Scientific Symposium FAIR Data Sciences for Green Life Sciences
12 December 2018 | Wageningen | the Netherlands

Conference Proceedings | DOI: https://doi.org/10.18174/FAIRdata2018.16279

A Recurrent Neural Network (RNN) based approach for reliably classifying land usage from satellite imagery

Kastanis, I.[1], P. Purwar[1], A. Steinecker[1], P. Schmid[1]
[1] CSEM SA, Switzerland
Corresponding author's e-mail: ikastanis@csem.ch

An approach for simplifying the process of Neural Network model creation in the domain of land usage classification is presented in this work. The proposed method is a complete pipeline for cleaning the data using minimal supervision and subsequently creating a crop specific pixel level temporal model. A first step in the domain of Big Data is controlling the integrity of the data for the purpose of model creation. Traditional methods are typically based on data specific heuristics to decide what information can be used for training the model. These approaches require significant effort in hand-crafted filtering methods, and often lack sufficient structure for formalisation and generalisation.

Images from the Sentinel-2A satellite together with parcel location and crop variety were collected forming temporal sequences. Using a Neural Network modelling approach, a method was developed to remove data that was not considered to be reliable. To decide which data will be used, an RNN based auto-encoder model was trained for each crop type. Input data was then projected in the latent space and analysed by means of clustering. The clusters in the latent space separate data that can be used reliably from data that is not representative of the crop variety. An expert can further refine these data groups by annotating boundary cases. In this manner the effort of the expert is optimised by focusing in targeted areas. Having selected the appropriate data, single variety classifier models were created by means of transfer learning and used for detecting whether an area contained a specific crop type or not. Experiments on different crop types have shown excellent results in distinguishing a variety from others while reducing the time and effort required for model creation. Given sufficient data from a multitude of crop varieties a multi-class model can also be trained. This approach can be expanded further to integrate ground sensors, aerial and ground imagery in order to achieve a finer spatiotemporal resolution. CSEM will offer land usage classification as a service through the DataBio[1] platform.

Using the proposed RNN based approach for detection and classification of vegetation types creates a new possibility in handling Big Data in the domain of Earth Observation with a multitude of potential applications in the areas of precision and smart agriculture, pre-emptive forest management, health monitoring and damage assessment. The previously described process of data selection and model creation offers a structured method of digitising expert knowledge in a data driven architecture. The presented pipeline significantly reduces the complexity of creating models by removing the need of hand crafted filtering, making it a cost-effective option for bringing neural network models to the market.

This work has received funding from the European Union's Horizon 2020 research and innovation programme within the project DataBio focusing on using Big Data for the bioeconomy under grant agreement No 732064.

[1] https://www.databio.eu