Article

Crop Mapping Using Random Forest and Particle Swarm Optimization based on Multi-Temporal Sentinel-2

Elahe Akbari [1], Ali Darvishi Boloorani [1,†], Najmeh Neysani Samany [1], Saeid Hamzeh [1], Saeid Soufizadeh [2] and Stefano Pignatti [3,*,†]

[1] Department of Remote Sensing and GIS, Faculty of Geography, University of Tehran, Tehran, Iran; akbari.elahe@ut.ac.ir (E.A.); ali.darvishi@ut.ac.ir (A.D.B.); nneysani@ut.ac.ir (N.N.S.); saeid.hamzeh@ut.ac.ir (S.H.)

[2] Department of Agroecology, Environmental Sciences Research Institute, Shahid Beheshti University, G.C., Tehran, Iran; s_soufizadeh@sbu.ac.ir

[3] Institute of Methodologies for Environmental Analysis, National Research Council (CNR IMAA), C.da S.Loja snc, 85050 Tito (Potenza), Italy

*   Correspondence: pignatti@imaa.cnr.it or stefano.pignatti@cnr.it

†   These authors contributed equally to this work.

check for updates

Abstract: Timely and accurate information on crop mapping and monitoring is necessary for agricultural resources management. Accordingly, the applicability of the proposed classification-feature selection ensemble procedure with different feature sets for crop mapping is investigated. Here, we produced various feature sets including spectral bands, spectral indices, variation of spectral index, texture, and combinations of features to map different types of crops. By using various feature sets and the random forest (RF) classifier, the crop maps were created. In aiming to determine the most relevant and distinctive features, the particle swarm optimization (PSO) and RF-variable importance measure feature selection methods were examined. The classification-feature selection ensemble procedure was adapted to combine the outputs of different feature sets from the better feature selection method using majority votes. Multi-temporal Sentinel-2 data has been used in Ghale-Nou county of Tehran, Iran. The performance of RF was efficient in crop mapping especially by spectral bands and texture in combination with other feature sets. Our results showed that the PSO-based feature selection leads to a more accurate classification than the RF-variable importance measure, in almost all feature sets for all crop types. The RF classifier-PSO ensemble procedure for crop mapping outperformed the RF classifier in each feature set with regard to the class-wise and overall accuracies (OA) (of about 2.7–7.4% increases in OA and 0.48–3.68% (silage maize), 0–1.61% (rice), 2.82–15.43% (alfalfa), and 10.96–41.13% (vegetables) improvement in F-scores for all feature sets). The proposed method could mainly be useful to differentiate between heterogeneous crop fields (e.g., vegetables in this study) due to their more obtained omission/commission errors reduction.

Keywords: crop mapping; feature selection; particle swarm optimization; random forest; multi-temporal Sentinel-2 image

1. Introduction

To ensure world food security and agricultural resources management and conservation, it is of utmost importance that decision makers [1,2] are able to attain reliable and timely information on agricultural production and cropland mapping with sufficient accuracies [3–5]. Furthermore, up-to-date information on agricultural land use is crucial for crop planning and management [6] as well

as retrieving site-specific agronomic information [7], e.g., biomass and yield estimation, crop phenology, plants stress monitoring, and soil productivity assessment [8].

Owing to the fact that agricultural lands are strongly affected by spatial and temporal changes within and between each vegetation season, crop mapping is a challenging task [9]. Due to the synoptic and repetitive nature of satellite observations, they could very well be used to provide time series data and spatially contiguous information on crop growth stages at regional to global scales [4,5]. Therefore, remote sensing data have been widely applied to crop mapping (e.g., [7,10–12]).

The primary challenge in the use of satellite images for crop type mapping is to apply a specific procedure designed to operate with a robust classifier and using suitable input data [9]. In most studies, RF has been used for land cover and crop mapping and their potential to tackle the classification problems are confirmed, but achieving higher accuracy in crop discrimination is still a challenging issue. To train supervised classifiers, various features (also named variables and attributes), which are commonly extracted to describe the functional and nonfunctional characteristics of an image, are used as input to the classification process that is highly influential on the accuracy of final maps and computational time. Determining an optimal feature subset by the smaller size among all available features in order to maintain accuracy, while reducing computational costs is a challenging issue in the literature. Furthermore, the challenge of adapting a multi-temporal data for crop discrimination has been demonstrated by researchers such as [13–16]. However, multi-temporal data entails a great number of features for each image acquisition date [14,16,17], and applying spectral and texture features to the original bands of images may increase the dimensionality of the data, where it detrimentally affects the accuracy of the classifier due to the Hughes effect [18]. Therefore, selecting the best subset of features is challenging. The efficiency of optical data for crop mapping and crop parameter retrieval is acknowledged by [1,15,16,19–26]. These studies have mostly assessed the effects of incorporating different features and thus different input data for improving classification accuracy. Some studies such as [1,16,20] applied feature selection methods to improve the classification accuracy and mitigate the effect of the "curse of dimensionality". The curse of dimensionality occurs when the number of features is large or highly correlated with one another, and may decrease the classification accuracy. Therefore, it is of ultimate significance to reduce data redundancy to discriminate different crops, particularly in complex land cover cases [16].

From the literature, crop mapping has been performed by different feature sets without considering the problem of the curse of dimensionality and also finding the highly informative features in multi-temporal data in most studies. Meanwhile, the studies on crop mapping using feature selection have primarily focused on the filter methods such as the RF-variable importance measure. Nonetheless, Gilbertson and van Niekerk (2017) [16] indicated that filters do not tend to perform well in downstream scenarios due to the lack of interaction with the classifier. Moreover, based on the literature, the use of a PSO method to perform feature selection is a novel contribution to the crop mapping field. However, compared with other heuristic and meta-heuristic algorithms, PSO, as a powerful feature selection procedure, has been widely used in other issues. PSO is preferred due to its simplicity, flexibility, efficiency, ability to avoid local minima [27–29], fewer parameters requirement, lesser computation requirement, and it also can converge more quickly for solving feature selection problems [30].

The Sentinel-2 satellite provides global coverage every five days in 13 spectral bands from 10 to 60 m spatial resolutions [31]. This sensor has shown a high potential to acquire high temporal and spectral resolution images, which is suitable for crop mapping and investigation [23]. Up to now, some studies have utilized Sentinel-2 data to crop types mapping, e.g., [23] used a simulated version of Sentinel-2 data (based on Landsat and SPOT images). This simulated data contains certain modifications to the original Sentinel-2 data particularly the absence of red edge bands. Pelletier et al. (2016) [23] suggested analyzing the contribution of spatial features in addition to spectral features and assessing the performance of the original Sentinel-2 data for crop mapping. Moreover, studies such as Griffiths et al. (2019) [26] and Vuolo et al. (2018) [25] used Sentinel-2 and harmonized Landsat-Sentinel datasets

to map crop types with solely spectral bands. They also concentrated on a deeper understanding of crop phenology periods to classification tasks.

The main goal of this research is to develop a classification-feature selection ensemble procedure with different feature sets for the classification of agriculture land covers on a per pixel basis. Hence, spectral bands, spectral indices, texture features, variation of spectral index (VSI), and their combinations were used for multi-temporal classification. Sentinel-2 A/B multi-temporal data were employed to apprehend the extent to which the data is capable of discriminating between different crop types. Selecting a robust classifier capable of managing a large number of variables or reducing the number of input variables in order to solely select the most informative variables can help solve the curse of dimensionality and the Hughes effect [19] and may maintain or increase the accuracy of classification. For this purpose, we perform and assess the RF classifier, which can handle large databases in mapping land cover. Furthermore, dimension reduction and efficient identification of the most informative features for crop mapping were performed via the RF-variable importance measure and particle swarm optimization (PSO) feature selection methods. Hence, we evaluate the merit of reducing dimensionality in crop discrimination with multi-temporal Sentinel-2 images. Finally, the optimal output of these classifier-feature selection methods in each feature set were combined by majority voting.

2. Materials and Methods

2.1. Study Area

The study area is Ghale-Nou county of Tehran, Iran (Figure 1). The area is flat, occupies about 21,400 ha, and is characterized by agricultural fields especially the silage maize. This area is one of the major areas of maize cultivation in Iran. The main cultivar of silage maize is the Zea maize 704 and 706 single cross based on the field research. In addition to the maize, alfalfa, vegetables, and rice are cultivated in this area. The main cultivar of alfalfa and rice are local such as Hamedani and Hashemi, respectively. Vegetables are cabbage, basil, carrot, celery, radish, eggplant, spinach, green bean, tomato, etc. The minimum and maximum temperature and average annual precipitation are –4 and 42 °C and 200 mm, respectively [32].
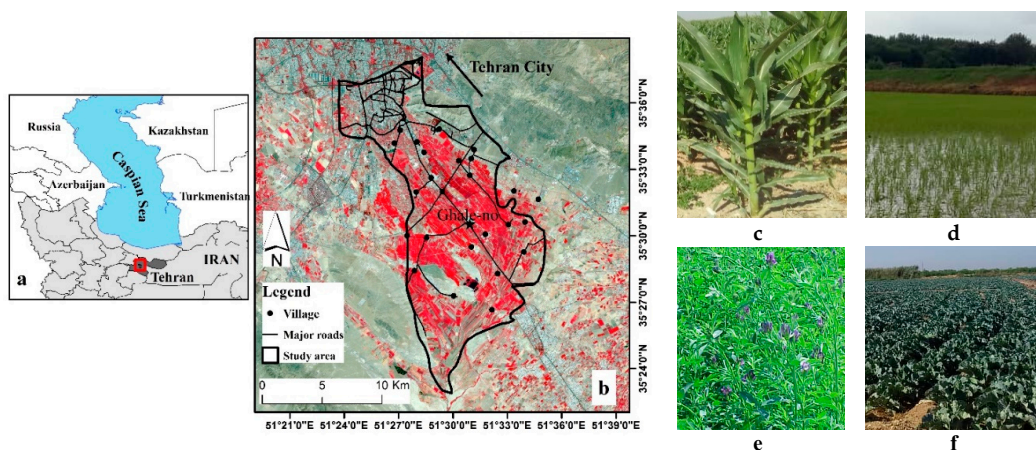

Figure 1. (**a**) Study area in the Tehran province. (**b**) Location of the study area in proportion to the surroundings of Tehran; Ghale-Nou county. Image in the background (band 8 (near infra-red (NIR)), band 4 (red), band 3 (green) as RGB (Red Green Blue) composition) are from Sentinel-2A at 21 August 2017, in almost near the flowering stage of certain crops (e.g., silage maize) of the region. (**c**) Silage maize. (**d**) Rice. (**e**) Alfalfa. (**f**) Vegetables field photos.

2.2. Field Data

Ground truth data were collected during crop growth periods in 2017 (from the end of July to the end of September). Figure 2 shows the satellite image acquisition dates and the growth cycle of

different crop types during field data collection in the study area. In summer, 155 fields at different locations in the entire study area were visited and these fields based on a stratified random sampling method were selected. For each field the crop type and the geographic coordinate were recorded. The crop type classes, number, and percent of the surveyed fields per training and validation samples are displayed in Table 1. These classes were selected based on their homogeneity/heterogeneity in the study area. Similarly, other land cover classes were surveyed including canebrake (23), bare soil (43), and water (15), these values referred to the number of surveyed areas. In situ data were randomly split into 70% and 30% for the model calibration and validation, respectively.
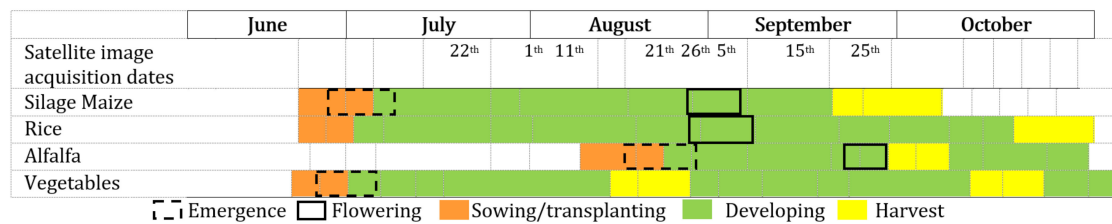
Figure 2. Satellite image acquisition date and phenology of different crop types in the study area. Rice has been transplanted and is not at the emergence stage. The time between sowing and harvest stages for vegetables and alfalfa is 40–50 days and again they grow and harvest.

Table 1. Number and area of surveyed fields in each crop type class per training and validation samples.

Crop Type	Number of Surveyed Fields	Area of Surveyed Fields (ha)	Area of Crop Type (ha)	Area of Surveyed Fields (% From Crop Type)	Area of Surveyed Fields (% From the Whole Study Area)
Rice	15	19.4	450.84	4.21	0.09
Silage Maize	64	338.6	5441.32	6.21	1.58
Alfalfa	33	54.1	695.72	7.76	0.25
Vegetables	43	69.3	914.40	7.55	0.32

2.3. Remote Sensing Imagery

Based on the phenological information of crops, eight Sentinel-2 A/B images were obtained from 22 July 2017 to 25 September 2017 (Figure 2). The image of 22 July was selected almost three weeks after the sowing stage of crops due to the emergence of vegetation cover in the satellite image and hence, preventing data redundancy. The level-1C product of Sentinel-2 images with cartographic geometry (UTM/WGS84 projection) [33] has been corrected for atmospheric and radiometric errors in order to produce the level-2A image through the Sentinel Application Platform (SNAP) software using a plugin called Sen2Cor. The Sen2Cor plugin applies atmospheric, terrain, and cirrus correction of the top of the atmosphere (TOA) level-1C input data and produces an ortho-image bottom of the atmosphere (BOA) corrected reflectance product as output [34]. In this study, we used the bands with a spatial resolution of 10 and 20 m: Band 2 (blue, 10 m), band 3 (green, 10 m), band 4 (red, 10 m), band 5–7 (red edge, 20 m), band 8 (NIR, 10 m), band 8a (NIR narrow, 20 m), and band 11–12 (SWIR, 20 m). Bands with a 10 m spatial resolution of Sentinel-2 A/B images have been resampled at 20 m by using a spatial cubic resampling method in the Environment for Visualizing Images (ENVI) software.

2.4. Feature Sets

In order to explore the ability of optical images of Sentinel-2A/B to distinguish land cover classes and to select the most relevant features for enhancing classification accuracy, multi-temporal features were produced based on the phenology of crops. Therefore, different feature sets and combinations of optical features as input data for crop mapping were analyzed (Tables 2 and 3). We did not repeat the time-consuming classification procedure for combinations of all feature sets and used another strategy that will be discussed in Section 2.5. For this study, besides spectral bands, texture features, and VSI,

six spectral indices were chosen from the literature in consideration of their classification performance or novelty. The ENVI and SNAP software were used for extracting the features.

A quantitative delineation of the visual characteristics of an image has been provided by texture features, such as smoothness, symmetry, roughness, etc. [35]. Texture features are diverse, ranging from simple statistics, such as the mean and standard deviation to statistics derived from the grey level co-occurrence matrix (GLCM) [19]. Texture features have been widely used as important indicators for improving classification accuracy, e.g., [6,36]. In this study, the gray level co-occurrence matrix (GLCM) introduced by Haralick et al. (1973) [37] as a general scheme for extracting textural information was used.

Table 2. Different feature sets and combinations of optical features for crop mapping.

No.	Feature Sets	Features	Number of Features
1	Spectral Band (SB)	Multi-temporal spectral bands	80 (8 × 10 bands)
2	Spectral Index (SI)	Multi-temporal NDVI [38], EVI [39], ARVI [40], RVI [41], NDWI [42], IRECI [43]	48 (8 × 6 index)
3	Texture (T)	Multi-temporal local moments (mean, variance), Haralik (CON, DIS, HOM, ENE, ENT, COR) [37] in each spectral band	640 (8 × 10 bands × 8 textures)
4	Variation of Spectral Index (VSI)	Multi-temporal VSI of NDVI, EVI, ARVI, RVI, NDWI, IRECI	48 (8 × 6 index)
5	Combination of FS 1, 3 (SB-T)		720
6	Combination of FS 3, 4 (T-VSI)		688
7	Combination of FS 1, 2, and 3 (SB-SI-T)		768
8	Combination of FS 1, 2, and 4 (SB-SI-VSI)		176

Abbreviations: Feature sets (FS), normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), atmospherically resistant vegetation index (ARVI), ratio vegetation index (RVI), normalized difference water index (NDWI), inverted red-edge chlorophyll index (IRECI), energy (ENE), contrast (CON), dissimilarity (DIS), homogeneity (HOM), entropy (ENT), and correlation (COR).

Table 3. Formula of features used in this study.

	Acronyms
$\text{NDVI} = \frac{\text{NIR}-\text{R}}{\text{NIR}+\text{R}}$ [38]	Red band (R), near infra-red (NIR)
$\text{EVI2} = 2.5 \times (\frac{\text{NIR} - \text{R}}{\text{NIR} + 2.4 \times \text{R} + 1})$ [39]	
$\text{ARVI} = \frac{NIR-((2 \times R)-Blue)}{NIR+((2 \times R)-Blue)}$ [40]	
$\text{RVI} = \frac{\text{NIR}}{\text{R}}$ [41]	
$\text{NDWI} = \frac{\text{Gr}-\text{NIR}}{\text{Gr}+\text{NIR}}$ [42]	Green band (Gr)
$\text{IRECI} = \frac{RE3-R}{RE1/RE2}$ [43]	Red edge 1 (RE1), red edge 2 (RE2), red edge 3 (RE3) are band 5, band 6, and band 7 of Sentinel-2, respectively.
$\text{ENE} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p^2\,(i,j)$ [37]	Gray levels of pixels in row (i) and column (j) in each matrix ($P\,(i,j)$)
$\text{CON} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p\,(i,j)\,(i-j)^2$ [37]	
$\text{DIS} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p\,(i,j)\left|i-j\right|$ [37]	
$\text{HOM} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{p\,(i,j)}{1+(i-j)^2}$ [37]	
$\text{ENT} = -\sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p\,(i,j) \ln p\,(i,\,j)$ [37]	
$\text{COR} = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p\,(i,j)\, \frac{(i-\mu x)(j-\mu y)}{\sigma x\,\sigma y}$ [37]	μx, μy, σx, and σy are the means and standard deviations for the row- and column-of matrix composed of $p(i,j)$

Abbreviations: Normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), atmospherically resistant vegetation index (ARVI), ratio vegetation index (RVI), normalized difference water index (NDWI), inverted red-edge chlorophyll index (IRECI), energy (ENE), contrast (CON), dissimilarity (DIS), homogeneity (HOM), entropy (ENT), and correlation (COR).
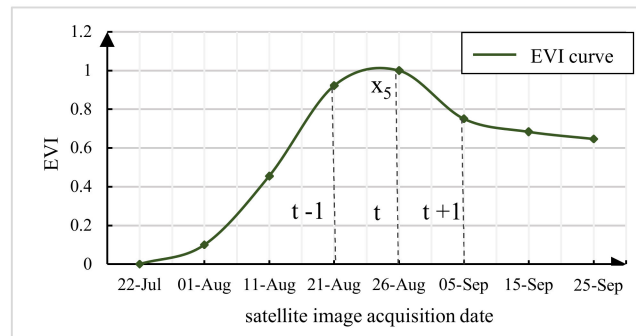
The VSI assessed the variation of spectral indices around each date, therefore, for all spectral indices, the variation factor has been produced and among all, just the enhanced vegetation index (EVI) was discussed for the sake of brevity. Figure 3 schematically illustrates the EVI feature changes

within the growing cycle of silage maize. Based on this figure, the VSI in x_5 at t for EVI was calculated using Equation (1).

$$VSI\_EVI_t = EVI_{t+1} - EVI_{t-1} \qquad (1)$$

where t is a specific date, the VSI were derived from the EVI values preceding ($t - 1$) and following ($t + 1$) the specific date.

Figure 3. Schematic illustration of the variation of spectral index (VSI) (for example VSI_EVI) calculation. Each point represents the silage maize EVI calculated based on the average of all pixel values in the silage maize class for Sentinel-2 multi-temporal data. t is a specific date at x_5 that intercepts the fifth value of EVI in this figure, $t - 1$ and $t + 1$ are the values preceding and following the specific date.
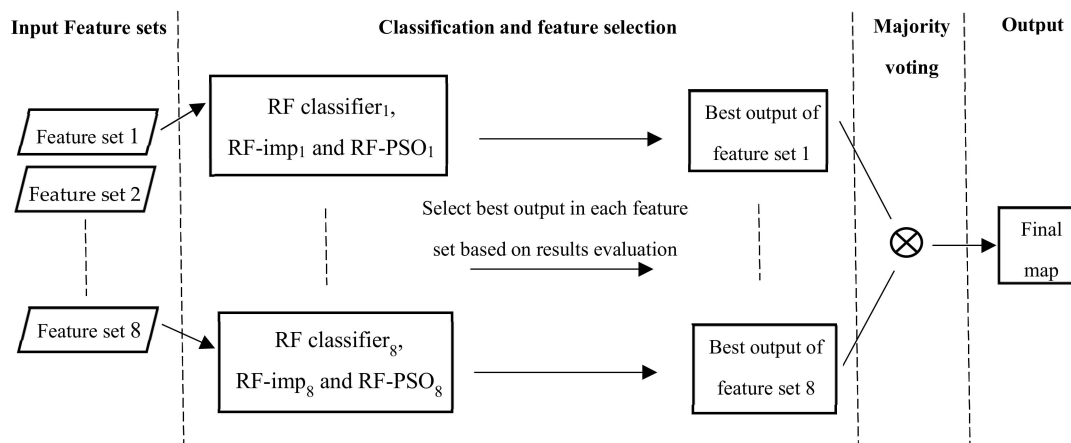
2.5. Classification-Feature Selection Ensemble Procedure

Several studies on remote sensing classification are indicating the increase of classification accuracy as the result of combining the outputs of independent classifiers. Two strategies have been adapted in this regard, (1) using a combination of different classifiers and (2) combining different outputs of one classifier [44]. This study proceeds to employ the latter strategy to compose an ensemble of independent classification results of optimal algorithm, which considers different features as input. In our work, (*i*) the RF classifier was performed and assessed in each feature set (RF classifier$_{1-8}$), (*ii*) two feature selection algorithms (i.e., RF-variable importance measure and PSO) were assessed with respect to the classification accuracy, (*iii*) the optimal algorithm/best output of RF, RF-variable importance measure, and RF-PSO methods in each feature set were selected based on the results evaluation, which were then, (*iv*) employed to generate a final map using majority voting (Figure 4). Majority voting is considered as one of the most popular combination approaches [45] and has been acknowledged in countless studies given its efficiency over other complex strategies. This method assumes a unique error for each classifier that is independent of the other classifiers' error [44]. Meanwhile, a combination of all feature sets as the input data to classifier leads to the computational complexity and time-consuming procedure. Our proposed method can solve this problem, by increasing the classification accuracy via combining different outputs of one classifier.

2.5.1. Image Classification

In this study, the random forests (RF) classifier was performed to classify the Sentinel-2 imagery for mapping crop types. RF showed priority than other classifiers because it is fast as well as ease of parameterization and robustness [20,23,25]. The RF method is an adaptable ensemble learning method, which combines K binary classification and regression trees (CART), wherein each tree is generated by applying an individual learning algorithm into subsets of the input variable sets that split using the Gini index as one of the attribute value tests [46]. The RF classification applies a random subset of input predictive variables rather than using the optimal ones in the division of every node of decision tree structure, consequently, the generalization error will be decreased [19]. By limiting the number of variables used for a split (i.e., subset), the computational complexity of the algorithm, and also the correlation between trees will be reduced [23]. Rodriguez-Galiano et al. (2012b) [20] and Pelletier et al.

(2016) [23] have studied the impact of different model parameterizations on RF results. The results indicated a limited influence of the model setting on classification accuracy. Accordingly, we used the default values for parameters and set the number of trees to 500. The categorical algorithm was set to the principle component analysis (PCA) to find the best split on a categorical predictor.

Figure 4. Flowchart of the classification-feature selection ensemble procedure. RF-imp and RF-PSO are feature selection methods, random forest (RF), RF-variable importance measure (RF-imp), RF-particle swarm optimization (RF-PSO), and the numbers 1–8 refer to the number of feature sets.

2.5.2. Feature Selection

Despite the apparent functionality of having more information in classification tasks, increasing the number of input features may increase the computational time and the "curse of dimensionality". Moreover, the features extracted from spectral bands in remote sensing may be highly correlated or redundant. Thus, it is logical to select the most relevant (most informative) features to reduce the effect of the curse of dimensionality [19]. Feature selection entails selecting a subset of features from the original dataset to create robust learning models and reduce computation costs and data redundancy [16]. In this study, the efficiency of two feature selection methods was evaluated, namely the RF-variable importance measure and PSO.

The RF classifier estimates the importance of variables and uses the informative ones to increase the accuracy of classification. Expectedly, dimension reduction will not cost the classification accuracy reduction in RF. In the RF-variable importance measure method, in order to assess the importance of a given variable (feature), RF measures the decrease in the accuracy by means of an out-of-bag (OOB) error estimation and the decrease in Gini index, while overlooking the given variable and keeping the rest constant. To determine the importance of each variable by OOB, the average of differences between the accuracies obtained through the modified OOB subset and the original ones are calculated. The decrease in the Gini impurity criterion for each variable is measured over all trees in the forest by the Gini importance index [19,20]. In this study, the recursive feature elimination method was not considered due to its computational complexity for a very high number of features.

The PSO algorithm is a simulation of social behavior that moves in a bird flock. This algorithm explores the entire search space, with each particle swarm seeking to exploit the optimum solution within the search space. In this study, the binary form of PSO was used to select features for classification. The search space for PSO implementation was considered as a feature space and the initial population is assigned randomly. The position of each particle is represented by a binary vector with the length equal to the number of the original features, which if its cells are set to 1 and 0, it means that the corresponding feature is selected or not, respectively. Moreover, the number of features that should be selected by the algorithm is defined. For each particle, a velocity vector as the same dimension of

particle position is generated by using a random float number generator. The changes and updates in particle velocity and position in each iteration can be explained by Equations (2) and (3) [27,47].

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_1 \left(Pbest_i(t) - x_{id}(t)\right) + c_2 r_2 \left(Gbest(t) - x_{id}(t)\right). \tag{2}$$

$$s(v_{id}(t+1)) = \frac{1}{1+e^{-vid}} \text{ if } r_3(t) < s(v_{id}(t+1)) \text{ then } x_{id}(t+1) = 1 \text{ else } x_{id}(t+1) = 0 \tag{3}$$

where c_1 and c_2 are acceleration constants (set as $c_1 = c_2 = 1.496$ in this study); r_1 and r_2 are random values between 0 and 1; w is the inertia weight (set as $w = 1$ in this study); $x_{id}(t)$ is the position of the ith particle within the d-dimensional search space representing a candidate solution of the problem at iteration t; $v_{id}(t)$ is the corresponding velocity of the ith particle; $Pbest_i(t) \in$ Rd is the best position during its past trajectory; $Gbest(t) \in$ Rd is the best global position over all trajectories by all particles; '$s(\cdot)$' is the sigmoid transformation function; and $r_3(t)$ is the uniform random number $\in [0,1]$.

In this study, in order to increase the exploration characteristic and find a new solution in the search space, we also considered the inertia weight damping ratio = 0.9997 for calculating w in each iteration ($w = w *$ damping ratio), where the value of w will decrease with the increase of the number of iteration. Furthermore, in our study, the quality of measure, or rather the cost function was selected in compliance with the corresponding confusion matrix and classification accuracy. Accordingly, the cost function was calculated by implementing the classifier in each iteration and using the K-fold cross validation method for training and validation subsets. The process of finding the global best values (optimal features) is iterative until the end criterion is met and therefore, these optimal features are the subset with the highest classification accuracy.

In our study, we sorted the obtained importance of features from the RF-variable importance measure method in descending order and the first half was kept. In order to compare and evaluate the performance of two feature selection methods, half of the features in each feature set was also selected for PSO. The image classification and RF-variable importance measure feature selection method have been implemented and coded by the random forest package and 'OOBPredictorImportance' property in MATLAB software. Meanwhile, the RF-PSO feature selection ensemble procedure has been developed in the MATLAB software.

2.5.3. Results Evaluation

From the database, 30% of in situ land cover data were used to assess the accuracy of the results. The overall accuracy (OA), Kappa coefficient (K), and F-score were calculated from the confusion matrix in MATLAB and Microsoft Excel software in order to evaluate the accuracy of the produced crop type maps. The F-score (Equation (4)) [36] is a per-class measure, which corresponds to the harmonic mean of the user's and the producer's accuracies. It reaches its highest and lowest accuracies at 1 and 0, respectively. The OA (%) and Kappa coefficient (Equations (5) and (6)) [48] are the probability that a pixel is classified correctly and measures the actual agreement between reference data and the classifier used versus the chance of a random classifier, respectively.

$$\text{F} - \text{Score} = 2 \times \frac{producer's \ accuracy \ \times \ user's \ accuracy}{producer's \ accuracy \ + \ user's \ accuracy} \tag{4}$$

$$\text{OA} = \frac{\sum_{i=1}^{r} Xii}{N} \tag{5}$$

$$\text{Kappa} = \frac{N \sum_{i=1}^{r} Xii - \sum_{i=1}^{r} (X i + \times X + i)}{N^2 - \sum_{i=1}^{r} (X i + \times X + i)} \tag{6}$$

where r (number of rows in the confusion matrix), Xii (the number of observations in row i and column i (on the major diagonal), $Xi+$ (total of observations in row i), $X+i$ (total of observations in column i), and N (total number of observations included in the matrix).

3. Results

3.1. RF Results

Here, the capacity of multi-temporal Sentinel-2 images was assessed to crops classification and mapping with different feature sets by seven land cover classes using RF. The OA, Kappa coefficient, and F-score of the crop maps for seven classes of maize, rice, alfalfa, vegetables, bare soil, water, and canebrake are shown in Table 4.

Table 4. Overall accuracy (OA, %), Kappa coefficient, user's accuracy (%), producer's accuracy (%), and F-score (%) of RF by different feature sets input data. Total mean and SD of F-score in each class were calculated for all feature sets.

Classes	SB			SI			T			VSI		
	UA	PA	F	UA	PA	F	UA	PA	F	UA	PA	F
Maize	90.96	100.0	95.27	91.57	98.70	94.99	87.65	97.32	92.23	91.57	96.82	94.12
Rice	100.0	99.19	99.59	100.0	98.40	99.19	100.0	98.40	99.19	100.0	96.85	98.39
Alfalfa	83.06	91.09	86.89	76.80	92.46	83.9	94.90	86.84	90.68	67.05	96.66	79.18
Vegetables	83.08	48.65	61.36	95.39	34.25	50.41	81.54	70.67	75.71	90.77	28.50	43.38
Bare Soil	98.06	94.24	96.11	98.70	96.51	97.59	96.43	96.28	96.35	99.35	95.63	97.46
Water	100.0	83.33	90.91	98.33	88.06	92.91	100.0	73.17	84.51	98.33	93.65	95.93
Canebrake	77.27	85.00	80.95	72.73	95.52	82.58	67.05	88.06	76.13	81.25	93.46	86.93
OA		90.8			89.57			91.41			88.14	
Kappa		0.882			0.867			0.889			0.85	

Classes	SB-T			T-VSI			SB-SI-T			SB-SI-VSI			Total Mean of F	Total SD of F
	UA	PA	F	UA	PA	F	UA	PA	F	UA	PA	F		
Maize	87.65	97.00	92.08	91.87	98.71	95.16	87.95	100.0	93.58	91.54	99.34	95.28	94.09	1.25
Rice	100.0	99.19	99.59	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.49	0.52
Alfalfa	92.11	91.48	91.79	94.90	88.34	91.49	92.81	85.84	89.18	83.76	90.48	86.98	87.52	4.03
Vegetables	87.69	59.38	70.81	84.62	70.51	76.92	87.69	63.33	73.55	95.39	46.27	62.31	64.31	11.51
Bare Soil	96.43	96.28	96.35	96.76	96.6	96.68	96.76	96.60	96.68	98.87	95.76	97.28	96.81	0.53
Water	100.0	73.17	84.51	100.0	75.00	85.71	100.0	75.00	85.71	100.0	89.55	94.48	89.34	4.45
Canebrake	77.84	91.33	84.05	71.26	94.66	81.31	73.30	95.56	82.96	75.00	95.65	84.08	82.40	3.14
OA		92.02			92.85			91.91			91.57				
Kappa		0.897			0.908			0.896			0.892				

Abbreviations: Spectral bands (SB), spectral index (SI), texture (T), variation of spectral index (VSI), user's accuracy (UA, %), producer's accuracy (PA, %), F-score (F, %), overall accuracy (OA, %), Kappa coefficient (Kappa), and standard deviation (SD).

Regarding the OA and Kappa coefficient, T-VSI (the highest OA and Kappa, 92.85% and 0.908%), followed by the SB-T, SB-SI-T, and SB-SI-VSI feature sets (91.57–92.02% in OA and 0.892–0.897% in Kappa) outperformed other feature sets and obtained better results in land cover classification. While using only the indices (VSI and SI feature sets) produced a relatively good result (88.14–89.57% in OA) and obtained the lower results in the classification process. The results of combining different feature sets were higher than when the feature set was used alone (0.61%, 1.22% between SB, T, and SB-T, 1.44%, 4.71% between T, VSI, and T-VSI, 1.11%, 2.34%, 0.5% between SB, SI, T, and SB-SI-T, 0.77%, 2%, 3.43% between SB, SI, VSI, and SB-SI-VSI feature sets in OA, respectively).

Meanwhile, the mean F-score values for all feature sets supported the efficiency of RF in land cover mapping. Rice, maize, and bare soil with the least F-score standard deviation for different feature sets (i.e., SD = 0.52–1.25) showed the least sensitivity to input data and was followed by canebrake and alfalfa (i.e., SD = 3.14 and 4.03, respectively). Rice fields mapping, given their homogeneous quality, was conducted with a relevantly higher accuracy and showed lower levels of sensitivity to input data for classification. However, the higher standard deviation of the F-score for vegetables (i.e., SD = 11.51) appeared to be indicative of the higher reactivity of RF towards different input feature sets. In addition, the lower F-score values for vegetables fields (mean F-score values for all feature sets = 64.31%) may indicate their heterogeneity quality.

3.2. Feature Selection Impacts on Classification Accuracy

Despite the high accuracy of the RF model to discriminate different classes, similar, or even more preferable results can also be attained through smaller feature inputs. Results from feeding the features obtained from the RF-variable importance measure and PSO approach to the RF model indicated an increase of about 0.22% to 1.06% and 0.66% to 2.22% in the overall accuracy, respectively, for all feature sets as opposed to when no feature selection algorithm was applied (Table 5). Classification results showed a similar increasing trend in F-score similar to overall accuracy when feature selection have been used, particularly for the main four crops in the region (i.e., maize, rice, alfalfa, and vegetables). Nonetheless, the superiority of the RF method with feature selection compared to without feature selection is not so clear such as the OA values when looking at individual F-score values for all feature sets. Accordingly, the average of difference values for all feature sets in each class were obtained (Figure 5) following the calculation of difference of F-score results between the RF with feature selection and RF without feature selection. The results showed that the RF-PSO algorithm outperformed the RF-variable importance measure method to a great degree, whereas the latter method proved better than the RF without a feature selection approach, i.e., the positive values were obtained for all classes (0.2% to 5.58% and 0.54% to 2.97% in the difference of F-score results, respectively) (Figure 5). The overall accuracy followed a similar trend to the F-score, showed the superiority of the RF-PSO model, followed by the RF-variable importance measure. Furthermore, the impact of feature selection on improving the classification accuracy was significantly higher for alfalfa, vegetables, and canebrake (2.56%, 5.58%, and 2.88% improvements in F-score between RF-PSO and RF than 1.22%, 2.97%, and 1.25% improvements in F-score between the RF-variable importance measure and RF) in comparison with the other classes. These three classes were considered heterogeneous classes in the study area and as mentioned in Section 3.1 and Table 4, they showed a higher standard deviation compared to other classes. However, with feature selection, these three classes were more precisely classified.

Table 5. F-score (%), overall accuracy (OA, %), and Kappa coefficient of the impacts of feature selection on the classification accuracy in different feature sets. Bold values correspond to the highest F-scores, Kappa coefficient, and OA.

Classes	SB			SI			T			VSI		
	RF	RF-imp	RF-PSO	RF	RF-imp	RF-PSO	RF	RF-imp	RF-PSO	RF	RF-imp	RF-PSO
Maize	95.27	93.76	**95.76**	94.99	95.06	**95.65**	92.23	91.71	**92.38**	94.12	94.17	**94.29**
Rice	99.59	99.59	**100.0**	99.19	99.59	**100.0**	99.19	99.59	**99.59**	98.39	95.72	**98.39**
Alfalfa	86.89	89.02	**89.16**	83.9	83.57	**87.17**	90.68	91.44	**91.51**	79.18	**80.87**	80.65
Vegetables	61.36	60.19	**61.69**	50.41	50.62	**56.38**	75.71	80.58	**81.75**	43.38	44.69	**46.15**
Bare Soil	96.11	96.28	**97.05**	97.59	97.58	97.51	96.35	96.29	**96.52**	97.46	97.37	**97.92**
Water	90.91	**93.02**	93.02	92.91	91.47	91.47	84.51	85.71	**86.33**	95.93	94.49	**96.77**
Canebrake	80.95	**86.62**	85.81	82.58	**87.12**	86.62	76.13	76.97	**79.09**	86.93	87.73	**87.73**
OA	90.8	91.4	**92.12**	89.57	89.97	**91.30**	91.41	91.85	**92.30**	88.14	88.36	**88.91**
Kappa	0.882	0.89	**0.899**	0.867	0.872	**0.889**	0.889	0.895	**0.900**	0.85	0.853	**0.859**
Classes	SB-T			T-VSI			SB-SI-T			SB-SI-VSI		
	RF	RF-imp	RF-PSO	RF	RF-imp	RF-PSO	RF	RF-imp	RF-PSO	RF	RF-imp	RF-PSO
Maize	92.08	93.98	**94.06**	95.16	95.28	**95.45**	93.58	94.93	**94.94**	95.28	94.72	**95.43**
Rice	99.59	99.59	**99.59**	100.0	100.0	**100.0**	100.0	99.59	**100.0**	100.0	100.0	**100.0**
Alfalfa	91.79	92.12	**93.20**	91.49	92.20	**92.80**	89.18	91.11	**92.97**	86.98	89.55	**93.08**
Vegetables	70.81	77.27	**78.32**	76.92	78.52	**81.16**	73.55	**81.43**	78.62	62.31	64.94	**74.99**
Bare Soil	96.35	**96.78**	96.69	96.68	96.04	96.43	96.68	96.59	**96.93**	97.28	**98.00**	97.42
Water	84.51	**88.89**	86.96	85.71	84.51	84.51	85.71	85.11	**87.59**	94.48	**97.56**	90.23
Canebrake	84.05	80.25	**85.27**	81.31	84.04	**84.42**	82.96	81.85	**88.55**	84.08	84.44	**84.59**
OA	92.02	92.85	**93.46**	92.85	93.09	**93.51**	91.91	92.85	**93.97**	91.57	92.63	**93.79**
Kappa	0.897	0.908	**0.916**	0.908	0.911	**0.916**	0.896	0.908	**0.922**	0.892	0.905	**0.92**

Abbreviations: Spectral bands (SB), spectral index (SI), texture (T), variation of spectral index (VSI), random forest (RF), random forest-variable importance measure (RF-imp), random forest-particle swarm optimization (RF-PSO), overall accuracy (OA, %), and Kappa coefficient (Kappa).

The computational time of SB, SI, VSI, and SB-SI-VSI was assessed and results showed that there was no serious difference of the computational time between the two feature selection methods.

The performance of the RF-variable importance measure method was faster than the RF-PSO algorithm, approximately about 20–60 s (about a few to 20 min expending for the RF-PSO in all feature sets). Moreover, the training time for the RF classifier required significantly a lesser time (about a few minutes expending in all feature sets) than the RF-PSO and RF-variable importance measure method (using a computer with Intel (R) Core(TM) i7-2640M CPU processor at 2.80 GHz and 8 GB of RAM in MATLAB). The difference of computational time by RF-PSO and RF-variable importance measure methods were lesser, while the accuracy of classification improved with the RF-PSO model, therefore, this feature selection method was chosen. An information of the features selected by these two models was discussed in Section 3.3.
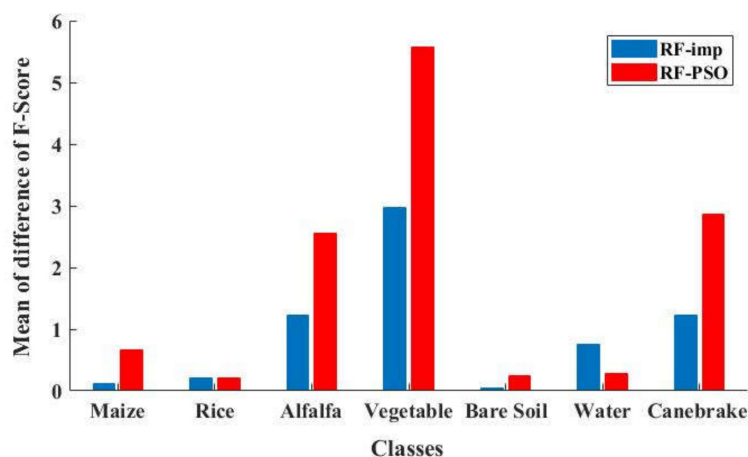

Figure 5. Mean of difference of F-score between RF-PSO and RF, RF-variable importance measure (RF-imp), and RF methods in each class for all feature sets.
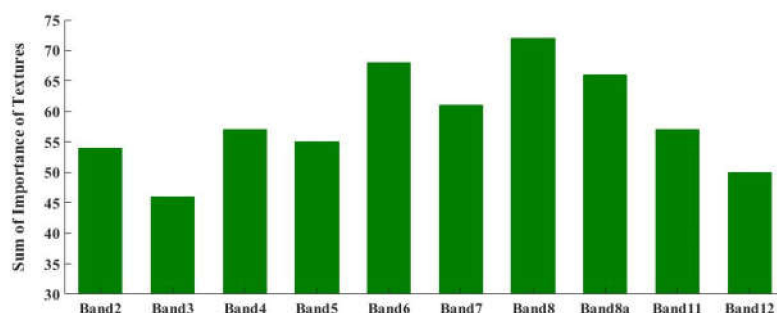
3.3. Importance of Features

We investigated the relative importance of each feature to classify crop types (Figure 6). For this purpose, the number of selected features in the feature sets in both PSO and RF-variable importance measure feature selection methods was summed and considered as the relative importance of each feature. Based on Figure 6, the red edge (band 7), SWIR (band 11, 12), and narrow NIR (band 8a) bands were selected as the most important spectral bands with the highest impact on the classes' separation. Moreover, among the spectral bands, the end region in the red edge band (band 7) of Sentinel-2 data (770–790 nm) was selected as the most significant and effective feature on classification. Visible bands (blue (band 2), green (band 3), and red (band 4) bands, respectively) had the least contribution to the crop classification.

NDWI and IRECI were selected as the most relevant features to separate different classes according to spectral indices and VSI, respectively (Figure 6). Whereas, the classification process has been gained the lowest impact from ARVI and EVI according to spectral indices and VSI, respectively.

Based on Figure 6, a similar behavior was observed for the more informative texture features in classifying the given classes and crop mapping for all bands, with the exception of band 6 (the middle red edge of Sentinel-2). Therefore, the "mean" feature for all bands as well as the "energy" feature for band 6 were selected as the most relevant features for classification, particularly for predicting the type of crops. Out of all bands, the "mean" feature in bands 5 (the first band of red edge), 11 (SWIR), 8 (NIR), and 8a (narrow NIR) had higher contributions to the classes separation, which in fact led to an increase in the intra-class separability. Consequently, after summing the entire selected texture features for each band, bands 8 (NIR), 6 (middle region in the red edge), 8a (narrow NIR), and 7 (end region in the red edge) had the highest contribution to the crop classification (Figure 7).
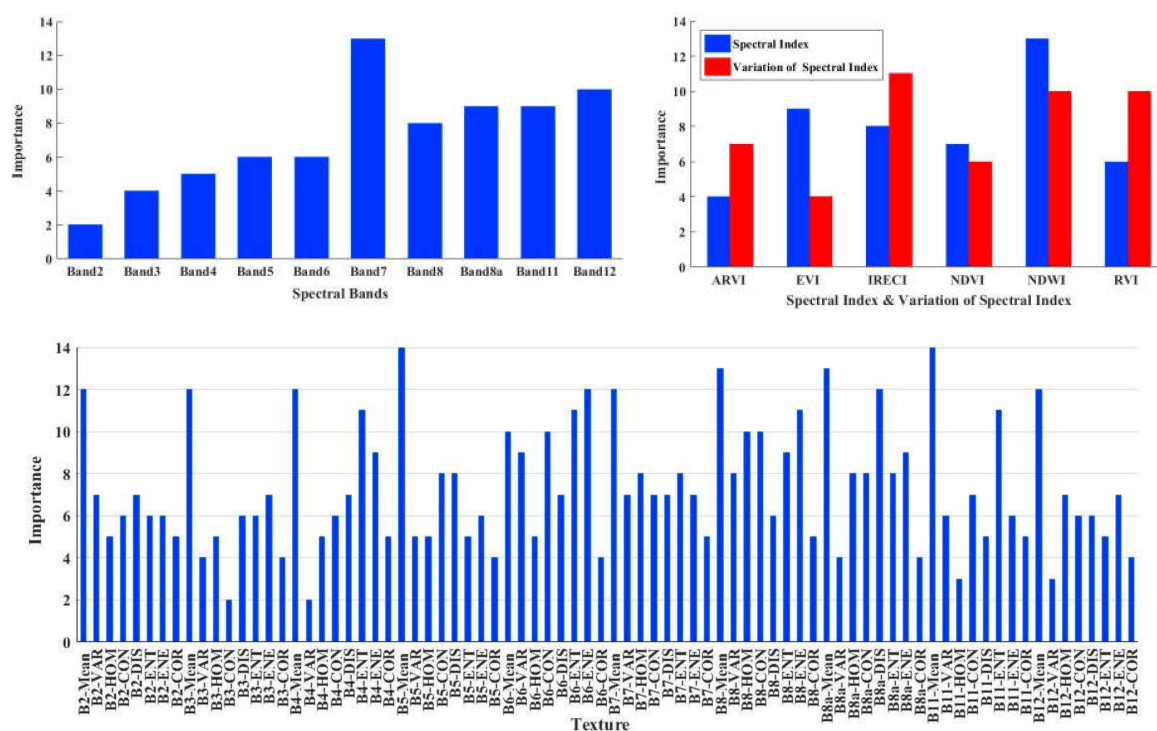
Figure 6. Importance of different features based on the summation of the number of features, which are selected by RF-variable importance and RF-PSO methods. B2–B12 are the Sentinal-2 bands from blue to SWIR, respectively (bands detail in [49]).

Figure 7. Importance of the texture feature set in each band according to the summation of all texture features selected per band.

3.4. Influence of Feature Sets on Crop Separability

The overall performance of the RF-PSO approach prevailed over the RF and RF-variable importance measure algorithms. Thus, in order to determine the contribution of each feature set on the separation of crops using the RF-PSO algorithm, the F-score was calculated and compared as a classification accuracy parameter (Table 6). Silage maize, rice, alfalfa, and vegetables were cultivated and harvested during summer in the study area. In the case of silage maize, the SB, SI, T-VSI, and SB-SI-VSI feature sets showed the highest F-score, indicating their suitability for its discrimination from other crops. Albeit, following these feature sets was SB-SI-T feature set, in which the difference in their F-scores was very low (difference between 0.49% and 0.76% in F-scores).

Nonetheless, in the case of rice, there appeared to be no significant difference between the F-scores obtained for single or combined feature sets. However, the VSI feature set showed a slightly lower F-score compared to the rest, therefore had a lower performance. Finally, for the alfalfa and vegetables classes, the results suggested that the combined feature sets and the texture feature were

more informative and led to better results in their separation from other crop types. Moreover, F-score analyses indicated the F-score values above 90% and 80% for both silage maize and rice and alfalfa, for all feature sets, respectively. In the case of alfalfa, all feature sets showed good results above 90% by considering the F-score values, except SB, SI, and VSI was below 80%. Despite the fact that the F-score for vegetables was below 70% for the SB, SI, and VSI feature sets, the remaining feature sets seemed to perform rather well.

Table 6. Impact assessment of different feature sets in separating crops using RF-PSO. Bold values correspond to the highest F-scores (%).

Feature Sets	Maize			Rice			Alfalfa			Vegetables		
	UA	PA	F	UA	PA	F	UA	PA	F	UA	PA	F
SB	91.87	100.0	**95.76**	100.0	100.0	**100.0**	85.85	92.73	89.16	95.39	45.59	61.69
SI	92.78	98.72	**95.65**	100.0	100.0	**100.0**	81.21	94.09	87.17	98.46	39.51	56.38
T	87.65	97.65	92.38	100.0	99.19	99.59	96.29	87.19	91.51	86.15	77.78	**81.75**
VSI	92.17	96.53	94.29	100.0	96.85	98.39	68.68	97.69	80.65	96.92	30.29	46.15
SB-T	90.39	98.05	94.06	100.0	99.19	99.59	95.36	91.13	**93.20**	86.15	71.79	**78.32**
T-VSI	91.57	99.67	**95.45**	100.0	100.0	**100.0**	97.22	88.77	**92.80**	86.15	76.71	**81.16**
SB-SI-T	90.36	100.0	94.94	100.0	100.0	**100.0**	95.13	90.91	**92.97**	87.69	71.25	**78.62**
SB-SI-VSI	91.27	100.0	**95.43**	100.0	100.0	**100.0**	95.13	91.11	**93.08**	96.92	61.17	74.99

Abbreviations: Spectral bands (SB), spectral index (SI), texture (T), variation of spectral index (VSI), user's accuracy (UA, %), producer's accuracy (PA, %), and F-score (F, %).

3.5. Majority Voting of the Classification-Feature Selection Ensemble Procedure

The majority voting approach was implemented by accounting the majority of results from the RF-PSO classifier for all feature sets. Results showed that the adapted procedure attains higher accuracy in comparison with RF using independent and single feature sets. As evident from Tables 4 and 7, the overall accuracy and Kappa coefficient for the proposed procedure were higher than the single RF method for all feature sets. Likewise, with regards to the class-wise accuracies, it was observed that the F-scores of the proposed procedure were higher than the single RF algorithm for all feature sets, except for water class. However, obtained F-scores for water class using the proposed procedure encountered slight reductions of 0.74% and 2.18% in comparison with SB-SI-VSI and VSI, respectively. In addition, however, the proposed procedure performed well compared to the ground truth information for this class (the obtained F-score is 93.75%). Accordingly, the primary objective of this study was to obtain the crop type map and the results showed consistent improvements to its classification accuracy (i.e., OA and F-score) of the proposed procedure in comparison with the RF algorithm with different feature sets.

Table 7. Confusion matrix of the classification-feature selection ensemble procedure. The values which have been displayed for each class are the pixels number of validation samples.

		Truth Data								
		Maize	Rice	Alfalfa	Vegetables	Bare Soil	Water	Canebrake	PA (%)	F-Score (%)
	Maize	305	0	0	0	0	0	0	100.0	95.76
	Rice	0	123	0	0	0	0	0	100.0	100.0
	Alfalfa	3	0	421	5	0	0	28	92.12	94.61
Classifier Results	Vegetables	4	0	12	60	0	0	1	77.92	84.51
	Bare Soil	20	0	0	0	610	0	0	96.83	97.76
	Water	0	0	0	0	8	60	0	88.24	93.75
	Canebrake	0	0	0	0	0	0	145	100.0	90.91
	UA (%)	91.87	100	97.23	92.31	98.71	100.0	83.33		
		OA = 95.51%			Kappa = 0.942					

Abbreviations: User's accuracy (UA, %), producer's accuracy (PA, %), overall accuracy (OA, %), and Kappa coefficient (Kappa).

Based on Table 4, the producer's accuracies of individual classes in the single RF algorithm for all feature sets were 98.49%, 99.0%, 90.40%, 52.70%, 95.99%, 81.37%, and 92.41% on average with

standard deviations equal to 1.30%, 1.10%, 3.44%, 15.95%, 0.8%, 8.3%, and 3.99% for silage maize, rice, alfalfa, vegetables, bare soil, water, and canebrake, respectively. Whereas, the user's accuracies of the individual classes were assessed 90.10%, 100%, 85.67%, 88.27%, 97.67%, 99.58%, and 74.5% on average with standard deviations equal to 1.96%, 0%, 9.99%, 5.26%, 1.21%, 0.77%, and 4.39% for silage maize, rice, alfalfa, vegetables, bare soil, water, and canebrake, respectively. Therefore, based on these values and as shown in Table 7, the user's and producer's accuracies of individual classes were also improved in the classification-feature selection ensemble procedure than the single RF algorithm for all feature sets by 1.77%, 0%, 11.56%, 4.04%, 1.04%, 0.42%, 8.83% and 1.51%, 1%, 1.72%, 25.22%, 0.84%, 6.87%, 7.59% on improvement of these average accuracies for silage maize, rice, alfalfa, vegetables, bare soil, water, and canebrake, respectively.

In the single RF algorithm for all feature sets, all the classes were classified correctly by considering the user's and producer's accuracies, however, vegetables and alfalfa were classified by the lower accuracies, which represent a high intra-class variability with the most standard deviation of accuracies. The producer's and user's accuracies in the classification-feature selection ensemble procedure showed that virtually all crops, except vegetables, had accuracies higher than 90% (Table 7). It should be noted that the producer's accuracy for vegetables (77.92%) was also reasonably accurate, indicating the low omission error for this crop. The user's accuracy for vegetables (92.31%) was also high, which itself specifies the rather acceptable performance of the proposed procedure in mapping vegetables with a low commission error, with only a few vegetables land parts misclassified. Bare soil, water, and canebrake had also high user's and producer's accuracies (higher than 80%). Therefore, the user's and producer's accuracies for all classes in the classification-feature selection ensemble procedure suggested a generally satisfactory performance of the method applied in crop mapping, especially for alfalfa and vegetables.

In addition, in order to compare the results of the classification-feature selection ensemble procedure and the single RF classifier, we also assessed a confusion matrix of the best output of single RF classifier among all feature sets (i.e., T-VSI feature set) (Table 8). The results showed the contribution of the classification-feature selection ensemble procedure in the classification by 2.67%, 0.6%, 0%, 3.12%, 7.59%, 1.08%, 8.04%, and 9.6% improvement in OA and F-score values of silage maize, rice, alfalfa, vegetables, bare soil, water, and canebrake, respectively. Meanwhile, considering the user's and producer's accuracies of individual classes indicated an improvement in accuracy by 0%, 0%, 2.33%, 7.69%, 1.95%, 0%, 12.07% and 1.29%, 0%, 3.78%, 7.41%, 0.23%, 13.24%, 5.34% for silage maize, rice, alfalfa, vegetables, bare soil, water, and canebrake, respectively. These results also confirmed the impact of the classification-feature selection ensemble procedure on improving the classification accuracy significantly for heterogeneous classes (i.e., alfalfa, vegetables, and canebrake with 3.12%, 7.59%, and 9.6% improvements in F-score, 2.33%, 7.69%, and 12.07% in user's accuracies, 3.78%, 7.41%, and 5.34% in the producer's accuracies, respectively) in comparison with the other classes, which further supported the previous results reported. The classification accuracy of water class was also improved by the classification-feature selection ensemble procedure compared to the single RF classifier with this feature set (i.e., the difference in F-score values was 8.04%).

Omission and commission errors with the classification-feature selection ensemble procedure were consistently lower than 10% for all classes, with the exception of vegetables and water by the omission error = 22.08% and 11.76 %, respectively and canebrake class by the commission error = 16.67%. The omission and commission errors of alfalfa, vegetables, and canebrake than other classes were highly decreased by the classification-feature selection ensemble procedure than the single RF classifier with T-VSI feature set (3.78%, 7.41%, and 5.34% decreases in omission error and 2.33%, 7.69%, and 12.07% decreases in commission error, respectively). These error reductions of alfalfa, vegetables, and canebrake have also been obtained by the proposed method than the single RF algorithm for all feature sets by 0.34–6.28%, 7.25–49.42%, and 4.35–15% decreases in omission error and 2.33–30.18%, 1.54–10.77%, 2.08–16.28% decreases in commission error, respectively.

Table 8. Confusion matrix of the best output of the RF classifier among all feature sets (i.e., T-VSI feature set). The values which have been displayed for each class are the pixels number of validation samples.

		Truth Data								
		Maize	Rice	Alfalfa	Vegetables	Bare Soil	Water	Canebrake	PA (%)	F-Score (%)
Classifier Results	Maize	305	0	0	0	0	0	4	98.71	95.16
	Rice	0	123	0	0	0	0	0	100.0	100.0
	Alfalfa	0	0	409	9	0	0	45	88.34	91.49
	Vegetables	5	0	18	55	0	0	0	70.51	76.92
	Bare Soil	20	0	0	0	598	0	1	96.6	96.68
	Water	0	0	0	0	20	60	0	75.0	85.71
	Canebrake	2	0	4	1	0	0	124	94.66	81.31
UA (%)		91.87	100.0	94.90	84.62	96.76	100.0	71.26		
OA = 92.845%					Kappa = 0.908					

Abbreviations: User's accuracy (UA, %), producer's accuracy (PA, %), overall accuracy (OA, %), and Kappa coefficient (Kappa).

One method that is equally important to implement is the factor of computation time. The computational time for majority voting of outputs of RF-PSO in the classification-feature selection ensemble procedure costs a few seconds (using a computer with Intel (R) Core (TM) i7-2640M CPU processor at 2.80 GHz and 8 GB of RAM in MATLAB software). Figure 8 illustrates the obtained crop map using the proposed procedure and the spatial distribution of training and validation samples of crops for visual comparison.
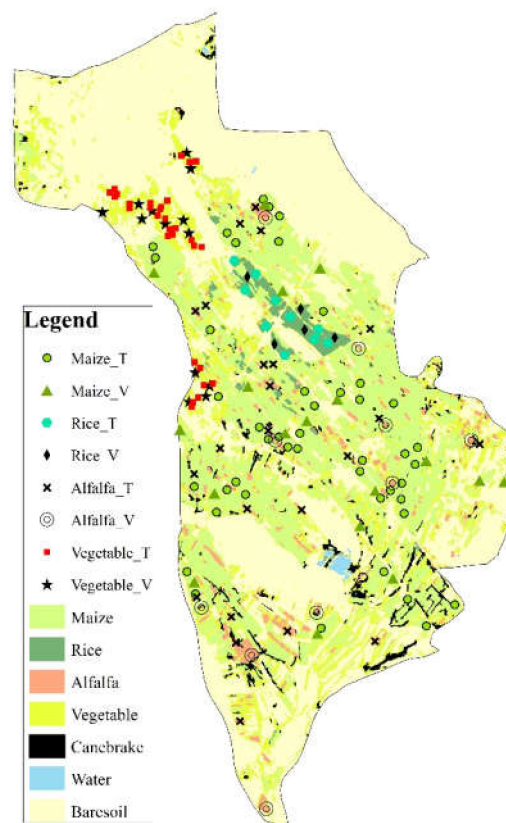
Figure 8. Crop mapping of the study area using the classification-feature selection ensemble procedure and the spatial distribution of training (T) and validation (V) samples of crops.

4. Discussion

Based on the results of OA, Kappa coefficient, and F-score values of different land cover classes for all feature sets, RF performance was efficient in crop mapping with multi-temporal Sentinel-2

images. T-VSI, followed by SB-T, SB-SI-T, and SB-SI-VSI feature sets outperformed other feature sets and obtained better results in crop mapping. The advantage of SB and T feature sets in combination with other feature sets for crop mapping has also been mentioned by Pelletier et al. (2016) [23] and Qian et al. (2017) [50], respectively. While vegetation indices were considered as only features in many multi-temporal crop mapping studies (such as Wardlow et al. 2007 [13]; Hao et al. 2016 [51]), incorporation of other features apart from vegetation indices could improve the classification accuracy and this result corresponded to the finding of [23]. Crop types have been shown with a positive sensitivity to input feature sets, where rice and vegetables classes were presented the least and with the highest reactivity of RF towards different input feature sets, respectively. Moreover, the lower F-score values for vegetables fields (mean and SD of F-score values for all feature sets = 64.31% and 11.51%, respectively) in RF for all feature sets was attained and may indicate its heterogeneity quality.

Despite the efficiency of RF to discriminate different classes, similar, or even more accuracy in the classification especially in heterogeneous classes maybe attainable through selecting a subset of relevant features. A similar increasing trend happened in F-score and overall accuracy when RF-PSO and RF-variable importance measure feature selection methods have been used. However, our results showed that RF-PSO could increase more classification accuracy than the RF-variable importance measure. Significantly higher improvements in accuracy have been observed for the alfalfa, vegetables, and canebrake than other classes, especially in the vegetables class. These three classes were considered heterogeneous classes in the study area, but with feature selection, these three classes were more precisely classified. Therefore, the dimensionality reduction with feature selection could improve crop differentiation when multi-temporal imagery was used. This result corresponded to the finding of Gilbertson J. K. and van Niekerk (2017) [23] and Rodriguez-Galiano (2012a) [19] by the RF-variable importance measure method with Landsat imagery and they mentioned it had a marginal effect on the results.

Results of features importance indicated the end region in the red edge (band 7), SWIR (band 11, 12), narrow NIR (band 8a) bands, whereas visible bands (blue (band 2), green (band 3), and red (band 4) bands, respectively) were selected as the most and least important spectral bands on land cover classes differentiation, respectively. Similarly, the end region in the red edge was selected as among the most sensitive spectral bands to leaf chlorophyll content (LCC) measuring by Verrelst et al. (2016) [52], on assessing sensitive bands of hyperspectral data to biophysical variables such as LCC and leaf area index (LAI). Moreover, the findings of Verrelst et al. (2016) [52] also indicated that the red edge region and NIR contribute to the highest impact of LAI, which was also confirmed by Kira et al. (2016) [53]. Therefore, in our study, this may be a satisfactory reason for why the mentioned bands were selected as the most important bands in classifying different classes, especially in regards to the classification of crops, since they can be used to determine the various characteristics of vegetation. Moreover, the blue and red bands were selected as the absorption bands due to LCC in the findings of Verrelst et al. (2016) [52]. Hence, the least sensitivity of these bands was also confirmed in our study where the visible bands have been determined as the least contribution to the crop classification.

NDWI and IRECI as well as ARVI and EVI were selected as the most and least relevant features to separate different classes based on spectral indices and VSI, respectively. Verrelst et al. (2016) [52] also evaluated the green band as among the most sensitive spectral bands to LCC. The NDWI is a combination of the NIR and the green band (Table 3), which both are sensitive to LAI and LCC, respectively. Whereas the IRECI is composed of the red edge and NIR bands (Table 3), which according to [52] are sensitive to the LAI and LCC biophysical parameters. Therefore, the spectral response of both NIR, green bands, and both red edge NIR bands in NDWI and IRECI could intensify sensitivity to LAI and LCC, respectively. Moreover, our findings of the most important spectral bands on the classes' separation (i.e., red edge and NIR) could indicate the most sensitivity of NDWI and IRECI to contribute more in crop classification. The finding of less sensitivity of ARVI and EVI were confirmed by the least important spectral bands (i.e., red and blue bands), because ARVI and EVI are a combination of the NIR and red/blue bands or both. The absorption characteristic of these bands could reduce the most

sensitivity feature of NIR to LAI. The more informative texture features in crop mapping were found in bands 5, 6, 7 (red edge), 11 (SWIR), 8 (NIR), and 8a (narrow NIR) in "mean" and "energy" features. These results also confirmed the results of [52], wherein the red edge and NIR bands were indicated as more sensitive bands to LAI and LCC, and thereby contribute more to the classes' separation.

The RF-PSO algorithm with optimal results of classification accuracy than the RF and RF-variable importance measure was used to determine the contribution of each feature set on the separation of crops. The SB, SI, T-VSI, and SB-SI-VSI feature sets had the highest capacity to identify silage maize. Among these feature sets, multi-temporal spectral bands (SB), obtained directly from remote sensing images, showed more F-score values in differentiation of other crops. This finding was compliant with the results by Zhong et al. (2014) [1] who employed Landsat images in order to attain crop maps for dissimilar feature sets of our study. No significant difference between the F-scores obtained for single or combined feature sets were obtained in rice. However, in the case of alfalfa and vegetables, and especially vegetables, single or combined feature sets could make a difference in F-score values of classification. In regards to vegetables, it is worth noting that the cultivated lands were highly scattered and heterogeneous, which may perhaps have led to the lower classification accuracy. Therefore, the homogeneity/heterogeneity quality of fields could cause obtaining the higher/lower classification accuracy, where if the fields were homogenous, there was no difference between the accuracy of single or combined feature sets in crop classification. However, for heterogeneous fields, the combined feature sets could improve the accuracy rather than single feature sets. Thus, the lower classification accuracy could be partially resolved by adding different features or features combining.

Finally, the classification-feature selection ensemble procedure by accounting the majority voting of results from the RF-PSO classifier for all feature sets was attained higher accuracy in comparison with the RF using independent and single feature sets, with regards to OA, Kappa coefficient, and F-score. Meanwhile, considering the user's and producer's accuracies of individual classes indicated an accuracy improvement of the classification-feature selection ensemble procedure than the single RF algorithm for all feature sets. Specifically, alfalfa and vegetables were classified by the lower classification accuracies in the single RF algorithm for all feature sets, however, the classification-feature selection ensemble procedure could increase their accuracies than the mean results of a single RF algorithm for all feature sets by 11.56% and 4.04% on user's accuracies and 1.72% and 25.22% on producer's accuracies, respectively. The best output of a single RF classifier among all feature sets (i.e., T-VSI feature set) compared with the results of the classification-feature selection ensemble procedure, could be used for assessing the impact of the proposed method on classification accuracy improvement. This comparison has also confirmed the contribution of the classification-feature selection ensemble procedure in improvement of the classification accuracy especially in heterogeneous classes (i.e., alfalfa, vegetables, and canebrake with 3.12%, 7.59%, and 9.6% improvements in F-score, 2.33%, 7.69%, and 12.07% in user's accuracies, 3.78%, 7.41%, and 5.34% in producer's accuracies, respectively).

More confusion between alfalfa and vegetables due to their similar biophysical characteristics led to misclassification pixels between these crops. The higher values of omission and commission errors for alfalfa and vegetables have been also indicated by Asgarian et al. (2016) [54] and Joseph 2005 [55], which confirmed that these crops are difficult classes to be spectrally separated. This major confusion of vegetables with alfalfa could be reduced by the classification-feature selection ensemble procedure. The error reduction for vegetables was obtained by the proposed method in comparison to the best output of the single RF classifier (i.e., T-VSI feature set) from 29.49% to 22.08% in omission error and from 15.38% to 7.69% in commission error. The similar improvement for vegetables classification by the proposed method indicated an error reduction from 29.33–71.5% to 22.08% in omission error and from 9.23–18.46% to 7.69% in commission error in all feature sets.

The superiority of the ensemble classifier over a single classification, according to majority voting, was also confirmed by Waske and Braun (2009); Waske et al. (2010); Löw et al. (2015) [44,56,57] for other strategies. Therefore, classification accuracy has been influenced by homogeneity/ heterogeneity

quality of fields and the proposed method could reduce the classification error specifically in heterogeneous classes.

5. Conclusions

The classification-feature selection ensemble procedure for land cover classification, particularly crop mapping, was assessed using multi-temporal feature sets consistent with the growth phenology of regional heterogeneous crops in Ghale-Nou county of Tehran, Iran. The results of Sentinel-2 data were indicative of the efficiency of RF in classifying different crops for all feature sets especially with the T and SB feature sets combining other feature sets. This study provides a new insight into the operation and effectiveness of feature selection algorithms in improving the classification accuracy. In fact, the present study illustrated the potency of incorporating the PSO approach and RF-variable importance measure as feature selection algorithms for improving classification accuracy of crop mapping using multi-temporal data; which increases the dimension of feature space. Our results revealed that, PSO outperformed the RF-variable importance measure in increasing the classification accuracy for all crop types. The relative analysis of features significance in crop mapping revealed that red edge, SWIR, narrow NIR bands, NDWI and IRECI indices, "mean" texture feature for the red edge, SWIR and NIR bands had a significant role in crop mapping. The results on which feature sets performed better to the classes' separability showed that spectral bands, combination feature sets, and texture had the highest accuracies for classifying maize, alfalfa, and vegetables crops, respectively. All feature sets showed a high performance in classifying rice and its separating from other classes.

Since each feature set performed better in separating specific crop types, it would appear that a fusion of different feature sets could assist in enhancing the classification accuracy and attaining a more preferable crop mapping. According to the results of the classification-feature selection ensemble procedure, which accounting by the majority votes of output results applying the RF-PSO for all feature sets in each pixel, it can be concluded that the ensemble procedure could well achieve more optimal results in comparison with the RF method by separate feature sets. Where results from the ensemble procedure compared to the RF method by separate feature sets witnessed an increase of 2.7% to 7.4% in overall accuracy in all feature sets. Likewise, with regards to the class-wise accuracies, it was observed that the F-scores of ensemble procedure compared to the RF algorithm by separate feature sets increased by about 0.48–3.68% (maize), 0–1.61% (rice), 2.82–15.43% (alfalfa), and 10.96–41.13% (vegetables) for all feature sets. In addition, the user's and producer's accuracies of each classes were confirmed by the superiority of the proposed method than RF by separate feature sets in classifying all classes especially in alfalfa and vegetables (increase of 11.56% and 4.04% on the user's accuracies and 1.72% and 25.22% on producer's accuracies than the mean results of a single RF algorithm for all feature sets, respectively). The contribution of the classification-feature selection ensemble procedure also confirmed improvement of the classification accuracy than the best output of a single RF classifier among all feature sets (i.e., T-VSI feature set) especially in heterogeneous fields (i.e., alfalfa, vegetables, and canebrake).

In this study, the similar biophysical characteristics between vegetables and alfalfa led to more confusion and misclassification pixels between these crops. The classification-feature selection ensemble procedure than the single RF classifier could improve the overall accuracy by about 2.7% to 7.4% for all feature sets, however, the proposed method mainly improves the class-wise accuracy and reduces the omission/commission errors in heterogeneous fields, which were misclassified as other classes (i.e., error reduction especially in the vegetables class from 29.33–71.5% to 22.08% in omission error and from 9.23–18.46% to 7.69% in commission error in all feature sets) with the assessed efficiency by multi-temporal Sentinel-2 imagery. Therefore, the proposed method could mainly be useful to differentiate between crops in the study area containing heterogeneous crops fields such as alfalfa and vegetables with similar biophysical characteristics, which are difficult classes to be spectrally separated.

Author Contributions: Conceptualization, E.A., A.D.B., N.N.S., S.H., and S.P.; data curation, E.A.; formal analysis, E.A.; investigation, E.A. and A.D.B.; methodology, E.A., A.D.B., S.H., and S.P.; resources, E.A.; software, E.A.;

References

1. Zhong, L.; Gong, P.; Biging, G.S. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 1–13. [CrossRef]

2. Yin, L.; You, N.; Zhang, G.; Huang, J.; Dong, J. Optimizing Feature Selection of Individual Crop Types for Improved Crop Mapping. *Remote Sens.* **2020**, *12*, 162. [CrossRef]

3. Meng, S.; Zhong, Y.; Luo, C.; Hu, X.; Wang, X.; Huang, S. Optimal Temporal Window Selection for Winter Wheat and Rapeseed Mapping with Sentinel-2 Images: A Case Study of Zhongxiang in China. *Remote Sens.* **2020**, *12*, 226. [CrossRef]

4. Massey, R.; Sankey, T.T.; Congalton, R.; Yadav, K.; Thenkabail, P.S.; Ozdogan, M.; Meador, A.S. MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types. *Remote Sens. Environ.* **2017**, *198*, 490–503. [CrossRef]

5. Song, X.; Potapov, P.V.; Krylov, A.; King, L.; Di, C.M.; Hudson, A.; Khan, A.; Adusei, B.; Stehman, S.V.; Hansen, M.C. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and fi eld survey. *Remote Sens. Environ.* **2017**, *190*, 383–395. [CrossRef]

6. Jin, H.; Mountrakis, G.; Stehman, S.V. Assessing integration of intensity, polarimetric scattering, interferometric coherence and spatial texture metrics in PALSAR-derived land cover classification. *ISPRS J. Photogramm. Remote Sens.* **2014**, *98*, 70–84. [CrossRef]

7. Papadopoulos, A.; Kalivas, D.; Theocharopoulos, S. Spatio-temporal monitoring of cotton cultivation using ground-based and airborne multispectral sensors in GIS environment. *Environ. Monit. Assess.* **2017**, *189*, 323–334. [CrossRef]

8. Villa, P.; Stroppiana, D.; Fontanelli, G.; Azar, R.; Brivio, P.A. In-Season Mapping of Crop Type with Optical and X-Band SAR Data: A Classification Tree Approach Using Synoptic Seasonal Features. *Remote Sens.* **2015**, *7*, 12859–12886. [CrossRef]

9. Bargiel, D. A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sens. Environ.* **2017**, *198*, 369–383. [CrossRef]

10. Song, Q.; Hu, Q.; Zhou, Q.; Hovis, C.; Xiang, M.; Tang, H.; Wu, W. In-Season Crop Mapping with GF-1 / WFV Data by Combining Object-Based Image Analysis and Random Forest. *Remote Sens.* **2017**, *9*, 1184. [CrossRef]

11. Inglada, J.; Vincent, A.; Arias, M.; Sicre, C.M. Improved Early Crop Type Identification By Joint Use of High Temporal Resolution SAR And Optical Image Time Series. *Remote Sens.* **2016**, *8*, 362. [CrossRef]

12. Qiu, B.; Fan, Z.; Zhong, M.; Tang, Z.; Chen, C. A new approach for crop identification with wavelet variance and JM distance. *Environ. Monit. Assess.* **2014**, *186*, 7929–7940. [CrossRef] [PubMed]

13. Wardlow, B.D.; Egbert, S.; Kastens, J. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sens. Environ.* **2007**, *108*, 290–310. [CrossRef]

14. Heinl, M.; Walde, J.; Tappeiner, G.; Tappeiner, U. Classifiers vs. input variables—The drivers in image classification for land cover mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 423–430. [CrossRef]

15. Zheng, B.; Myint, S.W.; Thenkabail, P.S.; Aggarwal, R.M. A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 103–112. [CrossRef]

16. Gilbertson, J.K.; Van Niekerk, A. Value of dimensionality reduction for crop differentiation with multi-temporal imagery and machine learning. *Comput. Electron. Agric.* **2017**, *142*, 50–58. [CrossRef]

17. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870. [CrossRef]

18. Bellman, R. *Dynamic Programming*, 2nd ed.; Dover Publications: Mineola, NY, USA, 2003.

19. Rodriguez-Galiano, V.; Olmo, M.C.; Abarca-Hernández, F.; Atkinson, P.M.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107. [CrossRef]

20. Rodriguez-Galiano, V.; Ghimire, B.; Rogan, J.; Olmo, M.C.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [CrossRef]

21. Shao, Y.; Lunetta, R.S. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS J. Photogramm. Remote Sens.* **2012**, *70*, 78–87. [CrossRef]

22. Waldner, F.; Canto, G.S.; Defourny, P. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 1–13. [CrossRef]

23. Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Dedieu, G. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sens. Environ.* **2016**, *187*, 156–168. [CrossRef]

24. Chen, Y.; Lu, D.; Moran, E.; Batistella, M.; Dutra, L.V.; Sanches, I.D.; Da Silva, R.F.B.; Huang, J.; Luiz, A.J.B.; De Oliveira, M.A.F. Mapping croplands, cropping patterns, and crop types using MODIS time-series data. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *69*, 133–147. [CrossRef]

25. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W. How much does multi-temporal Sentinel-2 data improve crop type classi fi cation ? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 122–130. [CrossRef]

26. Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sens. Environ.* **2019**, *220*, 135–151. [CrossRef]

27. Hafiz, F.; Swain, A.; Patel, N.; Naik, C. A two-dimensional (2-D) learning framework for Particle Swarm based feature selection. *Pattern Recognit.* **2018**, *76*, 416–433. [CrossRef]

28. Bharti, K.K.; Singh, P.K. Opposition chaotic fitness mutation based adaptive inertia weight BPSO for feature selection in text clustering. *Appl. Soft Comput.* **2016**, *43*, 20–34. [CrossRef]

29. Xue, B.; Zhang, M.; Browne, W.N.; Yao, X. A Survey on Evolutionary Computation Approaches to Feature Selection. *IEEE Trans. Evol. Comput.* **2015**, *20*, 606–626. [CrossRef]

30. Cervante, L.; Xue, B.; Zhang, M.; Shang, L. Binary Particle Swarm Optimisation for Feature Selection: A Filter Based Approach. In Proceedings of the 2012 IEEE Congress on Evolutionary Computation, Brisbane, Australia, 10–15 June 2012; pp. 1–8.

31. Drusch, M.; Bello, U.D.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's optical high-resolution mission for {GMES} operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [CrossRef]

32. IRIMO. Available online: www.irimo.ir (accessed on 30 September 2018).

33. ESA. Level-1C Product Types. 2019. Available online: https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-1c (accessed on 3 December 2019).

34. Zabala Ramos, S. *Comparison of multi-temporal and multispectral Sentinel-2 and Unmanned Aerial Vehicle imagery for crop type mapping Comparison of multi-temporal and multispectral Sentinel-2 and Unmanned Aerial Vehicle imagery for crop type mapping*; GEM thesis Ser; Lund University: Lund, Sweden, 2017.

35. Sun, X.H.; Qin, P. Texture analysis for remotely sensed imagery. In Proceedings of the Thematic Conference on Geological Remote Sensing, Pasadena, CA, USA, 8–11 February 1993; pp. 311–322.

36. Zhou, T.; Pan, J.; Zhang, P.; Wei, S.; Han, T. Mapping Winter Wheat with Multi-Temporal SAR and Optical Images in an Urban Agricultural Region. *Sensors* **2017**, *17*, 1210. [CrossRef]

37. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 610–621. [CrossRef]

38. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, in Technical Presentations, Section A, Washington, DC, USA, 10–14 December 1973; Freden, S.C., Mercanti, E.P., Becker, M., Eds.; Volume I, pp. 309–317.

39. Jiang, Z.; Huete, A.R.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [CrossRef]

40. Kaufman, Y.; Tanre, D. Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 261–270. [CrossRef]

41. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [CrossRef]

42. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]

43. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92. [CrossRef]

44. Waske, B.; Braun, M.H. Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 450–457. [CrossRef]

45. Kuncheva, L.I. *Combining Pattern Classifiers: Methods and Algorithms*; John Wiley & Sons: Hoboken, NJ, USA, 2004.

46. Breiman, L.E.O. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

47. Moradi, P.; Gholampour, M. A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy. *Appl. Soft Comput.* **2016**, *43*, 117–130. [CrossRef]

48. Petropoulos, G.P.; Kontoes, C.C.; Keramitsoglou, I. Land cover mapping with emphasis to burnt area delineation using co-orbital ALI and Landsat TM imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 344–355. [CrossRef]

49. ESA. SENTINEL-2 Radiometric Resolutions. 2019. Available online: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric (accessed on 3 December 2019).

50. Song, Q.; Zhou, Q.; Wu, W.; Hu, Q.; Lu, M.; Liu, S. Mapping regional cropping patterns by using GF-1 WFV sensor data. *J. Integr. Agric.* **2017**, *16*, 337–347. [CrossRef]

51. Hao, P.; Wang, L.; Zhan, Y.; Niu, Z.; Wu, M. Using historical NDVI time series to classify crops at 30m spatial resolution: A case in Southeast Kansas. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016. [CrossRef]

52. Verrelst, J.; Rivera, J.P.; Gitelson, A.; Delegido, J.; Moreno, J.; Camps-Valls, G. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 554–567. [CrossRef]

53. Kira, O.; Nguy-Robertson, A.L.; Arkebauer, T.J.; Linker, R.; Gitelson, A. Informative spectral bands for remote green LAI estimation in C3 and C4 crops. *Agric. For. Meteorol.* **2016**, *218*, 243–249. [CrossRef]

54. Asgarian, A.; Soffianian, A.; Pourmanafi, S. Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery. *Comput. Electron. Agric.* **2016**, *127*, 531–540. [CrossRef]

55. Joseph, G. *Fundamentals of Remote Sensing*; Universities Press: Himayatnagar, Hyderabad, India, 2005.

56. Waske, B.; Van Der Linden, S.; Benediktsson, J.A.; Rabe, A.; Hostert, P. Sensitivity of Support Vector Machines to Random Feature Selection in Classification of Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2880–2889. [CrossRef]

57. Löw, F.; Conrad, C.; Michel, U. Decision fusion and non-parametric classifiers for land use mapping using multi-temporal RapidEye data. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 191–204. [CrossRef]