

AUTOMATED CLASSIFICATION OF CROP TYPES AND CONDITION IN A MEDITERRANEAN AREA USING A FINE-TUNED CONVOLUTIONAL NEURAL NETWORK

Ruiz, L.A. ¹ *, Almonacid-Caballer, J. ¹, Crespo-Peremarch, P. ¹, Recio, J.A. ¹, Pardo-Pascual, J.E. ¹, Sánchez-García, E. ¹

¹ Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 València, Spain- (laruiz, jaialca, pabcrepe, jrecio, jepardo, elsnegar))@upv.es

Commission III, WG III/10

KEY WORDS: Crop classification, Convolutional neural network, VGG-19, Crop abandonment, Orthoimages

ABSTRACT:

Crop classification based on satellite and aerial imagery is a recurrent application in remote sensing. It has been used as input for creating and updating agricultural inventories, yield prediction and land management. In the context of the Common Agricultural Policy (CAP), farmers get subsidies based on the crop area cultivated. The correspondence between the declared and the actual crop needs to be monitored every year, and the parcels must be properly maintained, without signs of abandonment. In this work, Sentinel-2 time series images and 4-band Very High Resolution (VHR) aerial orthoimages from the Spanish National Programme of Aerial Orthophotography (PNOA) were combined in a pre-trained Convolutional Neural Network (CNN) (VGG-19) adapted with a double goal: (i) the classification of agricultural parcels in different crop types; and (ii) the identification of crop condition (i.e., abandoned vs. non-abandoned) of permanent crops in a Mediterranean area of Spain. A total of 1237 crop parcels from the CAP declarations of 2019 were used as ground truth to classify into *cereals*, *fruit trees*, *olive trees*, *vineyards*, *grasslands* and *arable land*, from which 80% were used for training and 20% for testing. The overall accuracy obtained was greater than 93% both, at parcel and area levels. Olive trees were the least accurate crop, mostly misclassified with fruit trees, and young vineyards were slightly confused with cereal and arable land. In the assessment of crop condition, only 9.65% of the abandoned plots were missed (omission errors), and 7.21% of plots were over-detected (commission errors), having a 99% of overall accuracy from a total of 1931 image subset samples. The proposed methodology based on CNN is promising for its operational application in crop monitoring and in the detection of abandonments in the context of CAP subsidies, but a more exhaustive number of training samples is needed for extension to other crop types and geographical areas.

1. INTRODUCTION

Crop classification using satellite and aerial imagery is a recurrent application in remote sensing, obtaining useful input for creating and updating agricultural inventories at different scales, yield prediction (Doraiswamy et al., 2005), generation of phenology maps, and land management. Different data sources and resolutions have been used for crop classification, such as low resolution MODIS (Moderate Resolution Imaging Spectroradiometer) images (250 m/pixel) (Wardlow et al., 2007) and mid-resolution Landsat images (30 m/pixel) (Devadas et al., 2012) for large-area crop mapping of extensive crops (cereal, soybean, alfalfa,...); four-band very high-resolution (VHR) images (Ozdarici-Ok et al., 2015), combining optical and synthetic aperture radar (SAR) images (Recio et al., 2011; Kussul et al., 2016), which provide information about vegetation structure and biochemical properties (Orynbaikyzy et al., 2019), or a more sophisticated integration of LiDAR data and hyperspectral airborne images for crop species classification (Liu and Bo, 2015).

In addition to the integration of different data sources, a variety of methodological approaches have been tested. Particularly efficient is the use of time-series datasets to characterise the temporal signature of crops along the year, allowing the classification of intensive crops with dynamic vegetation changes and smaller parcel size, such as the case of smallholder crop classification (Lambert et al., 2018). The emergence of

VHR sensors has supposed a transition from pixel-based to object-based classification methods, which are more suited for the extraction of contextual features to enhance classification. They have been tested with success by combining different mid-resolution (Peña-Barragán et al., 2011) and VHR data sources (Devadas et al., 2012; Abdullah Sohl et al., 2015; Liu and Bo, 2015). However, image segmentation for the definition of objects is a common source of errors working in agricultural landscapes, especially in spectrally heterogeneous crops, such as vineyards, fruit trees and orchards, where vegetation and soil coexist in the same parcel. In these cases, parcel-based approaches seem to work better, where objects are directly obtained from the agricultural database parcel boundaries (Ruiz et al., 2007). This approach has been applied in fragmented agricultural landscapes or permanent crops (Ruiz et al., 2009; Schmedtmann et al., 2015; Kussul et al., 2016).

In the context of the Common Agricultural Policy (CAP), farmers get subsidies based on the area cultivated. The correspondence between the declared and the actual crop needs to be monitored by the Governments of European countries every year. Furthermore, farmers should be subsidized only if the crop is properly maintained and there are no signs of abandonment. During the last years Copernicus data, in particular Sentinel-1 and Sentinel-2 images, are made freely available and their use is fostered by the European Union with the objective of enabling greater transparency and comparability of CAP results in different Member States (Kanjir et al., 2018).

* Corresponding author

In addition, national high-resolution image sets, such as aerial orthoimages from the National Programme of Aerial Orthophotography (PNOA) in Spain, even if they do not have a temporal dimension, can provide annual VHR four-band multispectral data very useful for the monitoring of the CAP declarations. In this context of CAP monitoring system, since the subsidies payment is made per agricultural administrative plot, parcel-based image classification seems to be a straightforward and efficient method in terms of completeness of feature extraction, noise reduction and processing time.

The use of satellite imagery time-series to specifically develop a CAP monitoring system had been tested first using series of Landsat ETM+ images following a parcel-based approach (Schmedtmann and Campagnolo, 2015). Later, Sitokonstantinou et al. (2018) compared parcel-based classification using Landsat-8 and Sentinel-2 images, obtaining superior performance of the latter due to the improved spatial, spectral and temporal characteristics, and Campos-Taberner et al. (2019) classified Sentinel-1 and Sentinel-2 images following a pixel-based approach. However, the use of VHR images has not been reported yet for this purpose.

Less attention has been paid to the identification of crop abandonment using remote sensing techniques, which is also a relevant issue to monitor in the context of the CAP. Alcántara et al. (2012) tested methods to map abandoned agriculture at broad scales with coarse-resolution satellite imagery (MODIS) using NDVI time series from 2003 to 2008 and phenology data, obtaining a classification overall accuracy of 65%. Yusoff and Muharam (2015) used Landsat images, crop phenology information and an object-oriented classification to detect crop abandonment at a finer scale, and Hermosilla et al. (2012) classified aerial high resolution images where some basic abandonment classes were included.

Crop classification has been done using different methods. Initially it was based on traditional statistical classification methods, later machine learning techniques were used, such as decision trees, random forest, support vector machine, etc. During the last several years, deep learning techniques based on the massive use of training samples are opening new perspectives. Among them, convolutional neural networks (CNN), which are able to learn automatically from raw images and to avoid the use of specifically designed features, definitely have a potential that needs to be explored in different applications. In this sense, Hu et al. (2018) proposed an improved CNN to automatically construct the training dataset and classify Landsat-8 images in generic land cover types, obtaining an overall accuracy improved by 5% and 14% compared to the support vector machine method and the maximum likelihood classification method, respectively. Chang et al. (2019) applied CNN to forest classification, Chen et al. (2019) to the classification of hyperspectral images, and Wang et al. (2019) to classify VHR imagery, also in generic land cover types.

However, CNN have not been studied yet for the specific classification of crops at parcel level using a combination of VHR, multispectral and time-series images. The goals of this work are: (i) to adapt a CNN to the classification of agricultural parcels in different crop types and to evaluate its accuracy for the classification of 4-band (visible and near-infrared) very high-resolution aerial imagery combined with Sentinel-2 time-series; and (ii) to apply this CNN to automatically detect crop condition (i.e., abandoned vs. non-abandoned) of permanent

crops (fruit trees, olive trees and vineyards) in a Mediterranean area.

2. MATERIAL

2.1 Study area and data

The study area is located in the Valencian region on Eastern Spain, in the municipalities of Utiel and Requena (Figure 1). This is a predominantly agricultural area with a majority of permanent crops, being the dominant crop the vineyard, then fruit trees and olive trees, and some other non-permanent crops, mainly cereals.

A total of 1237 crop parcels were selected from the CAP declarations of 2019 in these two municipalities (Figure 1), provided by the Department of Agriculture of the Valencian regional government. The parcels were selected from the predominant crops in the area: vineyard, fruit trees, olive trees, grassland and cereal. From the first three permanent crop types, some of them corresponded to abandoned parcels, as checked in the field during the 2019 CAP campaign. The Spanish Land Parcel Identification System from 2019, published by the Ministry of Agriculture, Fisheries and Food and updated every year, was used to extract the parcel boundaries.

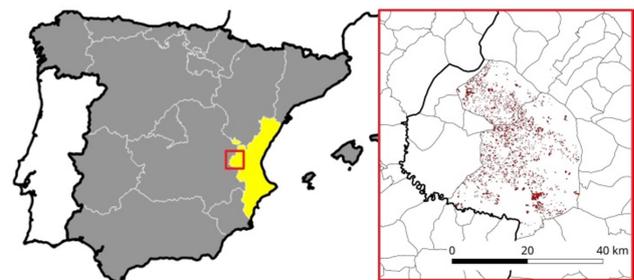


Figure 1. Location of the study area in the Valencian region (Spain): Overview (left) and detail with the distribution of the crop parcels depicted in red (right).

Two types of images were used in this study. On one side, VHR four-band (visible and NIR) orthoimages (0.25 m/pixel resolution) from the Spanish National Programme of Aerial Orthophotography (PNOA) acquired in June 2019 and provided by the Institut Cartogràfic Valencià (ICV). On the other side, a Sentinel-2 image time-series radiometrically corrected at level 2A (bottom of atmosphere) from the European Space Agency (ESA) repository, in the framework of the Copernicus European programme. The time-series was composed of 62 Sentinel 2A and 2B images acquired from September 2018 to September 2019 (Figure 2).

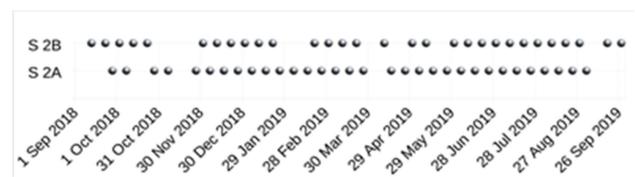


Figure 2. Temporal distribution of the Sentinel 2A and B time-series acquired from September 2018 to September 2019.

Since the Sentinel-2 features extracted for classification in this study were based on the Normalized Difference Vegetation

Index (NDVI), only bands 4 (red) and 8 (NIR) were used, both at 10 m/pixel of spatial resolution. The images were pre-processed at parcel level and only non-cloud pixels were used according to the mask band available in level 2A products of Sentinel-2. Removing remaining outliers followed two conditions. When the number of pixels remaining was less than 90% of the total pixels in the parcel, this image was not used for that parcel. On the other hand, if the NDVI difference of a parcel between a date and neighbouring dates was greater than their mean standard deviation, then this image was neither used for that parcel. In this way, we avoided anomalous pixel values in some dates at parcel level. The pre-processing was finished by obtaining the monthly NDVI average of the values per parcel, from September 2018 to September 2019.

3. METHODS

3.1 Definition of classes and sampling

The classes were defined attending to the main crops in the area and following the legend of the SIGPAC database. Thus, for the crop classification, in addition to classes *vineyard*, *fruit trees*, *olive trees*, *grassland* and *cereal*, *arable land* was also considered, representing those minority crops that at the moment of the orthoimage acquisition did not have vegetation. Given the differences in the *fruit tree* orchards and *vineyards* between the adult trees and the very young trees, two provisional classes were defined only for classification purposes, *young fruit trees* and *young vineyards*, composed of those parcels where the trees were just planted and they did not have yet a prominent vegetative activity (Figure 3).

For the classification of crop condition, only two classes were defined, *abandoned* and *non-abandoned*, corresponding to those permanent crops (*vineyards*, *fruit trees* and *olive trees*) that presented discontinuities in tree rows, abundance of weeds or lack of vegetative activity. Some examples of abandoned crops are also shown in Figure 3.

From the 1237 parcels available, 80% were randomly selected for training the classification process and the remaining 20% for testing, so the evaluation was done with 248 parcels fully independent from those used to train the neural network.

A restriction of the VGG-19 CNN used is that the size of the input images must be constant. Therefore, from our initial data set, a total of 11,836 image subsets of 128x128 pixels each were obtained from the initial sample of parcels available (1237). This means that the original orthoimage underlying each parcel was cut in 128x128 pixels image subsets (from now these will be referred as “samples”). Table 1 shows the number of parcels and samples used per class as training or testing sets.

Classes	N. Parcels		N. Samples	
	Training	Test	Training	Test
Cereal	96	21	1502	156
Fruit trees	369	96	2874	754
Olive trees	98	26	1012	461
Grassland	17	4	475	347
Arable land	46	10	352	167
Vineyard	363	91	3020	716
Total	989	248	9235	2601

Table 1. Distribution of the number of parcels and samples among the different classes, training and testing sets.

Therefore, a total of 9235 image subsets (samples) were used as input to train the CNN, as well as NDVI temporal features derived from Sentinel-2 time-series coming from 989 training parcels. The next two subsections describe the features extracted and the architecture of the CNN.

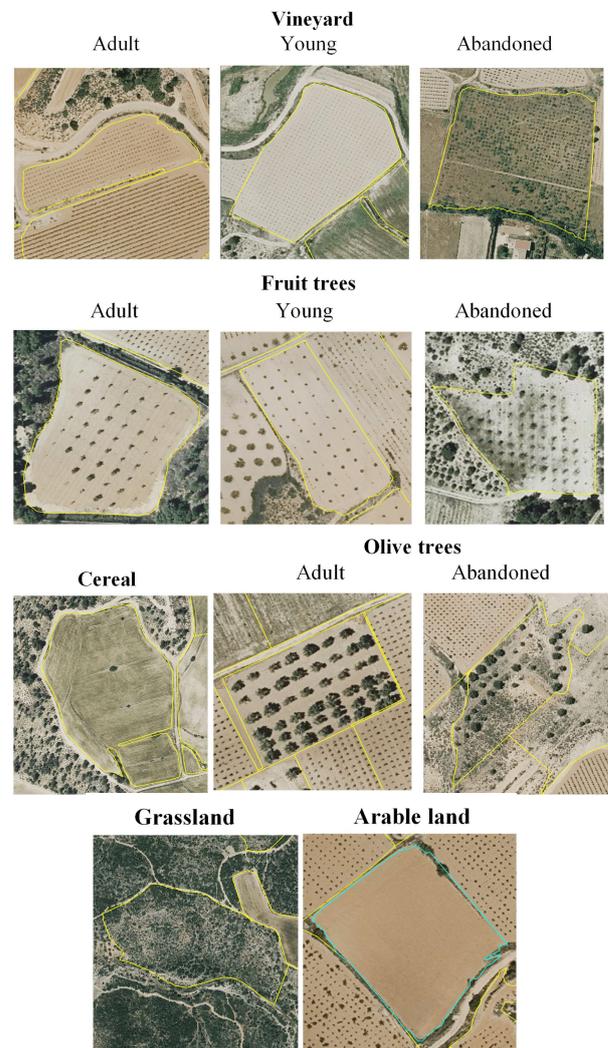


Figure 2. Examples of orthoimage subsets of parcels with the classes defined for crop classification and crop condition.

3.2 Feature extraction

After pre-processing of the Sentinel images, the NDVI was extracted at pixel level for all the images of the time series and its monthly average was computed. Then, the mean and standard deviation of NDVI values were obtained, building a temporal NDVI curve per parcel. In addition of the mean and standard deviation of the 13 monthly NDVI values -from September 2018 to September 2019 (i.e., 26 features) - the following phenology-related features were computed for every temporal NDVI curve per parcel (Figure 4):

- 1) Maximum value of NDVI
- 2) Minimum value of NDVI
- 3) Slope between maximum and minimum NDVI values
- 4) Slope from May to June
- 5) Integral of all the values of the NDVI

- 6) NDVI integration from November 2018 to February 2019
- 7) NDVI integration from March to May 2019
- 8) NDVI integration from May to August 2019

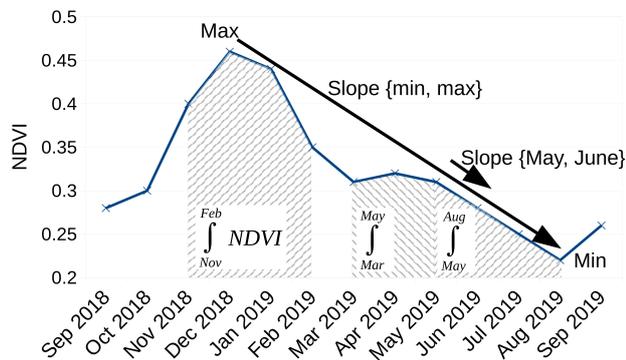


Figure 4. Graphical representation of a temporal NDVI curve with the phenology-related features computed per parcel from the Sentinel-2 images.

At the end, a total of 34 temporal and spectral features derived from Sentinel-2 NDVI curves were available per every parcel considered.

3.3 Description of the convolutional neural network (CNN)

We used the pre-trained VGG-19 CNN. This network has been trained on more than a million 224x224 3-band images from the ImageNet database (Russakovsky et al., 2015) in 1000 different classes. The pre-trained VGG-19 CNN was fine-tuned, keeping the first four convolutional blocks that correspond to more generic aspects of the image, such as edges, but modifying the weights of the last block and the top model by re-training the CNN with our specific training set. The number of neurons of the last layer was also changed from the original value, coinciding with the desired number of output crop classes. A learning rate of 0.001, optimizer SGD, decay 0.0001, momentum 0.9, application of Nesterov momentum and 10 epochs were used for the crop type classification CNN. In the case of the CNN for identification of abandonment, a learning rate of 0.001, optimizer RMSprop and 5 epochs was the optimum set of parameters tested.

We designed two CNN-combined and independent models, one for the classification of crop types and the other to identify abandoned and non-abandoned permanent crops (see Figure 5). The former consists of a multi-class image classification (i.e., eight classes), while the purpose of the latter is that, given a parcel, we could identify if this is abandoned or not (i.e., a binary classification problem). The overall architecture followed for the CNN was (Figure 5): the use of the pre-trained VGG-19 CNN for the convolution block, a NN for the top model of this block, in parallel another NN for the Sentinel-2 features, then both NN merge to end up in a hidden and a final layer, whose number of neurons depends on the image classification type.

More specifically, the pre-trained VGG-19 requires an input of three bands, however, our input images were composed of four bands (red, green, blue and infrared). Therefore, we added a previous convolution 2D with three filters and a kernel filter size of 1x1 to make input images compatible with VGG-19 input. Next, VGG-19 is composed of five convolutional blocks as follows: (i) two convolutions 2D with 64 filters each and a max pooling, (ii) two convolutions 2D with 128 filters each and a max pooling, (iii) four convolutions 2D with 256 filters each and a max pooling, (iv) four convolutions 2D with 512 filters each and a max pooling, and (v) four convolutions 2D with 512 filters each and a max pooling. Among the five convolutional blocks, we only set the last one as trainable, while the others kept the pre-trained values. After the convolutional blocks of VGG-19 we set a flatten and three hidden layers with 4096, 1500 and 300 neurons. Additionally, we applied a dropout of 0.5 in the connection between contiguous layers.

In order to introduce the Sentinel-2 derived features, we set a two-layer NN with 34 and 16 neurons, the first layer serving as input of the 34 features used. Again, we applied a dropout of 0.5 between contiguous layers. Then, we merged the 300 and 16 neurons from the convolutional blocks and Sentinel-2 NN, respectively, to create a fully-connected NN with a hidden layer of 50 neurons, and the output with a number of neurons equal to the eight classes for the crop type image classification, and one for the crop condition (i.e., abandoned or not) binary classification. In the combined top model we used in each hidden layer the ReLU activation, except for the output layer, where we used the softmax and sigmoid activation for the multi-class and binary classifications, respectively.

Finally, in the case of the multi-class classification we obtained the probability that an input belongs to that neuron or class. Therefore, the class with the higher probability is the one that is assigned to that input. For the binary classification, we obtained the probability that an input belongs to *non-abandoned*, so if the probability is below 50% it will belong to the class *abandoned*.

3.4 Evaluation

The evaluation of the crop classification results was performed at three levels: sample-based, parcel-based and area-based. The 20% of the total parcels (i.e., 248) was used as testing set. From these parcels, 2601 128x128 pixels image subsets were extracted for the sample-based evaluation. Since the payment of the CAP subsidies is applied by area, from the economic point of view it is important to know the accuracy of the classification at area level, so the losses due to errors can be quantified. In order to perform area-level evaluation the results obtained at parcel-level were weighted by the area occupied by each parcel.

The error matrices of the classification were computed at the three levels, then three standard indices were obtained: the overall accuracy, as the percentage of items correctly classified; the producer's accuracy, as the percentage of items belonging to a class correctly classified; and the user's accuracy, as the percentage of correct items classified in every particular class.

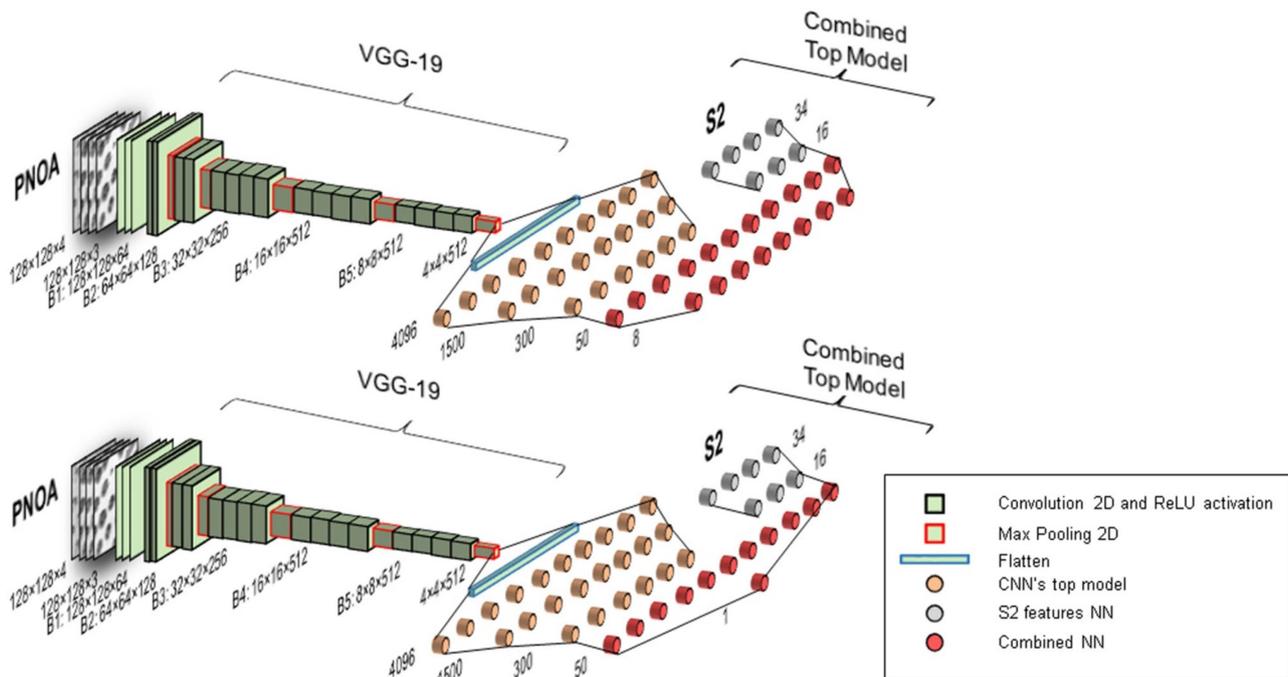


Figure 5. Architecture of the fine-tuned VGG-19 convolutional neural network used for crop type classification (above) and identification of crop abandonment (below) (PNOA refers to the input images used from the Spanish National Programme of Aerial Orthophotography).

4. RESULTS AND DISCUSSION

In the following two sub-sections we present and analyse the results obtained in the evaluation of crop type classification and in the detection of abandonment in permanent crops.

4.1 Crop classification

Table 2 and Figure 6 summarize the results obtained in the classification of six different crop types. Although initially eight classes were considered, after the classification, preliminary classes *young fruit trees* and *young vineyard* were eventually merged with classes *fruit trees* and *vineyard*, respectively.

Overall classification accuracies are greater than 93% at the three classification levels, but the area-based evaluation shows better results, as represented in Figure 6. This can be explained because larger parcels, which have more weight in this type of evaluation, are better classified than smaller parcels. Larger parcels are easier to characterize, since more image subsets are involved and the errors are easily masked out. Additionally, they usually have a more even distribution of plants as they are more modern and more homogeneous crop growing agricultural practices are applied, as opposite to older smallholding parcels.

The lowest producer's and user's accuracies correspond to class *olive trees*, which is mostly confused with *fruit trees*. There is an evident similarity of these two permanent crops, their spectral response along the year does not change much, so the Sentinel-2 derived features do not provide much information. On the other side, their structure, understood as the distribution of trees in the parcel, is very similar in both crops. Their main difference is related to the separation of trees in the plot, which in the case of olive trees is greater. Since the image subsets that

are introduced in the CNN are only 128x128 pixels, and the orthoimages spatial resolution is 0.25 m/pixel, the area covered could not be sufficient in some cases to characterize the structure of the olive trees. A potential solution for this could be to increase the size of the image subsets, as well as the introduction of some structural features derived from the VHR images in the CNN. In this sense, Balaguer et al. (2010) and Ruiz et al. (2011) proposed some structural features derived from the semivariogram and from the Hough transform of the parcel VHR images to classify tree crops, obtaining promising results. In the future, the inclusion of these types of features as input in the CNN could help to improve the discrimination of these two classes. The results of olive trees are slightly lower when evaluated per parcel with respect to per sample, but they improve when are evaluated per area. This reinforces the previous argument that small parcels are more difficult to characterize and, subsequently, to classify.

Other errors are related to the misclassification of *vineyards* with *cereal* and *arable land*. In the first case, the temporal features derived from Sentinel-2 image series should have more relevance in the classification, since the phenology and agricultural calendar of both crops are different. This leads to think that the 34 Sentinel-2 temporal features have considerably less influence than the VHR images, probably because they are introduced in the last part of the CNN, the so called top model, and only merge and interact with the neurons conveying information of the VHR images in a hidden layer at the end of the NN (see Figure 5). Future work should be focused on providing more relevance to those features in the classification by modifying accordingly the CNN structure. In the second case, the misclassification of *vineyards* with *arable land*, this basically happens with the very young vineyard plantations, where plants are still too small and the proportion of soil, like in arable land parcels, is much greater.

Classes	Samples (2601)		Parcels (248)		Area (393 ha)	
	PrAcc	UsAcc	PrAcc	UsAcc	PrAcc	UsAcc
Cereal	95.5%	82.8%	100%	87.5%	100%	86.3%
Fruit trees	93.4%	94.6%	96.9%	90.3%	98.8%	94.0%
Olive trees	89.6%	91.4%	61.5%	84.2%	83.7%	96.5%
Grassland	99.1%	96.4%	100%	100%	100%	100%
Arable land	98.2%	85.4%	100%	90.9%	100%	96.0%
Vineyard	94.0%	99.6%	95.6%	100%	95.7%	100%

Table 2. Crop type classification results using the proposed CNN, presented in terms of samples, parcels and area-based. (OvAcc: Overall Accuracy; PrAcc: Producer’s Accuracy; UsAcc: User’s Accuracy).

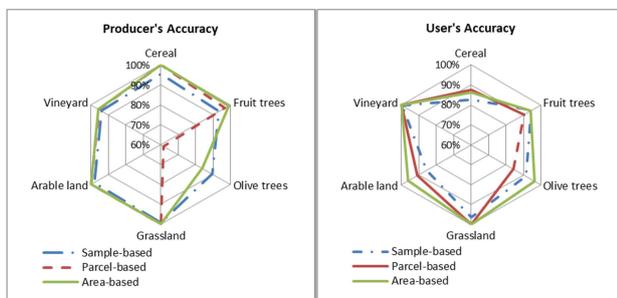


Figure 6. Comparison, in terms of producer’s and user’s accuracies, of the three types of evaluation performed at sample, parcel and area levels for crop classification using the described CNN.

As pointed earlier, several works reveal the complementarity of time series of multispectral and SAR images (e.g., Sentinel-2 and Sentinel-1) for crop monitoring. However, even if VHR aerial images had not been tested yet together with Sentinel-2 time series for crop classification and CAP monitoring, they have shown a great synergy in this study for the accurate classification of a combination of annual and permanent crops. Good results were reported for classification of annual crops, such as rice, corn, soybean (Xu et al., 2019), in addition to cereals, sunflower and even vineyards (Sitokonstantinou et al., 2018), but permanent crops such as fruit or olive trees have very subtle changes in their vegetation indices time series curves along the agricultural year, so the use of at least one VHR image can be decisive to increase the classification accuracy of these crop types, particularly when parcel sizes are small, such as the case of some areas in the Mediterranean region.

The use of CNN is opening new perspectives for crop classification, feature extraction is not necessary or may be drastically reduced, saving computing time and simplifying the classification procedure. However, compared to other methods, CNN needs to substantially increase the number of training samples to improve the results and have robustness (Chen et al., 2019). In addition, their architecture must be carefully designed to optimize the results, and when using different data sources like in our case (VHR orthoimages and Sentinel-2 time series),

the merging of these datasets should be tuned depending on the crop types in the area.

4.2 Identification of abandoned crops

Since not a sufficient number of abandoned parcels was available for a proper evaluation at parcel-level, the crop condition, i.e., the identification of abandoned crops, was evaluated only using the image subsets extracted from the original parcels. A total of 1931 samples were used, from which 114 were *abandoned* and 1817 *non-abandoned*, Table 3 shows the error matrix with only these two classes, as well as the overall accuracy and the accuracies of producer and user. Only abandoned parcels from classes *vineyards*, *fruit trees* and *olive trees* were used.

Since most of the parcels belong to *non-abandoned* class, the overall accuracy (99 %) is not a robust accuracy indicator in this case, being biased by the much greater proportion of samples from this class. However, producer’s and user’s accuracies of class *abandoned* are robust, providing information about the proportion of abandoned samples correctly detected, and the proportion of samples detected as abandoned that are correct, respectively. Interpreting the results, only 9.65% of the abandoned plots from any of the three classes considered were missed (omission errors), and 7.21% of plots were over-detected (commission errors).

Reference	Classified as	
	Abandoned	Non-abandoned
Abandoned	103	11
Non-abandoned	8	1809
PrAcc	90.35%	99.56%
UsAcc	92.79%	99.40%
OvAcc	99.02%	

Table 3. Results in the identification of abandoned crops. Only abandoned parcels from classes *vineyards*, *fruit trees* and *olive trees* were used. (OvAcc: Overall Accuracy; PrAcc: Producer’s Accuracy; UsAcc: User’s Accuracy).

These can be considered as good results for the operational application in the detection of abandonments in the CAP monitoring tasks. Even being aware that more classes should be tested, the three classes used in this study are probably the most relevant in the Mediterranean region of Spain, where a main part of these type of permanent crops are becoming economically unprofitable for traditional farmers, so the rate of abandonment is increasing. Thus, given a declared parcel for the CAP subsidies in one of these classes, using this CNN-based method abandoned parcels could be automatically detected with a low rate of errors.

Previous studies mapping abandoned agriculture over large areas with MODIS data obtained an overall classification accuracy of 65% (Alcántara et al., 2012). Even if these studies are not comparable due to the differences in spatial resolution used, our results are promising for the automated identification of abandonment in permanent crop parcels for operational use in CAP monitoring.

5. CONCLUSIONS

We fine-tuned a pre-trained convolutional neural network (VGG-19) with two different goals: the classification of crop types at parcel-level, and the identification of abandoned parcels with permanent crops. The input datasets were 4-band VHR aerial orthoimages and temporal and spectral features derived from Sentinel-2 time series images. After applying and testing it on a consistent set of parcels with known crop type and condition, the results show that the overall accuracy of the tested crop types is over 93%. Errors affect more to small parcels, olive trees (confused with fruit trees), and vineyards (confused with cereal and arable land). In order to increase the classification accuracy, larger input image subsets and the inclusion of some selected VHR image features may be needed to better capture the distribution of patterns of the trees in the parcel. In addition, a further modification of the CNN architecture to enhance the interaction of VHR images and Sentinel-2 time series derived features could increase the synergy of these two datasets.

The proposed CNN was able to distinguish between abandoned and non-abandoned permanent crops with a very low error rate, which is encouraging for the operational application in the detection of abandonments in the CAP monitoring tasks. However, future tests should be done in both applications by including more crop types and increasing the number of training samples, in order to improve robustness and applicability to different areas.

ACKNOWLEDGEMENTS

This research has been funded by the Conselleria d'Agricultura, Medi Ambient, Canvi Climàtic i Desenvolupament Rural, Generalitat Valenciana, through the nominative line S847000. The authors also thank the Institut Cartogràfic Valencià for providing very high resolution aerial imagery data of the study area.

REFERENCES

- Abdullah Sohl, M., Schlager, P., Schmieder, K., Rafique, H.M., 2015. Bioenergy crop identification at field scale using vhr airborne cir imagery. *Photogramm Eng Remote Sensing*, 81(8), 669-677. doi.org/10.14358/PERS.81.8.669
- Alcantara, C., Kuemmerle, T., Prishchepov, A. V., & Radeloff, V. C., 2012. Mapping abandoned agriculture with multi-temporal MODIS satellite data. *Remote Sens Environ*. https://doi.org/10.1016/j.rse.2012.05.019
- Balaguer, A., Ruiz, L.A., Hermosilla, T., Recio, J.A., 2010. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. *Comp and Geosc*, 36(2), 231-240. doi:10.1016/j.cageo.2009.05.003
- Campos-Taberner, M., García-Haro, F. J., Martínez, B., Sánchez-Ruiz, S., & Gilabert, M. A., 2019. A copernicus sentinel-1 and sentinel-2 classification framework for the 2020+ European common agricultural policy: A case study in València (Spain). *Agronomy*, 9(9). https://doi.org/10.3390/agronomy9090556
- Chang, T., Rasmussen, B. P., Dickson, B. G., & Zachmann, L. J., 2019. Chimera: A multi-task recurrent convolutional neural

network for forest classification and structural estimation. *Remote Sens*, 11(7), 1–29. https://doi.org/10.3390/rs11070768

Chen, C., Ma, Y., & Ren, G., 2019. A convolutional neural network with Fletcher-Reeves algorithm for hyperspectral image classification. *Remote Sens*, 11(11). https://doi.org/10.3390/rs11111325

Devadas, R., Denham, R.J., Pringle, M., 2012. Support vector machine classification of object-based data for crop mapping, using multi-temporal landsat imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B7, 185–190. doi.org/10.5194/isprsarchives-xxxix-b7-185-2012

Doraiswamy, P.C., Sinclair, T.R., Hollinger, S., Akhmedov, B., Stern, A., Prueger, J., 2005. Application of MODIS derived parameters for regional crop yield assessment. *Remote Sens Environ*, 97(2), 192-202. doi.org/10.1016/j.rse.2005.03.015

Hermosilla, T., Díaz-Manso, J.M., Ruiz, L.A., Recio, J.A., Fernández-Sarría, A., Ferradáns-Nogueira, P., 2012. Analysis of parcelbased image classification methods for monitoring the activities of the Land Bank of Galicia (Spain). *Appl. Geomat.*, 4(4), 245-255. doi.org/10.1007/s12518-012-0087-z

Hu, Y., Zhang, Q., Zhang, Y., & Yan, H., 2018. A deep convolution neural network method for land cover mapping: A case study of Qinhuangdao, China. *Remote Sens*, 10(12). https://doi.org/10.3390/rs10122053

Kanjir, U., Đurić, N., Veljanovski, T., 2018. Sentinel-2 based temporal detection of agricultural land use anomalies in support of common agricultural policy monitoring. *ISPRS Int. J. Geoinf.*, 7(10), 405. doi.org/10.3390/ijgi7100405

Kussul, N., Lemoine, G., Gallego, F.J., Skakun, S.V., Lavreniuk, M., Shelestov, A.Y., 2016. Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data, 9(6), 2500-2508. *IEEE J Sel Top Appl Earth Obs Remote Sens*. doi.org/10.1109/JSTARS.2016.2560141

Lambert, M.J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens Environ*. 216, 647-657. doi.org/10.1016/j.rse.2018.06.036

Liu, X., & Bo, Y., 2015. Object-based crop species classification based on the combination of airborne hyperspectral images and LiDAR data. *Remote Sens*. 7, 922-950. https://doi.org/10.3390/rs70100922

Orynbaikyzy, A., Gessner, U., Conrad, C., 2019. Crop type classification using a combination of optical and radar remote sensing data: a review. *Int J Remote Sens*, 40(17), 6553–6595. doi.org/10.1080/01431161.2019.1569791

Ozdarici-Ok, A., Ok, A.O., Schindler, K., 2015. Mapping of agricultural crops from single high-resolution multispectral images-data-driven smoothing vs. parcel-based smoothing. *Remote Sens.*, 7(5), 5611-5638. doi.org/10.3390/rs70505611

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens Environ.*, 115(6), 1301–1316. doi.org/10.1016/j.rse.2011.01.009

Recio, J.A., Ruiz, L.A., Hermosilla, T., Herrera-Cruz, V., Fernández-Sarría, A., 2011. Combination of TERRASAR-X and Optical Imagery for LU/LC Mapping using an Object-Based Approach. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-4/W19, 259-264. 10.5194/isprsarchives-XXXVIII-4-W19-259-2011

Ruiz, L.A., Recio, J.A., Hermosilla, T., 2007. Methods for automatic extraction of regularity patterns and its application to object-oriented image classification. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3/W49A), 117-121

Ruiz, L.A., Recio, J.A., Hermosilla, T., Fdez. Sarriá, A., 2009. Identification of Agricultural and Land Cover Database Changes Using Object-oriented Classification Techniques. In: *33rd International Symposium on Remote Sensing of Environment*, May 4 – 8, Stresa, Italy

Ruiz, L.A., Recio, J.A., Fernández-Sarría, A., Hermosilla, T., 2011. A feature extraction software tool for agricultural object-based image analysis. *Comp and Elec in Agric*, 76(2), 284-296. doi:10.1016/j.compag.2011.02.007

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis*, 211–252. doi.org/10.1007/s11263-015-0816-y

Schmedtmann, J., Campagnolo, M.L., 2015. Reliable crop identification with satellite imagery in the context of Common Agriculture Policy subsidy control. *Remote Sens.*, 7(7), 9325–9346. doi.org/10.3390/rs70709325

Sitokonstantinou, V., Papoutsis, I., Kontoes, C., Arnal, A.L., Andrés, A.P.A., Zurbano, J.A.G., 2018. Scalable parcel-based crop identification scheme using Sentinel-2 data time-series for the monitoring of the common agricultural policy. *Remote Sens.*, 10(6), 1–22. doi.org/10.3390/rs10060911

Wang, J., Shen, L., Qiao, W., Dai, Y., & Li, Z., 2019. Deep Feature Fusion with Integration of Residual Connection and Attention Model for Classification of VHR Remote Sensing Images. *Remote Sens*, 11(13), 1617. https://doi.org/10.3390/rs11131617

Wardlow, B. D., Egbert, S. L., & Kastens, J. H., 2007. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sens Environ*, 108(3), 290–310. https://doi.org/10.1016/j.rse.2006.11.021

Xu, L., Zhang, H., Wang, C., Zhang, B., & Liu, M., 2019. Crop classification based on temporal information using Sentinel-1 SAR time-series data. *Remote Sens*, 11(1), 1–18. https://doi.org/10.3390/rs11010053