

Article

Object-Based Image Classification of Summer Crops with Machine Learning Methods

José M. Peña^{1,2,*}, Pedro A. Gutiérrez³, César Hervás-Martínez³, Johan Six⁴,
Richard E. Plant² and Francisca López-Granados¹

¹ Institute for Sustainable Agriculture, IAS-CSIC, P.O. Box 4084, E-14080 Córdoba, Spain;
E-Mail: flgranados@ias.csic.es

² Department of Plant Sciences, University of California, Davis, CA 95616, USA;
E-Mail: replant@ucdavis.edu

³ Department of Computer Science and Numerical Analysis, University of Cordoba,
Campus de Rabanales, E-14071 Córdoba, Spain; E-Mails: pagutierrez@uco.es (P.A.G.);
chervas@uco.es (C.H.-M.)

⁴ Department of Environmental Systems Sciences, Swiss Federal Institute of Technology,
ETH-Zurich, CH-8092 Zurich, Switzerland; E-mail: jsix@ethz.ch

* Author to whom correspondence should be addressed; E-Mail: jmpena@ias.csic.es;
Tel.: +34-957-499-218; Fax: +34-957-499-252.

Received: 24 January 2014; in revised form: 16 May 2014 / Accepted: 19 May 2014 /

Published: 30 May 2014

Abstract: The strategic management of agricultural lands involves crop field monitoring each year. Crop discrimination via remote sensing is a complex task, especially if different crops have a similar spectral response and cropping pattern. In such cases, crop identification could be improved by combining object-based image analysis and advanced machine learning methods. In this investigation, we evaluated the C4.5 decision tree, logistic regression (LR), support vector machine (SVM) and multilayer perceptron (MLP) neural network methods, both as single classifiers and combined in a hierarchical classification, for the mapping of nine major summer crops (both woody and herbaceous) from ASTER satellite images captured in two different dates. Each method was built with different combinations of spectral and textural features obtained after the segmentation of the remote images in an object-based framework. As single classifiers, MLP and SVM obtained maximum overall accuracy of 88%, slightly higher than LR (86%) and notably higher than C4.5 (79%). The SVM+SVM classifier (best method) improved these results to 89%. In most cases, the hierarchical classifiers considerably increased the accuracy of the

most poorly classified class (minimum sensitivity). The SVM+SVM method offered a significant improvement in classification accuracy for all of the studied crops compared to the conventional decision tree classifier, ranging between 4% for safflower and 29% for corn, which suggests the application of object-based image analysis and advanced machine learning methods in complex crop classification tasks.

Keywords: agriculture; ASTER satellite images; object-oriented image analysis; hierarchical classification; neural networks

1. Introduction

The strategic management of summer irrigated crops, both herbaceous and woody crops, has relevant agro-environmental repercussions, e.g., as a means of avoiding the excessive use and the depletion of water resources, decreasing the rate of growth of GHG emissions and adopting soil conservation practices to guarantee that the future demands of agriculture are met [1]. To assist in the process of decision-making about the management of agricultural land or the implementation of agrarian actions, it is crucial to identify and monitor the distribution of crops at different spatial (parcel, county or region) scales every year. In many regions, this information is obtained regularly through farmer communications, which are subsequently partially checked in a program of ground visits to certain fields. This procedure is time-consuming, highly expensive and usually delivers inconsistent results because it generates reporting discrepancies and covers small areas or only very few accessible sampling fields [2]. An alternative and truly viable procedure to conduct a cost-effective follow-up of the crop field evaluation across large areas is to use affordable satellite imagery, which could also provide consistent data for studies involving multiple sampling times.

For decades, many investigations have addressed the topic of crop discrimination via remote sensing through the use of various sources of satellite imagery, such as Landsat, QuickBird or Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [3–5]. In this topic, several authors have observed that information concerning variations in crop calendar, crop patterns, crop management techniques and parcel sizes shall be incorporated to the classifier algorithms for a successful result [6,7]. These variations can be derived from the textural, contextual or, in some cases, morphological features of the images [3–5]. However, conventional methods based only on spectral pixel information lack these features, which limits their application for crop discrimination, mainly in heterogeneous parcels with high spectral variability and mixed pixels (e.g., in woody crops) [8]. Per-field classification can overcome these problems by merging adjacent pixels within each individual field into spectrally and spatially homogeneous objects created via a segmentation process and then classifying at the object level. This approach, known as object-based image analysis (OBIA), also incorporates the mentioned features to the classifier algorithm, which can lead to a drastic improvement of classification accuracy [9]. Recently, Castillejo-González *et al.* [8] enhanced pixel-based classifications up to 22% by applying OBIA in mapping ten different land use classes with QuickBird satellite imagery. Another source of errors is due to the high spectral similarity among different crops with common development patterns and growth calendars, e.g., herbaceous ones such

as tomato, corn, safflower or sunflower [5]. In these cases, confusion problems might be overcome by using multisource data composed of new discriminating features and by applying powerful methods, such as machine learning models, that enhance classification outputs [10].

One of the most widely studied fields of machine learning is undoubtedly supervised learning [11]. This type of learning seeks to derive a function (or, generally, a model) that can identify the class to which new examples belong based on a pre-labeled training set. Basically, the model is a function whose value depends on a set of parameters (coefficients for LR or weights for ANNs) and the input variables. These parameters are adjusted based on the training samples. Among others, the three most common tasks in supervised learning are binary classification, multi-class classification and regression. If the examples belong to only two categories, then the task is called binary classification. If there are multiple categories, then it is multi-class classification. Lastly, if the output value is a real number, the task is called regression. In contrast to this type of flat classification task, many important real-world classification problems are naturally cast as a hierarchical structure, in which the classes to be predicted are organized into a predefined hierarchy similar to a tree [12]. The advantage of using this type of technique is two-fold: the combined classifiers explore information about parent-child class relationships present in the hierarchy and, if the number of classes is large, the classifiers obtained will be less affected by the increase in problem complexity. For a highly interesting review of various hierarchical classification techniques and applications, we refer the reader to Silla and Freitas [13].

Among these models, nonparametric classifiers such as decision trees (DT), logistic regression (LR), artificial neural networks (ANN) and support vector machines (SVM) have been successfully used for remote-sensed image classification of agricultural landscapes because they can describe the intricate and complex nonlinear relationships that exist between canopy-level spectral information and crop conditions [14–17]. In this study, we evaluated the accuracy of the four cited methods for the identification and mapping of nine major summer crops from bi-temporal ASTER images captured in early and late summer in an agricultural region of California. MLP and SVM methods were selected as base classifiers given their popularity and their proven performance for pixel-based remote sensing problems [15], although only a very few investigations have tested their efficiency for object-based image analysis [16] or for a detailed classification of crop-fields. In a previous investigation over this region [5], classifications from a conventional decision tree algorithm reported a substantial degree of confusion occurred between certain herbaceous crops (e.g., alfalfa and safflower, among others) during the summer season, due primarily to similarities in the growth stage and cropping pattern during the period when the demand for crop irrigation in this agricultural area is intense. Although some level of miss-classification might be acceptable for an agricultural inventory, a thematic map of higher accuracy is required for the precise estimation of the area of each individual crop, e.g., to obtain further estimates of costs and predictions of possible water restrictions in summer. Therefore, we quantified the efficiency of the cited machine learning methods, either as single classifiers or combined in a hierarchical classification, for enhancing classification outputs both at the group level (woody or herbaceous) and at the level of individual crops. A first-stage binary classifier was trained to differentiate woody from herbaceous ones, and two second-stage multiclass classifiers (one for each group) were used to determine the specific crop. The classifiers were built from spectral and textural features derived from an OBIA framework, which allowed quantify the influence of these features in the accuracy of the classified map. Finally, classification errors due to quantity disagreement and

allocation disagreement [18] as affected by the type of crop were discussed in terms of their influence for management of agricultural lands.

2. Materials and Methods

2.1. Studied Cropland Area

This investigation was conducted with nine major crops growing during the summer of 2006 in the agricultural region of Yolo County, California, USA (center lat/lon coordinates 38.50 N, 121.50 W). This region has a Mediterranean climate characterized by hot, dry summers (daily maximum temperatures greater than 38 °C) and cool winters (daily maximum temperatures of 16 °C). The rainfall ranges between 200 and 460 mm, with the key rainfall period between November and March. The irrigation infrastructure of the region consists of three dams, two reservoirs and more than 300 km of canals and ditches that use gravity to deliver the water. Various systems, such as spraying, dripping or flooding, are commonly used to irrigate the fields. Yolo County has a total agricultural surface area of 177,000 ha (out of a total county area of 264,900 ha) with a very diverse cropping pattern representative of agriculture in the Central Valley of California. Listed in order of total crop area, the woody crops included in this study were walnut (5%), vineyard (4%) and almond (3%), and the herbaceous crops were alfalfa (19%), tomato (14%), rice (10%), sunflower (6%), safflower (5%) and maize (4%). These crops covered 70% of the county's cropland surface. The remaining 30% consisted of winter cereals (12%) and a combination of other types of hay and minor crops, for a total of as many as 130 different commodities. Apart from this cropland area, dry meadows represent approximately 20% of the county surface [19].

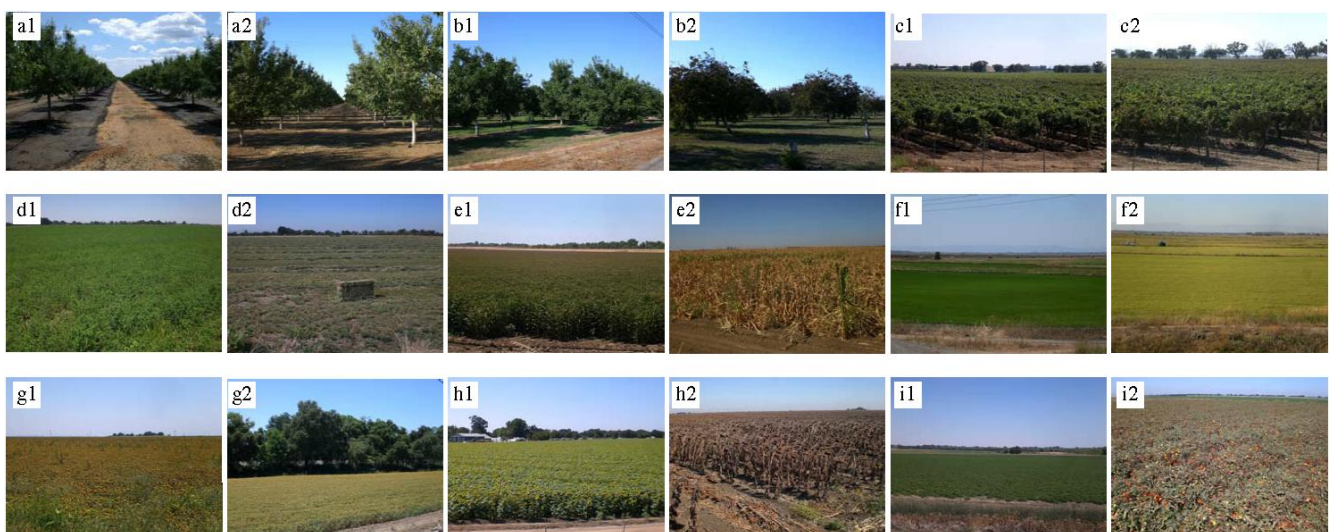
Several differences in crop field patterns were observed during field visits due to the type of crop management operation conducted by each farmer. In woody crops, the field pattern was affected by the tree size, separation between trees, plantation design and use of cover crops. In herbaceous crops, two different systems were observed: (1) conservation systems, in which field soil is protected by maintaining residue from the previous crop; and (2) traditional systems, in which tillage operations remove residues and leave the soil surface bare. In alfalfa cultivation, vegetation permanently covers the soil, although its canopy density changes several times within a year according to its cutting schedule (between six and ten times a year). This intra-crop variability increases the complexity of crop discrimination, justifying the need to explore powerful algorithms for performing a robust classification adapted to different crop patterns. A general view of the studied crops during the summer is shown in Figure 1.

2.2. Satellite Imagery and Derived Vegetation Indices

Based on a preliminary study of the crop calendars, we selected ASTER satellite images, obtained on 24 June and 12 September 2006, corresponding to key growth stages of the studied crops in early and late summer as previously reported by López-Granados *et al.* [20]. These authors successfully discriminated several types of crops at both extremes of the summer season by analyzing on-ground reflectance measurements. Either two or three scenes were acquired on each date to completely cover the studied zone. Each image was co-registered with ENVI 4.5 software (Research Systems Inc.,

Boulder, CO, USA) and then atmospherically corrected and radiometrically calibrated with its Fast Line-of-Site Atmospheric Analysis of Spectral Hypercubes (FLAASH) tool, which is based on MODTRAN4 radiative transfer code [21]. This process is recommended in multi-temporal studies and it involves the conversion of the original raw digital values to reflectance values. The visible (green and red) and near-infrared (NIR) bands, with 15 m spatial resolution, and all six shortwave infrared (SWIR) bands, with 30 m spatial resolution, were combined by resampling the 30 m images to a 15 m pixel size with the nearest-neighbor method.

Figure 1. General view of the nine studied crops in early (1) and late (2) summer: (a) almond, (b) walnut, (c) vineyard, (d) alfalfa, (e) corn, (f) rice, (g) safflower, (h) sunflower and (i) tomato.



Differences in spectral reflectance can be enhanced through the use of vegetation indices (VIs). In this investigation, ten VIs derived from ASTER wavebands were calculated for both studied dates and grouped according to the spectral region to which they belong (Table 1). The selected VIs have traditionally been used for monitoring variations in crop characteristics that can be crucial for discriminating summer crops [22]. For example, the Normalized Difference VI (NDVI) and the Green VI (VI_{green}) have commonly been used because they yield good relationships involving the type of vegetation, crop growth stage and assessment of crops [23]. Other VIs based on the short-wave infrared (SWIR) region, such as the Normalized Difference Tillage Index (NDTI) or the Normalized Differential Senescent VI (NDSVI), have been defined to enhance information obtained in the senescent crop stage or from crop residues [24]. The four studied machine learning models and the sixteen hierarchical configurations described in section 2.4 were constructed from an input dataset composed of the spectral VI described in Table 1 [23–30].

Table 1. Vegetation indices employed in this study.

Spectral Region/Vegetation Index (VI)	VI adapted to ASTER *	Reference
Visible Region		
Green Vegetation Index (VI _{green})	$(\text{Aster1} - \text{Aster2})/(\text{Aster1} + \text{Aster2})$	[23]
Visible-NIR Region		
Normalized Difference VI (NDVI)	$(\text{Aster3} - \text{Aster2})/(\text{Aster3} + \text{Aster2})$	[25]
Green NDVI (GNDVI)	$(\text{Aster3} - \text{Aster1})/(\text{Aster3} + \text{Aster1})$	[23]
Modified Chlorophyll Absorption in Reflectance Index (MCARI)	$[(\text{Aster3} - \text{Aster2}) - 0.2(\text{Aster3} - \text{Aster1})] \times (\text{Aster3}/\text{Aster2})$	[24]
NIR-SWIR Region		
Normalized Difference Index 5 (NDI5)	$(\text{Aster3} - \text{Aster4})/(\text{Aster3} + \text{Aster4})$	[26]
Normalized Difference Index 7 (NDI7)	$(\text{Aster3} - \text{Aster6})/(\text{Aster3} + \text{Aster6})$	[26]
Visible-SWIR Region		
Normalized Differential Senescent VI (NDSVI)	$(\text{Aster4} - \text{Aster2})/(\text{Aster4} + \text{Aster2})$	[27]
SWIR Region		
Normalized Difference Tillage Index (NDTI)	$(\text{Aster4} - \text{Aster6})/(\text{Aster4} + \text{Aster6})$	[28]
Lignin Cellulose Absorption (LCA)	$100 \times [(\text{Aster6} - \text{Aster5}) + (\text{Aster6} - \text{Aster8})]$	[29]
ASTER Shortwave Infrared Normalized Difference Residue Index (SINDRI)	$(\text{Aster6} - \text{Aster7})/(\text{Aster6} + \text{Aster7})$	[30]

* ASTER wavelength: ASTER1 (520–600 nm), ASTER2 (630–690 nm), ASTER3 (760–860 nm), ASTER4 (1600–1700 nm), ASTER5 (2145–2185 nm), ASTER6 (2185–2225 nm), ASTER7 (2235–2285 nm), ASTER8 (2295–2365 nm).

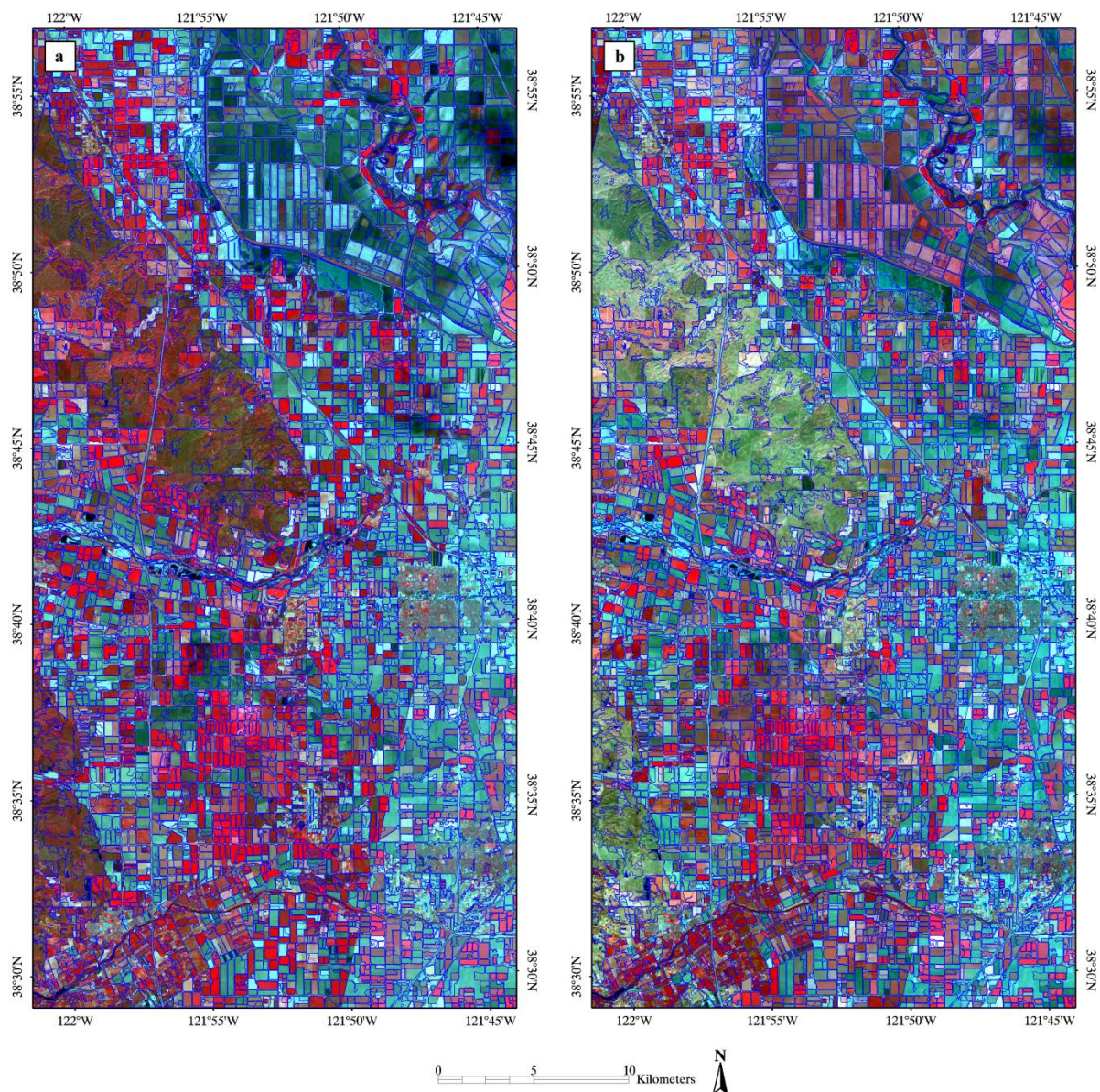
2.3. Image Segmentation and Definition of Object-Based Textural Features

Two consecutive segmentation processes were applied to the satellite images to delimit the crop field borders and generate the object-based framework (Figure 2). The multi-resolution segmentation algorithm implemented in the eCognition Developer 8 software (Trimble GeoSpatial, Munich, Germany) was used in both processes [31] by equally weighting the Green, Red and NIR bands. Values of 0.4, 0.6, 0.2 and 0.8 for scale 50 and 0.9, 0.1, 0.3 and 0.7 for scale 100 were used for color, shape, smoothness and compactness, respectively. The geometric quality of the segmentation outputs was estimated by applying the empirical discrepancy method, which consisted on contrasting the segmentation output with an already computed ground-truth reference and quantifying their discrepancy measures based on differences among certain spatial features [32]. In our investigation, we selected 350 ground-truth reference objects that represent fields from all the studied crop types and visually drew the field borders. Then, we assessed discrepancy in area, perimeter and shape index between the segmented objects and the real fields, reporting 94% average accuracy.

The object-based framework offers the possibility of computing the textural features that characterize each crop field. Object texture provides new information in addition to spectral data, and this information can enhance the power to discriminate heterogeneous classes [33], although it has the disadvantage of higher computational cost. In this investigation, the following textural features were evaluated: (1) the standard deviation (SD) of the green, red and near-infrared wavebands and (2) the homogeneity, dissimilarity and entropy based on the gray-level co-occurrence matrix (GLCM)

proposed by Haralick *et al.* [34]. The SD indicates the degree of local variability of the pixel values inside each object. The texture features based upon the GLCM were calculated by determining how often a pixel with the intensity (gray-level of the band considered) value i occurs in a specific spatial relationship to a pixel with the value j [31]. The homogeneity and dissimilarity features measure high or low object pixel uniformity, respectively, and the entropy feature is related to object pixel disorder [35]. Because other textural features based on the GLCM are more complex to interpret in the context of cropland classification, they were excluded to reduce the computation time required by the analysis [5].

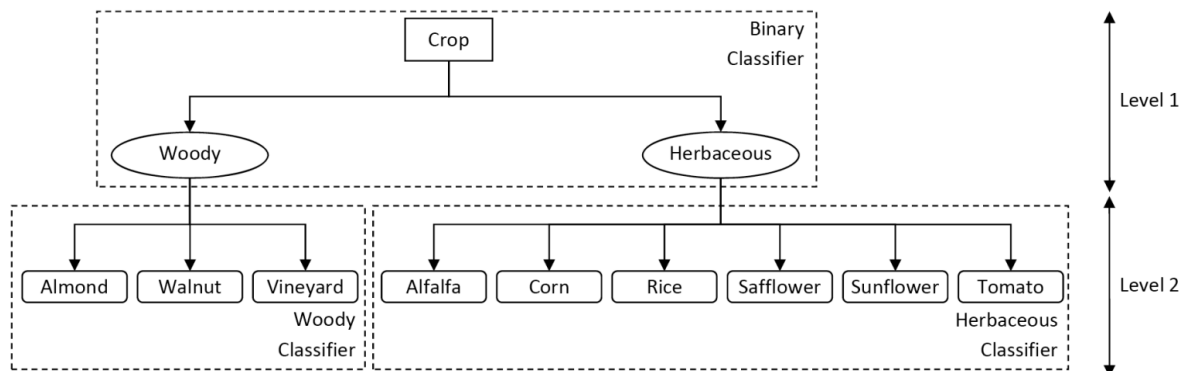
Figure 2. ASTER satellite images of the studied agricultural region in Yolo county (CA) after the segmentation process showing the crop-field borders: (a) color-infrared composition (green, red and near-infrared bands) of the image captured in early-summer, (b) color-infrared composition of the image captured in late-summer.



2.4. Machine Learning Methods

As previously discussed, a hierarchical classification technique was considered in order to count on the hierarchy information present in the target class variable, *i.e.*, the crop class. This hierarchy is shown in Figure 3, distinguishing the classes for which each classifier was trained.

Figure 3. Class hierarchy and classifier construction.



Two levels of classification were included:

(1) The first level was produced by a binary classifier, which was trained using the entire set of objects labeled as “woody” or “herbaceous”. The training dataset is described in Section 2.5.

(2) The second level was formed by two multi-class classifiers, one for each group of crops. Each classifier was trained using objects of the corresponding group (woody or herbaceous classifier). The labels for the three-class woody classifier were “almond”, “walnut” and “vineyard”, and the labels for the six-class herbaceous classifier were “alfalfa”, “corn”, “rice”, “safflower”, “sunflower” and “tomato”.

Once the three classifiers were obtained, the prediction was made by first applying the level-1 binary classifier to decide if the crop was woody or herbaceous and then using the corresponding specific level-2 classifier to decide a final class. To evaluate different configurations, four different types of algorithms were selected to construct level-1 and level-2 decision models, covering several of the most common machine learning classifiers:

(1) MultiLayer Perceptron (MLP): Artificial neural networks can associate complicated information with target attributes without any constraints on the sample distribution [36]. MLPs are the most common choice and correspond to a functional model whose hidden units are sigmoidal basis functions. We used the implementation included in the Weka machine learning tool [37], which applies a standard backpropagation process to optimize the different weights of the model. The number of hidden neurons, h , was obtained from the heuristic $h = (i + c)/2$, where i is the number of inputs and c is the number of classes. The number of iterations, i , was decided by a nested 5-fold cross-validation process, where i takes values in the following range $i \in \{50, 100, 150, 200, 250, 300\}$.

(2) Logistic Regression (LR): LR has become a widely used and accepted method of analysis for binary or multiclass outcome variables, as it can predict the probabilities associated with the states of the variables [17]. Various algorithms can be applied to obtain the LR model. For this study, we selected the Simple Logistic algorithm included in Weka. The Simple Logistic algorithm builds

multinomial logistic regression models with the LogitBoost algorithm (a boosting algorithm that performs forward stagewise fitting), first proposed by Friedman *et al.* [38] for fitting additive logistic regression models by maximum likelihood. The Simple Logistic algorithm adds one variable on each iteration, selecting the optimum number of iterations with a nested 5-fold cross-validation. This approach allows the pruning of certain inputs by stopping early in the analysis, thus avoiding overfitting.

(3) Support Vector Machine (SVM): It is a kernel learning method for classification problems in which linear separation is not possible in the input space [39]. SVM operates a non-linear transformation of the original input space X into a high dimensional feature space F , where optimal separating hyperplanes can be found. This is done by using a reproducing kernel function which plays a key role in the final performance of the classifier. In this paper, we consider the most general choice, the Gaussian kernel function. The kernel width associated to the Gaussian, σ , is adjusted by considering a nested 5-fold cross-validation process, where $\sigma \in \{10^{-3}, 10^{-1}, 10^1, 10^3\}$. The separating hyperplane is optimal when it maximizes the distance (margin) between the hyperplane and the closest points of the two classes (called support vectors), resulting in a good performance for the generalization set. Decision boundaries are smoothed to deal with the non separable case by introducing slack-variables, relaxing the hard-margin constraint. A cost parameter defined by the user, C , balances pressure on margin maximization and pressure on errors. Again, we set this parameters by considering a nested cross-validation with $C \in \{10^{-3}, 10^{-1}, 10^1, 10^3\}$.

(4) Classification trees: classification trees are decision trees, where the leaves represent classifications and the branches represent conjunctions of features that produce those classifications. C4.5 is an algorithm developed by Quinlan [40] to generate a decision tree and is an extension of the earlier ID3 algorithm. The C4.5 implementation of Weka has two main parameters to be adjusted: the confidence threshold for pruning, c , and the minimum number of instances per leaf, m . For pruning, both subtree replacement and subtree rising were considered. The two parameters mentioned were adjusted again considering a nested 5-fold cross-validation with the following ranges $c \in \{0.15, 0.25, 0.35, 0.45\}$ and $m \in \{1, 2, 3, 4\}$.

These four classifiers were considered for both levels, producing a total of sixteen hierarchical configurations: MLP + MLP, MLP + LR, MLP + SVM, MLP + C4.5, LR + MLP, LR + LR, LR + SVM, LR + C4.5, SVM + MLP, SVM + LR, SVM + SVM, SVM + C4.5, C4.5 + MLP, C4.5 + LR, C4.5 + SVM and C4.5 + C4.5. Additionally, they were applied as standard flat classifiers, constructing one classifier to predict the complete set of all possible crops. The source code for performing this hierarchical classification is available from the authors upon request.

2.5. Model Training and Evaluation

A total of 1007 crop fields (approximately 15% of the total arable area of the Yolo County) were randomly selected to train and evaluate the models. The number of fields of each crop was proportional to the total acreage in the studied zone, from a minimum of 30 fields for minor crops. The type of crop in each field was confirmed from maps provided by the Yolo agricultural commissioner's office and double-checked by visiting a subset of the fields during the summer season of the studied year. In each sampled field, the spectral and textural features described in the previous section were calculated after the segmentation process. We exported these features to .csv format by using the

export tool implemented in the eCognition software and next we opened and analyzed the .csv dataset in the WEKA software. The experimental design was conducted using a stratified 10-fold cross-validation procedure, leading to one confusion matrix for each fold. In 10-fold cross-validation, the data is randomly partitioned into 10 equal size samples. Of the 10 samples, a single sample is used to validate the model and the remaining 9 samples are used as to train the model. The cross-validation process is then repeated 10 times, with each of the 10 samples used exactly once as the validation data. A final confusion matrix was obtained by summing all of them. This external cross-validation is independent from the internal 5-fold cross-validation process used for parameter adjustment, which was done considering only the training set of the external one. The classification performance of each model was evaluated using this confusion matrix and calculating the Correct Classification Rate (CCR) of the overall classification and the minimum sensitivity (MS). On the one hand, the CCR is the percentage of parcels of all the crops classified as correct (dividing the correct parcels by all the parcels) and it represents the global accuracy of the classification task. On the other hand, the MS is the accuracy obtained for the worst classified crop, *i.e.*, it indicates the percentage of the parcels of this crop that was correctly classified. For example, a MS of 60% reveals that all crops were individually classified with accuracy higher or equal to 60%. Additional discussion and justification regarding the use of MS can be found in [40,41].

2.6. Classification Accuracy and Analysis of Crop-Field Disagreement

The method that reported the best result on overall CCR and the standard C4.5 decision tree classifier were studied in detail in terms of their classification performance of individual crops. We used C4.5 as basis for this comparison because the decision tree approaches has been commonly applied in similar land-use classification tasks [5,42,43]. In addition to the classification accuracy derived from the confusion matrices of each method, we also evaluated the two components of the classification errors: (1) due to quantity disagreement and (2) due to allocation disagreement. Pontius and Millones [18] illustrated these components by mean of a two-class pixel-based example, which we adapted to a multi-class object-based approach. In our investigation, the former component was the difference between the classified map and the ground-truth map with reference to the number of parcels of each crop, not considering their spatial positions. The latter component was the difference between both maps with reference to the spatial allocation among the parcels of each pair of different types of crop. This measurement indicated the number of parcels of each crop that were classified in the wrong place of the map and, simultaneously, matched to an equal number of misclassified parcels between each pair of crops considered. In other words, if the classified and the ground-truth maps have parcels attributed both to omission and commission errors (underestimated and overestimated classification, respectively) for two different crops, then these errors could be corrected by swapping the position of the parcels in a hypothetic better classification performance. This analysis is very relevant because differences in number and distribution of the crop-fields might affect decision-making process of management of agricultural lands.

3. Results and Discussion

The correct classification rate and the minimum sensitivity attained by each method are shown in Table 2. In order to assure the significance of the results, we performed a Friedman's statistical test [44], which is a nonparametric test to compare the effect of two factors. In our case, one factor was the algorithm considered (a total of 20 algorithms, from all flat and hierarchical classifiers) and the other was the fold of the dataset (with a total of 10 folds). Given that the results of the 10 folds are dependent, we considered Friedman's test instead of ANOVA. Six tests were applied, three tests for the CCR metric and other three for MS, where the three tests for each metric correspond to the three versions of the datasets (spectral features, textural features or both types of features). The CCR p -values were $1.139440e-24$, $1.160322e-23$ and $6.588730e-24$ with spectral, textural and complete sets of features, respectively, while MS p -values were $4.232993e-12$, $3.615609e-15$ and $5.071893e-14$. Consequently, the tests reported that performance differences obtained by the algorithms were significant with a high value of confidence.

Table 2. 10-fold cross-validation results for the various methods applied to all the studied crops: Correct Classification Rate (CCR) and Minimum Sensitivity (MS).

Method *	CCR (%)			MS (%)		
	Spectral (S)	Textural (T)	S + T	S	T	S + T
Standard Classification						
MLP	85	62	88	52	0	57
SVM	88	66	87	57	23	43
LR	86	63	86	48	7	47
C4.5	79	47	79	45	13	45
Hierarchical Classification						
MLP + MLP	87	63	87	59	3	61
MLP + SVM	88	66	87	60	23	37
MLP + LR	86	62	86	52	20	43
MLP + C4.5	81	50	83	45	3	32
SVM + MLP	88	63	87	67	3	53
SVM + SVM	89	66	86	60	23	37
SVM + LR	87	62	85	52	20	43
SVM + C4.5	82	50	82	45	3	29
LR + MLP	84	63	86	43	3	53
LR + SVM	85	65	86	47	17	37
LR + LR	83	62	85	45	13	43
LR + C4.5	79	50	82	41	3	32
C4.5 + MLP	82	60	82	51	7	51
C4.5 + SVM	84	63	82	57	20	27
C4.5 + LR	82	60	81	52	13	37
C4.5 + C4.5	78	48	78	45	3	26

* Multilayer Perceptron (MLP) neural network, Support Vector Machine (SVM), Logistic Regression (LR) and C4.5 decision tree; Within a column, the best method is in **bold** face.

3.1. Contribution of the Spectral and Textural Features

Classification accuracy varied if different datasets were used as inputs to build the models. Three different input datasets were prepared in parallel to compare the performance of the methods: (a) only spectral features; (b) only textural features and (c) both types of features. The contribution of the type of feature to the classification was evaluated by calculating the increase in accuracy resulting from the use of the combination of features (option c) rather than a single group of features (options a, b). The use of spectral features alone produced results that were almost as accurate or even more accurate than combining both kinds of features, except for MLP and MLP + C4.5. Overfitting could be the reason that adding textural features occasionally decreased the performance. It is possible that the models were learning overly complex decision surfaces that were specific to the training data and not generalizable to new samples.

The texture-based classifier yielded extremely poor overall accuracy compared with the spectral-based classifier, yielding CCRs ranging from 48% to 66% and from 78% to 89%, respectively. The greatest contribution of the textural features was found in the MLP method, in which the accuracy increased 3% relative to spectral-based performance.

3.2. Comparison of Standard Flat Classification Methods (MLP vs. SVM vs. LR vs. C4.5)

Among the standard classification methods, MLP, SVM and LR obtained similar CCR results for spectral and textural features, with overall accuracies between 87% and 85%, notably higher than 79% accuracy of C4.5 (Table 2). The classifiers based only on textural features attained the worst results, reporting CCR values between 47% for C4.5 and 66% for SVM. On the contrary, the best results were obtained by combining both spectral and textural features, reporting CCR values between 79% for C4.5 and 88% for MLP. The contribution of the spectral and textural features to each method is discussed in detail in Section 3.3. Although the C4.5 decision tree classifier yielded the worst results, it has the advantage of being a white-box model, thus allowing the interpretation of its parameters and of its set of rules. In contrast, both SVM and MLP are more complex models that make such interpretation difficult or impossible and can only be verified externally. Therefore, although the MLP, SVM, and LR models showed better classification power than the C4.5 model, the selection of each method will depend on the relative importance of classification performance and model interpretation. When all the crops and features were considered, MLP and SVM showed an improvement of 9% in overall accuracy. This value is sufficiently high to justify its use in the discrimination of summer crops.

3.3. Standard Flat vs. Hierarchical Methods

We also tested the hierarchical classification approach involving these classifiers and compared its performance with that of standard flat classifiers. Previous investigations have concluded that the combination of two or more classifiers can provide better classification accuracy than a single flat classifier [45]. The combination of SVM + SVM, with spectral features, attained the highest overall accuracy (89%) for the classification of all of the crops, although, as previously stated, the standard classification of SVM reported accuracy nearly as high (88%) (Table 2). The standard classifications of MLP and C4.5 were also slightly improved if SVM was integrated in the model, increasing MLP

accuracy from 85% to 88% and C4.5 accuracy from 79% to 84%. In contrast, hierarchical classification with C4.5 or MLP in the second level did not show, in general, a greater overall accuracy than standard flat classification methods. The reason for this result may be that these methods are not sufficiently complex to reflect additional relationships in the data resulting from isolating the different crop groups.

The MS results showed that hierarchical classification based on SVM + MLP (for spectral features) and on MLP + MLP (for spectral and textural features) markedly improved the accuracy of the most poorly classified class in comparison to standard flat classifiers. With spectral features, the MS of the SVM standard classification was only 57%, increasing to 67% when the hierarchical classification of SVM + SVM was performed. In the case of spectral and textural features, the MS of MLP was increased from 57% to 61% by the use of the level-2 MLP classifier. The maximization of MS values was important because all of the studied crops occupied a large area in this agricultural region.

The average computational time in seconds needed by the different algorithms/datasets was also calculated by using an Intel(R) Xeon(R) CPU E5405 at 2.00 GHz with 8 GB of RAM (Table 3). The highest computational time was always required by the MLP + MLP algorithm, followed by flat MLP. The lowest time was associated to LR and C4.5 methods or one of their variants. Textural features required more training time than spectral ones, and the combined set of features was always the most costly.

Table 3. Average computational time needed by the different methods for all datasets, including the optimization of the different parameters, training and test time.

Method *	Computational Time (seconds)		
	Spectral (S)	Textural (T)	S + T
Standard Classification			
MLP	102.29	125.17	173.03
SVM	27.79	35.66	48.15
LR	4.43	5.97	5.01
C4.5	5.22	8.63	11.26
Hierarchical Classification			
MLP + MLP	138.32	177.60	271.28
MLP + SVM	77.27	101.90	156.27
MLP + LR	61.38	81.16	126.59
MLP + C4.5	62.24	83.49	131.12
SVM + MLP	93.52	116.72	174.52
SVM + SVM	32.38	40.80	59.11
SVM + LR	16.56	20.23	29.84
SVM + C4.5	17.35	22.66	34.07
LR + MLP	80.39	100.78	148.94
LR + SVM	19.72	24.97	34.13
LR + LR	3.33	4.14	4.27
LR + C4.5	4.28	6.60	8.68
C4.5 + MLP	81.68	102.69	152.46
C4.5 + SVM	20.65	26.63	37.08
C4.5 + LR	4.66	6.09	7.82
C4.5 + C4.5	5.62	8.56	12.23

* Multilayer Perceptron (MLP) neural network, Support Vector Machine (SVM) and C4.5 decision tree.

3.4. Classification of Woody and Herbaceous Crops

Classification methods were also evaluated separately for woody crops (walnut, almond and vineyard) and for herbaceous crops (alfalfa, corn, rice, sunflower, safflower and tomato). All of these crops show vegetative growth during the summer, but both groups of crops have very different cropping patterns, management strategies and water requirements. The CCR and MS values for each group are shown in Table 4.

Table 4. 10-fold cross validation results of the different methods for woody and herbaceous crops: Correct Classification Rate (CCR) and Minimum Sensitivity (MS).

Method *	CCR (%)				MS (%)			
	Woody Crops		Herbaceous Crops		Woody Crops		Herbaceous Crops	
	S	S+T	S	S + T	S	S + T	S	S + T
Standard Classification								
MLP	74	80	88	90	63	61	52	57
SVM	79	80	91	88	67	67	57	43
LR	77	79	89	88	63	69	48	47
C4.5	66	69	82	82	55	49	45	45
Hierarchical Classification								
MLP + MLP	79	81	89	89	59	61	68	63
MLP + SVM	81	82	90	88	65	67	60	37
MLP + LR	79	79	88	88	59	65	52	43
MLP + C4.5	75	79	83	84	59	63	45	32
SVM + MLP	79	78	90	89	67	53	68	63
SVM + SVM	80	78	91	89	69	57	60	37
SVM + LR	79	76	89	88	67	55	52	43
SVM + C4.5	75	76	84	84	61	55	45	29
LR + MLP	69	76	88	89	43	53	68	63
LR + SVM	69	76	89	88	47	57	60	37
LR + LR	69	75	87	88	45	55	52	43
LR + C4.5	66	75	83	84	41	55	45	32
C4.5 + MLP	67	71	87	85	51	51	65	53
C4.5 + SVM	68	71	88	85	57	53	60	27
C4.5 + LR	67	70	86	85	53	53	52	37
C4.5 + C4.5	64	70	81	80	51	51	45	26

* Multilayer Perceptron (MLP) neural network, Support Vector Machine (SVM). Logistic Regression (LR) and C4.5 decision tree; Within a column, the best method is in **bold** face and the second one in *italics*.

For the woody crops, standard and hierarchical classifications based on MLP and SVM were the best methods, yielding CCR values greater than 80% for spectral and textural features. The MS results showed that standard classification with MLP and SVM produced very similar outcomes than the corresponding hierarchical classifiers. In herbaceous crops, standard classification with MLP and SVM and hierarchical classification based on both methods yielded the best results, with CCR values of approximately 90%. In contrast with the outcome for woody crops, the hierarchical classification of herbaceous crops with MLP + MLP and of SVM + MLP increased MS results up to 6% and to 20%,

respectively, compared to the standard classifiers. This group was composed of six different crops, with similar spectral and textural responses in most cases. With this number of classes and complexity, the selected hierarchical classifiers performed better, particularly for the worst classified crops, *i.e.*, those crops with the lowest accuracy. Consequently, the MS measure was more challenging for herbaceous crops than it was for woody crops in this investigation.

3.5. Classification Performance of Individual Crops

The C4.5 standard classifier and the method that best performed in this investigation, *i.e.*, hierarchical classification based on SVM + SVM, were compared at the level of individual crops by means of their confusion matrices (Tables 5 and 6) and the components of the classification errors due to quantity disagreement and allocation disagreement (Table 7).

Regarding classification accuracy, the SVM + SVM method improved the C4.5 results in all of the studied crops. The sensitivity (the percentage of ground crop-fields correctly classified for a given class) values increased from 2% in rice (e.g., the sensitivity changed from 96% for the C4.5 method, Table 5, to 98% for the SVM + SVM method, Table 6) to 29% in corn and the user's accuracy (the percentage of crop-fields labeled in the classified map that really corresponded to the given class) values increased from 5% in rice to 25% in safflower. In ascending order, the SVM + SVM method improved the classification of rice, safflower and vineyard approximately 3% on average, that of alfalfa, sunflower and tomato approximately 10% on average, that of almond and walnut approximately 15% on average and that of corn 29%. In general, greater improvement was reported for crops with a lower number of testing fields.

Table 5. Field-based confusion matrix based on the C4.5 standard classification with spectral features.

Ground- Truth Classes	User Classes											
	Woody Crops					Herbaceous Crops					Total Fields	Sensitivity
	Almond	Walnut	Vineyard	Alfalfa	Corn	Rice	Safflower	Sunflower	Tomato			
Almond	28	5	7	7	1			1	2	51	55%	
Walnut	3	71	5	23				1	7	110	65%	
Vineyard	3	3	46	6						58	79%	
Alfalfa	4	19	4	264	3	1		3	3	301	88%	
Corn		1	1	2	14	4	2	2	5	31	45%	
Rice				1	2	141		1	2	147	96%	
Safflower				1	2	1	17	2	7	30	57%	
Sunflower	3	1		3	1	1	6	85	15	115	74%	
Tomato		4		4	4	5	7	15	125	164	76%	
Total fields	41	104	63	311	27	153	32	110	166	1007		
User's Acc.	68%	68%	73%	85%	52%	92%	53%	77%	75%			

Overall Correct Classification Rate (CCR) = 79%; Overall sensitivity = 70%; Overall user's accuracy = 72%; Averaged Sensitivity of woody crops = 66%; Averaged User's Accuracy of woody crops = 70%; Averaged Sensitivity of herbaceous crops = 73%; Averaged User's Accuracy of herbaceous crops = 72%; Maximum and minimum values of Sensitivity and User's Accuracy are in **bold** face and in *italics*, respectively.

Table 6. Field-based confusion matrix based on the SVM + SVM hierarchical classification with spectral features.

Ground-Truth Classes	User Classes									Total Fields	Sensitivity
	Woody Crops				Herbaceous Crops						
	Almond	Walnut	Vineyard	Alfalfa	Corn	Rice	Safflower	Sunflower	Tomato		
Almond	35	8	5	1				1	1	51	69%
Walnut	2	91	5	6	1			1	4	110	83%
Vineyard	4	2	49	3						58	84%
Alfalfa		5	3	288		2		2	1	301	96%
Corn		1	1		23		2	3	1	31	74%
Rice	1			1		144		1		147	98%
Safflower					1	1	18	1	9	30	60%
Sunflower		1		3	1	1	1	98	10	115	85%
Tomato		4		1			2	10	147	164	90%
Total fields	42	112	63	303	26	148	23	117	173	1007	
User's Acc.	83%	81%	78%	95%	88%	97%	78%	84%	85%		

Overall Correct Classification Rate (CCR) = 89%; Overall Sensitivity = 82%; Overall User Accuracy = 86%; Averaged Sensitivity of woody crops = 79%; Averaged User's Accuracy of woody crops = 81%; Averaged Sensitivity of herbaceous crops = 84%; Averaged User's Accuracy of herbaceous crops = 88%; Maximum and minimum values of Sensitivity and User's Accuracy are in **bold** face and in *italics*, respectively.

Regarding classification errors, the total number of parcels affected either by quantity disagreement or allocation disagreement was 50 and 173 in the case of C4.5 and 42 and 84 in the case of SVM + SVM, respectively (Table 7). Both methods classified almost the same amount of almond, vineyard and corn parcels, and only minor differences of the number of parcels were observed in the other crops. However, although both methods got a similar amount of quantity disagreement (5% and 4% of the parcels, respectively), the allocation disagreement was notably higher in the C4.5 method (17%) than in the SVM + SVM method (8%), which affected to all the studied crops. For example, allocation disagreement of safflower was 11 parcels in the C4.5 method, but it was only 4 parcels in the SVM + SVM method. The former value was obtained as a result of adding 2 parcels of corn, 2 parcels of sunflower and 7 parcels of tomato, since the latter value was obtained as a result of adding 1 parcel of corn, 1 parcel of sunflower and 2 parcels of tomato, which are the number of the misclassified parcels that could be swapped between each pair of crops considered. The SVM + SVM method diminished the errors due to allocation of the C4.5 method from around 25% of reduction in the case of safflower and corn to 3% in rice and 1% in vineyard. The image classification performed by the SVM + SVM hierarchical method over the entire region of study is shown in Figure 4.

Table 7. Classification errors (number and percentage of parcels) due to quantity disagreement (QD) and allocation disagreement (AD) from the C4.5 standard classifier and the SVM+SVM hierarchical classifier with spectral features.

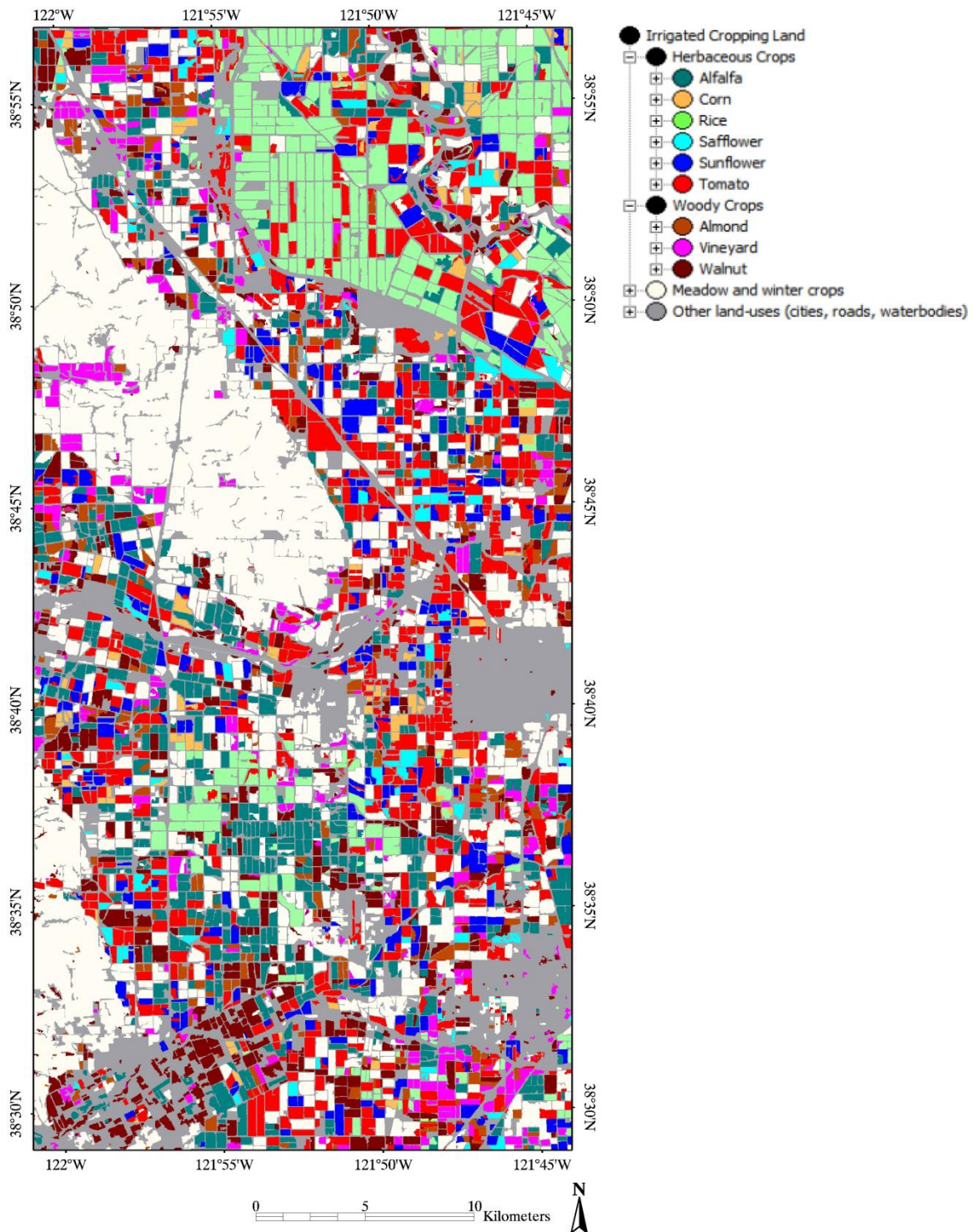
Crop	# Parcels	C4.5		SVM+SVM	
		QD	AD	QD	AD
Woody					
Almond	51	10 (20%)	11 (22%)	9 (18%)	6 (12%)
Walnut	110	6 (5%)	29 (26%)	2 (2%)	15 (14%)
Vineyard	58	5 (9%)	10 (17%)	5 (9%)	9 (16%)
Herbaceous					
Alfalfa	301	10 (3%)	36 (12%)	2 (1%)	12 (4%)
Corn	31	4 (13%)	11 (35%)	5 (16%)	3 (10%)
Rice	147	6 (4%)	6 (4%)	1 (1%)	2 (1%)
Safflower	30	2 (7%)	11 (37%)	7 (23%)	4 (13%)
Sunflower	115	5 (4%)	24 (21%)	2 (2%)	16 (14%)
Tomato	164	2 (1%)	35 (21%)	9 (5%)	17 (10%)
Total	1007	50 (5%)	173 (17%)	42 (4%)	84 (8%)

Major confusion was reported primarily between crops that belong to the same group, with the exception of woody crops and alfalfa (herbaceous). Because these crops permanently retain their greenish vegetation during the summer (Figure 1), their spectral differences were minor in the images examined, and this similarity reduced the probability of successful discrimination. In contrast, the canopy of the remaining herbaceous crops develops to the senescent yellowish stage in the late summer, resulting in increased spectral differences with the permanent crops. The classification performance was poorest for almond among the woody crops and safflower among the herbaceous crops. Almond trees were confused primarily with walnut orchards due to their similar aspect and plantation patterns or with vineyards in the cases of almond orchards characterized by small trees and affected to a certain extent by the soil background. The cases of confusion involving safflower could be explained by its similar spectral properties and vegetative development to sunflower, tomato and corn.

Despite a certain amount of misclassification, the key result of this study is summarized in two main points: (1) the classification accuracy and minimum sensitivity attained by the SVM + SVM method were extremely high for the crops that were found to be poorly classified in the conventional C4.5 decision tree method; and (2) the analysis of classification errors reported minimum differences between methods regarding the component of quantity disagreement (number of correctly classified parcels), but the SVM + SVM minimized the errors due to allocation disagreement. Moreover, we also notably enhanced classification of most crops tested in our previous investigation of the same agricultural region [5], although protocols of both cases were partly different. Regarding the number of classes, our findings were relevant in comparison to other remote sensing investigations that involved the classification of a smaller number of crops. For example, Conrad *et al.* [46] classified only three irrigated crops (cotton, rice, wheat) with 80% overall accuracy, and Simonneaux *et al.* [7] averaged 85% overall accuracy in the classification of three general land uses (annual crops, trees and bare soil). They also concluded that discrimination among different types of permanent crops (e.g., alfalfa, trees or woody crops) continue to represent major problems. In our investigation, the MLP + MLP method

increased the average sensitivity for woody crop classification by approximately 10% relative to the decision tree models.

Figure 4. ASTER satellite images classified by implementing the SVM+SVM hierarchical method in an OBIA framework.



In the context of our investigation, the selected classifiers constituted a powerful tool for the discrimination of nine major summer crops using remotely sensed data from only two dates obtained

with a timeframe of two and one-half months (from the end of June to mid-September). This low number of image acquisitions is also of great importance to avoid the need for an unaffordable number of images and the subsequent inconvenient bottleneck of image pre-processing and analysis at the very time that decisions must be made to support the programming and implementation of timely land management tools. However, additional analyses to test extrapolation to other regions are needed to evaluate the classifier robustness.

At this point, clear specifications are required regarding the improvement in accuracy of the thematic map, the computational requirements or the expertise for model interpretation involved in the modeling process and the desired objective. The computational requirements of the C4.5 and LR models were relatively low and nearly similar. On the contrary, the SVM + SVM and MLP + MLP hierarchical classification requires higher computational resources and greater expertise for model programming and interpretation than in case of the C4.5 models. From an agronomic perspective, if we aim to create a crop inventory map with a model whose implementation is easily interpreted, then a simpler and easier model as a decision tree, e.g., C4.5, would be the best choice because its accuracy of classification was adequate and the higher computational requirements, complexity and opacity of the other methods and hierarchical classification would not be justified. However, we may need to produce a very accurate thematic map with detailed classifications of summer crops because such a map may need to be prepared and ready for use in a number of scenarios. These scenarios include, e.g., demonstrating the illegal use of groundwater wells, modeling the greenhouse gas balance, decision-making and timely administrative procedures, use by the irrigation community, deriving costs for water use and modifying behavior or predicting restrictions in case of water scarcity. In such cases, the criteria for selecting a model should not be based on decreasing the computational requirements and complexity but on the accuracy of the classification. Then, according to our results, the SVM + SVM hierarchical classification, a more sophisticated and accurate model, would be highly recommended.

Another reason for the selection of more complex SVM and MLP models is that they can estimate the probabilities that an object belongs to any of the different crop types rather than the crisp classification used in the predictions obtained from C4.5 [37]. MLP and SVM outputs can easily be transformed into probabilities by applying the softmax transformation [17]. These probabilities can be used to establish confidence bounds for the predictions obtained from the models. These bounds can be expressed in terms that allow predictions to be discarded if two crop types are associated with similar probabilities. Alternatively, if the predictions are not discarded, they should at least be further evaluated by an expert.

4. Conclusions

Performance of four different machine learning methods (C4.5 decision tree, logistic regression, support vector machine and multilayer perceptron neural network), both as single classifiers and combined in a hierarchical classification, was evaluated for the task of classifying nine major summer crops in Central California in an object-based framework. A number of spectral (vegetation indices) and textural features derived from the segmentation of bi-temporal ASTER images were used as input dataset. As single classifiers, MLP and SVM obtained maximum overall accuracy of 88%, slightly higher than LR (86%) and notably higher than C4.5 (79%). In terms of overall accuracy, minor

differences were reported between the single and hierarchical classifications. However, the hierarchical classifiers based on the combination of MLP and SVM produced markedly better results if the accuracy for individual crops was considered. The best method (SVM + SVM) significantly improved classification accuracy of the C4.5 decision tree classifier for all the studied crops, ranging from 4% for safflower and rice to 29% for corn, and notably reduced the errors due to allocation disagreement between individual crops, from 17% to only 8%. The use of spectral features alone produced results that were almost as accurate as or even more accurate than in combination with textural features, suggesting minimum contribution of the textural features to the classification success.

Although an increased ability for discriminating summer crops was demonstrated with SVM and MLP in comparison to C4.5 decision tree, a higher computational time and model complexity of the best methods were also observed. Therefore, the selection of a unique classifier will depend on a desired balance between the improvement in the accuracy of the thematic map and the computational resources or expertise required for model interpretation.

The satisfactory results obtained across numerous fields in the studied agricultural region and for crops and cropping systems that are also representative of many other agricultural regions have strong implications because they opening opportunities to apply these classifiers to other areas worldwide, e.g., the Mediterranean Basin, although future analyses are necessary to evaluate their robustness.

Acknowledgments

This research was partly financed by the TIN2011-22794 project of the Spanish Ministerial Commission of Science and Technology (MICYT), FEDER funds, the P2011-TIC-7508 project of the “Junta de Andalucía” (Spain) and the Kearney Foundation of Soil Science (USA). The research of Peñá was co-financed by the Fulbright-MEC postdoctoral program, financed by the Spanish Ministry for Science and Innovation, and by the JAEDoc Program, supported by CSIC and FEDER funds. ASTER data were available to us through a NASA EOS scientific investigator affiliation.

Author Contributions

Conceived and designed the experiments: JMP PAG CHM JS. Performed the experiments: JMP PAG. Analyzed the data: JMP PAG. Interpretation of data: JMP PAG CHM FLG. Contributed reagents/materials/analysis tools: JMP PAG CHM JS REP FLG. Wrote the manuscript: JMP PAG CHM FLG.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Pacala, S.; Socolow, R. Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. *Science* **2004**, *305*, 968–972.

2. Peña-Barragán, J.M.; López-Granados, F.; García Torres, L.; Jurado-Expósito, M.; Sánchez-de la Orden, M.; García-Ferrer, A. Discriminating cropping systems and agro-environmental measures by remote sensing. *Agron. Sustain. Dev.* **2008**, *28*, 355–362.
3. De Wit, A.J.W.; Clevers, J.G.P.W. Efficiency and accuracy of per-field classification for operational crop mapping. *Int. J. Remote Sens.* **2004**, *25*, 4091–4112.
4. Pan, Y.; Hu, T.; Zhu, X.; Zhang, J.; Wang, X. Mapping cropland distributions using a hard and soft classification model. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4301–4312.
5. Peña-Barragán, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316.
6. Serra, P.; Pons, X. Monitoring farmers' decisions on Mediterranean irrigated crops using satellite image time series. *Int. J. Remote Sens.* **2008**, *29*, 2293–2316.
7. Simonneaux, V.; Duchemin, B.; Helson, D.; Er-Raki, S.; Olioso, A.; Chehbouni, A.G. The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *Int. J. Remote Sens.* **2008**, *29*, 95–116.
8. Castillejo-González, I.L.; López-Granados, F.; García-Ferrer, A.; Peña-Barragán, J.M.; Jurado-Expósito, M.; de la Orden, M.S.; González-Audicana, M. Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Comput. Electron. Agric.* **2009**, *68*, 207–215.
9. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16.
10. Lu, D.; Mausel, P.; Batistella, M.; Moran, E. Comparison of land-cover classification methods in the Brazilian Amazon Basin. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 723–731.
11. Cherkassky, V.S.; Mulier, F. *Learning from Data Concepts, Theory, and Methods*; Wiley-Interscience: Hoboken, NJ, USA, 2007.
12. Kocev, D.; Vens, C.; Struyf, J.; Džeroski, S. Tree ensembles for predicting structured outputs. *Pattern Recognit.* **2013**, *46*, 817–833.
13. Silla, C.; Freitas, A. A survey of hierarchical classification across different application domains. *Data Min. Knowl. Discov.* **2011**, *22*, 31–72.
14. Kimes, D.S.; Nelson, R.F.; Manry, M.; Fung, A.K. Attributes of neural networks for extracting continuous vegetation variables from optical and radar measurements. *Int. J. Remote Sens.* **1998**, *19*, 2639–2663.
15. Foody, G.M. Supervised image classification by MLP and RBF neural networks with and without an exhaustively defined set of classes. *Int. J. Remote Sens.* **2004**, *25*, 3091–3104.
16. Duro, D.C.; Franklin, S.E.; Dubé M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272.
17. Gutiérrez, P.A.; López-Granados, F.; Peña-Barragán, J.M.; Jurado-Expósito, M.; Hervás-Martínez, C. Logistic regression product-unit neural networks for mapping *Ridolfia segetum* infestations in sunflower crop using multitemporal remote sensed data. *Comput. Electron. Agric.* **2008**, *64*, 293–306.

18. Pontius, R.G.; Millones, M. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429.
19. Yolo County Agriculture Department. *Yolo County: Agricultural Crop Report 2006*; County of Yolo: Woodland, CA, USA, 2006.
20. Lopez-Granados, F.; Gomez-Casero, M.T.; Peña-Barragan, J.M.; Jurado-Exposito, M.; Garcia-Torres, L. Classifying irrigated crops as affected by phenological stage using Discriminant analysis and neural networks. *J. Am. Soc. Hort. Sci.* **2010**, *135*, 465–473.
21. Matthew, M.W.; Adler-Golden, S.M.; Berk, A.; Richtsmeier, S.C.; Levine, R.Y.; Bernstein, L.S.; Acharya, P.K.; Anderson, G.P.; Felde, G.W.; Hoke, M.L.; *et al.* Status of atmospheric correction using a MODTRAN4-based algorithm. *Proc. SPIE* **2000**, *4049*, 199–207.
22. Hatfield, J.L.; Prueger, J.H. Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management Practices. *Remote Sens.* **2010**, *2*, 562–578.
23. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87.
24. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; de Colstoun, E.B.; McMurtrey, J.E. Estimating corn leaf Chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239.
25. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. *Monitoring Vegetation Systems in the Great Plains with ERTS*. In *NASA Sp-351 I*; NASA: Washington, DC, USA, 12/1973; pp. 309–317.
26. McNairn, H.; Protz, R. Mapping corn residue cover on agricultural fields in Oxford County, Ontario, using thematic mapper. *Can. J. Remote Sens.* **1993**, *19*, 152–159.
27. Qi, J.; Marsett, R.C.; Heilman, P.; Biedenbender, S.; Moran, M.S.; Goodrich, D.C. RANGES improves satellite-based information and land cover assessments in southwest United States. *Eos Trans. Am. Geophys. Union* **2002**, *83*, 601–606.
28. Van Deventer, A.P.; Ward, A.P.; Gowda, P.H.; Lyon, J.G. Using Thematic Mapper data to identify contrasting soil plains to tillage practices. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 87–93.
29. Daughtry, C.S.T.; Hunt, E.R.; Doraiswamy, P.C.; McMurtrey, J.E. remote sensing the spatial distribution of crop residues. *Agron. J.* **2005**, *97*, 864–871.
30. Serbin, G.; Hunt, E.R., Jr.; Daughtry, C.S.T.; McCarty, G.W.; Doraiswamy, P.C. An improved ASTER index for remote sensing of crop residue. *Remote Sens.* **2009**, *1*, 971–991.
31. Trimble Germany GmbH. *eCognition Developer 8.7.2. Reference book*; Trimble Germany GmbH: Munich, Germany, 2013.
32. Ortiz, A.; Oliver, G. On the use of the overlapping area matrix for image segmentation evaluation: A survey and new performance measures. *Pattern Recognit. Lett.* **2006**, *27*, 1916–1926.
33. Dorigo, W.; Lucieer, A.; Podobnikar, T.; Čarni, A. Mapping invasive *Fallopia japonica* by combined spectral, spatial, and temporal analysis of digital orthophotos. *Int. J. Appl. Earth Obs. Geoinformation* **2012**, *19*, 185–195.
34. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 610–621.

35. Baraldi, A.; Parmiggiani, F. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 293–304.
36. Mather, P.M.; Koch, M. *Computer Processing of Remotely-Sensed Images: An Introduction*, 4th ed.; Wiley-Blackwell: Chichester, UK, 2011.
37. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Morgan Kaufmann Publishers: San Francisco, CA, USA, 2005.
38. Friedman, J.; Hastie, T.; Tibshirani, R. Additive logistic regression: A statistical view of boosting. *Ann. Stat.* **2000**, *28*, 337–407.
39. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297.
40. Quinlan, J.R. *C4.5: Programs for Machine Learning*; Morgan Kaufmann: San Mateo, CA, USA, 1993.
41. Gutiérrez, P.A.; Hervás-Martínez, C.; Martínez-Estudillo, F.J.; Carbonero, M. A two-stage evolutionary algorithm based on sensitivity and accuracy for multi-class problems. *Inf. Sci.* **2012**, *197*, 20–37.
42. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409.
43. Elnaggar, A.A.; Noller, J.S. Application of remote-sensing data and decision-tree analysis to mapping salt-affected soils over large areas. *Remote Sens.* **2009**, *2*, 151–165.
44. Hollander, M.; Wolfe, D.A.; Chicken, E. *Nonparametric Statistical Methods*, 3rd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014.
45. Steele, B.M. Combining multiple classifiers: An application using spatial and remotely sensed information for land cover type mapping. *Remote Sens. Environ.* **2000**, *74*, 545–556.
46. Conrad, C.; Fritsch, S.; Zeidler, J.; Rucker, G.; Dech, S. Per-field irrigated crop classification in arid central Asia using SPOT and ASTER data. *Remote Sens.* **2010**, *2*, 1035–1056.