Table 3. Field data and PRISMA and Sentinel-2 images used for crop type mapping of MKK, Grosseto, Maccarese, and Jolanda di Savoia sites for 2020–2023. The PRISMA images used for Group A and Group B were shown in green and red color, respectively. The images highlighted in bold were assumed as if they were acquired in the other Group.

Site	Year	N. of Fields	Area (ha)	Sentinel-2	PRISMA Image Date (Month/Day)
Maccarese	2021	92	1600	6	04/01, 05/17, 06/27, 09/04
	2022	165	1600	6	**01/16, 04/12, 05/29, 06/15, 07/14**
	2023	70	700	2	02/02, 03/21
Jolanda	2021	176	3700	6	**04/24, 06/04, 06/21**
	2022	210	4100	8	04/30, 05/12, 07/03, 08/01
	2023	105	2863	4	03/04, 04/07, 05/24, 07/03, 08/07
Grosseto	2020	6	20	1	07/31
MKK-Iran	2021	24	120	5	01/23, 03/11, 04/09, 05/14, 05/19, 06/23, 07/22
	2022	30	310	7	01/29, 02/27, 03/28, 05/08, 05/25, 07/05, 07/22
	2023	26	130	3	02/28, 03/17, 04/09

2.4. Pre-Processing and Processing EO Data

2.4.1. Pre-Processing

PRISMA L2D images were co-registered, with an Automated and Robust Open-Source Image Co-Registration Software (AROSICS v.1.11.0) algorithm using the Sentinel-2 image acquired at the closest date, to assure the co-registration (of about 0.5pixel of RMS), and then smoothed by a Savitzky–Golay filter (frame size of 7, 3rd degree polynomial). A total of 55 bands, including overlapping bands (band numbers 63–68), atmospheric absorption bands, and bands with SNR below 100, were excluded from processing. The image dimensions (composed by 173 bands after band removal) were reduced by the principal component analysis (PCA) approach and then PCs were normalized by computing the standard deviation of the image. An optimum of 15 and 5 PCs were selected for the PRISMA and Sentinel-2 image dimension reduction, respectively. To reduce the effects of the spectral mixing of the plants cultivated by farmers around fields, a negative 30 m buffer (1 pixel) at the edge of the fields was considered.

2.4.2. Machine Learning Classification Algorithms

Different machine and deep learning algorithms for crop type mapping using PRISMA and Sentinel-2 data were considered. To avoid errors that could happen between species with different cultivation seasons, we decided to develop separate classifiers for winter, summer, and perennial crop seasons. Each method is described in the following paragraphs.

Multiclass Naive Bayes (MNB): The MNB classifier is based on Bayes theorem and belongs to the family of generative learning algorithms, which means that it models the distribution of inputs for a given class or category. This approach assumes that the features of the input data are conditionally independent given the class, allowing the algorithm to make predictions quickly and accurately [51].

K-Nearest Neighbor (KNN): The KNN algorithm is an instance-based learning method that classifies elements based on the closest k training samples in the resource space. KNN is a non-parametric MLA that makes no assumptions regarding data distribution. This is important when classifying processes of land use change, for which there is little or no prior knowledge of data distribution. In KNN, the pixel whose class is unknown is assigned to the same class as its spectrally closest neighbors [52].

Random Forest (RF): The RF technique [53] is an ensemble classifier that employs many decision trees and combines their results to forecast an output. There are two main steps in building each decision tree in an RF: (1) creation of a new training set by random

selection of subsamples through bootstrapping sampling from a specific number of samples in the original dataset, and (2) random selection of features. These two random processes help to avoid overfitting [54], because RF averages all these multiple decision trees, trained on different parts of the same training set, and assigns each pixel to the class label that garnered the most "votes". The number of trees is obtained using out-of-bag data analysis, and the number of features is obtained using the square root of the features. The RF algorithm, in particular, can process high-dimensional data, works well with mislabeled data, and has high accuracy for large-scale studies [55].

Support Vector Machine (SVM): SVM was originally proposed as a linear binary classifier. Subsequently, by introducing suitable kernels, the algorithm can perform nonlinear classification. Previous studies show that the radial basis function (RBF) kernel works properly for classification [56]. The RBF kernel was used in this research, where *C* (cost) and gamma parameters were optimized. The optimal parameters were determined using the grid search method considering its shorter run time for finding the best combination of *C* and gamma parameters compared with other methods. *C* is the penalty coefficient that determines the allowed level of error or misclassification. Therefore, a larger *C* always achieves a better result in training, but the generated model has a risk of overfitting, which decreases the classification's generalization. The smaller the *C*, the higher the number of unclassified samples [57].

Hyperparameters tuning: The performance of the ML algorithm depends on hyperparameter adjustment. The grid search method was used in sequence to tune the hyperparameters of ML algorithms. In the grid search method, the domain of hyperparameters is discretized to a grid. Then, the performance of all possible combinations is assessed using statistical metrics (here, cross-validation (CV)). The set of hyperparameters that can maximize the average value in CV is selected as an optimal one for training the model. Table 4 presents the tuned values for the hyperparameters of the ML algorithms. For the other parameters of ML algorithms, the default settings were used. All the ML classification methods and hyperparameter optimization of the present work were implemented in the Classification Learner App in MATLAB 2023b.

Table 4. Hyperparameters tuning process for MNB, KNN, SVM, and RF algorithms.

MNB	Parameter	Distribution		Kernel type	
	Range	Gaussian, kernel		Box, Epanechnikov, Gaussian, Triangle	
	Tuned	Kernel		Gaussian	
KNN	Parameter	Distance weights	n_neighbors	Distance metric	
	Range	Equal, Inverse, Squared inverse	1–21	Euclidean, Minkowski, Spearman, Hamming, Jaccard	
	Tuned	Equal	7	Euclidean	
RF	Parameter	NumLearningCycles	Method	MaxNumSplits	MinLeafSize
	Range	10–500	Bag, LSBoost	1–50	1–50
	Tuned	100	LSBoost	15	10
SVM	Parameter	Gamma	C	Kernel type	
	Range	0.001–100	0.001–100	Gaussian, Linear, Quadratic, Cubic, RBF	
	Tuned	10	1.27	RBF	

2.4.3. Deep Learning Classification Algorithms

A convolutional neural network (CNN) usually consists of a sequence of convolution layers and pooling layers, followed by a fully connected layer. The main advantage of CNN to regular neural networks is that the input of CNN models is an image, which allows specific features to be encoded into the model. Those features then help to reduce the number of network parameters [58]. The advantages of the CNN algorithm are

automatic feature learning, multilayer feature learning, high accuracy, and high generalizability [59,60]. The CNN method is popular for working with high-dimensional features, performing hierarchical learning operations, and enabling automatic feature extraction. Two kinds of CNNs (1D and 3D) have been used for classification [61].

One-Dimensional Convolutional Neural Networks (1D-CNN): 1D-CNN uses 1D filters to record the input signal spectral pattern. The implementation of 1D-CNN combines pooling layers, fully-connected layers, and dropout. Convolutional filter widths of 3 and 5 were used. The convolutional layers have 32, 64, and 128 filters, respectively, and the channel number increases when going deeper. Pooling layers were fixed as max-pooling, with a window size of 2. The probability of dropping neurons was set to 10%. This model had one fully connected layer at the output end. The second last layer collected information from previous layers as a flat array, and the size was determined by considering the dimension of the input to the layer. The learning rate was set to 0.001 and the batch size was set to 256. To keep only the positive values, rectified linear unit (ReLU) was used (Figure 4).

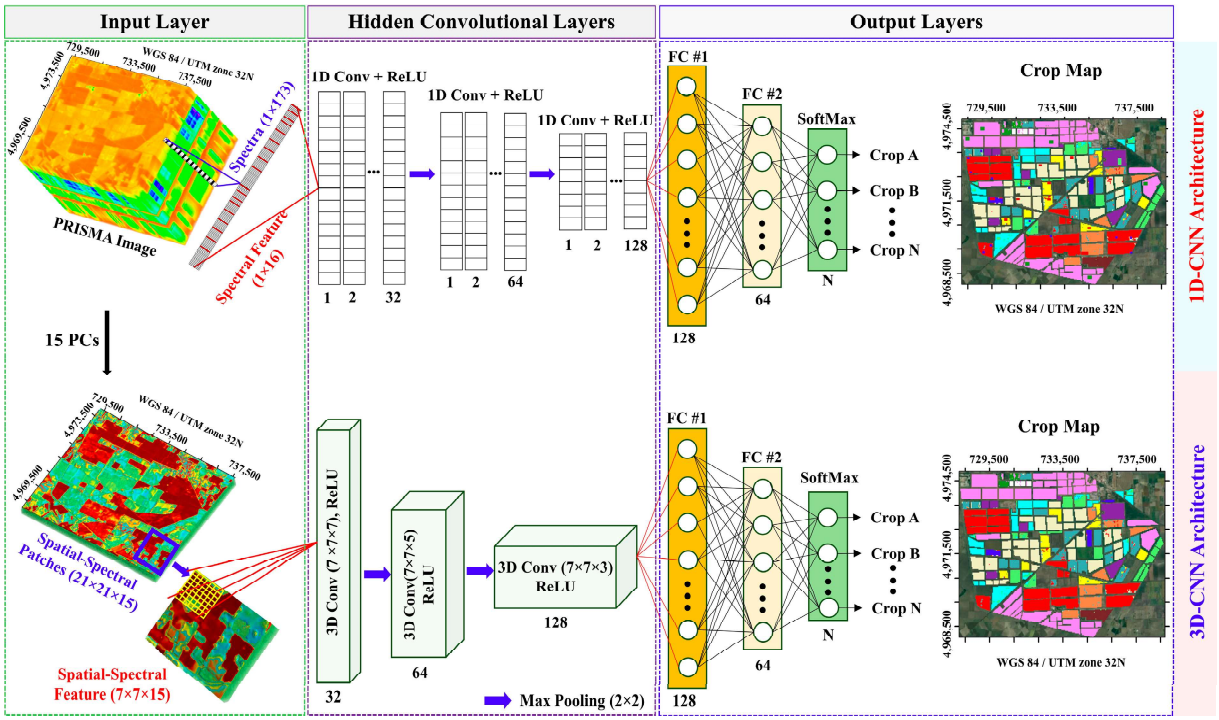

Figure 4. Architecture of the 1D-CNN (**top**) 3D-CNN (**bottom**).

Three-Dimensional Convolutional Neural Networks (3D-CNN): Other than spectral data, 3D-CNN takes advantage of the spatial information (which is not present in other methods) [60], which leads to increasing the computational cost. 3D-CNN executes successive sliding convolution operations in the input feature cube's plane and depth dimensions. Convolutional filter widths of 7, 5, and 3 were used, based on the smallest field of this study's sites. The convolutional layers had 32, 64, and 128 filters, respectively. For 3D-CNN, the network was trained by an initial learning rate of 0.001 for 100 epochs, a 256-batch size, momentum 0.9, learning rate factor 0.01, and Adam optimization. Different spatial window size was used to find an optimum window size. For each pixel, a 4D patch with size $m \times n \times d \times l$, where $m \times n$ refers to spatial window, d refers to spectral window, and l refers to label, were computed and used as input of CNN (Figure 4).

2.5. Classification Scenarios

Two scenarios were proposed for training and validation.

Same-farm TR/ACC: TRaining and ACCuracy assessment (TR/ACC) occurred in the same field. In this scenario, the pixels were randomly divided into TR (70%) and ACC sets (30%).

Cross-farm TR/ACC: To consider the spatial/temporal transferability of the trained model, TR/ACC data were selected from different farms. For sorghum, apple, and pear (Jolanda) and olive and almond (Maccarese) crops, which are not present across all farms, cross-year data from same-farm were used (Table 2).

2.6. Accuracy Assessment Scenarios

The same TR/ACC pixels were used for all the applied methods to avoid the effects of different partitioning of the dataset on the performance of methods. User accuracy (UA) and producer accuracy (PA), overall accuracy (OA), and Kappa coefficient were determined from the confusion matrix with respect to the ground validation data. The accuracy validation metrics are defined as follows:

$$\text{Overall accuracy (OA)} = \frac{\sum_{n=1}^{q} \mathrm{n}_{ii}}{n} \times 100 \tag{1}$$

$$\text{Kappa coefficient} = \frac{\left[n.\sum_{n=1}^{q} \mathrm{n}_{ii} - \sum_{n=1}^{q} \left(\mathrm{n}_{.i} \times n_{.i}\right)\right]}{\left[n^2 - \sum_{n=1}^{q} \left(n_{i.} \times n_{.i}\right)\right]} \tag{2}$$

$$\text{User Accuracy (UA)} = \frac{\mathrm{n}_{ii}}{n_{i.}} \times 100 \tag{3}$$

$$\text{Producer Accurac (PA)} = \frac{\mathrm{n}_{ii}}{n_{.i}} \times 100 \tag{4}$$

where n_{ii}, $n_{.i}$, and $n_{i.}$ are the values of the diagonal entry, the i-th row sum, and the i-th column sum in the confusion matrix, respectively; n is the total number of validation samples; and q is the number of rows of the confusion matrix.

3. Results

3.1. Spectral and Temporal Behavior of Species

Gaining knowledge on the spectral and temporal behavior of species in advance could lead to the design of a better procedure for classification. Figure 5 shows the simulated phenology of the dominant crop species of the Maccarese and Jolanda di Savoia farms (Table 1) derived from the Sentinel-2 normalized difference vegetation index (NDVI) time series data of 2023. As shown in Figure 5a, winter and durum wheat, barley, and triticale have high similarity in phenology patterns. Fava beans and peas start growing later than wheat and barley and are harvested at around the same time. Cardoon starts growing in autumn and has totally different phenology, thus it is expected to be discriminated easily from other species in early winter (Figure 5a). Maize, sorghum, rice, soybean, sunflower, and tomato species start to green-up from May but are harvested from September to October (Figure 5b). Peach, apple, and almond orchards start to green-up from January until October in this study's areas. Olive is an evergreen plant, and, even in winter, the NDVI is higher than 0.25. Simulation of alfalfa using the NDVI of Sentinel-2 was not successful, because alfalfa is cut 4–5 times between May and September, which makes the phenology pattern very different from other species. For this reason, instead of simulated phenology, the NDVI time series derived from Sentinel-2 is presented (Figure 5c).

In Figure 6, the PRISMA-derived spectral signature of the species (acquired in 2023) is shown. The similarity between the species' spectra, calculated using the cosine similarity method (cosine similarity function of MATLAB), is shown as a heat map in Figure 6e. The spectral signatures of winter and durum wheat, and barley acquired on 1 April 2021 at the Maccarese farm show high similarity '>95%') (Figure 6a), while the rest of the winter species (herbage and fava bean acquired on 1 April 2021 at the Maccarese farm and pea acquired on 24 April 2021 at the Jolanda di Savoia farm) show less similarity (Figure 6b).

For summer crops (maize, sorghum, rice, soybean, sunflower, and tomato acquired on 3 July 2022 at the Jolanda di Savoia farm), there is less similarity between species' spectral signatures, except between sorghum and maize. Rice shows a lower value compared with other summer species in the short-wave infrared (SWIR), which could be related to the presence of water in rice fields (Figure 6c). For perennial crops (olive, almond, and alfalfa acquired on 17 May 2021 at the Maccarese farm and pear and apple acquired on 4 June 2021 at the Jolanda di Savoia farm), pear shows less absorption in the red (630–680 nm) and higher reflectance in the 1300–2400 nm spectral range with respect to other perennial species (Figure 6d).

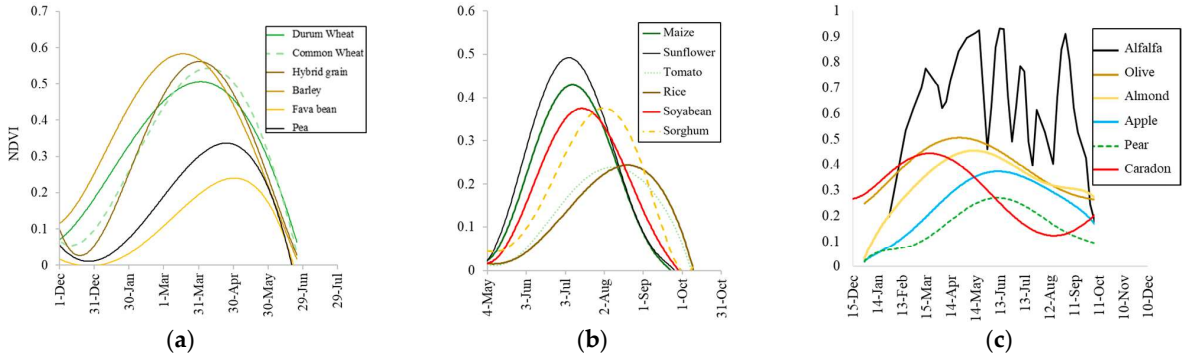

Figure 5. Plant simulated phenology derived from Sentinel-2 NDVI time series data of the (**a**) winter and durum wheat, barley, triticale, herbage, fava bean, and pea species; (**b**) maize, sorghum, rice, soybean, sunflower, and tomato species; (**c**) alfalfa, olive, almond, apple, pear, and cardoon species from the 2023 season.

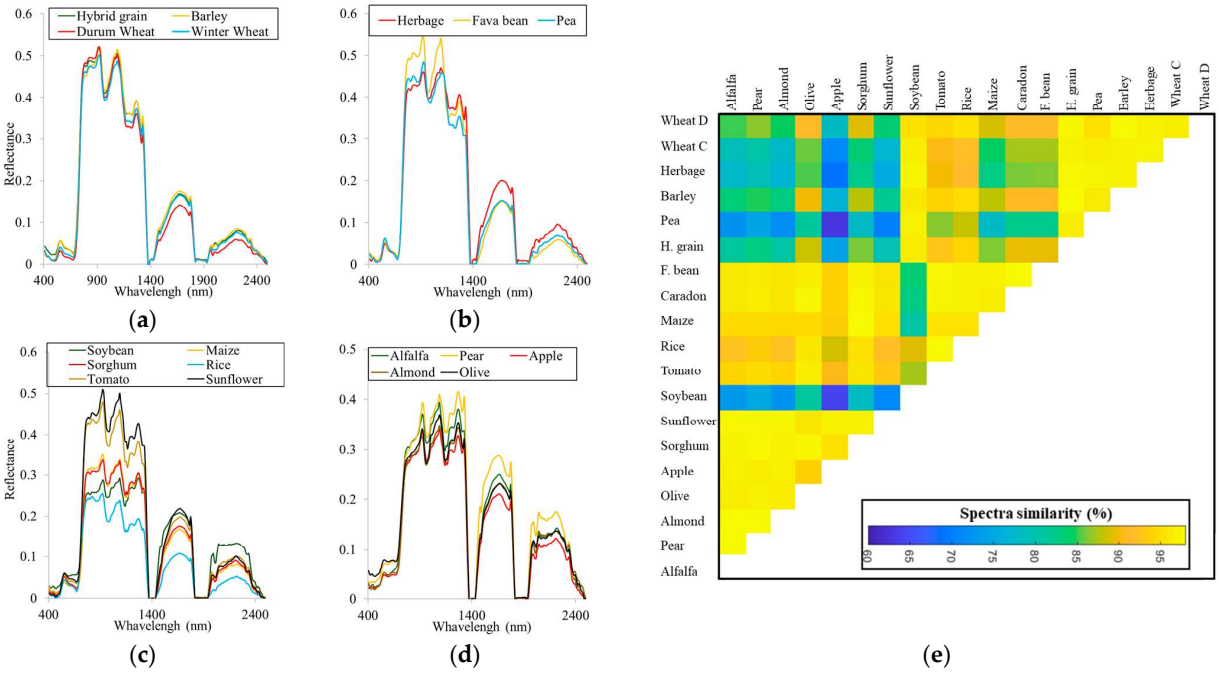

Figure 6. PRISMA-derived spectral signature of (**a**) winter and durum wheat, barley, and triticale acquired on 1 April 2021 at Maccarese farm; (**b**) herbage and fava bean acquired on 1 April 2021 at Maccarese farm and pea acquired on 24 April 2021 at Jolanda di Savoia farm; (**c**) maize, sorghum, rice, soybean, sunflower, and tomato species acquired on 3 June 2022 at Jolanda di Savoia farm; (**d**) olive, almond, and alfalfa acquired on 17 May 2021 at Maccarese farm and pear and apple acquired on 4 June 2021 at Jolanda di Savoia farm; and (**e**) similarity between the spectra of different species' PRISMA image from the four selected test sites.

Figure 7a shows the spectra of a specific pixel of the wheat fields from April and May's PRISMA images. In April, the wheat is in the heading stage and has strong chlorophyll absorption and high reflectance in the 800–1200 nm spectral range. In May, wheat is in the growth stage between flowering and ripening, exhibiting weak red absorption (less chlorophyll) and strong absorption in the 2100 nm spectral range (i.e., high lignin content). Based on the achieved primary result, the after-April and August images were eliminated from the winter and summer cultivation TR/ACC, respectively. Figure 7b shows the three different fields of alfalfa species at the Jolanda di Savoia farm in the acquired image on 3 July 2022. These fields are in different growing stages because of different harvesting dates and accordingly have different spectral reflectance.

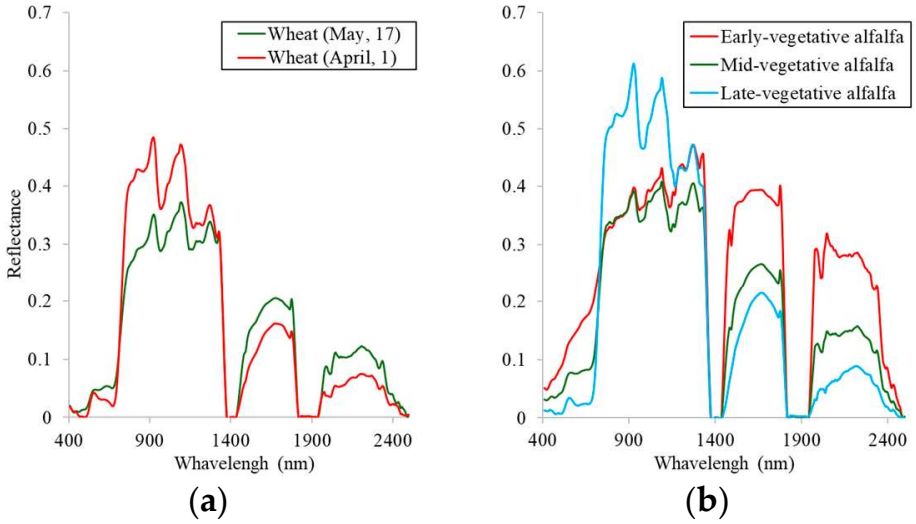

Figure 7. PRISMA-derived spectral signature of (**a**) winter and durum wheat species and (**b**) 3 different fields of alfalfa at Jolanda di Savoia farm in the acquired image from 3 July 2022.

3.2. Earliest Identifiable Time of Different Crops

Considering the earliest time that it is possible to apply crop type mapping, the classification was performed by the first available image in the season, and then all of the next available images were added one by one to the TR/ACC of the algorithms.

Figure 8 presents the OA of 1D-CNN algorithm applied to the time series of PRISMA images for each crop from 30 January (DOY 30) to 7 August (DOY 219). For winter (Figure 8a), cultivation DOY 81 is the first time the OA of all species reached 80% or higher. Wheat and triticale have lower OA until the end of the season, while the fava bean gains the best OA after DOY 102.

For summer (Figure 8b) cultivation, DOY 172 is the first time the OA of all species reached 80% or higher. Rice OA is reduced after DOY 190. Sorghum and soybean could be identified 20 days earlier than other species. For soybeans, this result depends on the fact that it is seeded before other species. For perennial (Figure 8c) cultivation, DOY 91 is the first time the OA of all species reached 80% or higher. Almond and cardoon could be identified 20 days earlier than other species.

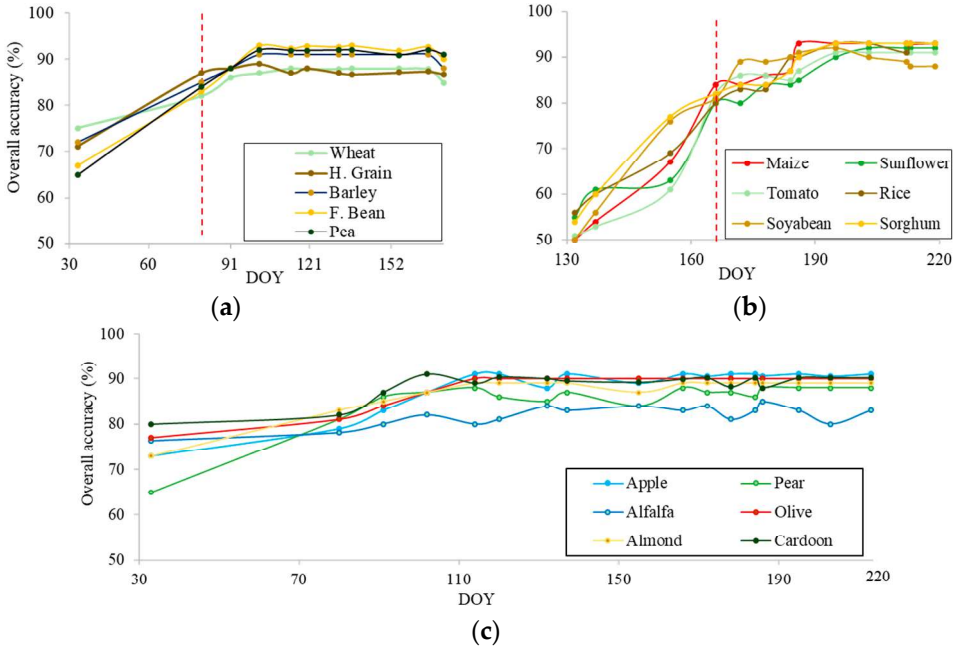

Figure 8. Per-species OA curves of early-season crop mapping for the Mediterranean crop calendar of (**a**) winter, (**b**) summer, and (**c**) perennial cultivations using 1D-CNN algorithm and PRISMA images. The vertical dashed red line represents the DOY when the OA first reached 80% or higher.

3.3. Classification Accuracy

The OA (for all cultivations) of classification using different machine and deep learning methods for PRISMA and Sentinel-2 images are shown as a bar graph in Figure 9. Figure 9 depicts that 1D-CNN and 3D-CNN methods were the most performing classifiers for all the seasons and sensors.

For same-farm TR/ACC of PRISMA, 1D-CNN and 3D-CNN provided overall accuracy higher than 95%, while for cross-farm TR/ACC, the OA was higher than 88%. 1D-CNN provided a 1–4% better result compared with 3D-CNN. Following these, CNN, RF (76% < OA < 79%), and SVM (73% < OA < 77%) algorithms showed the highest accuracy for cross-field validation. The accuracy of KNN (67% < OA< 73%) was slightly lower than RF and SVM, while the result of the MNB (OA < 53%) method was poor.

For Sentinel-2 image same-farm TR/ACC, 1D-CNN and 3D-CNN derived an overall accuracy higher than 88%, while for cross-farm TR/ACC, the OA was higher than 79%. 1D-CNN provided better results in winter and summer crops, while 3D-CNN provided better results for perennial crops. The RF (62% < OA< 66%) and SVM (60% < OA < 64%) algorithms had the highest accuracy for cross-farm TR/ACC after CNN methods. The accuracy of KNN (58% < OA < 64%) was a bit lower than RF and SVM, while the result of MNB (43% < OA < 47%) was poor.

The per-species classification PA and UA derived from PRISMA and Sentinel-2 images by the MNB, KNN, SVM, RF, and 1D-CNN and 3D-CNN classification algorithms are reported in Figure 10. Winter and durum wheat are merged in order to be counted as the same species to have a higher accuracy, especially in PRISMA images. For PRISMA image classification, among the other species, alfalfa showed the lowest UA using all the methods. The rest of the species did not show a stable trend in all methods. For PRISMA, alfalfa had the lowest UA, and maize showed the highest UA when in the cross-farm TR/ACC scenario. For Sentinel-2, alfalfa showed (similarly to PRISMA) the lowest UA, while winter wheat was characterized by the highest UA when TR/ACC in cross-farm (Figure 10).

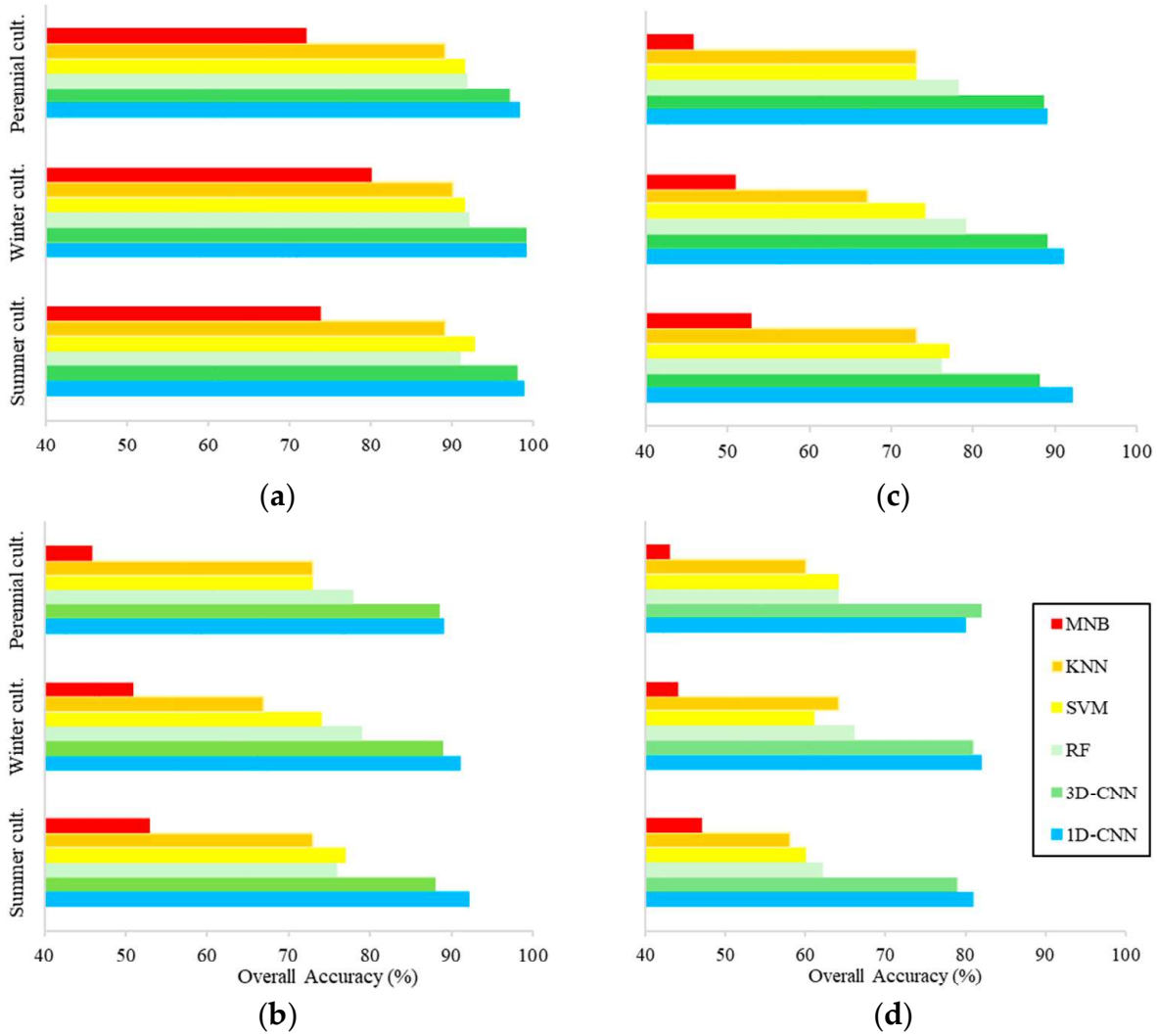

Figure 9. Overall accuracy (%) of the MNB, KNN, SVM, RF, 1D-CNN, and 3D-CNN classification methods for (**a**) same-farm TR/ACC-PRISMA, (**b**) cross-farm TR/ACC-PRISMA, (**c**) same-farm TR/ACC-Sentinel-2, and (**d**) cross-farm TR/ACC-Sentinel-2 accuracy assessment for all the selected sites.

Figure 11e,g show the classification maps for the year 2022 obtained using the 3D-CNN for the Jolanda di Savoia and MKK farms, respectively. Figure 11f shows the map of the Maccarese farm produced by the 1D-CNN method for the year 2021. Figure 11h shows the 1D-CNN method of the Grosseto farm for 2020. To avoid assigning two different labels to the same pixel, the fields used for both summer and winter cultivation (a few fields) were eliminated. The winter, summer, and perennial crop maps were merged to produce a comprehensive crop map of each field for one year. On all sites, it can be observed that most of the errors of omission and commission occurred in the pixels located at the margins of the fields. Classification errors in field edge pixels were higher than in central pixels, which was related to spectral mixing with neighboring fields with no similar cultivation. It was also related to the presence of trees/brushes surrounding the cultivated fields and water channels. A negative 30 m buffer (1 pixel) at the edges of fields resulted in 1% and 3% improvement in the OA of classification by 3D-CNN and SVM, respectively. This could have an impact on the usability of PRISMA in fragmented agricultural areas in the case that pansharpening techniques are not applied [62–64].