CROP CLASSIFICATION FROM SENTINEL-2 TIME SERIES WITH TEMPORAL CONVOLUTIONAL NEURAL NETWORKS

Sara Pérez-Carabaza

Ireland's Centre for Applied AI
University College Dublin, Dublin, Ireland

Vasileios Syrris, Pieter Kempeneers and Pierre Soille

European Commission,
Joint Research Centre (JRC), Ispra, Italy

ABSTRACT

Automated crop identification tools are of interest to a wide range of applications related to the environment and agriculture including the monitoring of related policies such as the European Common Agriculture Policy. In this context, this work presents a parcel-based crop classification system which leverages on 1D convolutional neural network supervised learning capacity. For the training and evaluation of the model, we employ open and free data: (i) time series of Sentinel-2 optical data selected to cover the crop season of one year, and (ii) a cadastre-derived database providing detailed delineation of parcels. By considering the most dominant crop types and the temporal features of the optical data, the proposed lightweight approach discriminates a considerable number of crops with high accuracy.

***Index Terms*—** Crop classification, Multi-temporal remote sensing images, Convolutional Neural Networks.

1. INTRODUCTION

This work deals with the Census Parcel-based Crop Classification (CPCC) problem, a sub-category of the land cover classification that addresses the systematic identification and mapping of regions on the Earth surface. The CPCC problem regards the detailed classification of agricultural areas, which are frequently organized by parcels. An agricultural parcel is defined as *"a continuous area of land on which a single crop group is cultivated by a single farmer"* [5]. CPCC is indeed a class of Parcel Crop Classification (PCC) problem where the parcel locations have been mapped thanks to the cadastral registration of state authorities and cartographic services. The CPCC problem has an important application related to the Common Agricultural Policy (CAP) which provides direct payments to farmers via the European Agricultural Guarantee Fund, and organizes actions to respond to market instabilities or environmental challenges. The CAP relies on a set of comprehensive administrative registers, ortho-photos and on-the-spot checks on subsidy applications managed by the Member States [5]. Although the majority of the declared information is valid, regular and costly on-the-spot checks of some of the registered parcels are necessary in order to ensure that financial aids are only paid for eligible agricultural areas [5]. The introduction of automated methods would be of great benefit due to: i) the reduction of the number of the costly on-the-spot checks, and ii) the better targeting of the field surveys on areas with a higher probability of incorrect assessment. These considerations have motivated the present work that proposes an effective CPCC method which builds upon the power of deep learning models and the availability of high spatial and temporal resolution satellite imagery provided by the Copernicus Programme. This novel CPCC schema exploits the temporal and spectral information from Sentinel-2 optical satellite imagery through a temporal convolutional neural network (TempCNN) in order to discriminate numerous different crop types.

The paper is organized as follows. Section 2 reviews related state-of-the-art methods. Section 3 describes the proposed CPCC model. Section 4 analyses its performance and Section 5 summarizes the main conclusions of the work.

2. RELATED WORK

In parcel crop classification problems, Satellite Image Time Series (SITS) have taken an outstanding role [8, 3]. Contrary to the classification of other land covers such as urban areas where temporal information have less discriminative power, the temporal information is very informative for crop classification and allows the models to learn in a relatively efficient way the different crop growing patterns [8]. Many of the state-of-the-art approaches to CPCC employ traditional machine learning methods, which are applied to the spectral bands values directly and to temporal phenological features extracted from the SITS as a combination of the information in the spectral bands (e.g. vegetation indices such as NDVI or EVI). For instance, [2] proposes a method based on a Decision Tree (DT) that considers as features the mean value of the pixels within each parcel for several spectral bands (Red, Green, Blue and NIR) and three vegetation indices. Besides, a Random Forest (RF) method has been proposed in [1] to classify 9 crop types in Khorezm (Uzbekistan) considering as features the NDVI and EVI mean and standard deviation val-

Table 1: Related machine learning approaches to CPCC.

Work	Satellite Imagery	Model	Classes
[1]	RapidEye	RF	9 crops
[2]	GeoEye-1	DT	12 crops
[3]	Landsat, S1	1D CNN, 2D CNN	11 crops
[4]	S1	RNN	11 crops
This work	S2	TempCNN	20 crops

ues of five RapidEye bands for each parcel. Recently, deep learning models such as Recurrent Neural Networks (RNN) are showing better performance to SITS classification than traditional machine learning methods [4, 3]. Following this approach, [4] proposes an RNN architecture for classifying 11 crop types in Camargue region (France) using Sentinel-1 (S1) images. In [3] the optical (Landsat) and S1 image time series are considered to classify 11 crop types through two CNN models: a 1D CNN that exploits the spectral patterns and a 2D CNN with convolutions acting on the spatial domain.

Table 1 summarizes basic features of the considered machine learning approaches for CPCC: type of satellite imagery used, type of model and number of crops classified. In this work we propose a CPCC model based on temporal convolutional neural networks that exploits the SITS spectral and temporal information of the parcels in order to classify 20 types of crops. This high number of crop types is a challenging task as the probability of having crops with similar spectral and temporal profiles that lead to misclassification errors is higher. The main reason we selected TempCNN for the proposed CPCC model was its ability to learn the temporal patterns in lower training times in comparison with other deep learning models that exploit the temporal dimension like RNN [8]. Temporal information is very substantial in crop classification as it allows to learn the characteristic growth patterns of the crops [4]. For instance, authors in [8] found that for crop seasonality the temporal features of SITS were more important than spatial patterns derived from medium resolution images learned by 2D CNN networks for a related PCC problem. This work aims to test the capacity of TempCNN in the challenging CPCC problem, and to the authors' knowledge this is the first work that tests the performance of TempCNN for crop classification with Sentinel-2 imagery.

3. MULTI-SPECTRAL TIME SERIES CLASSIFIER

This section describes the data processing methodology and the proposed CPCC system based on TempCNNs.

3.1. Data Processing Methods

In this work, we employ S2 products derived from the twin satellites S2A and S2B that provide high-resolution multi-spectral and multi-temporal imagery, accessible via the Copernicus Open Access Hub[1]. S2 products consist of 13

[1]https://scihub.copernicus.eu/

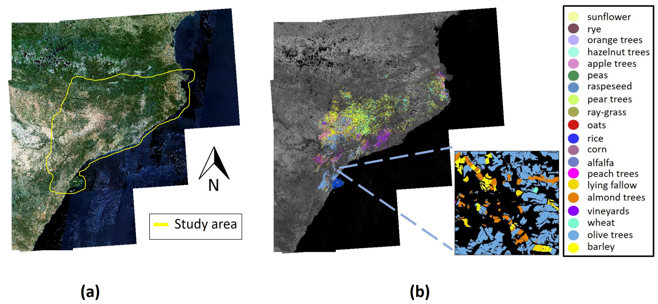

Fig. 1: (a) A true color S2 composite (b) Crop types over S2 grayscale composite.

spectral bands with a spatial resolution which differentiates among 10m, 20m and 60m, depending on the particular band. Due to the small size of some of the parcels under classification this work employs the four spectral bands at the highest (10 m) spatial resolution: blue B02 (490 nm), green B03 (560 nm), red B04 (665 nm) and near infrared (NIR) B08 (842 nm). The S2 products are 100×100 km^2 ortho-images in UTM/WGS84 projection, delivered in separate tiles according to the Military Grid Reference System (MGRS) in GML-JPEG 2000 format. In this work, we selected Level-2A S2 tiles that completely cover the land extent of the DUN-SINGPAG georeferenced database described afterwards and correspond to the crop season from 1 October 2017 to 30 September 2018. In order to mitigate the effect of cloud coverage a bi-monthly median composite is built, considering a selection of cloud-free pixel values of the S2 images that refer to a period of two weeks. The following categories from Scene Classification Layer (SCL) provided by Level-2A products were considered as cloudy or unfit for selection: no data, saturated or defective, dark area pixels, clouds and snow. In addition, missing values (for which no cloud-free pixel was available in the two weeks period) were interpolated using the nearest cloud-free pixels in time and a Savitzky-Golay filter was applied to the whole time series. Accordingly, a 24-length multi-variable time series where each time step corresponds to one S2 composite associated to a period of two weeks is built. Fig.1 (a) displays a true color representation of the resulting SITS corresponding to the first fortnight of July 2018 (where a yellow line delimits the study area).

The DUN-SIGPAC database[2] contains information about the crop type of the parcels in Catalonia, reported by farmers (as mandated by the European crop subsidy program) following the single agrarian declaration (Declaració Única Agrària, DUN) and made accessible by the Spanish Agricultural Land Geographic Information System (SIGPAC) public administrative database. The DUN-SIGPAC database is provided annually and contains several attributes for each parcel such as

[2]https://dadesobertessituam.opendata.arcgis.com/datasets

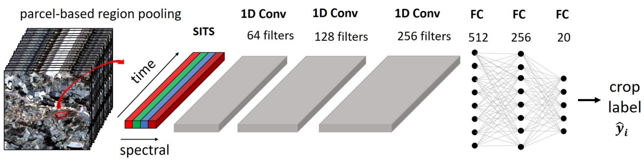

Fig. 2: Parcel crop identification model.

the crop type, location and parcel identifier. This work considers the 20 most frequent crops for classification and analysis, corresponding approximately to 80% of the total number of parcels as reported in the 2018 DUN-SIGPAC shapefile. The dataset under study corresponds to a total of 568,853 parcels, with a mean parcel extent of 0.18 Ha and among which olive trees is the most frequent crop. For the needs of the present study, we translated the parcel information (parcel identifier and crop type) of the shapefile into raster data projected to the Universal Transverse Mercator (UTM) coordinate system by considering a spatial resolution of 10 m that matches the spatial resolution of the selected S2 imagery. Fig. 1 (b) displays the discrete-valued map (over the S2 imagery visualized in greyscale) that represents the classes with a color map that associates different colors for each of the 20 crops. The remaining pixels that do not correspond to any of the classes under study (i.e. non-crop areas and other less frequent crop types) appear in black in the zoom-in region. The images with the parcel identifiers are used to compute the average values of the S2 SITS over all the pixels within each parcel (parcel region pooling).

3.2. CPCC Deep learning based Classifier

The proposed deep learning model is based on temporal convolutional neural networks, that is, deep neural network architectures composed of one or several 1D convolutional layers where 1D kernels convolve along the temporal dimension. Therefore, this type of CNN is suitable for time series classification problems like CPCC. When applying TempCNN to SITS, lower 1D convolutional layers aim at capturing small scale temporal variations, while deeper layers focus on overall seasonal patterns [8].

The proposed CPCC classifier predicts the most probable crop type among the 20 selected crops given the multispectral time series for each parcel. The classification process is sketched in Fig. 2. The CPCC model takes as input a vector with 4×24 values: the mean value of the pixels that constitute one parcel for every of the S2 Red, Green, Blue and NIR spectral bands and for the 24 15-day composites that cover an entire year. The input SITS feeds the TempCNN which assigns to each parcel a unique label $\hat{y}_i$ among the 20 considered types of crops. Then, the predicted crop label is assigned to all the pixels within parcel i. Fig. 2 also shows the TempCNN architecture used by the CPCC model, which is composed of three 1D convolutional layers and three dense layers with 512, 256 and 20 neurons respectively. The kernel size of the three 1D convolutional layers is set to 3 and the number of filters is equal to 64, 128 and 256 respectively. The activation function set to all the layers is the rectified linear unit (ReLU) with the exception of the last one which considers the softmax activation function. Besides, in order to avoid overfitting, the following regularization mechanisms were considered: dropout layers with a dropout rate 0.2 are added after the first two Fully Connected (FC) layers, batch normalization layers are considered after convolutional layers and early stopping mechanism is used during training; a stratified 5% of the training data have been used during the training for validation purposes. Finally, the categorical cross entropy as loss function and the Adam algorithm with a learning rate 10^{-4} have been chosen for the optimization process. The model was implemented using the Keras library on the JRC Big Data Platform [7]. Jupyter notebooks are available upon request.

4. EXPERIMENTAL RESULTS

For the evaluation of the performance of the CPCC system, 75% of the parcels have been assigned to the training set and the remaining 25% to the test set, following a stratified random sampling. The following metrics summarize the performance of the proposed CPCC classifier over the test set; overall accuracy: 0.845, weighted accuracy by the frequency of samples in each class: 0.848, macro-averaged F1-score: 0.863 and overall pixel accuracy (i.e. percentage of well classified pixels): 0.92. Hence, 84% of the parcels and 92% of the pixels from the test set have been correctly identified by the model. Besides, the similar value of the weighted accuracy (0.848) to the overall accuracy (0.845) indicates that the model is not biased to the more frequent classes. The confusion matrix obtained by the comparison of the reference crop types and the predicted labels with respect to the test set, is shown in Table 2. Each element c_{ij} of the confusion matrix corresponds to the number of crop types predicted as crop type i known to be crop type j. In Table 2, the recall (user's accuracy) values per class are placed at the last column, the precision (producer's accuracy) values per class and the F1-scores per class are placed at the last rows of the table. F1-scores per class close to 1 are a good indicator of the classification performance, as it implies a low number of false positive (high precision value) and false negative predictions (high recall value). Table 2 allows to observe how the classification performance depends on the crop type. The classification results for some of the crop types such as *corn* (recall equal to 0.93), *rice* (recall equal to 0.99) and *rapeseed* (recall 0.95) are very good, as the model classifies correctly more than 90 % of the crop types of the test set. Other crop types such as *lying fallow* and *almond trees* show lower classification performance. The case of *lying fallow* may be due

Table 2: Confusion matrix of the results obtained by CPCC system over DUN-SIGPAC dataset.

Predicted crops (columns); Reference crops (rows)

Label	BA	OT	WH	VI	AT	LF	PT	AL	CO	RI	OT	RG	PT	RA	PE	AT	HT	OT	RY	SU	Recall
Barley	**21783**	237	724	21	73	511	6	46	37	1	279	66	1	22	49	0	3	1	4	2	0.91
Olive trees	298	**26140**	79	402	1783	791	92	59	17	2	55	26	28	3	12	11	143	39	2	0	0.87
Wheat	1134	57	**12186**	1	24	141	0	24	33	0	112	36	1	15	20	1	1	0	11	2	0.88
Vineyards	19	444	1	**11177**	256	337	24	8	2	0	5	7	4	0	1	4	25	1	0	1	0.91
Almond trees	94	3351	15	223	**8867**	620	51	10	4	0	5	7	31	1	2	6	51	3	1	0	0.67
Lying fallow	495	2045	104	561	1072	**9240**	92	95	60	3	185	105	56	19	24	26	61	14	1	8	0.65
Peach trees	4	104	4	46	34	74	**3485**	3	4	1	5	1	56	0	0	24	7	4	0	0	0.90
Alfalfa	36	37	15	8	17	149	6	**5577**	28	0	88	140	7	6	3	4	2	2	12	0	0.91
Corn	88	12	25	5	1	49	6	21	**3585**	0	17	33	1	0	2	2	0	0	1	5	0.93
Rice	0	0	0	0	0	0	0	0	2	**818**	0	0	0	0	0	0	0	0	0	0	0.99
Oats	391	100	136	10	30	219	2	93	12	0	**4493**	139	0	14	5	0	4	0	6	1	0.80
Ray-grass	59	36	22	0	2	85	2	109	22	0	111	**3867**	3	4	0	1	1	1	1	1	0.89
Pear trees	1	45	2	11	19	39	83	14	4	0	1	0	**1480**	0	0	39	3	3	0	0	0.85
Rapeseed	24	6	24	0	3	19	0	5	1	0	8	13	0	**2005**	3	0	0	0	0	0	0.95
Peas (PE)	81	12	9	2	1	34	0	4	7	0	18	2	1	1	**1283**	0	0	0	0	1	0.88
Apple trees	3	24	3	12	12	26	29	13	2	0	2	1	34	0	0	**1083**	4	0	0	0	0.87
Hazelnut trees	2	225	2	38	61	45	1	2	0	0	0	2	3	0	0	1	**1953**	0	0	0	0.84
Orange trees	0	116	0	1	6	22	1	0	0	0	0	2	1	0	0	0	2	**380**	0	0	0.72
Rye	15	9	13	0	1	7	0	8	1	0	10	6	0	0	0	0	0	0	**225**	0	0.76
Sunflower	10	2	3	2	2	27	0	4	7	0	3	4	0	2	0	0	0	0	0	**208**	0.76
Precision	0.89	0.79	0.91	0.89	0.72	0.74	0.9	0.92	0.94	0.99	0.83	0.87	0.87	0.96	0.91	0.9	0.86	0.85	0.85	0.91	
F1-score	0.9	0.83	0.9	0.9	0.69	0.69	0.9	0.91	0.93	0.99	0.81	0.88	0.86	0.95	0.9	0.88	0.85	0.78	0.81	0.83	

to the confusion of terrains declared as unused with remaining crops in some of the parcels declared as *lying fallow*, whereas the lower figures in relation to *almond trees* are mostly due to the growth profile similarity with other trees like *olive trees*.

Furthermore, the proposed model is compared with an RNN architecture proposed in [4] for a crop classification problem with S1 imagery. This complex deep learning architecture (9,465,876 parameters in contrast to the 2,625,236 parameters of the proposed TempCNN model) shows lower performance on the DUN-SIGPAC dataset (overall accuracy of 0.819 vs the 0.845 of the proposed model) while demonstrating a much higher computational cost (2.7 h vs 0.5 h). Therefore, we can conclude that the proposed approach based on TempCNN is an efficient approach, reaching higher accuracy in lower training time.

5. CONCLUSIONS

This work proposes a deep learning approach for multi-spectral time series classification based on temporal convolutional neural networks. The proposed parcel crop classification system tackles the classification of a large number of crop classes, including crops with similar growth profiles, such as different types of fruit trees. The proposed CPCC system achieves an accuracy over 85% for the majority of the crop types. As future research, we will consider the inclusion of the spatial information through the use of 3D CNNs.

6. REFERENCES

[1] Conrad, C., Dech, S., Dubovyk, O., Fritsch, ... & Zeidler, J. (2014). Derivation of temporal windows for accurate crop discrimination in heterogeneous croplands of Uzbekistan using multitemporal RapidEye images. Computers and Electronics in Agriculture, 103, 63-74.

[2] García-Torres, L., Caballero-Novella, J. J., Gómez-Candón, D., & Peña, J. M. (2015). Census parcels cropping system classification from multitemporal remote imagery: A proposed universal methodology. PLOS ONE, 10(2).

[3] Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, 14(5), 778-782.

[4] Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., & Hossard, L. (2018). Deep recurrent neural network for agricultural classification using multi-temporal SAR Sentinel-1 for Camargue, France. Remote Sensing, 10(8), 1217.

[5] Owen, P. W., Milionis, N., Papatheodorou, I., Sniter, K., Viegas, H. F., Huth, J., Bortnowschi, R. (2016). The land parcel identification system: A useful tool to determine the eligibility of agricultural land. Special Report, 25.

[6] Rußwurm, M., & Korner, M. (2017). Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 11-19).

[7] Soille, P., Burger, A., De Marchi, D., Kempeneers, P., Rodriguez, D., Syrris, V., & Vasilev, V. (2018). A versatile data-intensive computing platform for information retrieval from big geospatial data. Future Generation Computer Systems, 81, 30-40.

[8] Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. Remote sensing of environment, 221, 430-443.