DOI: 10.26314/GARR-Conf19-proceedings-07

Adalitix: a Web Platform for Agricultural Data Analytics

Luca Coviello, Marco Cristoforetti, Cesare Furlanello

Fondazione Bruno Kessler

Abstract. Digital technologies ignited a revolution in the agrifood domain known as smart farming. In this work we present Adalitix, a cloud-based web platform for agricultural data analytics (including: collection, visualization and predictive modelling) for better crop management. We also describe two real smart farming cases, where Adalitix is used by vineyard technicians to estimate grape quality parameters and predict yield estimation directly in field. Data were acquired with a 2 Mpx smartphone camera and a handheld portable spectrometer with non-invasive techniques.

Keywords. Smart Farming, Agricultural Data, Web Platform, Deep learning, Complex Analytics

Introduction

The continuous development and adoption of connected IoT devices is impacting all aspects of our day life and environment. The aim is to build a smart society that takes decision based on the insights generated by analysing data collected with low cost sensors in different domains. With the always growing demand of food (+50% between 2012 and 2050), the few opportunities left to expand agricultural areas and the significant decline in yield growth (FAO, 2017), it is crucial to develop new technologies and infrastructures that help to produce more and safer food with less resources. Smart farming represents a unique opportunity for helping farmers to measure, monitor and analyse precisely the status of their crops in real time. Significant factors that slow down the adoption of digital technologies in the agrifood industry are the reluctance of growers to employ IT systems in field given their cost, resource demand and difficulty to use. In this work we propose Adalitix (Fig. 1), a web based solution to store, visualize and analyse agricultural data collected with in field sensors. Furthermore, it allows farmers to leverage cloud computing resources to automatically apply Artificial Intelligence (AI) models, e.g. deep learning, to forecast future crop development and improve decision making processes. Notably, Adalitix makes easy to publish the generated datasets, e.g. for research or data monetization purposes, with standardized metadata. Two use cases developed for viticulture will illustrate the benefits of Adalitix for agricultural technicians, enabling the adoption of AI techniques in a consistent and efficient framework.

1. Adalitix

1.1 Field Data Measurements

Agricultural field measurements can be recorded with different instruments or sensors and in different formats (e.g. images, temperatures, humidity, etc.). Adalitix was desig-

Torino, 4-6 giugno 2019 - ISBN: 978-88-905077-9-3 -

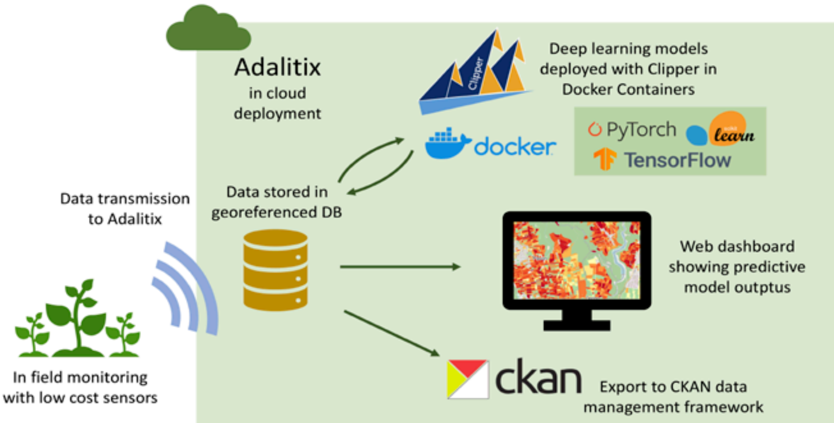


Fig. 1. Adalitix architecture components overview

ned to store various types of data (e.g. images, time series, geoTIFFs) with a particular focus on the spatial and temporal resolution, which directly impacts the analysis results. In Adalitix, data are organized into datasets and can be visualized or given as input to the analysis pipelines deployed on the system. Results are then stored and made available on the web platform.

1.2 Web Dashboard

The stored information can be visualized in the web dashboard to monitor field conditions and crop status and development. Moreover, data can be represented over the spatial and temporal axes to analyze short or long term changes. To this aim Adalitix displays the collected information on a map with the option to focus on particular areas or time spans. A dashboard enables selection and visualization of outputs generated from the analysis pipelines available on the platform. In particular, Adalitix functions are being designed to forecast the phenological development of plants and predict crop yield given other environment variables.

1.3 AI Analytics Pipelines

AI is a key enabling technology for extracting valuable insights from data, although requiring massive data and significant computing power for training models. The computational resources required at inference (use of trained models) are instead usually limited: models can be run on field on edge computing devices, smartphones or cloud systems at the cost of data transmission. Deep learning models and complex analytics are deployed in Adalitix with Clipper, an online prediction serving system (Crankshaw, Wang 2017). Clipper deploys every model and pipeline as an independent docker container with the required dependencies, accessible through a RESTful API. On the Adalitix web platform, users can upload pipelines on the stored datasets, manage results and make them available through the dashboard.

1.4 CKAN metadata and data indexing system

To simplify the data access and search process we adopted CKAN (Comprehensive

Knowledge Archive Network), an open source data management system that allows to publish, share, search and use data. CKAN provides a metadata repository and indexing system to easily retrieve, publish and export datasets for download or integration on other CKAN instances. Moreover, CKAN exposes its content and functionalities through RESTful APIs following the specifications defined by the Agency for Digital Italy (AgID) as national standard for data sharing between administrations. The CKAN Adalitix instance allows users to export its catalogue as a collection of metadata and descriptors.

2. Use Cases

Adalitix has been employed for noninvasive grape yield prediction and grape quality estimation, based on deep learning algorithms. Both classes of models are made available to vineyard technicians through the combined use of a smartphone and a handheld portable spectrometer directly in the field. Other IoT options are also available, e.g. connectivity by LoRaWan. For yield estimation, we adapted CSRNet (Li et al. 2018), a deep learning algorithms developed for the crowd counting context to estimate the number of grapes present in an image. The model reached an overall mean average percentage error MAPE = 10% (~17000 grape berries) with images taken from a medium distance. For quality estimation we developed a deep learning model applied to spectral signals (Zhang et al. 2019) (Ni et al. 2019) for sugar content estimation in grapes from ~2000 samples, reaching an R2 = 0.83 and a mean average error MAE = 0.77 brix degrees.

3. Conclusions and Future Work

Agriculture provides an elective application domain for AI combined with IoT technologies, mostly requiring low cost sensors, making the development of data platforms a crucial to increase productivity, value and quality of production. In this work we presented Adalitix, a web platform for storing, visualizing and analysing data collected with in field instruments connected to the Internet. The two AI solutions enabled by our platform for wine production illustrate how predictive models can introduce in smart farming ecosystems new tools to estimate quantity and quality. Further developments include upscale of data collection with low cost network technologies (e.g. Long Range Low power networks), enabling the use of Adalitix solutions with georeferenced time series data from smart devices deployed in remote areas with low Internet coverage.

4. Acknowledgements

The work presented here is sponsored by the Consortium GARR with the "Orio Carlini" research scholarship. FBK acknowledge support by CAVIT. Initial work has been developed within the FBK WebValley Data Science Project 2016 and 2017. Collaboration with CODIPRA Trento and MPA Solutions is also acknowledged.

References

FAO. (2017). The future of food and agriculture. Trends and challenges.

Crankshaw, D., Wang, X., Zhou, G., Franklin, M. J., Gonzalez, J. E., & Stoica, I. (2017).

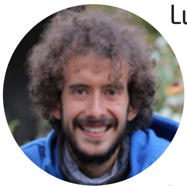
Clipper: A low-latency online prediction serving system. In 14th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 17), pp. 613-627.

Li, Y., Zhang, X., & Chen, D. (2018). Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1091-1100.

Zhang, X., Lin, T., Xu, J., Luo, X., & Ying, Y. (2019). DeepSpectra: An end-to-end deep learning approach for quantitative spectral analysis. Analytica Chimica Acta.

Ni, C., Wang, D., & Tao, Y. (2019). Variable weighted convolutional neural network for the nitrogen content quantization of Masson pine seedling leaves with near-infrared spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy (209) pp. 32-39.

Authors



Luca Coviello - coviello@fbk.eu

Luca Coviello received the M.Sc. in Data Science in 2018 from the EIT Digital Master School. He has been working at FBK since 2014, focusing on the application of machine learning to bioimaging and the development of cloud-based frameworks for large scale dataset analysis. He is now a GARR fellowship student at FBK, working on applying artificial intelligence to agriculture as a low-cost solution for improving crop yield.

Marco Cristoforetti - mcristofo@fbk.eu

Marco Cristoforetti graduated in theoretical physics in 2004 at the University of Milan and received his PhD in Physics at the University of Trento in 2007. Until 2014 he devoted himself to the study of non-perturbative QCD. Data scientist at FBK since 2014 applying Machine Learning and Deep Learning to data from heterogeneous sources such as social networks, agriculture and high energy physics. Member of the CERN ATLAS experiment.

Cesare Furlanello - furlan@fbk.eu

Cesare Furlanello is a data scientist, head of the FBK Unit for Predictive Models in Biomedicine and Environment. He is an expert in machine learning and bioinformatics applied to multi-modal health data (omics, bioimaging, medical phenotypes) with Deep Learning. Full professor habilitation in BioEngineering. Bioinformatics collaborator of US FDA/NCTR, RIKEN, and adjoint Faculty of the Wistar Institute in Philadelphia, he is the Past President of the MAQC Society for reproducibility and quality control over massive data.