Service-oriented Inference for Crop Classification using Satellite Images

P. Purwar, I. Kastanis, P. A. E. Schmid

The successful application of the crop classification method developed by CSEM in 2018 was followed by further verification to establish the robustness of the approach. Additional crop types were analyzed, and their results demonstrated high precision in classification. The maturity of the algorithm motivated the implementation of an online inference service integrated in the DataBio hub.

There are multiple aspects of a machine learning algorithm that need to be investigated and developed before it reaches a maturity that is attractive for the industry. For this purpose, the proposed approach [1] was evaluated further with additional crops and a service-oriented architecture was implemented.

The initial study of 2018 focused on developing a crop classifier on wheat parcels using a temporal sequence of satellite images. The two-step approach for building a classifier is further evaluated with two additional crops, maize and legumes. According to agricultural taxonomy, each crop has different sub-varieties. The classifier trained using parcels of wheat sub-variety W_1 was also used to classify other sub-varieties of wheat. A similar approach is used to create and test classifiers for maize and legumes. The results are shown in Figure 1. The details of training and testing can be found in [2].

Wheat			Maize		
Variety	Parcels	Accuracy	Variety	Parcels	Accuracy
W_1	1312	0.913	M_1	2003	0.859
W_2	882	0.959	M_2	1846	0.819
W_3	239	0.929			
W_4	244	0.951	Legumes		
W_5	380	0.918	L_1	1437	0.855
W_6	335	0.863	L_2	160	0.250
W_7	451	0.911	L_3	126	0.222
W_8	146	0.877	L_4	163	0.515

Figure 1: Classification accuracy for different crops and sub-varieties.

The classifiers trained using only one sub-variety, W_1 and M_1, were able to classify other unseen sub-varieties of wheat and maize with accuracy greater than 81%. This establishes that the classifiers learn generic features for the crop and do not suffer from overfitting issues. Contrary to wheat and maize, the performance for unseen sub-varieties of legumes is much lower. This is most likely due to the natural variation of the legumes, which contains a plethora of plants such as lentils, beans, peas, etc.

All three binary classifiers are deployed as on online service based on RESTful APIs. The CSEM server receives the parcel information and infers classification results at a pixel level. The average time for a single request takes less than 1 second to compute the probabilities for each pixel in a parcel. The CSEM server is capable of handling multiple requests from multiple clients. The service is integrated in the DataBio platform [3] with the GeoRocket [4] data store. GeoRocket displays the crop type for each parcel as declared by farmers as seen in Figure 2. When a parcel is selected, it generates a request to the CSEM server for classification.

Figure 2: Screenshot of the online DataBio platform.

In Figure 3, a parcel, highlighted in green, has been selected. The declared crop (maize) is shown in top-left corner. Each classifier verifies and analyzes the corresponding satellite data and displays results as a heat map. The results below show that 98% of the pixels in the selected parcel are maize. In addition, the aggregated results are shown as a tick or a cross for each classifier.

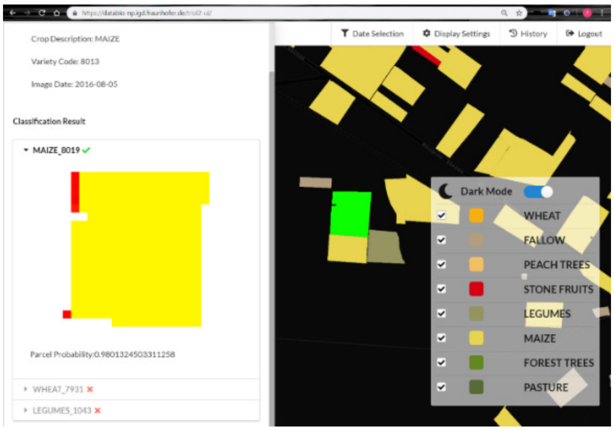

Figure 3: Screenshot of results obtained for selected parcel.

The proposed solution offers an efficient way for those looking to utilize Big Data in the agricultural domain for crop analysis. The two-step approach offers a structured and cost-efficient way of adding new types of crops. It is a cloud ready solution offering privacy through the use of JWT tokens, security, scalability, low latency and portability through Docker containerization.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732064.

[1] P. Purwar, I. Kastanis, M. Hoechemer, P. A. Schmid, "Recurrent Neural Network (RNN) based temporal classification of land usage using satellite imagery", CSEM Scientific and Technical Report (2018) 69.

[2] P. Purwar, S. Rogotis, F. Chatzipapadopoulus, I. Kastanis, "A reliable approach for pixel-level classification of land usage from spatio-temporal images", 2019 6th Swiss Conference on Data Science (SDS), Bern, Switzerland, 2019, pp. 93-94.

[3] https://www.databiohub.eu

[4] https://georocket.io/