UAV and ML-based Açaí Palm Monitoring in the Amazon: Facilitating YOLOv8 Usage in Agricultural Practices

Prince Nyarko[1], **Ilan Correa**[1], **Hervé Rogez**[2], **Aldebaro Klautau**[1]

[1]Núcleo de P&D em Telecomunicações,
Automação e Eletrônica (LASSE)
Espaço Inovação – Parque de Ciências e Tecnologia (PCT Guamá)
Estado Belém – PA – Brazil

[2]Centro de Valorização Agroalimentar de Compostos
Bioativos da Amazônia (CVACBA)
Espaço Inovação – Parque de Ciências e Tecnologia (PCT Guamá)
Estado Belém – PA – Brazil

prince.nyarko@itec.ufpa.br,{ilan,aldebaro}@ufpa.br,herverogez@gmail.com

Abstract. *This study explores the application of deep learning to detect and count açaí palm trees (*Euterpe oleracea*) in the Amazon using UAV imagery. A YOLOv8n model was trained on a georeferenced dataset collected across three municipalities at Northeast of the state of Pará, Brazil, achieving a mean average precision of 0.809 at IoU 0.5 (mAP50) and 0.312 at mAP50–95. The model was integrated into a web platform to enable non-experts, such as farmers and conservationists, to upload imagery and receive annotated results in real time. This facilitates more scalable and sustainable monitoring of açaí populations and supports forest conservation efforts.*

1. Introduction

The açaí palm tree (*Euterpe oleracea*) is a cornerstone of the Amazon rainforest ecosystem and a critical component of Brazil's agricultural economy. Known for its highly nutritious fruits (or drupes), the açaí juice has become a global superfood, driving increased demand for its cultivation and sustainable harvesting [IBGE 2019]. In Brazil, the açaí supply chain supports approximately 30,000 agro-extractive families, particularly in the states of Pará and Amapá, where it accounts for a significant portion of local income and employment [Euler 2020]. However, the detection and monitoring of açaí palm trees in dense and complex forest environments remain a significant challenge. Traditional methods, which rely on manual inspection and expert knowledge, are labor-intensive, time consuming, and prone to errors, making them unsuitable for large scale applications [Tan et al. 2022].

The need for accurate and efficient detection methods is further underscored by the growing pressures on the Amazon rainforest, including deforestation and climate change. According to the National Institute for Space Research (INPE), deforestation in the Amazon reached 13,235 km² in 2021, the highest level in over a decade [INPE 2021]. This has direct implications for açaí production, as the species thrives in biodiverse, intact ecosystems. Sustainable management of açaí palm populations is therefore essential, not only for economic reasons, but also for preserving the ecological integrity of the Amazon [Freitas et al. 2021].

Recent advancements in computer vision and deep learning offer promising solutions to these challenges [Osco et al. 2020]. Among these, the YOLO (You Only Look Once) family of models has emerged as a state-of-the-art solution for real-time object detection. YOLOv8, one of the latest versions of this model, combines high accuracy with computational efficiency, making it particularly well-suited for detecting açaí palm trees in dense and overlapping forest canopies [Ultralytics 2023]. Studies have demonstrated the effectiveness of deep learning models in similar agricultural applications, such as the detection of corn [Mendonça and Guedes 2024] and palm oil trees [Junos and Thannirmalai 2021].

In this work, we leverage the YOLOv8 model to develop a robust and efficient system for counting and detecting açaí palm trees. Trained on a diverse dataset of açaí palm images collected from the Amazon region, the model achieved a mean Average Precision (mAP50) of 0.809, successfully identifying trees under various environmental conditions, including dense foliage, occlusions, and varying lighting scenarios. To make this technology accessible to non-technical users, we have integrated the YOLOv8 model into a user-friendly web-based platform. This platform enables users to upload images, process them in real-time, and visualize detection results with annotated bounding boxes and confidence scores.

The integration of YOLOv8 with the web platform bridges the gap between advanced machine learning capabilities and practical, real-world applications. By providing an intuitive interface and real-time processing capabilities, the platform empowers farmers, researchers, and conservationists to monitor açaí palm populations efficiently. This not only supports sustainable harvesting practices but also contributes to the preservation of the Amazon rainforest ecosystem.

In the subsequent sections of this paper, we detail the methodology for training and evaluating the YOLOv8 model, describe the development and functionality of the web platform, and present the results of the system's performance.

2. Materials and methods

This section outlines the methodology for training the YOLOv8 model and developing a web platform to detect açaí palm trees. The approach is divided into three main stages, as illustrated in Figure 1.

Section 2.1 details the data acquisition process, including the use of UAVs to capture imagery and the methods applied for annotation. Section 2.2 explains dataset preparation, covering labeling, preprocessing, and enhancements to improve training quality. Section 2.3 describes the YOLOv8 model architecture, the training process, and the adjustments made to optimize performance. Section 2.4 presents the design and implementation of the web platform, highlighting its core features and how users can interact with it for detection and visualization tasks.

The proposed pipeline integrates UAV-based data collection, YOLOv8 training, and a web-based detection platform. It also incorporates OpenDroneMap (ODM) for scalable processing. This structure ensures reproducibility, scalability, and practical application in real-world forest monitoring scenarios.

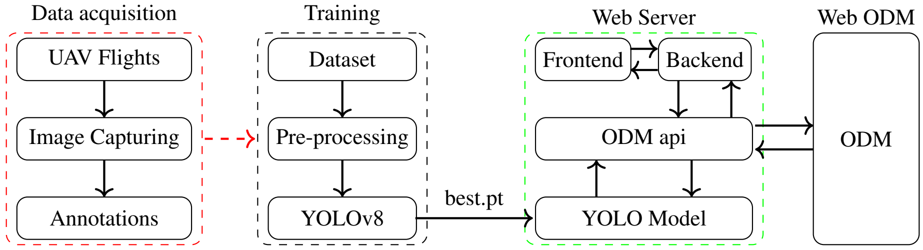

Figure 1. Pipelines of the proposed methodology for acai palm tree detection.

2.1. Data acquisition

We used the DJI Mini 4 Pro UAV to capture images to create the dataset imagery, utilizing multi-constellation GNSS (GPS/Galileo/BeiDou) for reliable geotagging. Its 48 MP (1/1.3" CMOS, f/1.7) camera ensured high-quality image capture, including in low light conditions.

Flight missions were conducted manually over açaí plantations in three municipalities: Abaetetuba, Santa Luzia do Pará, and São Sebastião da Boa Vista, all in the state of Pará, Brazil, which are located on the globe in Figure 2, chosen for their high palm density and environmental diversity. Operated at 30 meters, the drone followed a grid path with at least 70% image overlap, capturing 2,000 high-resolution images across varying weather and lighting conditions to ensure dataset diversity. With the integration of GPS

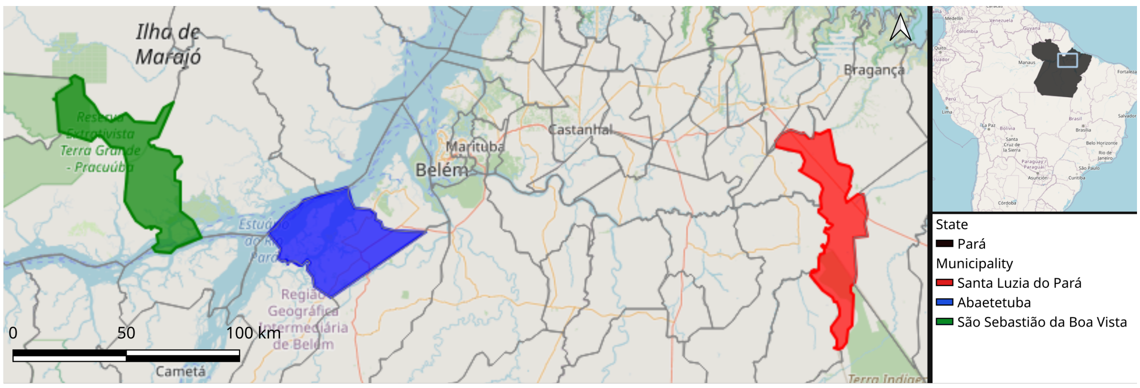

Figure 2. Map of the study area (Northeast of the state of Pará, Brazil).

coordinates in images enabled the generation of a georeferenced orthomosaic using OpenDroneMap (ODM). ODM is an open-source software platform designed for processing aerial drone imagery into detailed maps, 3D models, and point clouds [Castro et al. 2020]. This orthomosaic, combining overlapping images and GNSS data, provided a detailed map of the study area, as shown in Figure 3, crucial for precise tree localization, model validation, and advanced analyses like tree density and distribution estimation.

2.2. Dataset

The dataset preparation started with 612 manually selected images based on quality, clarity and non-repeating, from the total captured. Each image was manually annotated with precise bounding boxes for açaí palms (class "açai"), using zoom tools and grid overlays to ensure accuracy, especially for small, densely packed trees. All annotations underwent rigorous quality review, after which the dataset was systematically split into train-

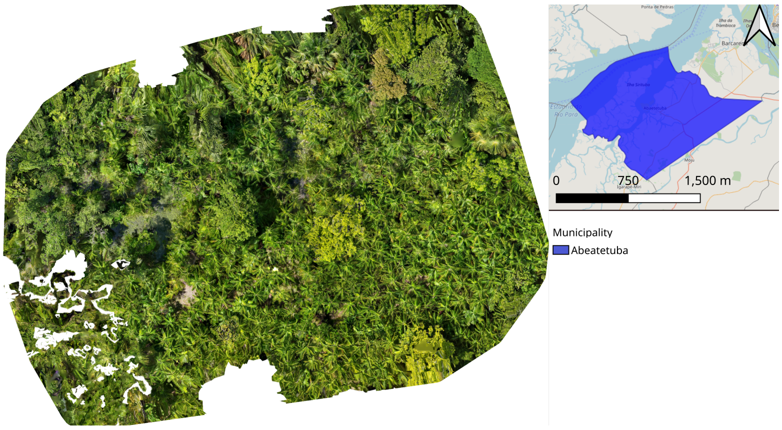

Figure 3. A orthomosaic generated by the ODM software.

ing (70%), validation (20%), and test (10%) subsets to ensure unbiased model evaluation. To enhance the dataset, we applied multiple augmentation techniques only no the training and test sets including rotations (90°, 180°, 270°), scaling (0.8×-1.2×), horizontal/vertical flips, brightness (±30%) and contrast (±20%) adjustments, and Gaussian noise injection, these steps expanded the dataset to 1,592 images. All images were resized to 640x640 pixels to ensure compatibility with YOLOv8. Finally, the dataset was converted to YOLOv8 format, with each image paired to its corresponding label file containing normalized bounding box coordinates and class identifiers for efficient model training.

2.3. YOLOv8 Architecture and Training

The *YOLOv8* model was chosen for its state-of-the-art performance in real-time object detection and its ability to effectively handle small objects, making it ideal for detecting açaí palm trees in the complex and dense environment of the Amazon rainforest. YOLOv8 expanded on the success of previous YOLO iterations and introduced new features and enhancements to further improve its performance and versatility. We employed YOLOv8n, a lightweight variant of the YOLOv8 architecture, selected for its balance between detection accuracy and inference efficiency suitable for web deployment. Although YOLOv11 has since been released, YOLOv8 was chosen due to its stable maturity at the time of development and its community support, especially within the Ultralytics ecosystem.

YOLOv8 follows a single-stage detection approach, meaning it predicts bounding boxes and classes probabilities directly from the input image in a single pass. At its core, YOLOv8 features a deep convolutional neural network (CNN) architecture known as Darknet. This backbone network extracts features from the input image, capturing relevant information at multiple spatial scales. This capability allows YOLOv8 to detect objects of varying sizes and complexities. Additionally, YOLOv8 replaces the C3 module with the C2f module. Unlike C3, which only utilizes the output of the last bottleneck, C2f concatenates the outputs of all bottlenecks, further enhancing feature representation which are shown in Figure 4 [Sohan et al. 2024].

The training process used transfer learning to fine-tune the YOLOv8 model, starting with YOLOv8n weights pre-trained on the COCO dataset, which contains diverse object categories. This allowed the model to specialize in detecting açaí palm trees, reducing training time and improving generalization. An input size of 640x640 pixels was chosen to balance accuracy and computational efficiency.

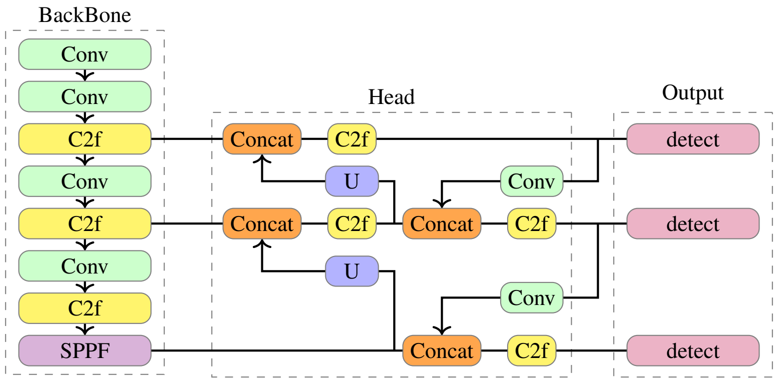

Figure 4. YOLOv8 architecture.

The model was trained exclusively on the "açaí" class to learn its distinctive features. Training was performed with a batch size of 16, a learning rate of 0.001, and for 50 epochs using the Adam optimizer. Loss functions monitored during training included box loss (L_{box}), class loss (L_{class}), and distribution focal loss (L_{DFL}). The total loss (L_{total}) is:

$$L_{total} = L_{box} + L_{class} + L_{DFL} \tag{1}$$

To evaluate model performance, the following standard object detection metrics were computed on the test set:

- **Precision**. The proportion of correctly predicted açaí trees among all predicted trees.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \tag{2}$$

- **Recall**. The proportion of actual açaí trees correctly identified by the model.

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \tag{3}$$

- **mAP@50**. Mean Average Precision calculated at an IoU threshold of 0.50, measuring detection performance with moderate overlap [Padilla et al. 2020].

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{4}$$

2.4. Web Development for Model Interaction

This section describes the development of the web application, which enables users to interact with the YOLOv8 model for açaí palm tree detection. The application consists of a frontend for user interaction (Section 2.4.1) and a backend for processing and communication with WebODM and the YOLOv8 model (Section 2.4.2). Together, these components provide a seamless and user-friendly platform for detecting and analyzing açaí palm trees.

2.4.1. Frontend

The frontend was developed using HTML, CSS, and JavaScript to create a clean and intuitive interface for users to interact with the application. HTML (HyperText Markup Language) was used to structure the content, defining elements such as the image upload form, progress indicators, and result display areas. CSS (Cascading Style Sheets)

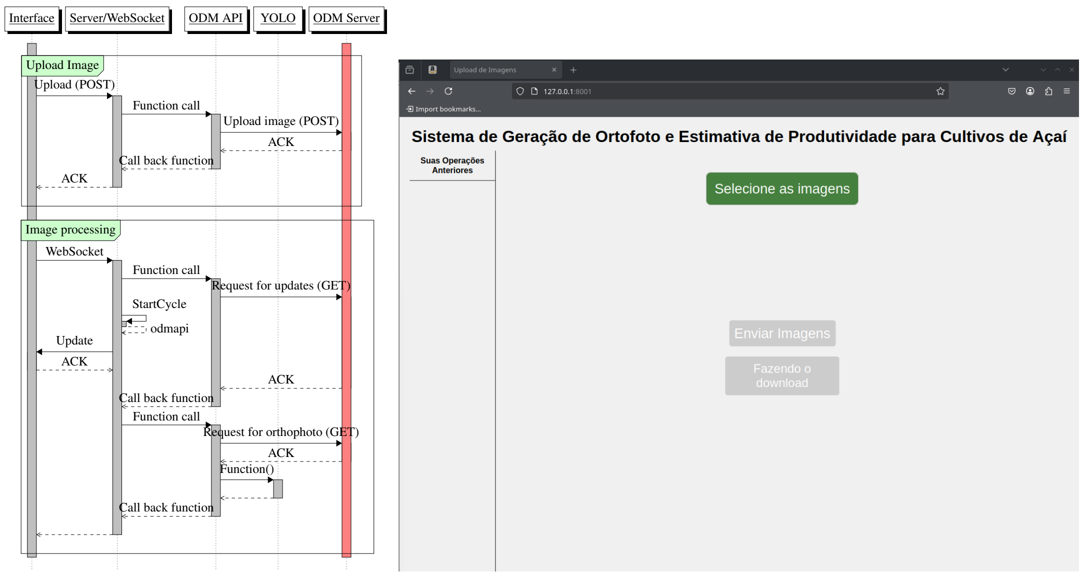

Figure 5. (a) UML sequence diagram and (b) Web interface for açaí palm tree detection.

was employed to style the interface, ensuring a visually appealing and responsive design that adapts to different screen sizes and devices. JavaScript added interactivity, enabling dynamic features such as real-time updates, file validation, and progress tracking.

Communication between the frontend and backend is handled through HTTP (HyperText Transfer Protocol) and WebSockets. When users upload images, the frontend sends them to the backend via HTTP requests, specifically using methods like POST to transmit data securely. During processing, WebSockets provide a persistent, bidirectional communication channel, allowing the backend to send real-time updates to the frontend. This integration ensures users receive immediate feedback, such as loading indicators and progress messages, without needing to manually refresh the page.

The primary role of the frontend is to allow users to upload images of açaí palm trees, monitor the processing status, and view detection results seamlessly. Designed with simplicity in mind, the frontend ensures that even non-technical users can navigate the platform effortlessly. Once processing is complete, the frontend displays the results with clear annotations, highlighting detected açaí palm trees and making it easy for users to interpret the output.

2.4.2. Backend Implementation

The backend forms the core of the application, integrating the frontend, WebODM, and the YOLOv8 model to deliver a seamless user experience. Built on FastAPI, a modern Python framework, the backend utilizes asynchronous operations to handle concurrent requests efficiently, ensuring high performance and scalability [Bansal and Ouda 2022].

When a user uploads an image, the backend transmits it to the WebODM server via the OpenDroneMap (ODM) API for processing. WebODM generates a high-resolution orthophoto, which is divided into 640×640 pixel blocks for analysis by the YOLOv8 model. This block size aligns with the model's input requirements, enabling precise and efficient detection of açaí palm trees. The total number of image subdivisions, N_{blocks}, is determined as:

$$N_{\text{blocks}} = \left\lceil \frac{\text{width of the orthophoto}}{640} \right\rceil \times \left\lceil \frac{\text{height of the orthophoto}}{640} \right\rceil, \tag{5}$$

where width of the orthophoto and height of the orthophoto represent the dimensions of the orthophoto, and $\lceil \cdot \rceil$ denotes the ceiling function, ensuring complete coverage of the image.

For each block, the YOLOv8 model performs object detection to identify açaí palm trees. Following detection, the productivity estimation algorithm, based on references published by Embrapa Amazônia [Neto et al. 2021], quantifies the productivity of the identified trees. Each detected tree is assigned an individual productivity value, calculated as:

$$P_{i,j} = n_{i,j} \times \alpha, \tag{6}$$

where $P_{i,j}$ is the productivity for block, $n_{i,j}$ is the number of detected açaí palm trees in block , α is the productivity factor (e.g., 5.0 kg per tree), as defined by Embrapa Amazônia. The total productivity P_{total} for the entire orthophoto is the sum of productivities across all blocks:

$$P_{\text{total}} = \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} P_{i,j}, \tag{7}$$

where N_x and N_y represent the number of blocks in the horizontal and vertical directions, respectively.

After processing, the results from individual blocks are merged to generate a comprehensive detection output, complete with annotations, total number of açaí palm trees detected, and total productivity estimated per year. The backend also employs WebSockets to provide real-time updates to users, ensuring transparency and engagement during long-running tasks. Combined with asynchronous task processing and robust error handling, the backend delivers a reliable and scalable solution for açaí palm tree detection and productivity estimation.

3. Results

The results of this work are divided into two main components: the qualitative Analysis of YOLOv8 Model in detecting açaí palm trees (Section 3.1) and the functionality of the web-based platform designed for user interaction with the model (Section 3.2). Together, these components form a comprehensive solution for accurate detection and user-friendly application.

3.1. Qualitative Analysis of YOLOv8 Model

The YOLOv8 model was trained on the dataset created in Section 2.2. Training was performed on an Ubuntu 23.10 system with an Intel i7-5930K, 32GB RAM, and an NVIDIA RTX 3060 (12GB VRAM), using PyTorch and Ultralytics for implementation. The GPU accelerated training, while TensorBoard tracked metrics (loss, mAP, learning rate). Dataset images and labels followed YOLO formatting (Section 2.2) and it was evaluated on the test subset defined during dataset preparation. As detailed in Section- 2.3, model performance was assessed using standard object detection metrics.

These results in Figure 6 demonstrate that the model can accurately detect and localize açaí palm trees across a range of complex scenarios. The high precision score of 0.827 indicates a low rate of false positives, while the strong recall value of 0.787 reflects the model's ability to identify the majority of açaí trees present in the imagery. Additionally, the mAP@50 score confirms the model's effectiveness in producing bounding boxes that closely match the ground truth.

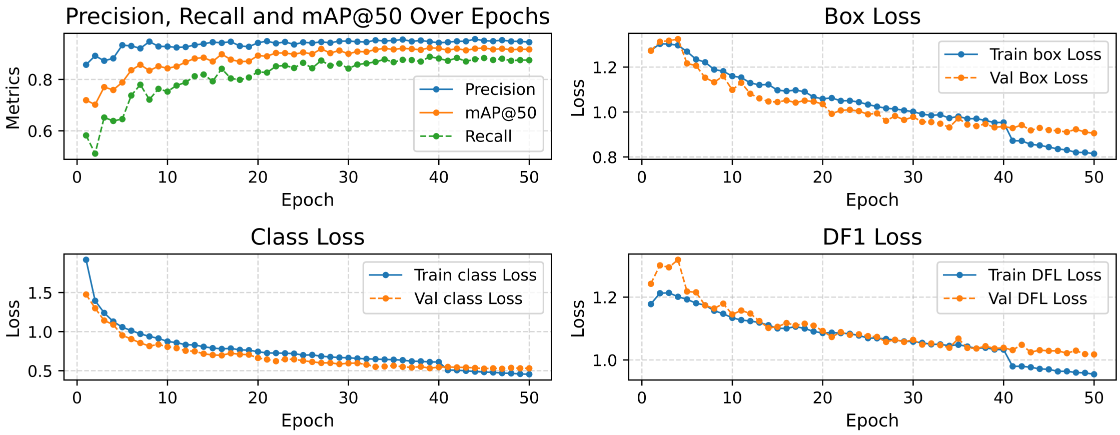

Figure 6. Precision-recall curve and training metrics for the YOLOv8 model.

The training process exhibited steady convergence, with decreasing loss curves across all monitored components (L_{box}, L_{class}, and L_{DFL}), as visualized in Figure 6. To evaluate real-world applicability, annotated images were analyzed in challenging scenarios, such as overlapping canopies and cluttered backgrounds. The model consistently produced accurate bounding boxes and reliable confidence scores. Figure 7 showcases examples of the model's detections.

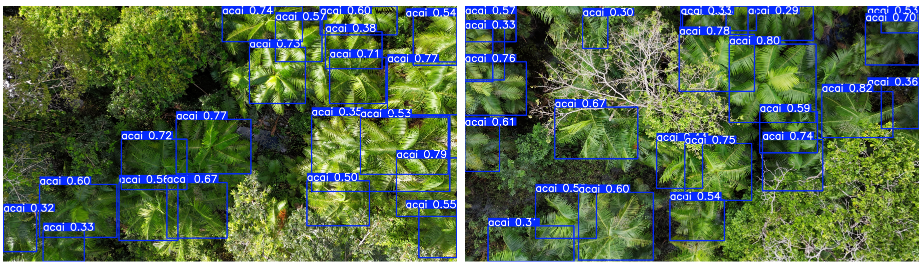

Figure 7. Açaí palm tree detection using the YOLOv8 model. Bounding boxes indicate detected trees, with confidence scores displayed.

3.2. Web Platform Evaluation

The web-based platform was developed to provide an intuitive and accessible interface for interacting with the YOLOv8 model. It enables users to upload images and visualize detection results. Rigorous performance and scalability testing confirmed the platform's ability to efficiently handle multiple concurrent user requests, ensuring reliability across varying workloads.

Figure 8 showcases the platform's clean and user-friendly interface, which includes image upload, detection results with bounding boxes, and download options. The design prioritizes usability, enhancing the overall user experience. By integrating the YOLOv8 model with the web platform, a comprehensive solution for açaí palm tree detection is established, enabling users to perform detection tasks efficiently without requiring technical expertise.

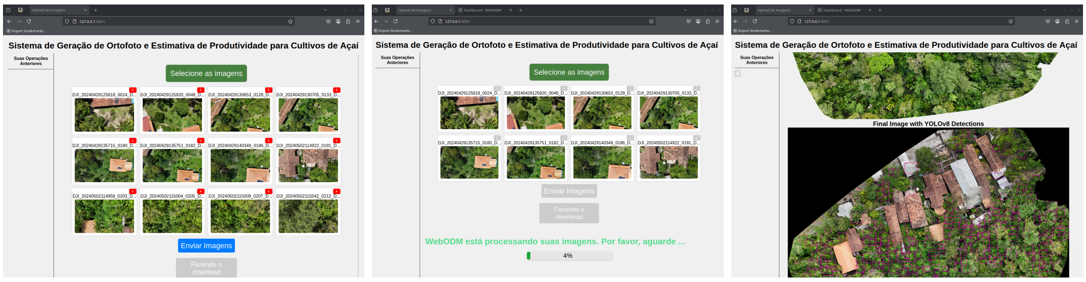

Figure 8. Web Interface for Açaí Palm Tree Detection. The interface includes image upload, detection results, and download options.

4. Conclusion and future directions

This study introduced a system that combines UAV imagery with the YOLOv8n object detection model to identify and count açaí palm trees in the Amazon region. The model achieved a mean Average Precision of 0.809 (mAP50) and 0.312 (mAP50–95), demonstrating its effectiveness under challenging conditions such as dense canopy and occlusion. A web-based platform was developed to allow non-expert users to upload images and receive real-time feedback and detection results, making the technology more accessible for use in agriculture and conservation.

While the system shows strong potential, it has several limitations. It has not yet been tested with end users, such as farmers or environmental agents, and no comparisons have been made with alternative models like YOLOv7 or YOLOv11. Additionally, the dataset and source code are not yet publicly available, which limits reproducibility.

Future work will focus on conducting usability studies with local stakeholders, benchmarking performance against other detection models, improving precision at higher IoU thresholds, and releasing the dataset and platform code under an open license. These steps will enhance the tool's reliability, transparency, and impact for sustainable resource monitoring in the Amazon.

Acknowledgments

This article was developed in the project INOVAÇAÍ, which was funded by the Amazonian Foundation for the Support of Studies and Research (FAPESPA), process number 2023/737637. This work was also partially funded by CNPq (Grant #. 422439/2023-1) and by the Sao Paulo Research Foundation (FAPESP) (Grant #. 2023/00811-0). We thank these institutions for their support.

References

Bansal, P. and Ouda, A. (2022). Study on integration of fastapi and machine learning for continuous authentication of behavioral biometrics. In *2022 International Symposium on Networks, Computers and Communications (ISNCC)*, pages 1–6.

Castro, P., Borges, J., Klautau, S., Pinheiro, S., Shibata, L., Lins, S., and Klautau, A. (2020). Uav-based crop monitoring through multispectral image processing using open-source tools. In *Proceedings of the XXXVIII Brazilian Symposium on Telecommunications and Signal Processing (SBrT)*, pages 22–25, Florianópolis, SC.

Euler, A. M. C. (2020). Aspectos socioeconômicos e ambientais do açaizeiro no estado do amapá. *II Jornada de Botânica e Ecologia, II Jornada Amapaense de Botânica*, pages 254–258.

Freitas, M. A., Magalhães, J. L., Carmona, C. P., Arroyo-Rodríguez, V., Vieira, I. C., and Tabarelli, M. (2021). Intensification of açaí palm management largely impoverishes tree assemblages in the amazon estuarine forest. *Biological Conservation*, 261:109251.

IBGE (2019). Produção da extração vegetal e da silvicultura. *Instituto Brasileiro de Geografia e Estatística*.

INPE (2021). Monitoramento do desmatamento na amazônia legal por satélite. *Instituto Nacional de Pesquisas Espaciais*.

Junos, M. H. and Thannirmalai, S. (2021). Performance evaluation of improved yolo models for fruit detection. In Editor, A., editor, *Smart Agriculture Technologies*, pages 123–134. Springer.

Mendonça, A. U. and Guedes, E. B. (2024). Classificacao e deteccao inteligentes de graos para agricultura digital na cultura do milho. In *Laboratório de Sistemas Inteligentes, Escola Superior de Tecnologia, Universidade do Estado do Amazonas*, pages 1–10. Universidade do Estado do Amazonas.

Neto, J. T. F., de Carvalho, J. E. U., do Socorro Padilha de Oliveira, M., and do Nascimento, W. M. O. (2021). Estimativas de produtividade. Content fully migrated on: 12/20/2021.

Osco, L. P., dos Santos de Arruda, M., Junior, J. M., da Silva, N. B., Ramos, A. P. M., Érika Akemi Saito Moryia, Imai, N. N., Pereira, D. R., Creste, J. E., Matsubara, E. T., Li, J., and Gonçalves, W. N. (2020). A convolutional neural network approach for counting and geolocating citrus trees in uav multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 160:97–106.

Padilla, R., Netto, S. L., and da Silva, E. A. B. (2020). A survey on performance metrics for object-detection algorithms. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 237–242.

Sohan, M., Sai Ram, T., and Rami Reddy, C. V. (2024). A review on yolov8 and its advancements. In Jacob, I. J., Piramuthu, S., and Falkowski-Gilski, P., editors, *Data Intelligence and Cognitive Informatics*, pages 529–545. Springer Nature Singapore, Singapore.

Tan, X. J., Cheor, W. L., Yeo, K. S., and Leow, W. Z. (2022). Expert systems in oil palm precision agriculture: A decade systematic review. *Journal of King Saud University - Computer and Information Sciences*, 34(2):1321–1334.

Ultralytics (2023). Yolov8 documentation. Accessed: 2024-10-01.