the ANN (88% mean OA), achieving the higher accuracy rates in the different combinations of features. Both SVM and RF obtained a similar mean OA (81%) but presenting different performances, depending on the dataset used for training. SVM achieved higher rates when using features from the RGB data (80% and 78%, respectively for the SVM and RF), while RF performed better when using all features excluding the CSM (82% and 79%, respectively for RF and SVM). When using only features from the multispectral data and all features, both classifiers had a similar performance. Kappa values follow a similar trend to the OA, being more pronounced in the ANN. Regarding the F1 score of each class in all combinations of features, generally the higher rates were achieved in the grapevine class, followed by the soil class. This trend is common to all classifiers. Generally, other vegetation class presents a higher F1 score than the shadow class in all classifiers, when considering the mean value. Still, when the CSM is not considered, the opposite happens. Regarding the precision and recall metrics by class type, ANN achieves the higher rates in all classes for recall and for grapevine, shadow and other vegetation classes in precision. RF has the highest precision rate for the soil class (0.92). A similar value is achieved in the grapevine class, but this classifier presents the lower precision value in the other two classes. The SVM classifier presented a similar precision rate. However, recall results are higher when compared to those from RF for the grapevine class, and lower in the remainder classes.

When using height values from the CSM, there is an overall increase in classifiers performance. Indeed, a mean value of 6% increase is observed on the OA, when considering all classifiers trained with all features and trained without the CSM. The most noticeable increase was verified in the ANN for all vineyards (mean of 9%). It is even the greater in vineyards A and C. Moreover, a grapevine classification decrease is verified when CSM is not considered, regardless the classifier or features used.

The proneness of inter-class misclassifications can be assessed through the analysis of precision and recall metrics. The mean value of this metrics when considering all the studied vineyards and using each dataset, was analysed and it is presented in Table 4. When using the RGB dataset, precision values above 0.9 were obtained in all classifiers for the grapevine class. The same was verified in the recall values of SVM and

Table 4

Mean performance of the different classifiers and dataset combinations. Best performance in each parameter highlighted in bold.

Model	Dataset			
	RGB	MSP	All	No CSM
Overall accuracy (Kappa)				
SVM	0.80 (0.72)	0.82 (0.75)	0.84 (0.78)	0.79 (0.71)
RF	0.78 (0.70)	0.82 (0.75)	0.84 (0.78)	0.82 (0.75)
ANN	0.85 (0.79)	0.82 (0.87)	0.94 (0.92)	0.86 (0.80)
F1 score (Soil class)				
SVM	0.70	0.86	0.87	0.87
RF	0.82	0.79	0.87	0.89
ANN	0.85	0.88	0.96	0.93
F1 score (Shadow class)				
SVM	0.70	0.71	0.75	0.76
RF	0.77	0.70	0.76	0.86
ANN	0.84	0.80	0.92	0.86
F1 score (Other vegetation class)				
SVM	0.78	0.78	0.81	0.73
RF	0.67	0.83	0.82	0.75
ANN	0.77	0.85	0.92	0.81
F1 score (Grapevine class)				
SVM	0.93	0.89	0.89	0.81
RF	0.87	0.89	0.87	0.80
ANN	0.93	0.91	0.96	0.83

ANN classifiers. As for other classes, lower precision values were obtained. This is the case of the other vegetation class in the SVM and RF classifiers, where values equal or lower than 0.7 were obtained, meaning that samples from this class were classified as being part of other classes. The recall value of this class when using a RF classifier was similar, indicating that both FN and FP classifications are prone to occur. Low recall values were observed in soil and shadow classes for the SVM and in shadow and other vegetation classes for both RF and ANN, showing that samples from other classes are prone to be classified in these classes.

When analysing the datasets using multispectral features, it is observed that the shadow class presents the lower precision in all classifiers, with more relevance in SVM and RF. This means that non-shadowed areas can eventually be classified as shadows. Precision values higher than 0.9 were achieved in all classifiers for soil and grapevine classes. A similar trend is observed for recall values, except for the soil class in the RF. Low recall values are observed in other vegetation class for SVM and soil class in the RF.

By using all features, a general improvement is verified. Indeed, all classes show a precision value higher than 0.9 in the ANN. Moreover, the recall values are also above this value in soil, other vegetation, and grapevine classes. All shadow samples were correctly classified. However, the recall value for this class demonstrates that some samples from other classes are being classified as shadows. Precision values equal or above 0.9 were registered for soil and grapevine classes in SVM and RF classifiers. With the SVM classifier, the shadow class presents the lowest precision and recall values. The same is verified in the RF classifier, which also shows that samples from other vegetation classes were classified in other classes.

When excluding the use of CSM, recall values support that samples from other classes are being classified as grapevines (lowest recall value in all classifiers). Regarding precision values, samples from other vegetation class are being classified in other classes. Soil class seems to maintain the performance (less difference in values when to the usage of all features). The shadow class presents good precision results in both RF and ANN. Still, there are cases of other classes being classified as shadow. Indeed, this class seems to be prone to more misclassification cases with the SVM classifier.

3.2.2. Individual performance per classifier and dataset

Considering performance per dataset and studied vineyard, ANN presents the highest OA when using RGB features (average of 85%), achieving a 92% in vineyard B, 88% in vineyard A and 76% in vineyard C. SVM follows with 87% in vineyard B and 86% in vineyard A. As for vineyard C, RF presented the second higher OA with 74%, while SVM achieved 65%. In vineyards A and B, the RF classifier achieved 77% and 84% OA, respectively. Kappa values follow a similar trend.

When using only multispectral features, ANN achieved the higher overall accuracies in vineyards A and B: respectively, 86% and 85%. In vineyard C, both ANN and RF classifiers achieved 89% OA, but with a slightly higher kappa value for ANN. A difference of 1% in the overall performance is verified when comparing SVM and RF in vineyard A (85% and 84%, respectively) and B (73% and 72%, respectively). As for Chianti's vineyard, the SVM achieved 87% OA.

The combination of RGB and multispectral features without using height values from the CSM shows a decrease in ANN's performance, when compared with the accuracy obtained using RGB or multispectral features with the CSM: 91%, 84% and 83% for vineyard B, A and C, respectively. RF achieved a better performance than SVM in vineyards B (respectively, 84% and 77%) and C (respectively, 82% and 79%). In vineyard A, SVM achieved 82% OA, while RF 79%.

By using all features (normalized RGB and multispectral bands, all vegetation indices and CSM), ANN performs above 90% in all studied vineyards: 92% in vineyard C, 95% in vineyard A, and 96% in vineyard B. As for SVM, the overall performance decreased in vineyard A, when comparing it to the model trained with the dataset using only RGB features. However, the higher OA was achieved in vineyard C (88%). Regarding RF and similarly to SVM results, a decrease is also verified in all studied vineyards, when compared to the results obtained using the multispectral features in vineyards A and C (83% and 88%, respectively), and to the results using only RGB features in vineyard B (80%).

3.3. Vineyard classification using different datasets

Classifier models trained with the different datasets were applied to the three studied vineyards. Shapefiles generated from the mean-shift algorithm were used after the estimation of the same features. A total of 17,364 objects were generated in vineyard A (1.24 ha), 4991 in vineyard B (0.37 ha), and 38,897 in vineyard C (1.34 ha). Areas within orange squares highlighted in Fig. 1 were used to visually demonstrate differences in the classification results.

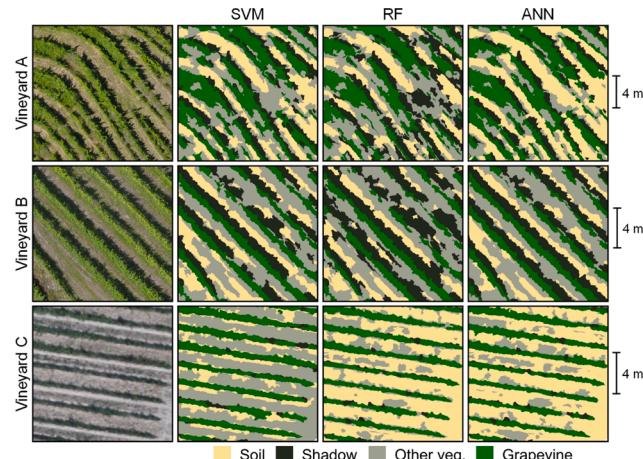


Fig. 4. Vineyard classification results from each classifier using RGB features. RGB representation in the first column.

3.3.1. Classification with RGB features

Results when considering the dataset composed of RGB data (Fig. 4) present different classification outcomes among both vineyards and classifiers. While RF models show a clear over estimation of shadows in all studied vineyards, in vineyards A and C some areas had grapevines estimated as shadows. SVM presents grapevines classified as both soil and other vegetation in vineyards A and B. Moreover, there was also an over estimation of other vegetation in soil parts in vineyard C, as well as soil and other vegetation classified as shadows. ANN presents shadows classified as other vegetation and an over classification of soil in other vegetation areas in vineyards A and C. In vineyards A and B, some grapevine branches and grapevines with a lower vigour were classified as other vegetation. Furthermore, an over estimation of shadows in grapevine objects is verified in vineyard C.

3.3.2. Classification with multispectral features

Normalized bands from the multispectral data, vegetation indices and CSM were used in another dataset. Fig. 5 presents classification results in a portion of the studied vineyards. Most grapevines seem to be correctly detected. However, in vineyard A some inter-row vegetation was classified as grapevine with less predominance on the predictions obtained from the ANN classifier. In vineyard B, some thinner grapevine parts were classified as other vegetation in all ML classifiers. The same happens in vineyard A in plants with lower vigour, which were classified as shadows or other vegetation. This is more noticeable in RF and SVM results. In vineyard C, grapevine objects were classified as being shadows when resorting to RF and SVM models. Furthermore, there are some vine rows that were almost totally classified as shadows in all classifiers. An over classification of shadows is verified in vineyard B when using SVM and RF ML classifiers. These misclassifications included objects from soil and other vegetation classes. Both SVM and RF present over classification of other vegetation in soil parts.

3.3.3. Classification with RGB and multispectral features

Vineyard classification results when using all features are presented in Fig. 6. There is a visible misclassification of shadows in the results of vineyard B in both the SVM and RF classifiers. This occurs mainly in areas with soil and other vegetation. Incorrect classification of shadows is also noticeable in RF results from vineyard A. Grapevines classified as other vegetation are observed both in vineyards A and B in all classifiers, within the slope areas between grapevine rows. However, this seems to occur less in the classification performed with ANN. In vineyard C there are cases where non-grapevine objects were classified as being

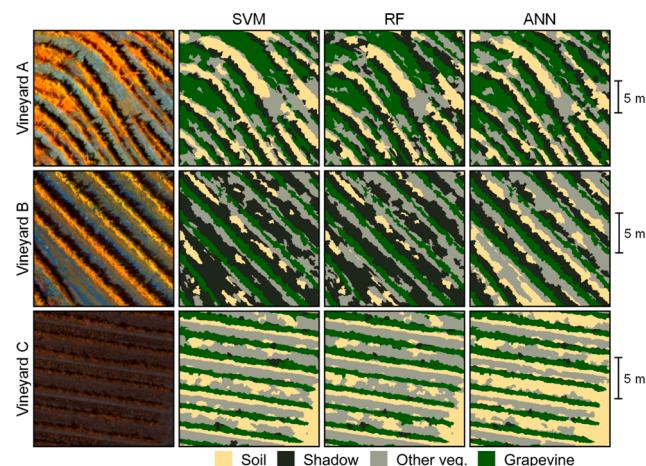


Fig. 5. Vineyard classification results from each classifier using multispectral features. False colour representation (near infrared, green, red) in the first column.

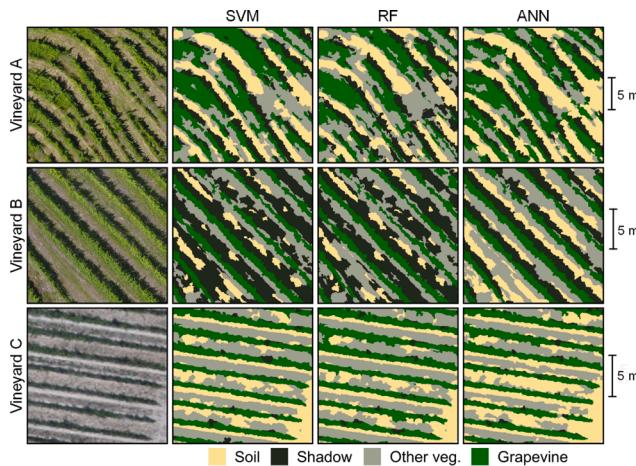


Fig. 6. Vineyard classification results from each classifier using all features. RGB representation in the first column.

grapevines in both SVM and RF results. In these two classifiers, some soil parts were also classified as being other vegetation, in vineyard C. Some grapevine objects were also classified as shadow in the results from the RF classifier.

Models trained with the features from both sensors but excluding height values from the CSM resulted in an overall classification of grapevines as other vegetation (Fig. 7). In fact, SVM seems to present a lower occurrence on these cases when compared to the other selected classifiers. Over classification of shadows seem to be reduced in vineyard B for RF but persisted in the SVM classification results.

The area predicted in each class per model in the studied vineyards is presented in Fig. 8. Classification of vineyard A using the ANN approach presents a higher incidence of soil and grapevines. The other two classifiers show a larger occupation area in other vegetation and shadow classes. In vineyard B, ANN showed a higher area for soil, other vegetation and grapevine classes, when compared to the area predicted by the other two classifiers. As for shadows, SVM classifier predicted a higher area, while the lower area was predicted by the ANN (1663 m^2 and 833 m^2 , respectively). Moreover, RF showed the lowest estimated area of soil (469 m^2) and grapevines, with 666 m^2 . As for vineyard C, all classifiers present a small presence of shadows: 4% to 5%. RF shows the higher amount of vegetation (grapevine and other vegetation), followed by the SVM, while more soil was predicted using the ANN.

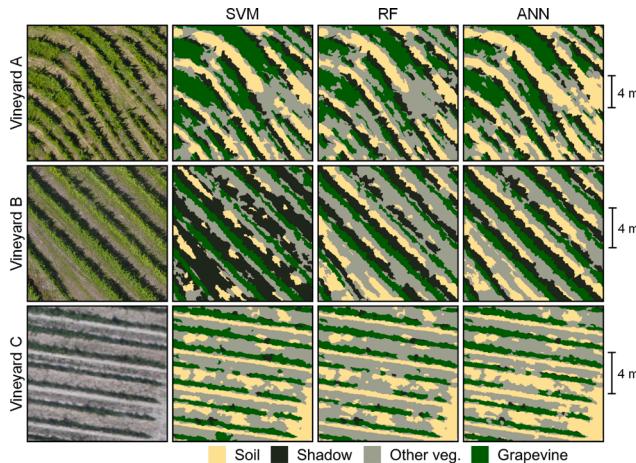


Fig. 7. Vineyard classification results from each classifier using all features except height from the CSM. RGB representation in the first column.

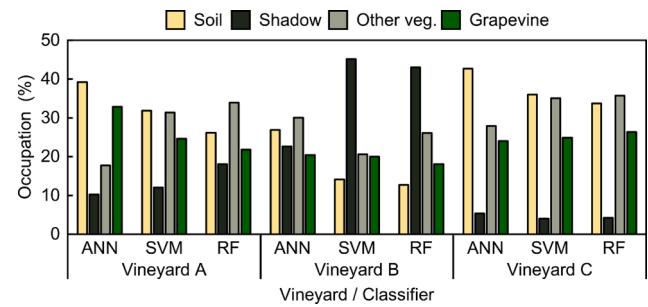


Fig. 8. Class occupation area in each vineyard when using all features in the three classifiers.

3.4. Grapevine classification using dense point clouds and image thresholding

The automatic point cloud classification method available in Pix4Dmapper Pro was applied during dense point cloud computation of the RGB imagery from the three studied vineyards (Fig. 9a). For a visual comparison, classification results from the ANN using all features (Fig. 10), can be used. Point cloud classification (Fig. 9a) shows connected vine rows in all vineyards. This has a higher predominance in vineyard A. As the rows have a difference in slope in this vineyard, vegetation between grapevine rows was classified as “high vegetation”. This is also true in four rows from vineyard B. In vineyard C, there are some parts of grapevine rows that were misclassified and a clearly noticeable area in the north-western part that contains soil classified as “high vegetation”. Nevertheless, small portions of grapevines in vineyards A and B that were under classified by the ANN, were over classified in the point cloud classification.

When using NDVI (Fig. 9b) for grapevine vegetation detection (with threshold values of 0.29 in vineyard A, 0.53 in vineyard B and 0.31 in vineyard C), results are not totally reliable. Most of the grapevine vegetation can be detected, but other vegetation present in the vineyard is also considered (more noticeable in vineyards A and C), or an under estimation can happen (vineyard B). On the other hand, by considering a height threshold applied to the CSM (Fig. 9c), results improve. However, this can lead to under estimation of grapevine vegetation, alongside with other erroneous parts being considered as grapevine (vineyard A). Furthermore, grapevine rows appear thicker than they really are (vineyard B) and less vigorous plants are not being considered (vineyard C).

4. Discussion

4.1. Features characterization

Features used in this study presented distinct behaviours among the classes, whilst normalized bands had a less noticeable difference. Vegetation indices showed a clearer difference. Indeed, most of them presented higher values for grapevines, followed by other vegetation and shadows. The soil class had the lowest values. This trend is observed in the NIR and green normalized bands, but not in GRVI.

When observing the mean values for each vineyard (Fig. A1), differences among the sensors are perceptible. Both vineyards A and B have similar trends in the normalized bands. In fact, almost no differences were verified for bands derived from the RGB sensor, while multispectral bands reveal differences in the NIR band. With regard to vegetation indices, an overall similar tendency is verified when comparing vineyards A and B. As for vineyard C, where a different setup (UAV and sensors) was used, a distinct behaviour, when compared to vineyards A and B, is noticeable. While the normalized values of the RGB bands present similar values, the NIR band of the multispectral sensor presents

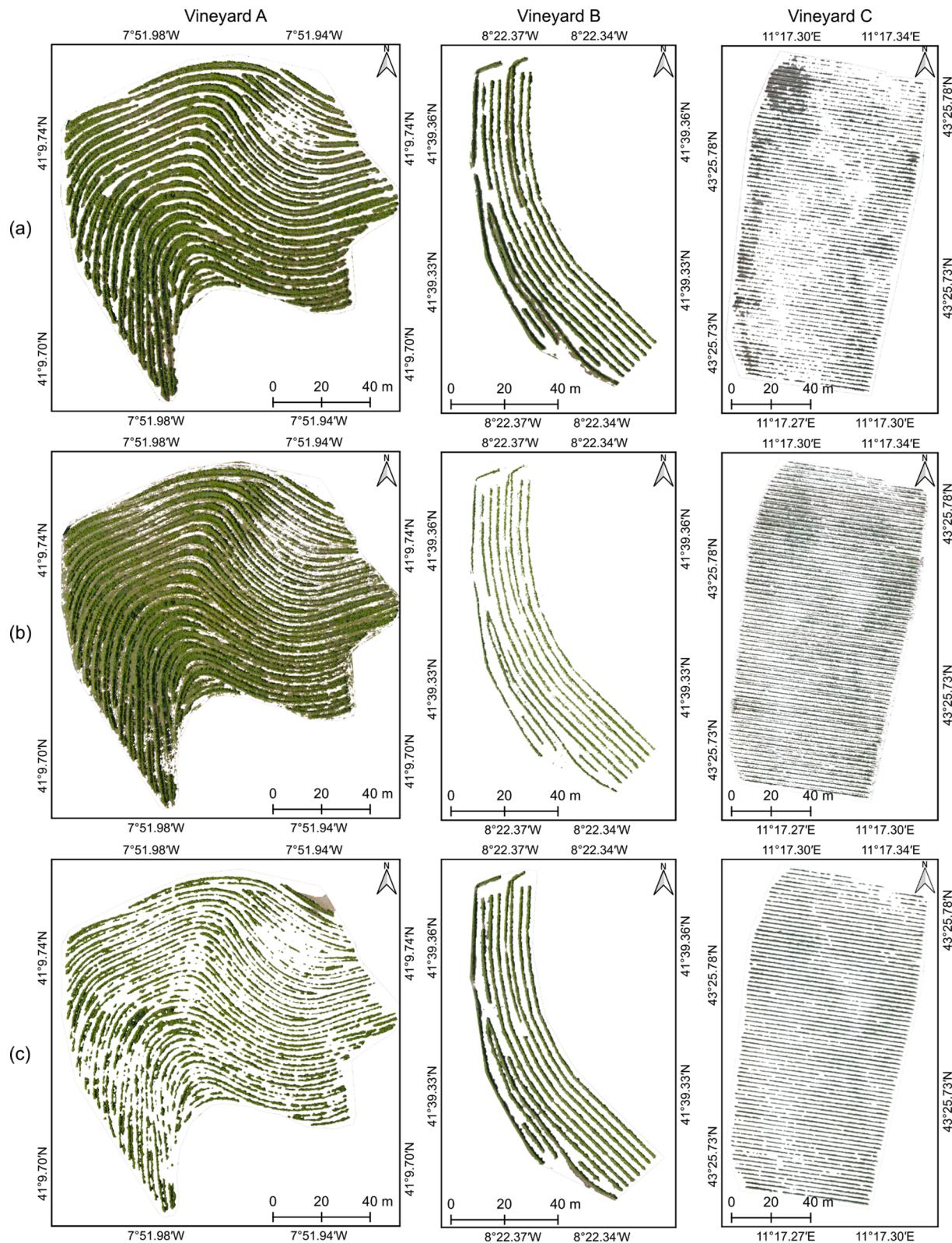


Fig. 9. Vineyard classification results for the three analysed vineyards: (a) using automatic point cloud classification, (b) by using image thresholding based in the normalized difference vegetation index (c) in the crop surface model.

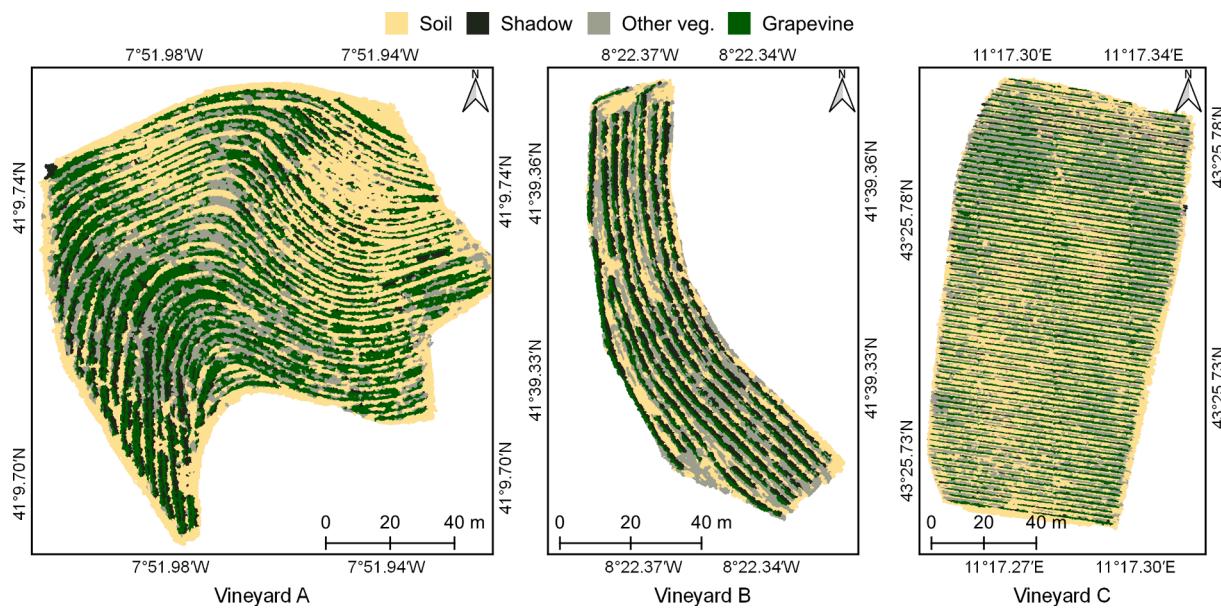


Fig. 10. Complete overview of the vineyard classification results of the artificial neural network classifier when using all features.

Table 5

Published studies addressing vineyard classification (*) and grapevine segmentation (**). Methods, imagery type (RGB, MSP: multispectral, CIR: colour infrared), and the best achieved results are provided, using overall accuracy (OA), when present.

Input data	Spatial resolution (m)	Methodology	Results	Reference
RGB from Aircraft CIR	0.30	Object oriented procedure (OOP) Ward's modified method (WMM)	OPP 84% OA WMM 82% OA	Puletti et al. (2014) **
	0.06	Hough space clustering	95% correct row detection	Comba et al. (2015) **
CIR	0.04	Local histogram equalization Total least squares Histogram filtering	Mean sensitivity and precision of 0.971	Nolan et al. (2015) **
RGB	0.02	Contour recognition Skeletonization techniques - Threshold on vegetation indices (G% and 2G_RBi) (2 classes) - ANN, RF and k-means clustering on three classes	G% = 96% 2G_RBi = 98% K-means = 64% ANN = 97% RF = 94%	Poblete-Echeverría et al. (2017) * and **
RGB	0.04 to 0.05	Binarization of RGB vegetation index and CSM, morphological operations	94% OA	Pádua et al. (2018) **
RGB	0.01	DSM-based OBIA procedure	OA > 93.6% (mean of 95.3%)	De Castro et al. (2018) **
RGB	<0.01	K-means clustering and CLARA algorithm on the TriangularGreenness Index (TGI)	CLARA correctly detected the vegetation.	Fuentes-Peñaillillo et al. (2018) *
RGB and MSP	0.03	Different combination of bands	Error percentageRGB: HSV-S (8%), HSV-G (100%), K-means (32%), DSM (14%)MSP: HSV-NRG (44%), HSV-RGN (11%), K-means (26%), DSM (5%)	Cinat et al. (2019) *
RGB	0.02	HSV colour spectrum K-means clustering Top-hat filtering for DSM HSV and Lab colour spectrums; ExG; Image thresholding	OA HSV = 88%	Karatzinis et al. (2020) **
RGB and CIR	0.01	Deep learning segmentation using SegNet	Lab = 86% ExG = 91%	Kerkech et al. (2020b) *
RGB and CIR	0.01	DL segmentation using SegNet; Pixel-wise (PW) and grapevine level (GL) classifications	88% OA PW: 90% OA GL: 95% OA Using both RGB and MSP masks	Kerkech et al. (2020a) *
RGB	0.02	Same methodology, RGB imagery. Red, green and blue pixel values also used as features.	SVM: 45% RF: 88% ANN: 86%	Pádua et al. (2020) *

higher values for the shadow class. RGB vegetation indices show higher values for the grapevine class. The same happens for the multispectral vegetation indices, excluding of GNDVI.

By cross-comparing results between the multispectral sensor used in vineyard C and those obtained by the sensor used in vineyards A and B, a conclusion is that there are differences in the range of values and in the classes order. The shadow class presents the second highest value in most of them, while the soil class presents the lowest values.

Differences on vineyards A and B values can be justified by the planting style (Section 2.1) (Costa et al., 2015; Magalhães, 2008). In vineyard B, plants can reach a height of almost 3 m. However, the feature variance observed among sensors (vineyards A and B compared to vineyard C) leads to the need of model training, especially for the multispectral imagery.

4.2. Classification performance

By analysing the classifiers' performance, it can be concluded that the best overall performance is achieved when using ANNs. This classifier presents an OA 7% higher and a kappa coefficient 0.09 higher than the remainder classifiers, when considering the mean values of all vineyards and datasets used in this study. Misclassification errors seem fewer than when using SVM and RF, as it can be seen by analysing classification results (Figs. 4, 5, 6, and 7). Both RF and SVM present more pronounced misclassification errors in all datasets: shadows classified as soil (Figs. 4, 5, 6, and 7); soil classified as other vegetation (Fig. 4); and other vegetation classified as grapevine (Fig. 7). The over classification of shadows is clearly noticeable in the occupation percentage area from both SVM and RF (Fig. 8). Low accuracy of shadow detection in multispectral imagery is also reported in Kerkech et al. (2020a), when compared to RGB imagery, these results were justified due to low ground brightness.

Regarding the features used in the classification process, there were significant improvements in the RF classifier, when exclusively using multispectral-based features. On the other hand, the opposite happened when ANN and SVM were used. Indeed, this can be concluded from the interpretation of the qualitative (Figs. 4 and 5) and quantitative (Figs. A2 and A3) results. Hence, results show that RGB dataset presents a higher importance, since higher accuracies were achieved when compared with the multispectral dataset. However, the combination of data from both sensors resulted in the best performance, reinforcing the importance of its use for specific applications where classification quality is decisive. Kerkech et al. (2020a) reached a similar conclusion when applying a deep learning approach in the classification process by combining the classification masks of both RGB and multispectral sensors. In Poblete-Echeverría et al. (2017), an ANN also outperformed a RF in the classification of grapevine, soil and shadows. Moreover, when using vegetation indices as features, a performance increase was also noticeable.

As for the importance of the CSM in vineyard classification, the mean height information proved to be crucial: OA increased in all vineyards and classifiers. Features derived from DEMs also proved to be of high importance in the classification of grassland communities and agro-silvopastoral systems (Melville et al., 2019; Pádua et al., 2019). In fact, feature ranking approaches can help understand which features have an higher contribution (Su et al., 2018).

In a previous study, an RGB dataset was used for vineyard classification and RF and ANN presented an OA above 85% (87.8% and 86.1%, respectively), while SVM obtained only 45.3% (Pádua et al., 2020). The study used the same feature types as the RGB dataset but the "raw" digital numbers of each band were also included. This can negatively influence results, since these features are correlated with the normalized

band values. As such, it proves the originality of this study, which adds value to the implemented novelties.

Table 5 summarises a list of studies found in literature and their main results. The vast majority intended to distinguish grapevines from other vineyard elements, or only considered three classes (grapevine, soil, and shadows). Furthermore, it should be noted that these studies were not carried out in the same type of vineyards: steep slope terrain, presence of terraces, curved vine rows or excess of inter-row vegetation.

Regarding grapevine detection studies, a comparison can be established by considering the F1 score results for the grapevine class. Indeed, the mean F1 score obtained in this study for the ANN is of 0.83 (all classifiers where above 0.80), when not considering the CSM. This value is lower than most reported in published studies found in literature (Comba et al., 2015; de Castro et al., 2018; Karatzinis et al., 2020; Nolan et al., 2015; Pádua et al., 2018; Poblete-Echeverría et al., 2017; Puletti et al., 2014). However, when including CSM height information, F1 score increases. In the RGB dataset, F1 score value increases to 0.93 and to 0.96 if multispectral data is also used (all features), which is well above most of the referred published results. This can be an indicator that if the models were retrained to classify between grapevine and non-grapevine objects (2 classes), performance could overcome most of the previously achieved results, if applied to commercial vineyards.

Unsupervised clustering approaches as K-means shown unreliable results in several studies that applied it (Cinat et al., 2019; Fuentes-Peñailllo et al., 2018; Poblete-Echeverría et al., 2017). Indeed, it tends to over-estimate grapevines, and showed problems in shadow estimation. Using other colour spaces, as HSV (hue, saturation and value), ended up in unstable classification results between vineyards (Cinat et al., 2019).

Common threshold operations achieved good results in other studies (Table 5) when applied to commercial vineyards (Burgos et al., 2015; Campos et al., 2019; Karatzinis et al., 2020; Pádua et al., 2019; Poblete-Echeverría et al., 2017). However, due to the nature of the vineyards used in this study and to the data quality, these approaches tend not to be completely successful, as shown in Fig. 9. Other vegetation is detected along with grapevines when using NDVI. As for CSM, it is dependent on the quality of the DTM, DSM, and photogrammetric processing parameters (Burgos et al., 2015). The combination of both of these outputs can lead to a 94% accuracy detection (Pádua et al., 2018) in regular vineyards. However, if applied to more complex scenarios, as in vineyard A, where a considerable number of grapevines was not considered and some soil and inter-row vegetation was accounted in the height threshold, it would be quite challenging.

4.3. Potential application scenarios and challenges to overcome

The proposed approach can distinguish different classes in a vineyard and to estimate the occupation area of each class (Fig. 8), helping the farmers/winegrowers to obtain an overall context of the vineyard. Indeed, it shows potential to be extended for other uses, such as in detecting grapevines' disease where shadow identification is critical to discard shadowed grapevine parts, since those can change plants' reflectance. Therefore, only non-shadowed pixels are analysed (Albetis et al., 2017). Areas deemed as "other vegetation" can be used in the same manner to discriminate invasive weed species among other cover crops, as presented in De Castro et al. (2020). Furthermore, elements classified as grapevine vegetation can be used to apply other classification approaches in their pixels (Burgos et al., 2015).

The approach presented in this study also can be expanded outside vineyards' scope, since the evaluated classes are commonly found in other contexts involving different crops and tree species. It is a solid approach for land use classification. Moreover, it can be applied in

multi-temporal datasets for classification of land use and land cover (Talukdar et al., 2020).

Data quality improvements must be considered. In fact, there can be inter-class contamination on the objects that result from the OBIA procedure. In fact, a small number of pixels related to other classes can be present in some objects, potentially causing classification errors. To overcome this issue and improving the classification, Mathews (2014) reported that the use of median object values would produce more reliable results than the mean value, due to the possible inclusion of small portions of soil or shadows. Furthermore, De Castro et al. (2018) stated that by increasing spatial resolution the robustness of the OBIA procedure can improve.

Deep learning semantic segmentation is another approach to be considered for future work. Indeed, this type of techniques will also perform well in this type of study. However, this type of approach is more demanding in terms of data (quantity, labelling, etc.) and computational resources. The amount of data required for training should be significantly increased: more vineyards should be included, and to increase the robustness of the convolutional networks, data at different growth stages should also be considered. The methods presented in the manuscript were intended to be tested with as fewest as possible data and number of samples. By using polygons to classify and test the different approaches it was possible to verify the model suitability for different datasets and to avoid a complete labelling of the whole study area.

Multispectral crop monitoring, mainly based in vegetation indices calculation combined with segmentation techniques, represents an excellent solution for vegetative vigour characterization and also to monitor phytosanitary status (Rodríguez et al., 2021), which must not be disregarded. However, as described in Matese and Di Gennaro (2021), there is a large number of factors that can affect the quality of spectral data, such as stable light conditions, sun radiation angle and radiometric correction expertise. In that context, the advance in RGB technology and photogrammetric software automation, opens an extremely interesting perspective on RGB sensors, when compared to multispectral. RGB sensors provide a higher resolution at a lower cost, especially considering that most commercial UAVs come equipped out of the box with a RGB camera, typically 16 MP or better. Moreover, a direct consequence of the higher absolute resolution is the opportunity to obtain the same resolution of a multispectral camera but at a higher flight quote, boosting vineyard monitoring in terms of more surface/time coverage (Di Gennaro and Matese, 2020). RGB cameras can provide not only a wide number of consolidated vegetation indices, but also geometric analysis of the canopy directly linked to vegetative growth and biomass, relatively independent from light conditions. Thus, RGB cameras represent an effective tool for a large community of users to perform easy aerial data acquisition for crop segmentation, using inexpensive instrumentation (<1000.00 €), overcoming any issues related with the need of spectral know-how on radiometric correction steps.

5. Conclusions

Three vineyards with distinct topography and within different wine regions were studied in an OBIA vineyard classification procedure. UAV-based RGB and multispectral data was acquired using different combinations of UAVs and sensors. Data was then used in a photogrammetric processing pipeline to compute different outcomes, creating datasets for ML composed of a set of features from both sensors. Three supervised ML classifiers were assessed to classify vineyard elements into four classes. The best performing ML approach ANN outperformed SVM and RF. Both SVM and RF showed misclassification errors, especially in the estimation

of shadows. RGB-based features shown a greater overall contribution for an accurate classification, when compared with multispectral data. However, when using data from both sensors, there was an improvement in classification. The absence of CSM from the dataset contributed to poor classification results. When comparing the results from the approach proposed in this study with already published classification methods (point cloud classification, threshold of NDVI and CSM), it proved to be stable and effective in cases where other methods presented distinct behaviours. Thus, it can be applied to vineyard classification processes based on UAV-based data. Nevertheless, performance of the selected classifiers can improve with a fine tuning of hyper parameters.

CRediT authorship contribution statement

Luís Pádua: Conceptualization, Methodology, Formal analysis, Software, Writing – original draft. **Alessando Matese:** Writing – original draft, Methodology. **Salvatore Filippo Di Gennaro:** Validation, Writing – original draft. **Raul Morais:** Visualization, Investigation, Writing – original draft. **Emanuel Peres:** Conceptualization, Software, Writing – original draft, Writing – review & editing. **Joaquim J. Sousa:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge Castello di Fonterutoli owned by Marchesi Mazzei Spa Agricola for having hosted the experimental activities. The authors also want to thank Andrea Berton (IGG-CNR) for UAV management and data acquisition.

Funding

This research activity was supported by National Funds by FCT - Portuguese Foundation for Science and Technology, under the project UIDB/04033/2020, and by E-Crops Project - Technology for Sustainable Digital Agriculture - ARS01_01136, within the National Operational Programme on Research and Innovation 2014-2020 (Italy).

Appendix A

Fig. A1 presents the mean values of the samples for each class in the three studied vineyards. Mean-values from the RVI in vineyard A for soil, shadow, other vegetation and grapevine classes are 2.08, 6.83, 6.97, 13.50, respectively. For vineyard B, values of 0.97, 1.62, 2.60, and 4.02 were obtained, and for vineyard C the values are 1.54, 7.39, 3.75, and 4.83. Regarding the mean height obtained from the application of CSM for the soil, shadow, other vegetation and grapevine, they were respectively 0.04 m, 0.23 m, 0.04 m and 1.15 m, for vineyard A, 0.05 m, 0.36 m, 0.05 m, and 1.92 m, for vineyard B, and 0.26 m, 0.31 m, 0.05 m, and 1.30 m, for vineyard C.

Fig. A2 presents the classifiers results for each vineyard regarding the F1 score, overall accuracy and kappa coefficient. Precision and recall values are presented in **Fig. A3**.

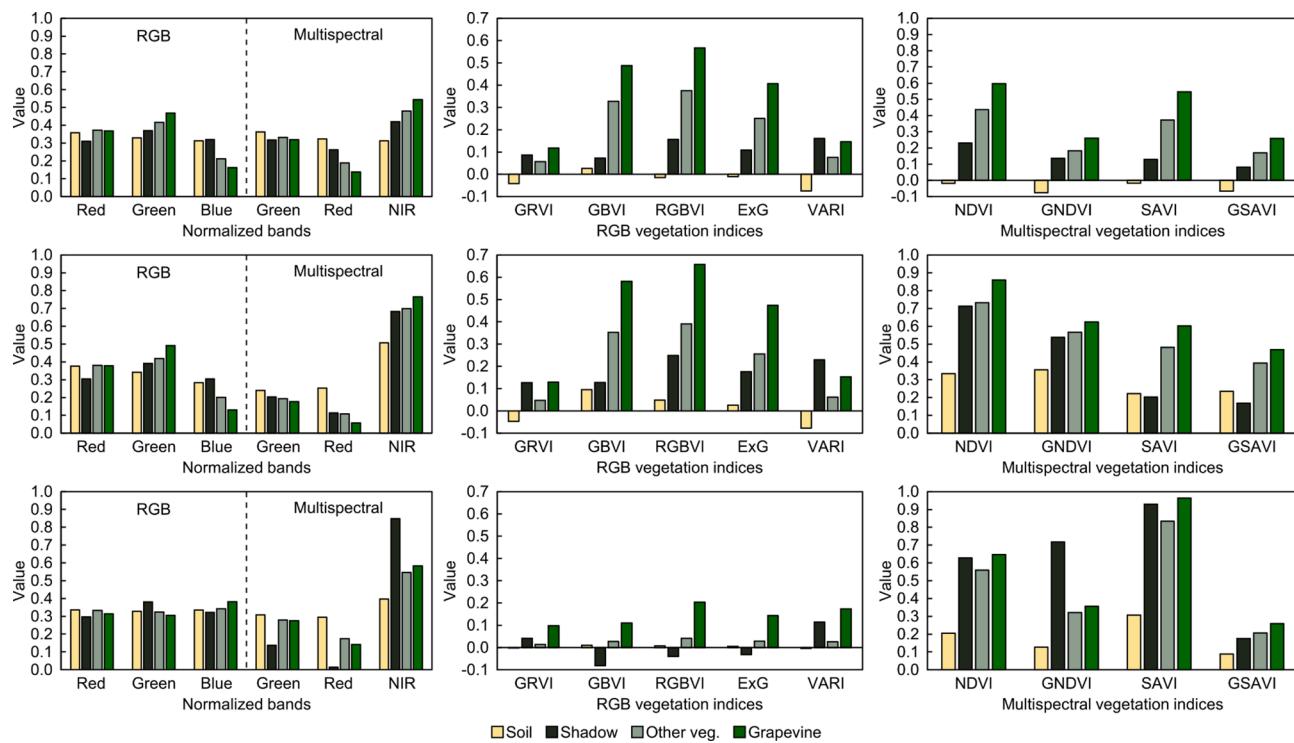


Fig. A1. Overall mean value of the square samples in each vineyard for the normalized bands of both sensors (left column), the RGB vegetation indices (centre column) and the multispectral vegetation indices (right column). Top-row vineyard A, centre row vineyard B and in the bottom row vineyard C.

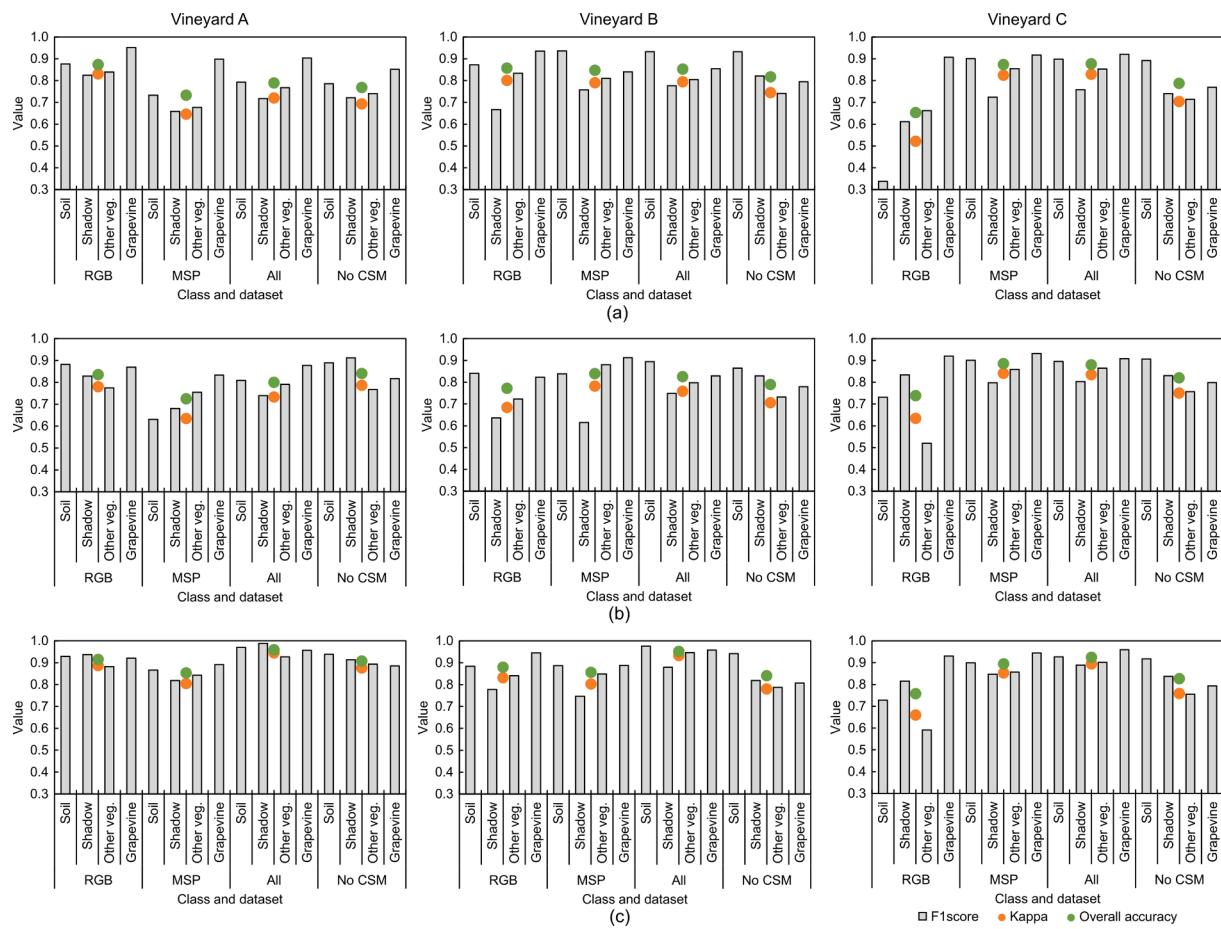


Fig. A2. Overall accuracy, F1 score and kappa value for the three studied vineyards and different datasets, when training the classifiers: (a) support vector machine; (b) random forest; and (c) artificial neural network.

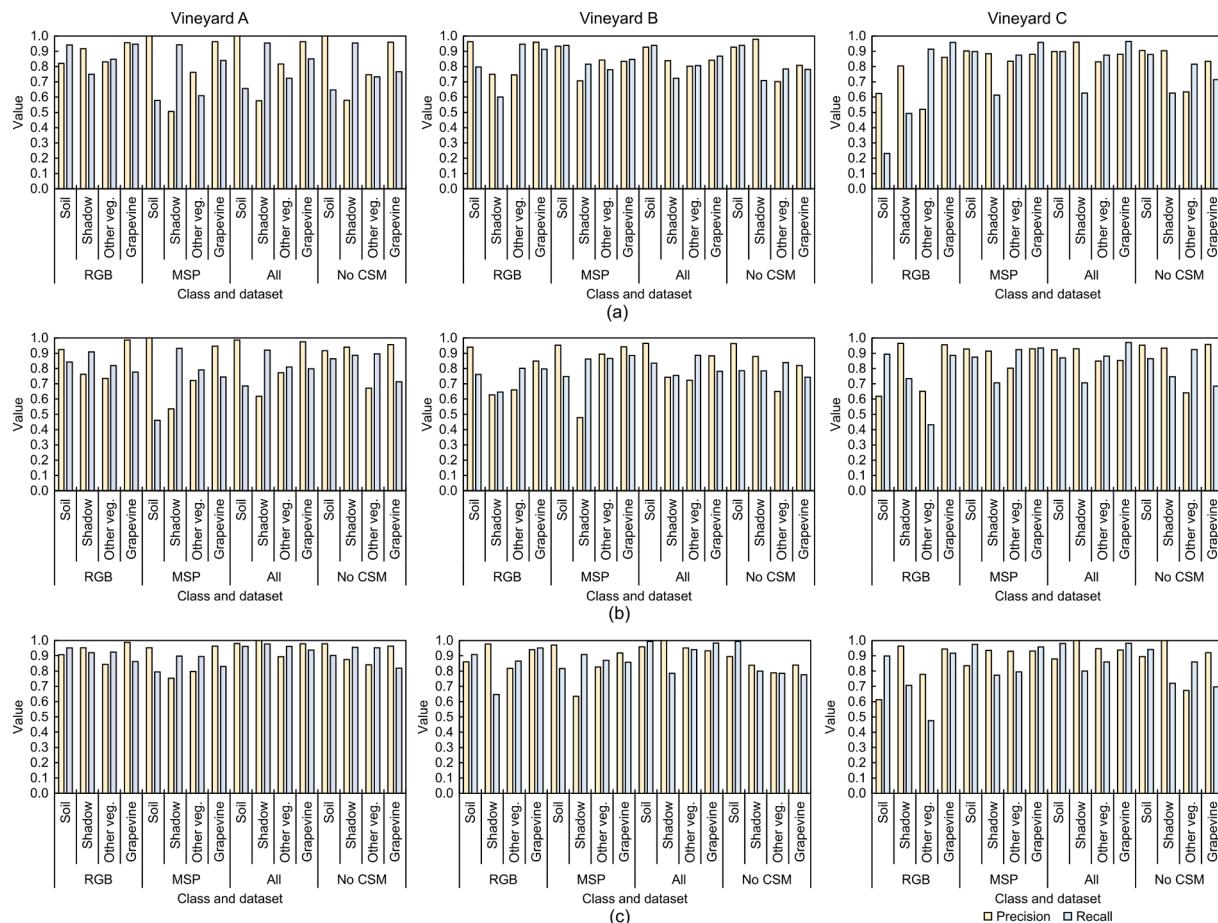


Fig. A3. Precision and recall values for the three studied vineyards and different datasets, when training the classifiers: (a) support vector machine; (b) random forest; and (c) artificial neural network.

References

- Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., Sousa, J.J., 2017. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sens.* 9, 1110. <https://doi.org/10.3390/rs911110>.
- Albertis, J., Duthoit, S., Guttler, F., Jacquin, A., Goulard, M., Poilvè, H., Féret, J.-B., Dedieu, G., 2017. Detection of Flavescence dorée Grapevine Disease Using Unmanned Aerial Vehicle (UAV) Multispectral Imagery. *Remote Sens.* 9, 308. <https://doi.org/10.3390/rs9040308>.
- Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35 (2-3), 161–173.
- Becker, C., Rosinskaya, E., Häni, N., D'Angelo, E., Strecha, C., 2018. Classification of aerial photogrammetric 3D point clouds. *Photogramm. Eng. Remote Sens.* 84 (5), 287–295.
- Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G., 2015. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Observ. Geoinform.* 39, 79–87. <https://doi.org/10.1016/j.jag.2015.02.012>.
- Bramley, R.G.V., 2001. Progress in the development of precision viticulture - variation in yield, quality and soil properties in contrasting Australian vineyards.
- Burgos, S., Mota, M., Noll, D., Cannelle, B., 2015. Use of very high-resolution airborne images to analyse 3D canopy architecture of a vineyard. *Int. Arch. Photogramm. Remote Sens. Spatial Inform. Sci.* 40, 399.
- Campos, J., Llop, J., Gallart, M., García-Ruiz, F., Gras, A., Salcedo, R., Gil, E., 2019. Development of canopy vigour maps using UAV for site-specific management during vineyard spraying process. *Precis. Agric.* 20 (6), 1136–1156.
- Chang, C.-C., Lin, C.-J., 2011. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. <https://doi.org/10.1145/1961189.1961199>.
- Cinat, P., Di Gennaro, S.F., Berton, A., Matese, A., 2019. Comparison of Unsupervised Algorithms for Vineyard Canopy Segmentation from UAV Multispectral Images. *Remote Sens.* 11, 1023. <https://doi.org/10.3390/rs11091023>.
- Comba, L., Biglia, A., Ricauda Aimonino, D., Gay, P., 2018. Unsupervised detection of vineyards by 3D point-cloud UAV photogrammetry for precision agriculture. *Comput. Electron. Agric.* 155, 84–95. <https://doi.org/10.1016/j.compag.2018.10.005>.
- Comba, L., Gay, P., Primicerio, J., Ricauda Aimonino, D., 2015. Vineyard detection from unmanned aerial systems images. *Comput. Electron. Agric.* 114, 78–87. <https://doi.org/10.1016/j.compag.2015.03.011>.
- Costa, R., Fraga, H., Malheiro, A.C., Santos, J.A., 2015. Application of crop modelling to portuguese viticulture: implementation and added-values for strategic planning. *Ciências Téc. Vitiv.* 30, 29–42. <https://doi.org/10.1051/ctv/20153001029>.
- de Castro, A.I., Jiménez-Brenes, F.M., Torres-Sánchez, J., Peña, J.M., Borrà-Serrano, I., López-Granados, F., 2018. 3-D Characterization of Vineyards Using a Novel UAV Imagery-Based OBIA Procedure for Precision Viticulture Applications. *Remote Sens.* 10, 584. <https://doi.org/10.3390/rs10040584>.
- de Castro, A.I., Peña, J.M., Torres-Sánchez, J., Jiménez-Brenes, F.M., Valencia-Gredilla, F., Recasens, J., López-Granados, F., 2020. Mapping *Cynodon dactylon* Infesting Cover Crops with an Automatic Decision Tree-OBIA Procedure and UAV Imagery for Precision Viticulture. *Remote Sens.* 12, 56. <https://doi.org/10.3390/rs12010056>.
- Di Gennaro, S.F., Matese, A., 2020. Evaluation of novel precision viticulture tool for canopy biomass estimation and missing plant detection based on 2.5D and 3D approaches using RGB images acquired by UAV platform. *Plant Methods* 16, 91. <https://doi.org/10.1186/s13007-020-00632-2>.
- Fuentes-Penalillo, F., Ortega-Farfías, S., Rivera, M., Bardeen, M., Moreno, M., 2018. Using clustering algorithms to segment UAV-based RGB images, in: 2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA). In: Presented at the 2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA), pp. 1–5. <https://doi.org/10.1109/ICA-ACCA.2018.8609822>.
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7).
- Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87. [https://doi.org/10.1016/S0034-4257\(01\)00289-9](https://doi.org/10.1016/S0034-4257(01)00289-9).
- Guimaraes, N., Pádua, L., Marques, P., Silva, N., Peres, E., Sousa, J.J., 2020. Forestry Remote Sensing from Unmanned Aerial Vehicles: A Review Focusing on the Data,

- Processing and Potentialities. *Remote Sens.* 12, 1046. <https://doi.org/10.3390/rs12061046>.
- Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X).
- Inglaada, J., Christophe, E., 2009. The Orfeo Toolbox remote sensing image processing software. In: 2009 IEEE International Geoscience and Remote Sensing Symposium. Presented at the 2009 IEEE International Geoscience and Remote Sensing Symposium, p. IV-733-IV-736. <https://doi.org/10.1109/IGARSS.2009.5417481>.
- Jurado, J.M., Pádua, L., Feito, F.R., Sousa, J.J., 2020. Automatic Grapevine Trunk Detection on UAV-Based Point Cloud. *Remote Sens.* 12, 3043. <https://doi.org/10.3390/rs12183043>.
- Karatzinis, G.D., Apostolidis, S.D., Kapoutsis, A.C., Panagiotopoulou, L., Boutilis, Y.S., Kosmatopoulos, E.B., 2020. Towards an Integrated Low-Cost Agricultural Monitoring System with Unmanned Aircraft System. In: 2020 International Conference on Unmanned Aircraft Systems (ICUAS). Presented at the 2020 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 1131–1138. <https://doi.org/10.1109/ICUAS48674.2020.9213900>.
- Kawashima, S., Nakatani, M., 1998. An Algorithm for Estimating Chlorophyll Content in Leaves Using a Video Camera. *Ann. Bot.* 81, 49–54. <https://doi.org/10.1006/anbo.1997.0544>.
- Kerkech, M., Hafiane, A., Canals, R., 2020a. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Comput. Electron. Agric.* 174, 105446. <https://doi.org/10.1016/j.compag.2020.105446>.
- Kerkech, M., Hafiane, A., Canals, R., Ros, F., 2020b. Vine Disease Detection by Deep Learning Method Combined with 3D Depth Information. In: El Moataz, A., Mammass, D., Mansouri, A., Nouboud, F. (Eds.), *Image and Signal Processing, Lecture Notes in Computer Science*. Springer International Publishing, Cham, pp. 82–90. https://doi.org/10.1007/978-3-030-51935-3_9.
- Khaliq, A., Comba, L., Biglia, A., Ricauda Aimonino, D., Chaberge, M., Gay, P., 2019. Comparison of Satellite and UAV-Based Multispectral Imagery for Vineyard Variability Assessment. *Remote Sens.* 11, 436. <https://doi.org/10.3390/rs11040436>.
- Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A., 1995. Color Indices for Weed Identification Under Various Soil, Residue, and Lighting Conditions. *Trans. ASAE* 38, 259. <https://doi.org/10.13031/2013.27838>.
- Magalhães, N., 2008. *Tratado de viticultura: a videira, a vinha eo terroir*. Chaves Ferreira.
- Mahmon, N.A., Ya'acob, N., 2014. A review on classification of satellite image using Artificial Neural Network (ANN). In: 2014 IEEE 5th Control and System Graduate Research Colloquium. Presented at the 2014 IEEE 5th Control and System Graduate Research Colloquium, pp. 153–157. <https://doi.org/10.1109/ICSGRC.2014.690871>.
- Matese, A., Di Gennaro, S.F., 2021. Beyond the traditional NDVI index as a key factor to mainstream the use of UAV in precision viticulture. *Sci. Rep.* 11, 2721. <https://doi.org/10.1038/s41598-021-81652-3>.
- Matese, A., Di Gennaro, S.F., Berton, A., 2016. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. *Int. J. Remote Sens.* 38 (8-10), 2150–2160.
- Matese, A., Toscano, P., Di Gennaro, S.F., Genesio, L., Vaccari, F.P., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R., Gioli, B., 2015. Intercomparison of UAV, Aircraft and Satellite Remote Sensing Platforms for Precision Viticulture. *Remote Sens.* 7, 2971–2990. <https://doi.org/10.3390/rs70302971>.
- Mathews, A.J., 2014. Object-based spatiotemporal analysis of vine canopy vigor using an inexpensive unmanned aerial vehicle remote sensing system. *JARS* 8, 085199. <https://doi.org/10.1117/1.JRS.8.085199>.
- Melville, B., Lucieer, A., Aryal, J., 2019. Classification of Lowland Native Grassland Communities Using Hyperspectral Unmanned Aircraft System (UAS) Imagery in the Tasmanian Midlands. *Drones* 3, 5. <https://doi.org/10.3390/drones3010005>.
- Michel, J., Youssefi, D., Grizonnet, M., 2014. Stable mean-shift algorithm and its application to the segmentation of arbitrarily large remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 53 (2), 952–964.
- Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* 66, 247–259. <https://doi.org/10.1016/j.isprsjprs.2010.11.001>.
- Nolan, A., Park, S., Fuentes, S., Ryu, D., Chung, H., 2015. Automated detection and segmentation of vine rows using high resolution UAS imagery in a commercial vineyard. Presented at the Proceedings of the 21st International Congress on Modelling and Simulation, Gold Coast, Australia, pp. 1406–1412.
- Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybernet.* 9, 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>.
- Pádua, L., Adão, T., Hruška, J., Guimarães, N., Marques, P., Peres, E., Sousa, J.J., 2020. Vineyard Classification Using Machine Learning Techniques Applied to RGB-UAV Imagery. In: IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, pp. 6309–6312. <https://doi.org/10.1109/IGARSS39084.2020.9324380>.
- Pádua, L., Guimarães, N., Adão, T., Marques, P., 2019. Classification of an Agrosilvopastoral System Using RGB Imagery from an Unmanned Aerial Vehicle. In: Moura Oliveira, P., Novais, P., Reis, L. (Eds.), *Progress in Artificial Intelligence*. EPIA 2019. Springer.
- Pádua, L., Vanko, J., Hruška, J., Adão, T., Bessa, J., Sousa, A., Peres, E., Morais, R., Sousa, J.J., 2018. Vineyard properties extraction combining UAS-based RGB imagery with elevation data. *Int. J. Remote Sens.* 39, 5377–5401. <https://doi.org/10.1080/01431161.2018.1471548>.
- Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J.J., Peres, E., Morais, R., 2017. UAS, sensors, and data processing in agroforestry: a review towards practical applications. *Int. J. Remote Sens.* 38, 2349–2391. <https://doi.org/10.1080/01431161.2017.1297548>.
- Panagiotidis, D., Abdollahnejad, A., Surový, P., Chitecupo, V., 2017. Determining tree height and crown diameter from high-resolution UAV imagery. *Int. J. Remote Sens.* 38, 2392–2410. <https://doi.org/10.1080/01431161.2016.1264028>.
- Poblete-Echeverría, C., Olmedo, G.F., Ingram, B., Bardeen, M., 2017. Detection and Segmentation of Vine Canopy in Ultra-High Spatial Resolution RGB Imagery Obtained from Unmanned Aerial Vehicle (UAV): A Case Study in a Commercial Vineyard. *Remote Sens.* 9, 268. <https://doi.org/10.3390/rs9030268>.
- Puletti, N., Perria, R., Storchi, P., 2014. Unsupervised classification of very high remotely sensed images for grapevine rows detection. *Eur. J. Remote Sens.* 47, 45–54. <https://doi.org/10.5721/EuRS20144704>.
- Rodríguez, J., Lizarazo, I., Prieto, F., Angulo-Morales, V., 2021. Assessment of potato late blight from UAV-based multispectral imagery. *Comput. Electron. Agric.* 184, 106061. <https://doi.org/10.1016/j.compag.2021.106061>.
- Rouse Jr., J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring Vegetation Systems in the Great Plains with ERTs. *NASA Special Public.* 351, 309.
- Salamí, E., Barrado, C., Pastor, E., 2014. UAV Flight Experiments Applied to the Remote Sensing of Vegetated Areas. *Remote Sens.* 6, 11051–11081. <https://doi.org/10.3390/rs61111051>.
- Sozzi, M., Kayad, A., Marinello, F., Taylor, J., Tisseyre, B., 2020. Comparing vineyard imagery acquired from Sentinel-2 and Unmanned Aerial Vehicle (UAV) platform. *OENO One* 54, 189–197. <https://doi.org/10.20870/oeno-one.2020.54.1.2557>.
- Sripada, R.P., Heiniger, R.W., White, J.G., Meijer, A.D., 2006. Aerial Color Infrared Photography for Determining Early In-Season Nitrogen Requirements in Corn. *Agron. J.* 98, 968–977. <https://doi.org/10.2134/agronj2005.0200>.
- Su, J., Liu, C., Coombes, M., Hu, X., Wang, C., Xu, X., Li, Q., Guo, L., Chen, W.-H., 2018. Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Comput. Electron. Agric.* 155, 157–166. <https://doi.org/10.1016/j.compag.2018.10.017>.
- Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., Rahman, A., 2020. Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sens.* 12 (7), 1135.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).