TEBAKA: Territorial Basic Knowledge Acquisition. An Agritech Project for the Italy

Lorenzo Epifani, Antonio Caruso

Dept. of Mathematics and Physics "Ennio de Giorgi", University of Salento, Lecce, Italy

Abstract—Tebaka, proposal, impact, some work. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Index Terms—digital agriculture, precision agriculture, satellite images, deep learning, machine learning, sensor networks and iot

I. INTRODUCTION

The enabling technologies of distributed sensor networks, internet of things, remote sensing using multi-spectral cameras on satellites, drones and aerial platforms, together with smart decision support systems are key contributors for the emerging field of *Digital Agriculture* (DA) or *Agriculture 4.0* [1]. This is an *umbrella term* used to designate the transformation of farming that includes digitalization and automation of farming tasks. This area of research encompass a new vision of the farming industry as a cyber-physical farm, and includes all technological areas such as: big-data platforms, machine learning and data-driven models to relate the observations of the phenological stages of plants with field yield and crop quality.

The high-level goal is to support agronomist with *Precision Agriculture* management practices, to better solve the challenges and demand of an increasing world-population, rising production costs (like energy), labor shortages, climate and environmental changes (like water resource scarcity). It leverage extensive knowledge acquisition using remote sensing technologies, and on-the-field deployment of sensors and IoT networks, to improve knowledge acquisition and situated situation awareness, to maximize resource usage and crop yield while minimizing associated costs.

In the United States, the use of yield-maps and soil maps are adopted only by 5 to 25 percent of total U.S. farms; on the other hand, automated guidance, has increased sharply, with the advance in deep-learning applied to computer vision, and it is applied to over 50 percent of the cultivated areas. A good survey of the trend in the technology evolution of a *Smart*
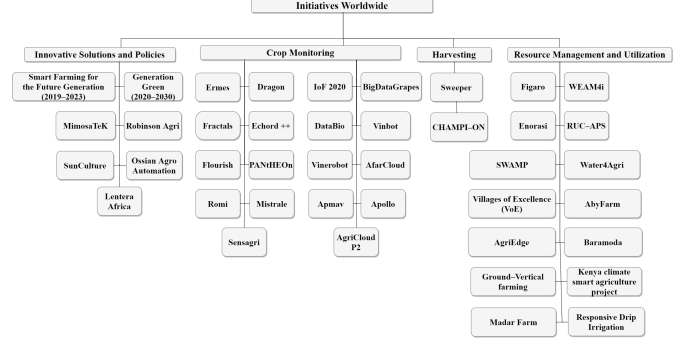

Figure 1: A classification of active European Union funded project in the area of Digital Agriculture (from [3])

Farm in USA is in [2]. USDA (The U.S. Department of Agriculture) is investing up to 2.8 billion in 70 selected projects under the first Partnerships for Climate-Smart Commodities funding pool, as a part of the Inflaction Reduction Act.

In Europe, the European Community has funded many project that integrate ICT technologies and Digital Argriculture, we see in Figure 1 from [3] a collection of active research project and startups, classified in different areas, such as: *Crop Monitoring* (the largest one), *Harvesting*, optimization of *Resource Managements* and *Innovative Policies and Solutions* with more than 41 iniziatives. In Table 2 another list of worldwide projects that are actually funded, with a increasing role for African and Middle-East countries.

In this paper, we present the activities of TEBAKA: a national project funded by PON-FESN (Europen Union Social Fund) dedicated to develop the adoption of Digital Agriculture in the south part of Italy (Mezzogiorno) and in particular in Apulia (Puglia). The project span four years, from 2020 to 2024 and with an overall budget of 8.5 Million Euro it put together more than 20 partners from industry, two university, several national research institutions of Italian CNR, and private partners with a mixed background (someone more oriented to ICT services, someone more related to agronomy). We discuss the project in the following section, but we want to summarize here the key ideas involved. The project goals is to develop a new generation of decision support system for agronomists for the culture of grain, vine and olive: for this goal, we conduct a large campaign of data acquisition, with measurements taken directly on the culture and soil, spanning two full year. Moreover, all measurements will be taken at the same time of satellite observation (with Sentinel 2), drone

Project/Company	Country or Countries	Objectives
Villages of Excellence (VoE) (2021–2023) [157]	India and Isreal	Improve the productivity and quality of horticulture; Increase income of farmers
Nosho Navi 1000 (2014–2016) [158]	Japan	Large-scale smart rice farming
AbyFarm [145]	Singapore	IoT, blockchain and machine learning for urban farming to ensure sufficiency of crops
Smart Farming for the Future Generation (2019–2023) [159]	Vietnam and Uzbekistan	Development of policies, enhancement of skills and knowledge. Support farmers through post-harvest handling and markets
AgriEdge [160]	Morocco	Precision agriculture platform for practices such as efficient use of water and fertilizer, monitoring weather data, and satellite imaging.
Generation Green 2020–2030 [144]	Morocco	Introduction of new technologies for sustainable agriculture; Installation of over 100,000 solar pumps for irrigation
Baramoda [161]	Egypt	Sustainable agriculture; Maximize the efficiency of agri-waste management
Ground–Vertical farming [147]	Lebanon	Improve efficiency of agricultural yield; Reduce water consumption and costs.
Robinson Agri [162]	Lebanon	Greenhouse technology; Distribution of agricultural resources
Kenya Climate Smart Agriculture Project (KCSAP) [163]	Kenya	Improve productivity in case of climate change risks and provide response in emergencies
MimosaTek [164]	Vietnam	IoT platforms for precision farming solutions
Ossian Agro Automation [146]	India	Wireless automation systems (Nano Ganesh) for irrigation in rural areas
SunCulture [143]	Kenya	Solar-powered irrigation systems (RainMaker2)
Madar Farm [141]	United Arab Emirates	Ensure food and water security
Responsive Drip Irrigation [142]	United States of America	Smart irrigation
Lentera Africa [165]	Kenya	Technology enabled farmer advisory services; Farm inputs and equipment

Figure 2: A list of worldwide projects related to the area of Digital Agriculture (from [3])

observations equipped with a multi-spectral camera and with aircraft. This four levels of observation are unique in all the projects presented in the previous discussion, they allow a *pyramidal representation* at different spatial resolution of the phenomena observed, and in different time. On top of this raw observations, we develop data-driven models (mostly based on *neural networks*) that extract features from images related with the specific conditions of the annual crop life cycle.

The rest of the paper is organized as follow: in II we review the project structure and management organization, the partnership and roles, the research goals and actual status; in III we present some interesting original results obtained, in two different areas: in III-A we describe a web application and the related backend, to support agronomist in large field to collect samples from different gps locations with the aids of smartphones and/or tablets, this is also strictly related to a well know problem about the *optimal deployment of a sensor network in a field*; in III-B we explain why in the overall pipeline of analysis it is important to properly do a *semantic segmentation* of the plant crown (foliage) and present a semi-supervised approach to the problem; in IV we collected some of the most relevant works that shape the area, and finally we drawn in V the conclusions.

II. TEBAKA PROJECT DESCRIPTION

Tebaka is a "*platform*" solution for crops of grain, vines and olive oil, all of which are significantly present in the sud of Italy, and will operate through stages of knowledge development and design / realization of prototype solutions that is articulated in four years (2020/2024) with the four major phases described in the following.

DATA ACQUISITION:
- Analysis of the territory of Puglia, identification of the areas and parcels cultivated to observe, definition / authorization of the operational missions and their verification.
- Definition of payload and cargo platforms of sensors to be used (on already operational satellites and on aircraft and land systems to integrate / develop), development of a control room for mission management and operational integration with a data acquisition and processing center.

- Definition of data acquisition criteria from various platforms / pay loads during the various missions envisaged, verification of operational adequacy in the identified territories and basic processing of acquired data through representation in thematic maps.

ANALYSIS AND MODELLING PHASE:
- Definition of the theoretical models of the annual production cycles of the agricultural crops considered, of the fusion criteria of data from multisensor sources, of the criteria for the analysis of the data acquired in the missions, of the criteria for the realization of the data driven models with the use of artificial intelligence logic and deep learning, of criteria for their integration with theoretical models, and the implementation of integrated and validated basic models.
- Development of Intensive Data Acquisition Activity from Payloads / Platforms for refinement of models, realized by means of large amounts of data processing, definition of criteria for identifying potential critical factors and development of models / support decision systems for their management, development of the design and the prototype of the complete system.

PROTOTYPE DEVELOPMENT:
- Realization of the prototype of an integrated platform composed of technological systems (satellites, manned and unmanned aircrafts, manned and unmanned land vehicles and fixed stations, payloads of sensors with different characteristics and performance, storage environments of large amounts of data and applications in the field of artificial intelligence for data manipulation;
- The building of data driven models; the development of models and relationships between the observed phenomena elements realized with machine learning logic, data analysis, decision support systems) in order to develop a product / service for the world of crops with different types of extension on the territory (horizontal for grain, vertical for vine and mixed for the olive tree, others with comparable characteristics) to be offered to various actors operating in the world of innovative agriculture such as Public Organizations, Category Associations, Service Providers, Agricultural Companies d medium / large national and international levels, possibly complementing the proposed solution with other complementary existing or developing ones.

The proposed project has as its ultimate goal the creation of an innovative product / service prototype for agriculture, in particular for grain, olive and vine crops, oriented to provide instruments and knowledge for the acquisition and to generate information for management decisions, in particular to provide guidance in the initial stages of potential and / or critical manifestations and to guide the best-performing actions. The concept of working with models of annual production cycles also allows to better estimate the final industrial result, and therefore economic, of the productions in progress.

The constitution of the advisory board presented is the basis for both orienting / validating the proposed solution and disseminating the product / service proposal to the agricultural

world that is increasingly recognizing and utilizing intelligent agriculture solutions for the optimization of missions of the various actors involved. During the development of the project, a comparison will be made between the participants and external stakeholders to evaluate the possibility / opportunity to realize a spin-off that, through the acquisition of the platform, can to industrialize, maintain and use it as a basis for providing advice / services on the market.

It is primarily usefull for large territorial use but still maintains its validity for smaller areas, encouraging in this regard the concept of joint use among multiple actors in order to benefit from innovative solutions through affordable costs. The solution enables both private and public entities to design high value added services. For private sector it can be usefull for consultancy while for the public actors it can support policies and actions in order to favour the development of the agricultural production.

The articulation of activities, and hence the results that can be achieved for each one, also creates value in both the realization / improvement of system components (better bid ability in their business markets) and the development of internal knowledge and capabilities to be used also for development of other products / services / solutions. The greatest knowledge gained from the university and research entities can also offer better and consulting skills in the market, also contributing to the enhancement of the territory.

III. Project Activities

A. An online Data Collection Application for Agronomist on the Field

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

B. Semi-Supervised Semantic Segmentation of Plant Crown

Segmentation is a recurrent task in literature regarding precision agriculture: the information provided by a pixel-wise classification map can by used to solve different domain specific tasks ([4]–[6]), e.g. biovolume estimation, tree crowns mapping, estimate the distribution of the popuilation for different species , monitoring changes in green areas. We introduce this task, in TEBAKA, in order to monitor the grow of plants and in particular the *volumetric leaf area* and its evolution across time.

The two main remote sensing technologies that have been used over the years are satellite and UAV imagery, each one having his pros and cons. Even if satellite imagery has also been used successfully by several authors[7]–[9], the unmanned aerial vehicles (UAVs) have proven to be a game-changing technology for various reasons. The main one is that for the majority of agriculture challenges, there are strict resolution requirements that satellite imagery cannot satisfy. These requirements can be meet by UAV imagery while keeping costs acceptable. Satellite imagery is the best choice only if we have to deal with large scale observable phenomena[10], [11], since specific patterns can be observed only from a certain scale onward. For these reasons, we decided to choose UAV imagery as data source.

In agronomic, there are many well known indexes that are used to assign a score to each pixel. These indexes are often linear (or non-linear) combination of spectral channels, and have proven to be highly correlated to specific characteristics available in areas that have and index score above a certain threshold. The Normalized Difference Vegetation Index (NDVI) is an index, between −1 and 1, that can be used to assess whether or not the target pixel contains live green vegetation. Even if NDVI thresholding can be used to produce a segmentation map of a UAV image, there are many issues that make it a non-optimal choice. Different objects with the same spectrum obtains the same score, even if they belong to different classes. Moreover, NDVI does not consider local properties and patterns. Image analysis through deep learning has proven to provide state of the art solutions for instance/semantic segmentation in different domains. For this reason, this method has been widely used in precision agriculture in the last decade.

However, supervised deep learning methods require large, hand-labelled datasets, that are carefully prepared by domain experts, in fact, in all the works examined, and reviewed in the next section, the labelling is carried out by agronomists. In our work, we propose an original solution for the automatic labeling of UAV images containing vineyards and olive trees. We tested this method, by using this data-set to train a well-known semantic-segmentation neural network model, the entire system can be considered a self-supervised model, with the advantage that the self-labelling pipeline, can be used to label any plant species without the costly manual labour of a domain expert (in this case an agronomist). To the best of our knowledge, there is no other work using a similar solution in this domain, in particular we are not aware of a single method that can be used on more a single culture type.

The self supervised pipeline is a system that takes as input an hyper-spectral orthomosaic and returns the set $\mathcal{D}$:

$$\mathcal{D} = \{\, (x_i, y_i) : i \in \{1...N\} \,\}$$

where x_i represents the i-th tile of the input orthomosaic, y_i

is the ground truth mask for x_i and N is the total number of collected tiles i.e. the size of the dataset. Masks and tiles have the same spatial resolution, but masks have only one channel, i.e. they are bitmaps, with the value of each pixel be equal to 0 (background) or 255 (vegetation) (we still use 8 bit for each pixel). This representation is useful for visualisation purposes. During the development of the machine learning model, we represented 0 and 255 with the one-hot-encoding.

Only two classes have been assigned to speed up prototyping, but their number can be easily extended, i.e. by labelling each species with a different value. The self supervised pipeline (Figure 3) works as follow:

- The NDVI is computed pixel-wise from the multispectral input orthomosaic and stored as a single-channel image. The NDVI formula is:

$$NDVI = \frac{NIR - R}{NIR + R}$$

  Where NIR (Near Infra Red) and R (Red) bands are respectively stored in the 6th and the 10th channels of our source orthomosaic. L: aggiungere descrizione NIR e motivo formula NDVI
- The NDVI map is the input to a Tunable Denoising Module that is part of the semi-supervised model. This component (shown in Figure 4) is a stack of different well known image processing techniques executed sequentially. The module is executed by fixing a set of parameters that govern its operation. In our experiments, the values of these parameters were found by drawing inspiration from the search criteria for hyperparameters used in machine learning (e.g. grid search inside intervals considered reasonable). This module produces the pixel-wise ground truth map for the entire hyperspectral source orthomosaic. The tunable parameters of this module are 5 (highlighted in red in previous image). Values th_1 and th_2 governs the two tresholding stages, while sw, tw and str respectively represents the search window size, the template window size and the strength of the fast non local means denoising method. For the dilation and the erosion, we used kernels containing only 1 with size 3×3 and 5×5. An erosion follower by a dilation is called opening. The opening is a smoothing filter and the specific behavior depends on the kernel size and content. An opening preserve as much as possible regions having similar structure respect to the kernel, while removing those different. Our kernels with all 1's allow to remove isolated groups of pixels, while their dimension governs the minimum size that groups of pixels must have in order to be preserved. Kernel size and content could be governed introducing some parameters, although in our experiments we achieved good results with static kernels. In our case, they serve to remove groups of isolated pixels classified as plants
- The orthomosaic/ground truth are both split in many tiles of a fixed size. We pairwise store each input tile and its corresponding ground truth tile. The size of the tiles was chosen consistently with the size of the input expected
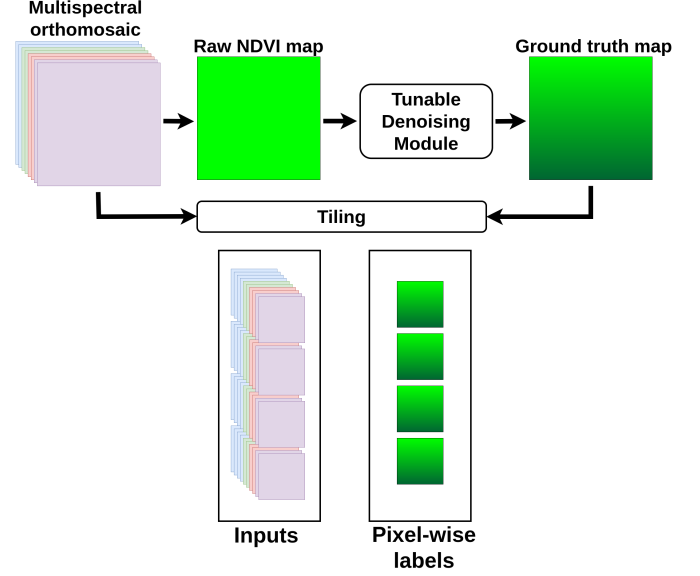

Figure 3: Graphical representation of the entire self-supervised pipeline.

from the semantic segmentation deep learning model we trained.

Before storage, pairs of tiles containing insufficient information are discarded. Specifically, if the pixels in the truth map describing vegetation are fewer than a given threshold, then the pair is discarded. In our experiments, we labeled 906 tiles (net of rejected ones) starting from 5 UAV orthomosaics acquired in 2 different areas of interest and spanning an area of ..km^2.

Flights for the acquisition were performed in a temporal window spacing from the 26 May 2022 to 30 Sep 2022. The semantic segmentation model we chose for training is the U-Net [12], with some modifications, for example making it capable of handling the 10 channels for the images in input.

We chose U-Net over other similar architectures for several reasons: it has fewer parameters than competing networks. Moreover, it was originally proposed to solve a task of similar complexity (biomedical image segmentation) with excellent results, which were then replicated on use cases very similar to ours. Lastly, in these publications (and also in the original work) the U-Net was trained from scratch with datasets of comparable size to ours (often even smaller).

We performed several training runs in order to cross validate different hyperparameters: decay rate, learning rate, number of epochs, and batch size. ok, and we are happy ;), so ....discuss the result of the training in fig 5.

ok, well written, not bad, continue, put also something in the related work.

IV. RELATED WORKS

In we have...contribute here

V. CONCLUSION

In this paper we presented the TEBAKA project, and discuss its relevance in the overall scenario of national or international

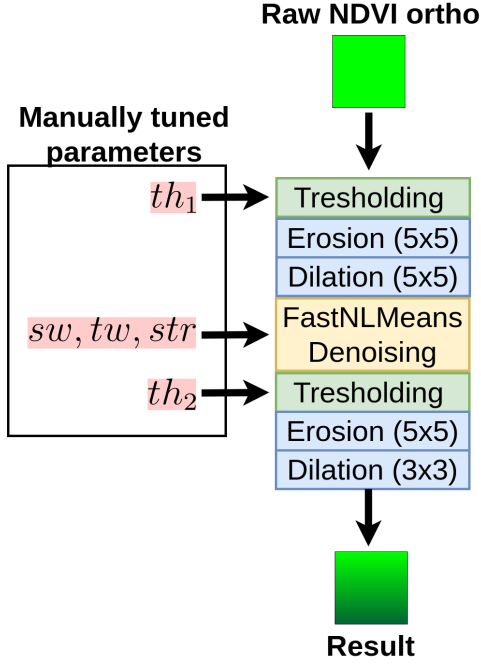Figure 6: Overlay between the input tiles and the result pro

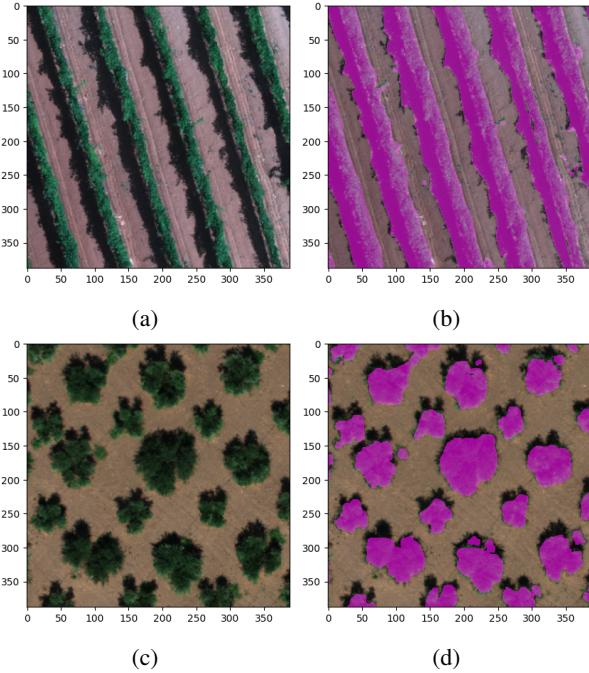Figure 4: Schema of the Tunable Denoising Module.

Figure 5: (a) Input tile, vineyard; (b) Overlay between the previous input tile and the result produced by our self supervised pipeline; (c),(d) are the same but for olive trees
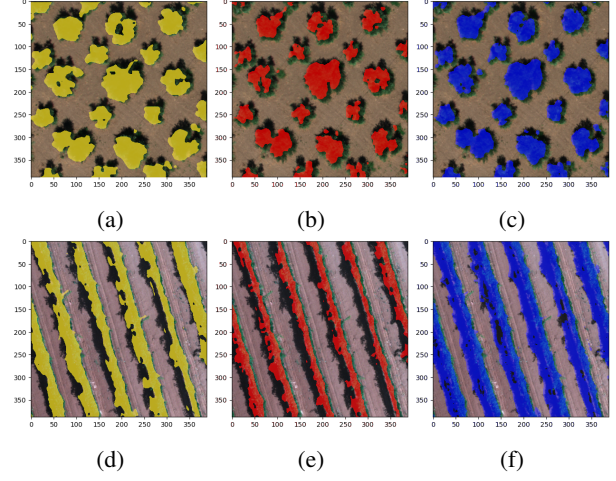
project related to Precision Agriculture. We emphasized the originality of the project, in particular how it lives at the intersection of communications problems related to satellite data management and IoT deployment on the field, and computational problems, i.e. new advanced data-driven ML models that better integrate all knowledge of the domain. We presented a case of a particular task, *image semantic segmentation* of the crown of olive trees or rows of grape plants and shown how to implement an original semi-supervised network that produce the segmentation with unexpected accuracy. In the next year of the project we will obtain new data from the second phase of data acquision on the filed, to validate the models and improve their correspondace with data measured directly from the culture. There is still a large gap between the research and goals of TEBAKA and the actual industrial adoption of precision agriculture from real actors of the industry, and we will work on disseminating the project result in order to inprove this scenario.

REFERENCES

[1] M. De Clercq, A. Vats, and A. Biel, "Agriculture 4.0: The future of farming technology," *Proceedings of the World Government Summit, Dubai, UAE*, pp. 11–13, 2018.

[2] J. McFadden, E. Njuki, and T. Griffin. "Precision agriculture in the digital era: Recent adoption on us farms." (2023), [Online]. Available: https://www.ers.usda.gov/webdocs/publications/105894/eib-248.pdf?v=9219.8.

[3] S. Yadav, A. Kaushik, M. Sharma, and S. Sharma, "Disruptive technologies in smart farming: An expanded view with sentiment analysis," *AgriEngineering*, vol. 4, no. 2, pp. 424–460, 2022, ISSN: 2624-7402.

[4] A. Safonova, E. Guirado, Y. Maglinets, D. Alcaraz-Segura, and S. Tabik, "Olive tree biovolume from uav multi-resolution image segmentation with mask r-cnn," *Sensors*, vol. 21, no. 5, 2021, ISSN: 1424-8220.

[5] Z. Ye, J. Wei, Y. Lin, *et al.*, "Extraction of olive crown based on uav visible images and the u2-net deep learning model," *Remote Sensing*, vol. 14, no. 6, 2022, ISSN: 2072-4292.

[6] S. Egli and M. Höpke, "Cnn-based tree species classification using high resolution rgb image data from automated uav observations," *Remote Sensing*, vol. 12, no. 23, pp. 1–17, 2020.

[7] M. Rußwurm and M. Körner, "Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders," en, *ISPRS International Journal of Geo-Information*, vol. 7, no. 4, p. 129, 2018, ISSN: 2220-9964. (visited on 06/15/2022).

[8] R. C. Daudt, B. L. Saux, and A. Boulch, *Fully Convolutional Siamese Networks for Change Detection*, en, Oct. 2018. (visited on 07/12/2022).

[9] D. Ienco, R. Gaetano, C. Dupaquier, and P. Maurel, "Land Cover Classification via Multi-temporal Spatial Data by Recurrent Neural Networks," en, *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1685–1689, Oct. 2017, ISSN: 1545-598X, 1558-0571. (visited on 06/15/2022).

[10] V. A. Gurumurthy, R. Kestur, and O. Narasipura, *Mango Tree Net – A fully convolutional network for semantic segmentation and individual crown detection of mango trees*, en, Jul. 2019. (visited on 10/26/2022).

[11] E. Guirado, J. Blanco-Sacristán, E. Rodríguez-Caballero, *et al.*, "Mask R-CNN and OBIA Fusion Improves the Segmentation of Scattered Vegetation in Very High-Resolution Optical Sensors," en, *Sensors*, vol. 21, no. 1, p. 320, Jan. 2021, ISSN: 1424-8220. (visited on 10/26/2022).

[12] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, May 18, 2015. arXiv: 1505 . 04597[cs]. (visited on 07/13/2022).

[13] Y. Li, J. Ma, and Y. Zhang, "Image retrieval from remote sensing big data: A survey," *Information Fusion*, vol. 67, pp. 94–115, 2021, ISSN: 1566-2535.

[14] T. Adão, J. Hruška, L. Pádua, *et al.*, "Hyperspectral imaging: A review on uav-based sensors, data processing and applications for agriculture and forestry," *Remote Sensing*, vol. 9, no. 11, 2017.

[15] Y. Lu and S. Young, "A survey of public datasets for computer vision tasks in precision agriculture," *Computers and Electronics in Agriculture*, vol. 178, p. 105 760, Nov. 2020, ISSN: 01681699. (visited on 03/13/2023).

[16] A. Bouguettaya, H. Zarzour, A. Kechida, and A. M. Taberkit, "Deep learning techniques to classify agricultural crops through UAV imagery: A review," *Neural Computing and Applications*, vol. 34, no. 12, pp. 9511–9536, Jun. 2022, ISSN: 0941-0643, 1433-3058. (visited on 03/13/2023).