Applying a U-Net Convolutional Neural Network for Mapping Banana Crops in the Atlantic Forest Region of Brazil Using CBERS-4A High Spatial Resolution Imagery

Hugo do Nascimento Bendini [1], Mateus de Souza Miranda [2], Thales Sehn Körting [3]

[1] Department of Fisheries Resources and Aquaculture (DERPA), Faculty of Agrarian Sciences of Vale do Ribeira (FCAVR), State University of Sao Paulo (UNESP), Registro, Brazil – hugo.bendini@unesp.br
[2] Artificial Intelligence Laboratory for Aerospace and Environmental Applications, Applied Computing, National Institute for Space Research, Brazil – mateusmirandaa2@hotmail.com
[3] Remote Sensing Postgraduate Program (PGSER), Earth Sciences General Coordination (CGCT), Brazil's National Institute for Space Research (INPE) – thales.korting@inpe.br

Keywords: Remote sensing; Convolutional Neural Networks; U-Net; CBERS-4A; Crop mapping; Banana crops.

Abstract

Mapping banana crops in heterogeneous tropical landscapes remains challenging due to spectral similarity with surrounding vegetation, fragmented smallholder systems, and complex land-use mosaics. This study applies a deep learning approach, using a U-Net model, on high spatial resolution CBERS-4A imagery to map banana crops in Brazil's Ribeira Valley, a subtropical region with high rainfall and heterogeneous land cover. Reference data were created through manual interpretation of satellite imagery supported by field knowledge. Representative image tiles were selected and divided into smaller patches for model training, validation, and testing. The U-Net model was trained with standard optimization techniques and evaluated using common semantic segmentation metrics. On the validation set, it achieved strong performance (accuracy 0.91, F1-score 0.84, AUC-ROC 0.96, AUC-PR 0.92). Performance was maintained or improved on the independent test set (accuracy 0.91, F1-score 0.86, AUC-ROC 0.97, AUC-PR 0.93), indicating good generalization. with high agreement between predicted and reference data. Most errors occurred at boundaries between crops and natural vegetation. Additional validation using official agricultural statistics confirmed consistency at the municipal scale. The approach demonstrates that high-resolution imagery combined with deep learning can effectively map banana crops in the region and offers a promising tool for agricultural monitoring and land-use planning in complex environments. The code, trained models, and data are publicly available at https://github.com/hnbendini/banana-unet-mapping.

1. Introduction

Bananas, a globally important fruit, provide employment, support rural livelihoods, and contribute to food security in tropical and subtropical regions, including Latin America and sub-Saharan Africa, yet generate residues and require fertilizers and pesticides. If mismanaged, it can cause environmental and phytosanitary impacts (Varma; Bebber, 2019). The Atlantic Forest, one of the world's most biodiverse biomes, retains about 21% of its area in Brazil's Ribeira Valley, home to traditional communities, family farmers, and major national banana-producing areas (Fundacao SOS Mata Atlantica; INPE, 2023). In this region, cultivation ranges from smallholder to monoculture systems; while agroforestry enhances ecosystem services, monocultures, especially in riparian permanent preservation areas or near conservation units, can cause environmental impacts if not properly managed (Ribeiro et al., 2009; Soares et al., 2025).

The use of remote sensing (RS) in banana cultivation has evolved considerably over the past four decades, driven by advances in sensor technologies, spatial and temporal resolution, and the integration of artificial intelligence (AI) techniques, including machine learning (ML) and deep learning (DL). These technological advancements have enabled a shift from coarse-scale area mapping based on spectral indices toward high-precision detection of individual plants and early disease diagnosis.

Pioneering studies by Pedroni (2003) utilized Landsat TM imagery to map commercial export farms, though spectral mixing remained a challenge. Simultaneously, Synthetic Aperture Radar (SAR) proved essential for tropical regions due to its cloud-penetrating capability and sensitivity to the high-water content and volumetric roughness of banana canopies. Early efforts by Beaulieu et al. (1999) and Leclerc & Hall (1999) demonstrated that commercial crops could be identified by their distinct brightness and rectangular shapes in radar imagery. Verhoeye and De Wulf (1999) proved that multitemporal ERS-1 SAR could achieve considerable accuracies in binary classifications by exploiting the stability of the crop's backscatter signature.

Johansen et al. (2009) used SPOT-5 and Object-Based Image Analysis (OBIA) to incorporate textural and spatial information, effectively distinguishing bananas from other orchard crops. The introduction of Unmanned Aerial Vehicles (UAVs) began to bridge the gap between satellite mapping and field-level monitoring (Johansen et al., 2014). High-resolution UAV data and the application of Convolutional Neural Networks (CNNs) moved beyond simple mapping to plant-level health assessment and individual plant counting (Neupane et al., 2019). Ye et al. (2020) demonstrated that machine learning classifiers (SVM, RF, and ANN) could identify Fusarium wilt (TR4) with high accuracy using UAV-based multi-spectral imagery. Their work highlighted the critical role of the red-edge band in disease detection and suggested that spatial resolutions higher than 2 meters are necessary to identify phytosanitary anomalies reliably. Modern approaches integrate Sentinel-1 (radar) and Sentinel-2 (optical) images to overcome persistent cloud cover while providing high-frequency updates. Studies by Aksu & Demir (2021) and Alabi et al. (2022) have shown that this fusion is a robust method for extracting banana crops within heterogeneous and fragmented smallholder systems.

Recent studies using dense Sentinel-2 time series (with temporal resolution around 5 days) combined with ML methods have shown strong potential for large-scale mapping, yet the 10-meter spatial resolution often limits detection in smallholder and intercropped systems typical of regions like Brazil’s Ribeira Valley (Soares et al., 2025; Guzmán-Álvarez et al., 2022). Despite technological progress, banana crops remain challenging due to several inherent factors. The structural and phenological complexity of banana crops, including asynchronous growth stages, flexible and dynamic canopies, and intercropped or mixed farming systems, introduces significant spectral variability, complicating both pixel-based and object-based classification.

Platform and data limitations further constrain accuracy, as medium-resolution satellite imagery often produces mixed pixels in small plots, cloud cover limits optical data acquisition in tropical climates, and UAVs, though offering high-resolution imagery, face operational, regulatory, and logistical limitations when applied over large areas. Additionally, specific applications, such as productivity estimation and early disease detection, remain difficult. Spectral similarities between healthy and diseased plants, overlapping canopies, and subtle initial symptoms often lead to low detection rates with conventional ML approaches. Furthermore, error propagation in hierarchical classification systems hinder model generalization, particularly in heterogeneous agricultural landscapes.

Deep Learning (DL) is a subfield of machine learning based on artificial neural networks with multiple layers, designed to automatically learn hierarchical representations from data. Unlike traditional approaches that rely on handcrafted features, DL models perform end-to-end learning, where both low-level and high-level features are jointly optimized during training. This hierarchical representation learning enables the extraction of increasingly abstract patterns, ranging from simple edges and textures to complex spatial structures. Such capability has made DL particularly effective in tasks involving high-dimensional data, including image analysis, natural language processing, and, more recently, remote sensing applications (LeCun et al., 2015; Goodfellow et al., 2016).

In the context of remote sensing, the rise of DL has represented a paradigm shift, enabling the transition from pixel-based classification approaches to models capable of learning complex spatial feature hierarchies. Unlike traditional machine learning algorithms, which often struggle to distinguish crops with spectral signatures like native vegetation, Convolutional Neural Networks (CNNs) leverage convolutional filters to capture contextual, textural, and geometric information. these architectures, U-Net stands out due to its encoder–decoder structure and skip connections, which allow precise recovery of spatial details and object boundaries, making it particularly suitable for semantic segmentation tasks in heterogeneous landscapes (Ronneberger et al., 2015; Ma et al., 2019).

U-Net shows strong promise for complex heterogeneous visual patterns on diverse environments but remains underexplored for banana mapping. High-resolution imagery captures the fine-scale texture and structure essential for identifying perennial crops in such settings, and pan sharpening further improves spatial detail when a finer panchromatic band is available. In the CBERS-4A (China-Brazil Earth-Resources Satellite) satellite, the 2 meters WPM (Wide Swath Panchromatic and Multispectral) band enables detailed pansharpened products, when combined to multispectral data from WPM (with 8 meters spatial resolution) providing free national coverage with at least one cloud-free scene per year in many regions and offering a practical source for operational perennial-crop mapping.

This study evaluates whether a U-Net CNN applied to high-resolution CBERS-4A WPM imagery can accurately detect banana crops across heterogeneous municipalities in the Ribeira Valley. To our knowledge, pansharpened CBERS-4A WPM imagery has not yet been used for this task.

2. Study Area

The study was conducted in the Ribeira Valley, southwestern São Paulo State, Brazil, focusing on the municipalities of Registro and Cajati (Figure 1). The region is characterized by a humid subtropical climate (Köppen classification Cfa), with high annual rainfall ranging from 1500 to 2000 mm and relative humidity between 60% and 100%. The landscape is highly heterogeneous, comprising wetlands, river floodplains, hillslopes, and gently undulating terrain. Smallholder farming predominates, with banana (*Musa* spp.) plantations distributed across both lowland humid areas and higher-elevation slopes. Remnants of Atlantic Forest are still present throughout the valley, and agroforestry systems integrating banana with perennial crops such as peach palm (*Bactris gasipaes*) are common. This environmental and physiographic diversity makes the Ribeira Valley an ideal setting for evaluating agricultural mapping methods under complex landscape conditions.

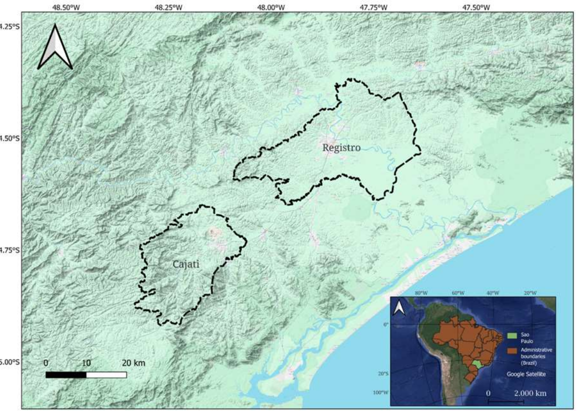

Figure 1. Geographic location of the study area in the Ribeira Valley, southwestern São Paulo State, Brazil, indicating the municipalities of Registro and Cajati. The region is situated within a highly heterogeneous tropical landscape composed of wetlands, floodplains, hillslopes, and agricultural mosaics dominated by banana cultivation.

3. Methodology

3.1 Reference Data and Sample Generation

Reference samples were generated through manual interpretation of high-resolution satellite imagery. Label delineation was primarily based on visual inspection of Google Earth imagery (Google LLC), supported by the CBERS-4A scene to ensure

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-907-2026

spectral and temporal consistency. The CBERS-4A WPM imagery[1] (path/row 204/144, L4 level) was acquired on 9 December 2022 and includes visible and near-infrared bands pansharpened to 2 meters spatial resolution using the Brovey Transform. Acquisition dates between datasets were carefully cross-checked to minimize inconsistencies due to phenological variation, seasonal effects, or land-use changes. Field knowledge and selective field observations were additionally used to validate interpretation patterns and improve labeling reliability.

A set of representative image tiles (approximately 1.2 × 1.2 km) was selected across the study area to capture spatial heterogeneity. Tile selection was designed to ensure diversity in land-cover conditions, including areas with dense banana crop coverage as well as regions containing only background (non-banana) classes. Within each tile, exhaustive binary annotation was performed, assigning each pixel to either banana (foreground) or non-banana (background). Visual interpretation was guided by characteristic spatial patterns observable in high-resolution imagery, including canopy texture, planting row geometry, and spectral homogeneity (see Figure 7).

3.2 Patch Extraction and Preprocessing

The final labelled tiles were subdivided into training samples using a sliding-window approach with a patch size of 64 × 64 pixels and an overlap of 16 pixels between adjacent windows. The overlap improves the representation of boundary and transitional regions. Each extracted patch contains the digital number representation of four spectral bands (blue, green, red and infrared) and was independently normalized using min–max scaling to the range [0, 1], defined as:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where x represents pixel values within a patch, and x_min and x_max denote the minimum and maximum values within that patch. For constant-value patches, normalization was skipped to avoid division by zero. This per-patch normalization reduces radiometric variability between scenes and improves training stability for the U-Net model.

3.3 Dataset Composition and Split Strategy

The final dataset consisted of 2,467 image patches derived from CBERS-4A WPM imagery. Each patch has a spatial dimension of 64 × 64 pixels with four spectral bands. The dataset exhibited a moderate class imbalance, with an average banana (positive class) proportion of approximately 0.33 across all subsets.

A two-stage split strategy was applied to ensure robust evaluation, resulting in a final split of 60% training, 20% validation, and 20% testing. All model tuning and early stopping were performed exclusively using the validation set, while the test set remained completely unseen until final evaluation. All subsets preserved a consistent distribution of positive pixels, with banana prevalence of approximately 33.2 (training), 30.9 (validation), and 32 (test), indicating balanced sampling across splits. In terms of spatial coverage, the training, validation, and test sets correspond to approximately 6M, 2M, and 2M pixels, respectively.

[1] National Institute for Space Research (INPE), *Satellite Image Catalog*, https://www.dgi.inpe.br/catalogo/explore (accessed September 30, 2025).

3.4 Model Architecture and Training

A custom U-Net convolutional neural network was implemented using Keras API and TensorFlow for binary semantic segmentation of banana crops. The architecture follows a symmetric encoder-decoder design with skip connections that link feature maps between corresponding spatial resolutions, enabling the preservation of fine-grained spatial information that is often lost during downsampling. Figure 2 presents the U-Net architecture adapted for banana crops segmentation using CBERS-4A multispectral imagery (64×64×4 input patches).

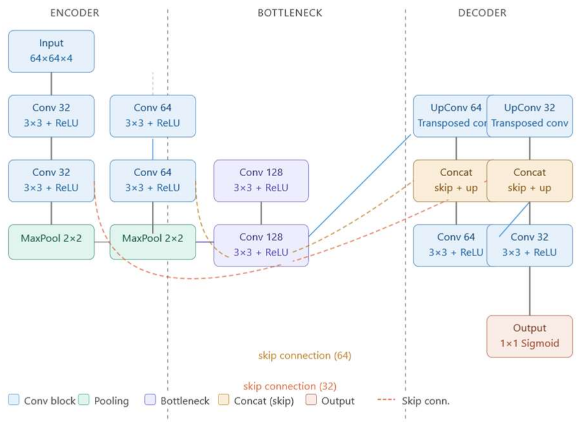

Figure 2. U-Net architecture adapted for banana crops segmentation using CBERS-4A multispectral high spatial resolution imagery (64×64×4 input patches).

The encoder path is composed of two convolutional stages. The first block applies two consecutive 3 × 3 convolutional layers with 32 filters and ReLU activation, followed by a max-pooling operation that reduces the spatial resolution by a factor of two. The second block repeats this structure with 64 filters, allowing the network to progressively learn more complex and abstract feature representations. A deeper bottleneck layer is introduced at the lowest spatial resolution, consisting of two 3 × 3 convolutional layers with 128 filters and ReLU activation, which encodes high-level contextual information about the input scene. The decoder path symmetrically restores spatial resolution using transposed convolution (Conv2DTranspose) layers for upsampling. At each decoding stage, feature maps from the corresponding encoder layer are concatenated via skip connections, ensuring the recovery of spatial detail and boundary information. Each decoder block contains two 3 × 3 convolutional layers with ReLU activation applied after concatenation, refining the fused feature representations.

The final output layer consists of a 1 × 1 convolution with sigmoid activation, producing a pixel-wise probability map for the binary classification of banana versus non-banana areas. The model was trained using the Adam optimizer with default hyperparameters and a binary cross-entropy loss function. Training phase was conducted for up to 50 epochs with a batch size of 8. To improve generalization and reduce overfitting, early stopping was applied based on validation Intersection over Union (IoU), with a patience of 10 epochs. Model checkpointing was used to retain the best-performing model according to validation IoU, which was selected for final evaluation on the independent test set. The code development was developed in the Brazil Data

Cube for providing the Geospatial Cloud Computing Data Science Python environment (Python Software Foundation, 2024; Queiroz et al., 2024; Ferreira et al., 2020) and the pipeline optimization was assisted by Claude Sonnet 4.6 (Anthropic, 2026).

3.5 Evaluation Metrics

Model performance was evaluated using metrics of accuracy, precision, recall, F1-score, Intersection over Union (IoU), and the area under the ROC and Precision–Recall curves. The IoU metric measures the overlap between predicted and ground truth regions and is defined as:

$$IoU = \frac{TP}{TP+FP+FN} \tag{2}$$

where TP (true positives) represents correctly predicted banana pixels, FP (false positives) represents incorrectly predicted banana pixels, and FN (false negatives) represents missed banana pixels.

In this study, a probability threshold of 0.5 was used to convert predicted probability maps into binary classifications. Let $y_i \in \{0,1\}$ denote the ground truth label and $\hat{y}_i \in [0,1]$ the predicted probability for pixel i. The binarized prediction is defined as:

$$\tilde{y}_i = \begin{cases} 1 & \text{if } \hat{y}_i > 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The IoU over all N pixels is then expressed as:

$$IoU = \frac{\sum_{i=1}^{N} y_i \tilde{y}_i}{\sum_{i=1}^{N} (y_i + \tilde{y}_i - y_i \tilde{y}_i)} \tag{4}$$

where $N$is the total number of pixels. This formulation is consistent with the implementation used during evaluation, with safeguards against division by zero.

3.6 Inference, Reconstruction, and Area-Based Validation

During inference, predictions were generated on image patches and reconstructed into the original spatial layout. Overlapping patches (50% overlap) were used to reduce edge artifacts, and predictions in overlapping regions were merged using pixel-wise averaging to produce a seamless probability mosaic. The final output was georeferenced by mapping predictions back to the original image grid.

To further assess the reliability of the mapped results at the administrative scale, an area-based validation was performed. The final probability maps were thresholded at 0.5 to generate binary classification maps, which were then vectorized to compute the total area of predicted banana crops per municipality.

These estimates were compared against official agricultural statistics from the SIDRA/IBGE database[2] for the corresponding reference year. This external comparison provides a consistency check beyond pixel-level accuracy metrics, allowing evaluation of whether the model produces reliable crop extent estimates at the municipal scale, which is particularly relevant for operational agricultural monitoring applications.

4. Results and Discussion

4.1 Training Behaviour and Model Convergence

The U-Net model exhibited stable convergence during training, with rapid improvements in the initial epochs followed by gradual stabilization. Figure 3 shows the training and validation learning curves over epochs. The loss curves decrease consistently without divergence, while the IoU metrics improve progressively, confirming stable convergence and increasing segmentation performance throughout training.

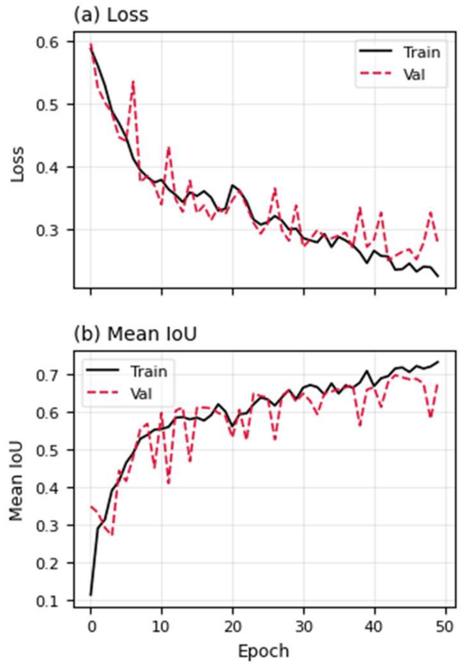

Figure 3. Training and validation learning curves of the U-Net model over epochs, showing (left) binary cross-entropy loss and (right) mean Intersection over Union (IoU).

The Intersection over Union (IoU) increased from 0.35 in early epochs to a maximum of 0.6975 at epoch 44, which was selected as the final model using early stopping and checkpointing. Training loss decreased consistently from approximately 0.59 to below 0.23, while IoU steadily increased, reaching values above 0.70 in later epochs. The relatively close behaviour between training and validation metrics indicates limited overfitting and effective regularization through dropout and validation-based model selection.

4.2 Quantitative Performance on Validation and Test Sets

The final model was evaluated using both validation and independent held-out test datasets at a fixed probability threshold of 0.5. On the validation set, the model achieved an accuracy of 0.9064, precision of 0.8735, recall of 0.8151, F1-score of 0.8433, IoU of 0.7290, AUC-ROC of 0.9592, and AUC-PR of 0.9180.

[2] IBGE’s Automatic Data Retrieval System (SIDRA), Table 1613. Available at: https://sidra.ibge.gov.br/tabela/1613

On the independent test set, all metrics were maintained or improved, with accuracy of 0.9133, precision of 0.8717, recall of 0.8551, F1-score of 0.8633, IoU of 0.7595, AUC-ROC of 0.9675, and AUC-PR of 0.9339.

Figure 4 presents the Precision–Recall and Receiver Operating Characteristic (ROC) curves computed on the test dataset, where the optimal decision threshold was selected by maximizing the F1-score and the ROC curve illustrates the trade-off between sensitivity and false positive rate.

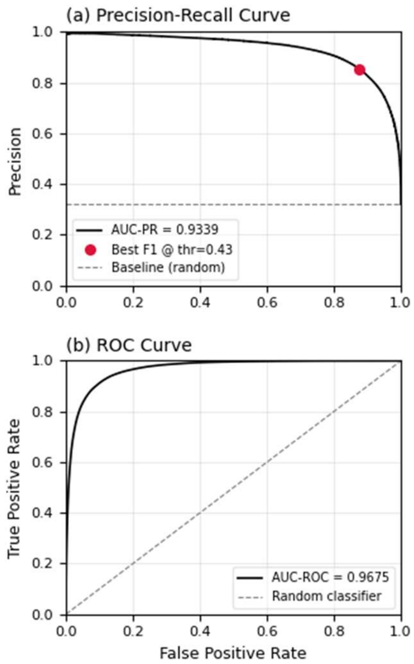

Figure 4. Model performance evaluation using (left) Precision–Recall curve and (right) Receiver Operating Characteristic (ROC) curve computed on the test dataset.

The close agreement between validation and test metrics, alongside high AUC values on both sets, indicates strong generalization ability without significant overfitting.

4.3 Pixel-Level Confusion Analysis

Figure 5 presents the pixel-wise confusion matrix for banana and non-banana classes on the independent test dataset.

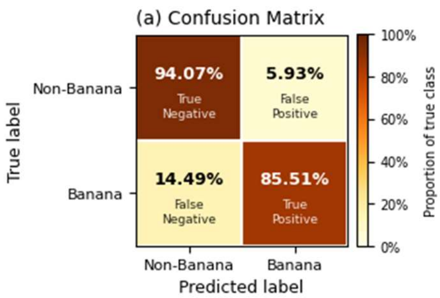

Figure 5. Pixel-wise confusion matrix for banana and non-banana classes on the independent test dataset (threshold = 0.5), summarizing true positives, true negatives, false positives, and false negatives.

Pixel-wise evaluation on the test set revealed 1,293,774 true negatives (63.94%) and 554,160 true positives (27.39%), with 81,589 false positives (4.03%) and 93,901 false negatives (4.64%). The relatively balanced distribution of errors indicates no strong bias toward over- or under-prediction. The slightly higher omission error (false negatives) suggests a mild tendency to miss some banana pixels, which is expected in heterogeneous agricultural environments where spectral similarity exists between banana crops and surrounding vegetation.

4.4 Spatial Error Patterns

Figure 6 highlights the spatial distribution of banana crops in the study area. Misclassifications were primarily concentrated in forested areas and along crops boundaries.

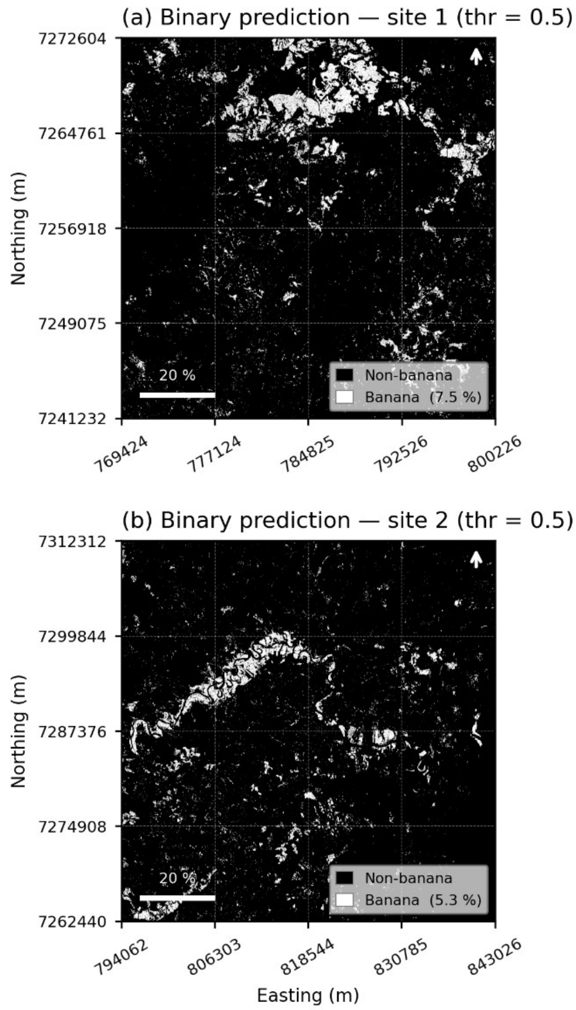

Figure 6. Spatial distribution of banana crops in the study area: (a) reference labels, (b) U-Net predicted map, and (c) false-colour CBERS-4A composite for visual comparison. The threshold (thr = 0.5) is related to the Equation 3. The maps are presented in the UTM projection, zone 22S.

Figure 7 shows the visual assessment of banana crop mapping results across four selected insets.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-907-2026 | © Author(s) 2026. CC BY 4.0 License.

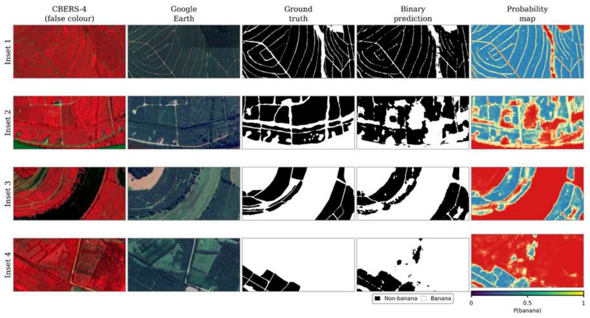

Figure 7. Visual assessment of banana crop mapping results across four selected insets. Each row shows a different data source or model output: CBERS-4 false-colour composite, Google Earth imagery, ground truth reference, binary prediction map, and probability map. Insets were selected to represent contrasting landscape contexts.

These errors are associated with spectral similarity between dense natural vegetation and banana canopies, as well as mixed pixels in transitional zones. Error distribution also varied across property types, ranging from smallholder farms to larger commercial crops, reflecting differences in spatial structure, management intensity, and canopy organization.

4.5 Area-Based Validation and External Consistency

Figure 8 shows the comparison between banana crops area estimated by the proposed U-Net model and official agricultural statistics (IBGE SIDRA) for the municipalities of Registro and Cajati, Sao Paulo state.

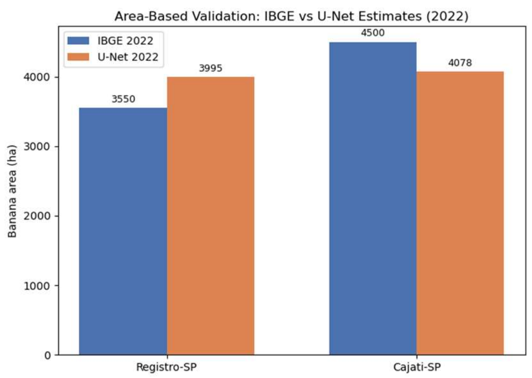

Figure 8. Comparison between banana crops area estimated by the proposed U-Net model and official agricultural statistics (IBGE SIDRA) for the selected municipalities.

The comparison showed good agreement between mapped and reported banana crop areas at the municipal level. This confirms that the model provides not only accurate pixel-level classification but also reliable aggregated estimates of crop extent at administrative scales. The results indicate strong agreement at the administrative scale, supporting the external consistency of the mapping approach.

4.6 Discussion and Implications

The proposed U-Net model demonstrated strong performance for banana crops mapping using CBERS-4A WPM imagery, achieving a test IoU of 0.7595 and F1-score of 0.8633. These results indicate that convolutional neural networks can effectively capture spatial and spectral patterns of banana cultivation even under moderate class imbalance and highly heterogeneous tropical landscapes.

A key contribution of this study is the demonstration of the potential of pansharpened CBERS-4A WPM imagery (2 m spatial resolution) to resolve fine-scale agricultural structures that are typically difficult to detect using medium-resolution satellite data. This capability is particularly relevant in regions such as the Ribeira Valley, where banana production is characterized by a mosaic of smallholder plots and larger commercial crops embedded within complex land-cover transitions.

Unlike widely used national mapping initiatives such as MapBiomas (Souza et al., 2020), which rely primarily on Landsat and Sentinel imagery with coarser spatial resolution, the proposed approach reduces misclassification errors associated with spectral mixing between banana crops, pasture, and native vegetation. Coarser-resolution datasets often overestimate agricultural extent in fragmented landscapes due to pixel heterogeneity and limited ability to capture narrow or irregular crop geometries. In contrast, the use of high-resolution CBERS-4A data enables more precise delineation of crop boundaries and improves discrimination between perennial crops and surrounding forest or secondary vegetation.

This result highlights the strategic importance of CBERS-4A as a free, nationally available Earth observation resource, reinforcing its potential for operational agricultural monitoring at national scale. Beyond crop mapping, the approach has direct implications for policy and management applications, including support for geographic indication systems for fruit production, land-use certification, and environmental compliance monitoring. In particular, accurate mapping of banana crops can contribute to agricultural health surveillance systems, such as the monitoring of Fusarium wilt Tropical Race 4 (TR4), a highly destructive disease that can persist in abandoned crops and spread through contaminated soil and planting material (Ploetz et al., 2015). The ability to delineate smallholder areas with higher spatial precision is therefore critical for supporting sanitary surveillance agencies and guiding targeted intervention strategies.

Despite these advantages, limitations remain, particularly related to the patch-based training strategy, which does not fully guarantee spatial independence between samples. No spatial block cross-validation was applied, which may lead to optimistic performance estimates and should be addressed in future studies. Additional improvements could be achieved by integrating multi-temporal and multi-sensor data, such as Sentinel-2 and SAR imagery, to enhance robustness across phenological stages and reduce residual spectral ambiguity between vegetation types.

Although the model performed well overall, residual confusion persists in specific landscape contexts, particularly in transition zones between agricultural fields and forest formations, along crop field edges, and in wetland environments (Figure 7). These areas are characterized by mixed pixels, gradual spectral transitions, and high structural complexity, which increase ambiguity between banana crops and surrounding vegetation types.

These findings corroborate the results of Soares et al. (2025), who demonstrated that even with high-frequency Sentinel-2 data, the spectral similarity between banana and peach palm in the Atlantic Forest's diversified landscapes remains a significant source of misclassification. Addressing these limitations will require a more detailed understanding of local landscape structure, as well as more comprehensive field validation campaigns to better characterize boundary conditions and refine labeling consistency in ecotonal areas.

Another important limitation relates to temporal generalization. The current study is based on a single-date CBERS-4A WPM image and therefore does not evaluate the model’s ability to generalize across different seasons or years. Given the strong phenological variability of perennial crops and surrounding vegetation, temporal transferability remains an open question. In addition, because the method relies on pansharpened high-resolution imagery, its robustness across different CBERS scenes with varying illumination, atmospheric conditions, and sensor acquisition characteristics is still uncertain. Further research is needed to assess whether the model can maintain stable performance when applied to different regions, acquisition dates, or sensor configurations without retraining.

Nevertheless, the results clearly indicate strong potential for broader application. Transferring the methodology to other major banana-producing countries, such as Ghana and Democratic Republic of the Congo in Sub-Saharan Africa and Ecuador in South America, could be highly valuable. These regions represent globally significant production areas but also exhibit strong landscape heterogeneity and similar challenges related to smallholder fragmentation and spectral overlap with natural vegetation (Alabi et al., 2022; Selvaraj et al., 2019). However, successful transfer would require careful recalibration, incorporation of local reference data, and likely adaptation to regional environmental conditions.

While the proposed approach demonstrates encouraging performance and operational relevance, it should be viewed as a foundational step toward scalable high-resolution agricultural monitoring systems, rather than a fully generalized solution. Continued research combining multi-temporal analysis, multi-sensor integration, and expanded field validation will be essential to fully unlock its potential at national and international scales.

5. Final Considerations

This study demonstrates that U-Net applied to high spatial resolution CBERS-4A WPM imagery can effectively map banana crops in the heterogeneous landscapes of the Ribeira Valley, Brazil. The proposed workflow is computationally efficient, reproducible, and capable of generating seamless regional-scale maps with high spatial consistency. The robust performance observed across diverse landscape conditions highlights the method’s potential for scaling to larger geographic extents.

Future developments should prioritize the integration of multi-temporal data to capture the complex phenological dynamics of banana crops, utilizing deep learning architectures, such as 3D-CNNs or Vision Transformers (ViTs), specifically designed for time-series analysis. From an operational perspective, these high-frequency monitoring frameworks are critical for sanitary agencies to implement proactive biosafety policies. Furthermore, robust mapping supports the legal and economic framework of Geographical Indications. By establishing precise, verifiable spatial records of production boundaries, remote sensing safeguards the authenticity of regional products and enhances the added value of local brands. Finally, these tools can potentially be used to promote environmental governance by assisting in the alignment of agricultural growth with land-use regulations, such as the management of riparian buffers and conservation units. This ensures that the banana sector can demonstrate compliance with international sustainability standards, fostering a balance between productive expansion and natural resource preservation.

Acknowledgements

Acknowledgements of support of Brazil Data Cube for providing the Geospatial Cloud Computing Data Science environment. The authors also thank the São Paulo Research Foundation (FAPESP, grant no 2023/09118-6), and the Brazilian National Council for Scientific and Technological Development (CNPq, grant no 302205/2023-3).

Declaration of Generative AI Use

During the preparation of this manuscript, the authors used OpenAI’s ChatGPT and Google’s Gemini to refine the linguistic structure and improve the clarity of the original text. These tools were used to sharpen the expression of the authors' ideas and ensure grammatical precision in English. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the final version of the manuscript.

References

Aksu, D., Demir, N., 2021: Extraction of banana fields from satellite images. *Proceedings of the International Symposium on Applied Geoinformatics (ISAG2021)*, Riga, Latvia, Dec. 2–3.

Alabi, T.R., Adewopo, J., Duke, O.P., Kumar, P.L., 2022: Banana mapping in heterogenous smallholder farming systems using high-resolution remote sensing imagery and machine learning models with implications for banana bunchy top disease surveillance. *Remote Sensing*, 14(20), 5206. doi.org/10.3390/rs14205206.

Anthropic, 2026. Claude AI Software, Version 3.5. Anthropic PBC. claude.ai (10 Jan 2026).

Beaulieu, N., Leclerc, G., Imbach, A.C., et al., 1999: The contribution of RADARSAT-1 SAR imagery to monitor land use in coastal areas of Costa Rica and Nicaragua. *Proc. 20th Canadian Symp. on Remote Sens.*, Ottawa, ON.

Ferreira, K.R., Queiroz, G.R., Vinhas, L., et al., 2020: Earth Observation Data Cubes for Brazil: Requirements, Methodology and Products. *Remote Sensing*, 12(24), 4033. doi.org/10.3390/rs12244033.

Fundacao SOS Mata Atlantica, INPE, 2023: Atlas dos Remanescentes Florestais da Mata Atlântica—Período 2021–2022. Technical Report, Fundação SOS Mata Atlântica, São Paulo, Brazil, pp. 1–61. https://cms.sosma.org.br/wp-content/uploads/2023/05/SOSMAAtlas-da-Mata-Atlantica_2021-2022-1.pdf (20 August 2025).

Goodfellow, I., Bengio, Y., Courville, A., 2016: Deep Learning. *MIT Press*.

Guzmán-Álvarez, J.A., et al., 2022: Uso de sensores remotos en la agricultura: aplicaciones en el cultivo del banano. *Agronomía Mesoamericana*, 33(3). doi.org/10.1515/am.v33i3.48279.

Johansen, K., Phinn, S., Taylor, M., 2009: Mapping banana plantations from object-oriented classification of SPOT-5 imagery. *Photogrammetric Engineering & Remote Sensing*, 75(9), 1069–1081. doi.org/10.14358/PERS.75.9.1069.

Johansen, K., Sohlbach, M., Sullivan, B., Stringer, S., Peasley, D., Phinn, S., 2014: Mapping banana plants from high spatial resolution orthophotos to facilitate plant health assessment. Remote Sens., 6(9), 8261-8286. doi.org/10.3390/rs6098261.

LeCun, Y., Bengio, Y., Hinton, G., 2015: Deep learning. *Nature*, 521, 436–444. doi.org/10.1038/nature14539.

Leclerc, G., Hall, C.A.S., 1999: Remote Sensing and Land Use Analysis for Agriculture in Costa Rica. In: Hall, C.A.S. (Ed.), *Quantifying Sustainable Development: The Future of Tropical Economies*. Academic Press, San Diego, pp. 247–265.

Ma, L., Liu, Y., Zhang, X., et al., 2019: Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.*, 152, 166–177. doi.org/10.1016/j.isprsjprs.2019.04.015.

Neupane, Bipul, Horanont, T., Hung, N.D., 2019: Deep learning-based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV). PloS one, 14(10), e0223906. doi.org/10.1371/journal.pone.0223906.

Pedroni, L., 2003: Improved classification of Landsat Thematic Mapper data using modified prior probabilities in large and complex landscapes. International Journal of Remote Sensing, 24(1), 91–113. doi.org/10.1080/01431160304998.

Ploetz, R.C., 2015: Management of Fusarium wilt of banana: A review with special reference to tropical race 4. *Crop Protection*, 73, 7–15. doi.org/10.1016/j.cropro.2015.03.001.

Python Software Foundation, 2024: Python Language Reference, Version 3.x. *Python.org*.

Queiroz, G.R., Ferreira, K.R., 2024: BDC-Lab: Ambiente de ciência de dados geoespaciais do Brasil Data Cube. *Anais IV SIBIDE*, 70–72.

Ribeiro, M.C., Metzger, J.P., Martensen, A.C., et al., 2009: The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? *Biol. Conserv.*, 142(6), 1141–1153. doi.org/10.1016/j.biocon.2009.02.021.

Ronneberger, O., Fischer, P., Brox, T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. *Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 234–241. doi.org/10.1007/978-3-319-24574-4_28.

Selvaraj, M.G., Vergara, A., Montenegro, F., Alonso, J.R., Ladbury, P., Ebert, A.W., Brown, N., 2019: AI-powered banana diseases and pest detection. Plant Methods, 15(1), 92. doi.org/10.1186/s13007-019-0475-z.

Soares, V.B., Parreiras, T.C., Furuya, D.E.G., Bolfe, É.L., Nechet, K.D.L., 2025: Mapping Banana and Peach Palm in Diversified Landscapes in the Brazilian Atlantic Forest with Sentinel-2. Agriculture, 15(19), pp. 2052. doi.org/10.3390/agriculture15192052.

Souza Jr, C.M., Shimbo, J.Z., Rosa, M.R., et al., 2020: Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth Engine. *Remote Sensing*, 12(17), 2735. doi.org/10.3390/rs12172735.

Varma, V., Bebber, D.P., 2019: Climate change impacts on banana yields around the world. *Nature Climate Change*, 9(10), 752–757. doi.org/10.1038/s41558-019-0559-9.

Verhoeye, J., De Wulf, R., 1999: An image processing chain for land-cover classification using multitemporal ERS-1 data. *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Hamburg, Germany, vol. 2, pp. 1179–1181.

Ye, H., Huang, W., Huang, S., Cui, B., 2020: Identification of Banana Fusarium Wilt Using UAV-based Multi-spectral Imagery and Support Vector Machine. *Remote Sensing*, 12(1), 141. doi.org/10.3390/rs12010141.