IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10642298

DETECTION OF DIFFERENT CROP GROWTH STAGES BY APPLYING DEEP LEARNING OVER SENTINEL-2 IMAGES OF BAHÍA, BRAZIL

Yineth Viviana Camacho-De Angulo [1,2], *Darwin Alexis Arrechea-Castillo* [3,4], *Yessica Carolina Cantero-Mosquera* [4], *Yady Tatiana Solano-Correa* [2,4], *Mauro Roisenberg* [1]

[1]Universidade Federal de Santa Catarina. Florianópolis - SC - Brasil, 88040-900.
[2] Universidad Tecnológica de Bolívar. Cartagena, Colombia, 130010.
[3] International Center for Tropical Agriculture. Americas Hub – Palmira, Colombia, 763537.
[4] Universidad del Cauca. Popayán, Colombia, 190001.

ABSTRACT

This study delves into the agricultural landscape of Bahia, Brazil, employing the Mask R-CNN deep learning model with satellite imagery to detect three crop growth stages (early, mid-growth and maturity stage). This model is suited to the region's complex terrain and diverse crop patterns, providing accurate instance segmentation crucial for monitoring crop development. Remarkable results have been achieved with a limited dataset of just 54 images for training, underscoring the model's efficiency in scenarios where extensive data collection is challenging. The validation metric chosen for this study is the Intersection over Union (IoU), preferred for its ability to quantify the pixel-wise overlap between the predicted and actual segmentations, offering a clear measure of accuracy in spatial contexts. An IoU of 90% was obtained, demonstrating Mask R-CNN's robustness and potential for precision agriculture in challenging environments.

Index Terms— Crops Growth, Deep Learning, Image Instance Segmentation, Remote Sensing, Mask R-CNN.

1. INTRODUCTION

In precision agriculture, the seamless integration of cutting-edge technologies has ushered in a new era of data-driven decision-making. One of the pivotal aspects in this domain is the accurate monitoring and assessment of crop growth stages, a task traditionally reliant on labor-intensive field observations [1, 2]. Timely and precise information about crop growth enables farmers to make informed decisions regarding irrigation, fertilization, pest control, and harvesting [3]. The state of Bahia, Brazil, with its diverse agroecological zones and varied climatic conditions, presents a microcosm of the challenges and opportunities inherent in sustainable agriculture [4]. From the soybean fields that stretch across the Cerrado to the cocoa plantations nestled in the Atlantic Forest; understanding and managing the diverse growth stages of crops is critical for optimizing yields and ensuring the resilience of Bahia's agricultural sector [5]. The dynamic agricultural landscapes of this state encapsulate the equilibrium between traditional practices and technological innovation [6]. The most advanced techniques for detecting crops, crop types and crop growth stages involve the integration of deep learning algorithms with high-resolution satellite imagery, even though there exist simpler approaches that have also shown to be robust and precise. An example of this can be seen in [7] and [8], where authors exploited Sentinel-2 (S2) Satellite Image Time Series (SITS) in order to accurately map both crop fields and crop-types by exploiting automatic techniques based on image processing and standard neural networks and achieving accuracies over 90%. In the context of automatic detection of crop growth stages, in [9] a model was trained for Planet images and by exploiting regression analysis using the Gradient Boosted Decision Tree Technique (GBDT) to make predictions of cabbage growth stages from 0 to 70 days, with accuracies ranging between 51-74%.

This research presents an application of deep learning (DL) techniques for the detection, classification and instance segmentation of crop growth stages in Bahía, Brazil by exploiting S2 images. Similar works are not found over the area (to the best of authors knowledge).

2. STUDY AREA AND DATA

This research is based in the State of Bahía, located in the northeastern region of Brazil. The state is a highland region with dramatic landscapes, caves, waterfalls, and hi-king trails [4]. Bahia has a diverse economy, with contributions from agriculture, industry, and services. It is a major producer of commodities such as cocoa, sugar cane, and soybeans. The studied area is shown in Fig.1 and is covered by nearly 11 S2 tiles. The selected images were acquired on December 15, 2023. To validate the results, the labeling of three crop growth stages was made: early, mid, and mature; considering the presence of soybean and rice crops in the area. This task was done by photointerpretation over the 54 small areas highlighted by green squares in Fig.1 and by making use of a RGB image. These images offer a spatial resolution of 10 m/pixel, which aligns perfectly with the extensive nature of the agricultural fields in the region, where individual crops typically exceed 100 ha. This spatial resolution strikes an ideal balance, offering sufficient detail for accurate crop growth mapping, while avoiding the impracticalities associated with processing higher-resolution imagery over such large spatial scales. Conversely, images with lower spatial resolution would fail to capture the necessary details for effective analysis.

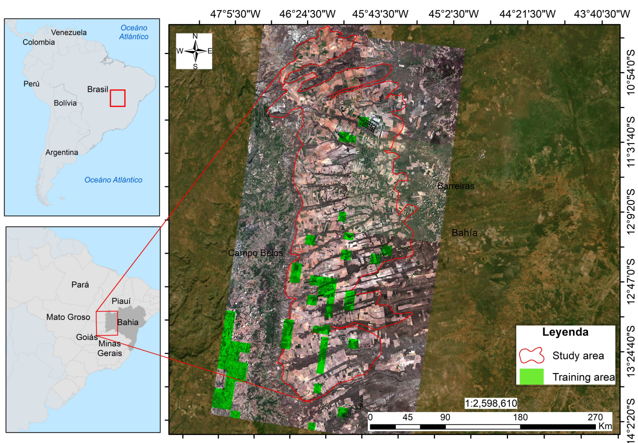

Fig. 1: Study area location in Bahía, Brazil; right panel: real color image with training areas highlighted in green.

3. PROPOSED APPROACH FOR CROP GROWTH STAGES DETECTION

This study follows the workflow provided in Figure 2. The methodology comprises several interconnected stages:

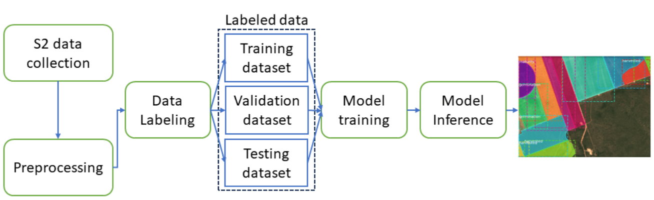

Fig. 2: General proposed methodology.

3.1. S2 Data Collection and Pre-processing

S2 images were systematically chosen and downloaded from the Copernicus Data Space Ecosystem. The original S2 tiles, with 10980x10980 pixels, were partition into smaller, more manageable segment,s of 1024x1024 pixels. This segmentation facilitated the efficient processing of the data, ensuring that the robustness of the Mask R-CNN framework did not result in prohibitively heavy computational loads. This is done for all the images required to analyze the studied area. At the same time, homogenization stages are applied in order to guarantee similar spectral responses across different images.

3.2. Data Labeling

Images were annotated by means of photointerpretation and using the Computer Vision Annotation Tool (CVAT) [10], an open-source web-based tool that is known for its user-friendly interface and powerful functionalities. CVAT was specifically chosen for this task due to its efficiency in handling large datasets and its ability to provide precise annotation tools which are vital for identifying the nuanced differences between various crop growth stages. Annotations were performed by delineating polygons indicative of distinct crop growth stages. For model training, 42 images were annotated with crop growth stages. The validation set contained 120 images for parameter tuning and overfitting prevention. Additionally, 5 images were annotated for the testing dataset to assess the model's generalization to new data.

3.3. Model Training

A variation of the Mask R-CNN frameworwk [11] was employed using the training and validation dataset. This DL model is adept at instance segmentation, which is crucial for distinguishing between different growth stages on a per-pixel basis. During training, we leveraged transfer learning techniques, using a pre-trained model to initialize the weights, which we then refined using our agricultural dataset.

3.4. Model Inference

In the final stage, the trained Mask R-CNN model was deployed to infer the growth stages on the testing dataset. The model inference provided us with instance segmentation objects, indicating the classified crop growth stages. The model's performance was validated using the Intersection over Union (IoU) metric, which is particularly effective in spatial validation scenarios as it quantifies the overlap between predicted segments and ground truth labels.

4. PRELIMINARY RESULTS AND DISCUSSIONS

A successfully Mask R-CNN model was trained for crop growth stage detection, achieving promising results with a relatively small dataset, an indication to the effective use of transfer learning. Despite the complexity of the task and the complexity of the DL framework, the implementation of transfer learning allowed the model to learn from only 54 images: 42 for training and 12 for validation. This distribution ensured that overfitting was minimized during the training phase. During the inference stage, 5 additional unseen images were used. The inference outcomes were encouraging, as the model demonstrated a high level of accuracy in the predictions made. The visual result of 3 out of the 5 test images are detailed in Figure 3. These preliminary findings suggest that the model has generalized well, offering robust predictions even with a limited dataset.

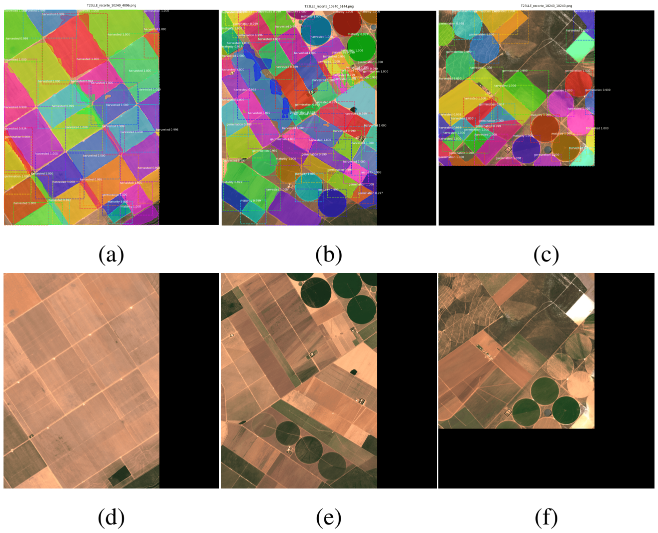

Fig. 3: Results of the model inference and corresponding RGB images for selected test areas.

Figure 3 shows the instance segmentation process for the three mapped crop growth stages, yielding visually good outcomes. These qualitative assessments are supported by the IoU metric. The IoU metric provides a ratio of overlap between the predicted segmentation and the ground truth, which is a reliable indicator of precision in spatial analysis. Figure 4 presents the IoU results for the five evaluated images. The specific outcomes for the early growth, mid-growth, and mature stages were 0.92, 0.90, and 0.94, respectively. This gives an average IoU of 0.92 which indicates a substantial degree of overlap and alignment with the actual growth stages, thus evidencing the model's generalization capabilities. This score, with a complex model like Mask R-CNN and a limited training dataset, underscores the efficacy of the transfer learning approach and the model's robustness in accurately classifying and segmenting different stages of crop growth.

The proficiency of the model, trained on a modest dataset of images, has shown commendable results, which is particularly noteworthy given the constraints of the training set. The analysis was conducted using only the RGB bands provided by S2, without leveraging the additional seven spectral bands available at 10m and 20m. This approach suggests that there is substantial room for improvement in model performance. Furthermore, widely recognized vegetation indices like the Normalized Difference Vegetation Index (NDVI) were not used. NDVI is crucial in precision agriculture for its ability to provide insights into plant health, vigor, and biomass, which are closely tied to the growth stages of crops [8]. Incorporating NDVI [12] and other spectral bands could significantly enhance the model's sensitivity to the subtle differences between growth stages, potentially enabling the mapping of a broader range of crop development phases. However, it is important to maintain a degree of modesty regarding these

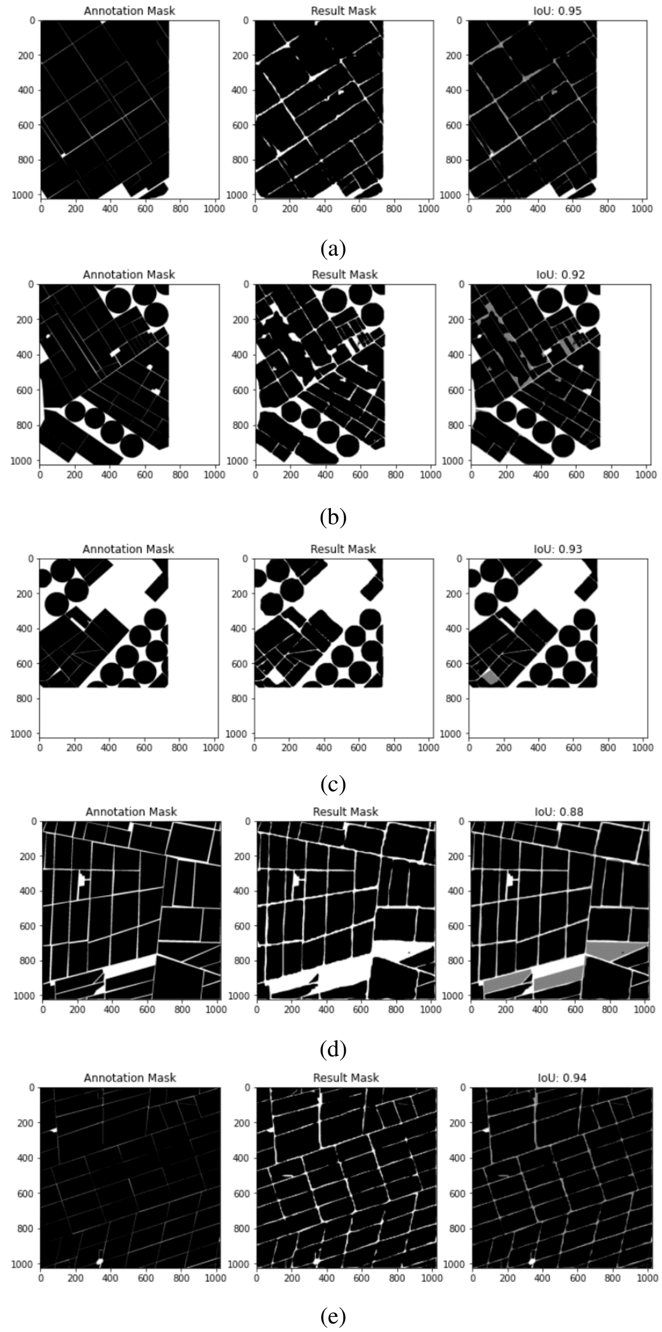

Fig. 4: Intersection over Union (IoU) results for the five test areas.

findings. The inference drawn from a limited set of just 5 images, while promising, requires further improvements and tests. Future work should not only consider expanding the dataset to include a wider variety of images but also explore the integration of additional quantitative metrics. Dissecting the IoU scores for each individual crop growth stage could yield deeper insights and further validate the model's discriminatory power. The preliminary results are encouraging, since they represent a foundation upon which further refinement of the model's capabilities can be built. The potential for incorporating additional spectral data and vegetative indices, along with a broader dataset for training, validation and inference, opens the floor for future enhancements and assessments of the model's performance.

5. CONCLUSIONS

A framework for detecting different growth stages, together with sementing crop fields, has been presented. The proposed framework is applied over S2 images, and explores three different crop growth stages (early, mid-growth and maturity stage). While maps showing the three stages are not directly provided (due to space constrains), an average IoU score above 90% reflects the model's high accuracy (early growth = 92%, mid-growth = 90%, and mature = 90%), especially considering the limited size of the training dataset, which consisted of only 54 small images. This high IoU score, despite the small dataset, highlights the model's capability to learn from a constrained number of samples, showcasing the potential for deploying DL in precision agriculture with limited resources. The results shown in this research suggest that the model has generalized well, offering robust predictions and underscore the potential for applying advanced DL techniques in precision agriculture, particularly in scenarios where data acquisition is challenging or resource-constrained. Despite the dataset's limitations, the model's predictions have proven to be robust and reliable. The results reinforce the viability of utilizing sophisticated DL methodologies in the context of precision agriculture. Moreover, the model's success achieved without the aid of additional spectral bands or vegetation indices like NDVI (which are typically integral to such tasks) opens avenues for further enhancements.

6. ACKNOWLEDGMENTS

This research was performed with support from the Ministry of Science Technology and Innovation, MINCIENCIAS, in Colombia, through the announcement No. 933 of 2023 for national doctoral programs with a territorial, ethnic and gender focus in the context of the Mission-Oriented Polic, the project ”4D Integrated Quantitative Characterization of Reservoirs” (Contract No. FEESC/PETROBRAS/4600671602); and the CGIAR Initiative on Accelerated Breeding, thanks all donors that globally support the work of the Initiatives through their contributions to the CGIAR system.

7. REFERENCES

[1] D. S. Bullock, M. Boerngen, H. Tao, B. Maxwell, J. D. Luck, L. Shiratsuchi, L. Puntel, and N. F. Martin, “The data-intensive farm management project: changing agronomic research through on-farm precision experimentation,” *Agronomy journal*, vol. 111, no. 6, pp. 2736–2746, 2019.

[2] K. Malhotra and M. Firdaus, “Application of artificial intelligence in iot security for crop yield prediction,” *ResearchBerg Review of Science and Technology*, vol. 2, no. 1, pp. 136–157, 2022.

[3] R. P. Sishodia, R. L. Ray, and S. K. Singh, “Applications of remote sensing in precision agriculture: A review,” *Remote Sensing*, vol. 12, no. 19, p. 3136, 2020.

[4] A. Robinson, *Bahia: The Heart of Brazil's Northeast*. Bradt Travel Guides, 2010.

[5] E. Benami, *Shaping the Producer's Problem: Essays on Land-Use Zoning and Certification in the Sustainability of Brazilian Oil Palm and Coffee*. Stanford University, 2018.

[6] K. G. Engås, J. Z. Raja, and I. F. Neufang, “Decoding technological frames: An exploratory study of access to and meaningful engagement with digital technologies in agriculture,” *Technological Forecasting and Social Change*, vol. 190, p. 122405, 2023.

[7] Y. T. Solano-Correa, K. Meshkini, F. Bovolo, and L. Bruzzone, “Automatic large-scale precise mapping and monitoring of agricultural fields at country level with sentinel-2 sits,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 3131–3145, 2022.

[8] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, “A semi-supervised crop-type classification based on sentinel-2 ndvi satellite image time series and phenological parameters,” in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 457–460, 2019.

[9] D.-C. Hsiou, F. Huang, F. J. Tey, T.-Y. Wu, and Y.-C. Lee, “An automated crop growth detection method using satellite imagery data,” *Agriculture*, vol. 12, no. 4, p. 504, 2022.

[10] CVAT.ai Corporation, “Computer Vision Annotation Tool (CVAT).”

[11] W. Abdulla, “Mask r-cnn for object detection and instance segmentation on keras and tensorflow.” https://github.com/matterport/Mask$_RCNN$, 2017.

[12] T. N. Carlson and D. A. Ripley, “On the relation between NDVI, fractional vegetation cover, and leaf area index,” vol. 62, no. 3, pp. 241–252.