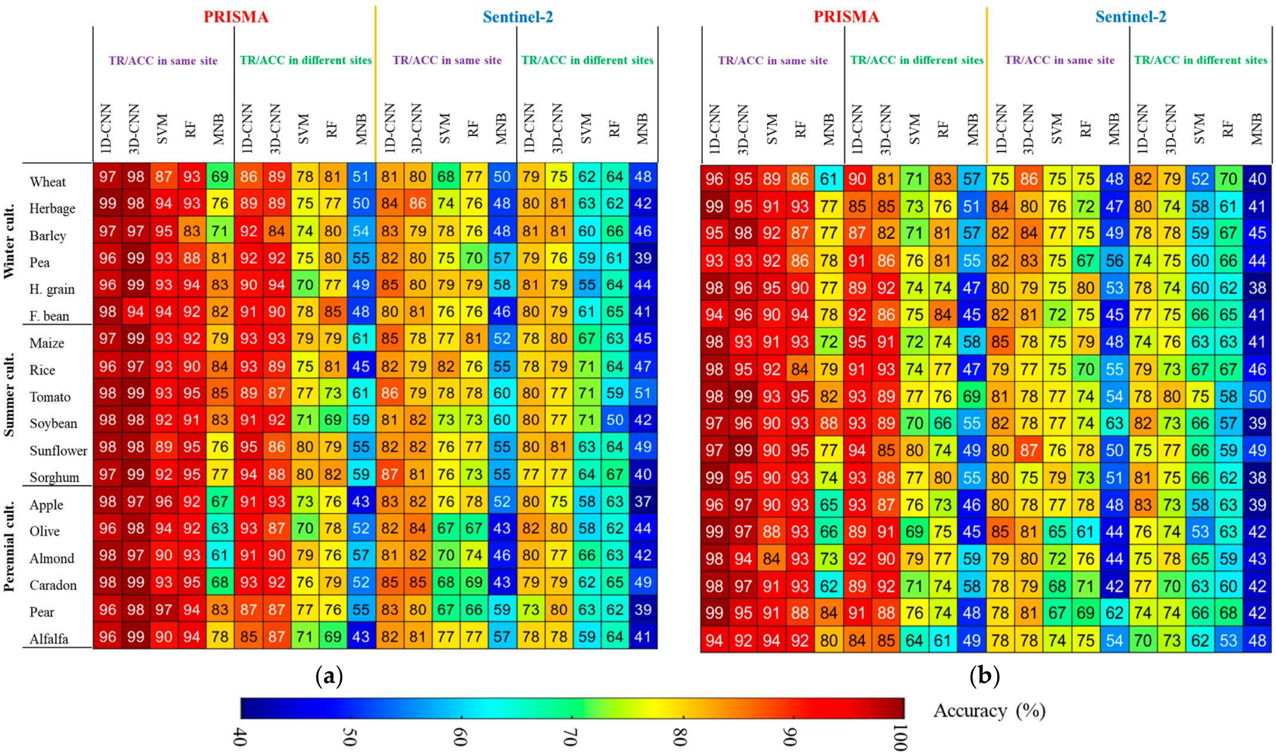

Figure 10. Per-species classification (**a**) PA and (**b**) UA derived from PRISMA and Sentinel-2 images for the MNB, KNN, SVM, RF, 1D-CNN, and 3D-CNN classification methods for winter, summer seasons, and perennial cultivations for all the selected sites.

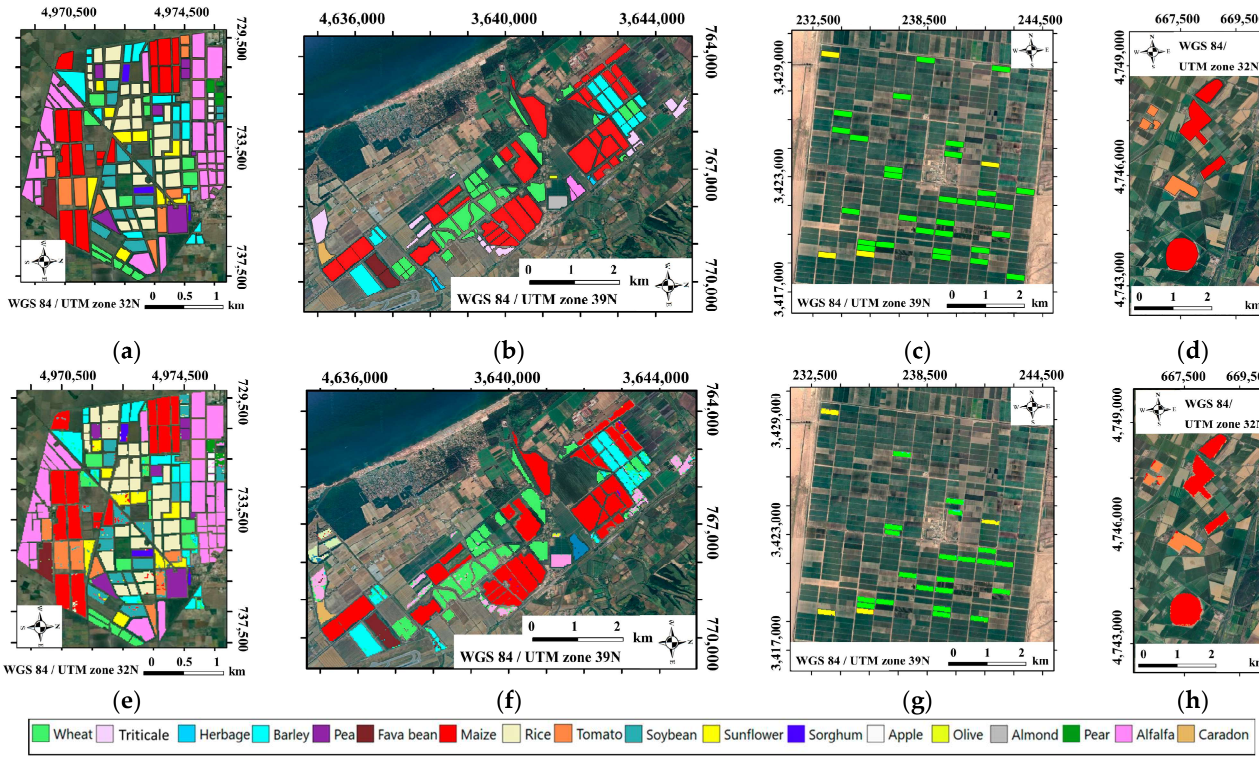

Figure 11. Ground truth of (**a**) Jolanda di Savoia, (**b**) Maccarese, (**c**) MKK, and (**d**) Grosseto. The crop type maps for the entire growing season produced from PRISMA image using (**e**) 3D-CNN of Jolanda di Savoia farm for 2022, (**f**) 1D-CNN method of Maccarese farm for 2021, (**g**) 3D-CNN of MKK farm for 2022, and (**h**) 1D-CNN method of Grosseto farm for 2020.

4. Discussion

4.1. Effects of Reflectance Temporal Variation and Field Heterogeneity on Classifier Performances

In this study, we split the cultivar into three seasonal categories (winter, summer, and perennial). Based on the plant phenology provided by the Sentinel-2 NDVI time series (Figure 5), it is evident that the same-season crops show limited differences in the growing phenology. As an example, the higher accuracy in the detection of cardoon and olive orchards (Figure 8c) occurred in the early season images (related to the earlier growing stage) (Figure 5c), while for peas, the later phenological growing stage led to a one-month delay in comparison with other winter-season species (Figure 5a).

Among the species, alfalfa showed the lowest accuracy as a direct factor related to the multiple harvesting patterns (Figure 5c) and the difficulties in obtaining PRISMA images suitable to catch its growing behavior. This is also depicted in Figure 6c, where alfalfa shows low spectral similarity with the winter crops while showing high similarity with respect to the summer crops growing in accordance with the harvesting cycle. As mentioned in the study [65], temporal resolution is more important than spectral resolution for alfalfa (the high capability of Sentinel-2 time series data for alfalfa mapping is also highlighted in the literature).

The spectral signatures of the analyzed crops do not show clearly differences between species (Figure 6e). These high spectral similarities lead to increasing the classification omission and commission errors between barley, wheat, and triticale species. The reason is related both to the similarity between the growing calendar (Figure 5) and their spectral behavior (Figure 6).

This result agrees with the work of Wilson et al., 2015 [65], that taking advantage of the growing season maximizes the accuracy in the discrimination of soybeans, canola, wheat, oats, and barley using hyperspectral reflectance. Furthermore, in accordance with Buchhart et al. [66], the crop temporal spectral signatures change according to the growing season, as pointed out by Figure 7. This is connected to the change in the plant's biophysical/biochemical characteristics (e.g., canopy structure, chlorophyll, and water content [67]), in accordance with the different growth stages [68]. Methods like feature selection applied to spectral data [69] could be useful to minimize the influence of growth stages on crop identification accuracy, but this still needs to be further analyzed in accordance with the study of Sun et al. [70] and Graeff and Claupein [71], which observed that reflectance temporal changes pattern should be considered within the algorithm configuration. This issue can be better faced taking advantage of the increasing availability of hyperspectral satellites offered by EnMAP, EMIT, and DESIS, and upcoming missions like CHIME and PRISMA-SG. At this scope, the conventional CNNs or trending sequence models, like recurrent neural networks (RNNs) and transformers, can hardly address (i.e., simultaneously and efficiently) the spectral/temporal variability issues. For future work, the authors' suggestion is to tackle this variability issue by implementing state space models like SpectralMamba [72].

4.2. Effects of Field Heterogeneity on Classifier Performances

The spectral differences that can occur within the same field could be related to the different soil properties in term of texture and organic carbon (soil fertility) that could affect the growing rate and then the biophysical and biochemical properties of plants [73]. These changes in soil properties have a clear impact on the spectral signature that could lead to an incorrect classification. This has been observed clearly at the Jolanda di Savoia site, where texture within a field could vary significantly where paleo riverbeds determine sharp variations in the soil fertility [74] as they are related to sharp passages from clay to silty soils (Figure 12c). Figure 12a shows the 2022 crop map identifying the wheat, while Figure 12b shows the 3D-CNN classification results, leading to an erroneous classification of the sparse area classified as barley instead of wheat with a few sparse pixels of pea that, as shown in Figure 6e, are characterized by a high similarity impacting the covariance used by the classification algorithms. Figure 12c shows the PRISMA true-color RGB of an image acquired on 30 April 2022 from the Jolanda di Savoia farm. Two different areas, depicted in

Figure 12a, with red and yellow boxes, correspond to "sparse" and "dense" wheat selected to compare their reflectance, which is shown in Figure 12d. The spectral difference related to sparse and dense vegetation appears to be important along all the spectra, a difference large enough to confound the algorithm in the crop species detection.

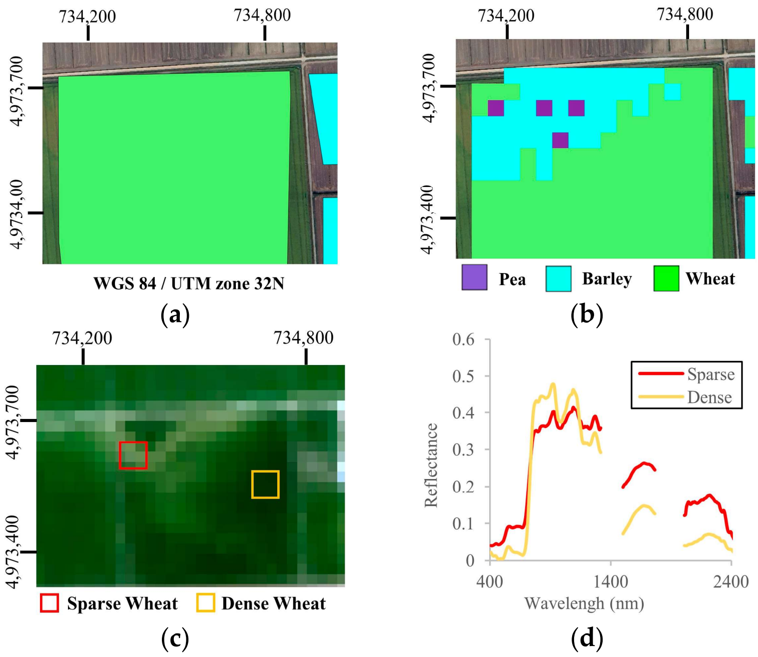

Figure 12. (**a**) Ground truth, (**b**) classification result of a non-homogeneous field at Jolanda di Savoia farm with high misclassification in a field cultivated with durum wheat, (**c**) PRISMA RGB (bands 32-21-10) acquired on 30 April 2022, (**d**) spectral reflectance of sparse and dense wheat.

4.3. Early-Stage Crop Mapping Using ML and DL Algorithms

Regarding the performances of the different algorithms tested, for all cultivations, (Figure 9) both CNN (1D and 3D) methods showed to be the best performing classifiers, achieving OAs higher than 88% and 79% for PRISMA and Sentinel-2, respectively. In all the cases we considered, 1D-CNN produced higher OA than 3D-CNN when applied to the PRISMA imagery. Similarly, for the Sentinel-2 time series, with the exception of the cross-farm TR/ACC scenarios applied to the perennial cultivation, the 1D-CNN provided higher OA than the 3D-CNN (Figure 9).

Following 1D-CNN and 3D-CNN, the RF (OA higher than 76% and 62% for PRISMA and Sentinel-2, respectively) and SVM (OA higher than 73% and 60% for PRISMA and Sentinel-2, respectively) algorithms gained the next level of accuracy. The accuracy of KNN (OA higher than 67% and 58% for PRISMA and Sentinel-2, respectively) was lower than RF and SVM, while the result of the MNB (OA higher than 46% and 43% for PRISMA and Sentinel-2, respectively) method performed not well, showing lower accuracy (Figure 9).

If we analyze the classification performance along the available image time series, assuming an OA of 80% as an acceptable threshold for a suitable crop mapping, then, at DOY 81, all winter species can be recognized by PRISMA images (Figure 8a). The triticale showed the maximum OA earlier than wheat and barley in accordance with the different winter development among crops (Figure 8a). Wheat and triticale had lower OA until the end of the season (Figure 8a). The spectral similarity between barley triticale and winter wheat was high (Figure 6e), which led to confusion for the classifiers. The fava beans gained the best OA after DOY 102, as it was related to later growth start with respect to the other species (Figure 7a).

For summer cultivation, DOY 172 was the first time for which all species could be separated by PRISMA images (Figure 8b). Rice OA reached the maximum earlier and then started to reduce after DOY 190. Sorghum and soybeans could be identified 20 days earlier than other species, while maize and sorghum showed the better identification after DOY 185 (Figure 5b). For perennial cultivation, DOY 91 was the first time for which they could be separated by PRISMA images (Figure 6c) and, among them, almonds could be identified 20 days earlier than other species. Cardoon showed the highest user accuracy until DOY 91 (OA up to 92%). This was also related to the fact that cardoons started to grow before other species (Figure 5b), which led to having less similarity in reflectance during the first growing stages.

We were able to estimate the cultivated area for maize, tomato, and cardoon two months ahead of harvest. For soybeans, wheat, barley, triticale, and peas, the estimation could be carried out about three months before the harvest, while, for rice, sunflower, and sorghum, we could estimate their cultivated areas about four months before the harvest. This information is very useful for a range of applications, including monitoring for compliance with agricultural policies such as the EU Common Agricultural Policy (CAP), in terms of surfaces declared by farmers for crops. Additionally, since the proper identification of cultivated areas is a fundamental requirement for yield forecast analysis at the regional scale, this information is extremely valuable for entities dealing with both food security and food commodity trading, as well as insurance. This supporting information allows to detect at the earliest the critical regions that could suffer food crises, allowing local governments and international humanitarian organizations to plan remediation activities [75]. Moreover, it is crucial to have current information about the planting locations and potential supplies of cereals (maize, wheat, and rice) and legumes (soybeans, peas) months before harvest in order to predict more accurately yields and expected food prices [76].

4.4. The Effects of Pixel/Field Size on the 3D-CNN

As shown in Figure 9, the 1D-CNN and 3D-CNN show better performance in comparison with other classical ML algorithms and the 1D-CNN shows a slight superiority to 3D-CNN. The superiority of 1D to 3D-CNN was not expected, because 3D-CNN takes advantage of the spatial information besides spectral data, while the 1D-CNN method is only based on spectral data and considers each pixel as independent. To better understand the effect of field size on the performance of these two methods, the fields were classified based on their width into three categories (small when <150 m, moderate in the range 150–250 m, and large when >250 m). The Kappa coefficient of 1D-CNN and 3D-CNN methods for these categories are shown in Figure 13. The Kappa coefficients for large field classifications using 3D-CNN and PRISMA and Sentinel-2 images were 0.839 and 0.74, respectively. For small fields, the Kappa coefficients were reduced to 0.75 and 0.66, respectively. The Kappa coefficient of the 1D-CNN method did not show a significant change between different categories for PRISMA- and Sentinel-2 image-derived maps.

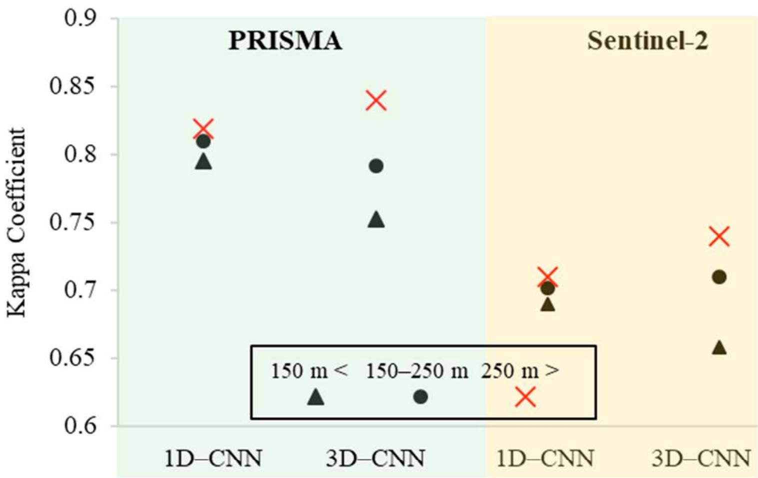

Figure 13. Kappa coefficient of small (<150 m), moderate (150–250 m), and large (>250 m) fields provided by PRISMA and Sentinel-2 images.

The limited results of 3D-CNN with respect to the 1D-CNN were hence strictly related to the small field size and, since most of the Maccarese and Jolanda di Savoia fields had widths lower than 200 m (Figure 1d), this made it difficult for the 3D-CNN algorithm to perform well. It is, therefore, expected to overcome this limited performance of the 3D-CNN in the case of future pansharpening of the hyperspectral at 30 m to 5 m resolution offered by the co-registered PRISMA PAN or by new missions like PRISMA-2 (having a higher spatial resolution mode).

4.5. Cross-Farm TR/ACC for Unavailable Data Situation

Comparison of the OA results between the same-farm and the cross-farm TR/ACC scenarios (Figure 9) showed significant differences in the classifiers' performances. Regarding PRISMA data (reported in panels (a) and (b)), apart from a general decrease in accuracy, it should be noted that the CNN algorithms presented a limited reduction of about 6% in OA, while for the SVM, RF, and MNB algorithms, the OA was reduced by 17%, 15%, and 22%, respectively. The stability of CNNs in the cross-farm TR/ACC scenario seems quite promising and deserves consideration in future studies. The results (Figure 9) show that the classifiers' performances in terms of accuracy were lower with respect to the cross-farm TR/ACC compared with the same-farm scenario. With respect to the same-farm scenario, just a limited reduction of 6% OA was observed by using the CNN algorithms, while OA was significantly reduced by 17%, 15%, and 22% for SVM, RF, and MNB, respectively. The reason behind the stability of CNNs in comparing the ML algorithms in the cross-farm TR/ACC scenario needs additional consideration.

The transfer learning scheme based on deep learning approaches helps to gain crop mapping results in transfer sites, even though abundant historical labels of crop type are required. This study states an innovative data set offered by the combination of hyperspectral imagery and known related ground truth. As a preliminary exploitation of the hyperspectral time series potential, we decided to use trivial cross-farm approaches to simply mimic more complex learning approaches testing different TL types and theories (e.g., fine-tuning, feature extraction, few-shot learning, domain adaptation).

Moreover, this study highlights some points that are to be considered for further study. Dimension reduction/feature extraction being performed at present with a trivial PCA could be enhanced with more sophisticated feature extraction methods, better preserving the spectral information contents, or balanced with techniques like stacked auto-encoder (SAE). Furthermore, the self-attention mechanism could be used to focus on the most informative features. Future work should explore the capability of developed foundation models (e.g., SAM, GPT, DINO) in excluding the feature extraction step within the present ML workflow.

5. Conclusions

In this study, we exploited the first available PRISMA time series collected by the hyperspectral satellite over the more densely surveyed agricultural test sites located in Europe and Iran. PRISMA time series were processed to tune and test ML and DL classification approaches for large-scale crop mapping in agriculture, with particular attention paid to early crop mapping. The comparison of the performance of six classification methods—MNB, KNN, RF, SVM, 1D-CNN, and 3D-CNN—on PRISMA and Sentinel-2 time series confirmed the powerfulness of the PRISMA times series when combined with the 1D-CNN algorithm in the production of an accurate crop mapping and in the definition of a precise early crop map product. All common crops from three agricultural sites in Italy and from a region with strong diverse agronomy in South Iran were in fact accurately detected. 1D-CNN showed the highest accuracy in the crop type identification, followed by 3D-CNN, RF, SVM, and KNN as the next accurate methods (89%, 91%, and 92% for winter, summer, and perennial cultivations, respectively). Classification accuracy remained comparable whether a cross-farm or single-farm training scenario was utilized (reduction of 6% in AO). The product accuracy obtained by using the Sentinel-2 time series was substantially lower