Assessing the Performance of Deep Learning Networks for Real-Time Deforestation Segmentation in the Amazon and Atlantic Biomes

Giovana A. Benvenuto[1], Wallace Casaca[1], Rogério Negri[2], Marilaine Colnago[1]

[1]Instituto de Biociências, Letras e Ciências Exatas
Universidade Estadual Paulista (IBILCE/UNESP)
São José do Rio Preto – SP – Brasil

[2]Instituto de Ciência e Tecnologia
Universidade Estadual Paulista (ICT/UNESP)
São José dos Campos – SP – Brasil

giovana.a.benvenuto@unesp.br, wallace.casaca@unesp.br

rogerio.negri@unesp.br, marilaine.colnago@unesp.br

Abstract. *Deforestation poses significant ecological and societal challenges, while advances in satellite imagery and deep learning have enhanced monitoring precision and scalability. This study evaluates ten deep learning models for deforestation segmentation, including U-Net, ResNet, FCN, and YOLO variants, assessing accuracy and computational efficiency, focusing on the Amazon and Atlantic Forest biomes. The results highlight that U-Net and ResNet50 achieve the highest accuracy, while YOLOv8 and YOLOv11 offer an optimal balance between speed and performance. The findings contribute to the model selection for real-time deforestation detection, supporting conservation and environmental decision-making in underexplored areas like the Atlantic Forest.*

1. Introduction

Deforestation and its environmental and societal consequences is a long-standing issue, yet its accelerated growth has reached alarming levels. In the Brazilian Amazon Biome, deforestation increased by 56.6% between 2019 and 2021 compared to the period 2016 to 2018 [Cabral et al. 2024]. Similarly, the Atlantic Forest, one of the most threatened biodiversity hotspots in the world, has already lost a significant portion of its original forest cover [Maurenza et al. 2024]. Beyond biodiversity loss, deforestation impacts air quality and disrupts rainfall cycles, not only in affected regions, but also in distant areas, with far-reaching consequences across various sectors of society [Rodrigues et al. 2024].

Deforestation requires that authorities implement effective measures to prevent and mitigate the causes of forest loss, which are often linked to human activities. Remote Sensing (RS) has been widely utilized as an effective large-scale tool for facilitating the comprehensive monitoring of forested regions. When integrated with AI-driven solutions, such as Deep Learning (DL) networks, RS technology has significant potential for the automatic detection of deforestation in remote regions, enabling continuous monitoring and the generation of alerts to support timely and informed decision-making.

Building on this potential, in this study, we conduct a comprehensive evaluation of deep learning architectures for deforestation detection, systematically comparing ten state-of-the-art and emerging models. Our analysis covers both well-established DL segmentation networks, such as U-Net [Ronneberger et al. 2015], FCN [Long et al. 2015] and ResNet [He et al. 2016], and modern detection architectures, including YOLOv8 [Jocher et al. 2023], YOLOv9 [Wang and Liao 2024], and YOLOv11 [Jocher and Qiu 2024]. Unlike previous works that often rely on models trained for general-purpose land cover classification, our study specifically focuses on deforestation segmentation, optimizing models for this critical environmental challenge while enhancing their reliability and applicability in real-world monitoring. Furthermore, while various segmentation models have been applied to the Amazon and other biomes outside Brazil, deforestation in the Atlantic Forest has received limited attention in prior research. Indeed, by querying the Scopus database with the terms Atlantic Forest, deforestation, and segmentation, only five journal articles were retrieved, none of which specifically focused on this biome. Moreover, no previous study has conducted a comparative analysis of DL segmentation models for deforestation in this context.

Beyond evaluating the accuracy of the DL networks, we also assess computational efficiency, an essential factor for large-scale monitoring applications. By analyzing the trade-offs between assertiveness and processing cost, we provide new insights into the practical deployment of these models, particularly in scenarios with limited computational resources. Additionally, by bridging the gap between traditional and modern segmentation architectures, this work contributes to a more refined understanding of model selection for deforestation mapping, offering a benchmark for future studies and practitioners with a particular focus on two Brazilian biomes: the Atlantic Forest and the Amazon.

This paper is organized as follows. Section 2 reviews related work, while Section 3 describes the methodology, including the datasets, deep learning methods, and experimental setup. Section 4 presents and discusses the results obtained on each dataset. Finally, Section 5 provides the conclusions and outlines directions for future research.

2. Related Work

The advancement of RS technology has enabled the systematic application of DL techniques to address various complex issues, including the general detection of deforestation. In [Md Jelas et al. 2024], an overview of 22 studies in this domain was provided. Among the analyzed works, the authors highlight the use of remote sensing imagery from satellites as a significant advantage, and conclude that there is still substantial potential for further exploration in this field. In the general literature, various comparative studies have explored the performance of well-known models, as evidenced by the following papers.

Bem et al. [de Bem et al. 2020] assess three DL models, SharpMask, U-Net, and ResUnet, against Machine Learning (ML) approaches, Random Forest and Multilayer Perceptron, to detect and segment changes in the Amazon biome. The results indicate that the DL models outperformed their ML counterparts, with ResUnet being the most accurate within the studied domain. Adarme et al. [Ortega Adarme et al. 2020] also compare DL-based methods to identify regions affected by deforestation in the Amazon and Cerrado biomes. In their study, four methods were investigated: Early Fusion (EF), Siamese Network (SN), Convolutional Support Vector Machine (CSVM), and the classic SVM,

used as a baseline. The study area was divided into four tiles for training and testing the models. In the best-performing tile, the method that achieved the highest F1-score for the Amazon biome was EF, with an average overlapping area of 63.3%.

Torres et al. [Torres et al. 2021] analyze the U-Net, ResUnet, SegNet, FC-DenseNet, MobileNetV2 and Xception networks. Among these, ResUnet proved to be the most cost-effective, balancing computational efficiency and segmentation performance. It achieved F1-scores of 70.7% and 70.2% for the Landsat-8 and Sentinel-2 image datasets, respectively. Similarly, Careli et al. [Careli et al. 2024] compares the performance of two deep learning segmentation frameworks, YOLOv8 and the Mask R-CNN, by testing different setups for both. The evaluation was conducted on a dataset compiled from publicly available samples via web scraping, along with a second set of AI-generated images.

Kovačovič et al. [Kovačovič et al. 2025] examine the performance of three DL models for the detection and segmentation of various elements in satellite images, including forests, fields, roads, buildings, and lakes, with high precision. By constructing their own dataset, they evaluated the YOLOv5, YOLOv8, and Mask R-CNN networks, concluding that the small variant of YOLOv8 was the most effective. The YOLOv8 network was also evaluated in the study proposed by Silpalatha and Jayadeva [Silpalatha and Jayadeva 2025], who developed a model based on this architecture and compared its performance with several DL frameworks, such as DeepLabV3+, U-Net++, Attention U-Net, and R-CNN. Their evaluation was conducted using the WHDLD (World High-Resolution Land Development) dataset, which comprises high-resolution satellite images from different geographical regions.

The literature review highlights the effectiveness of deep networks in addressing the challenge of segmenting deforested areas, showcasing a wide range of models applied in different scenarios. However, the performance of newer YOLO network variants, such as YOLOv9 and YOLOv11, remains unexplored in the deforestation literature. In addition, the DL models are conveniently trained for classes other than deforestation, and we do not look specifically at this problem. Moreover, existing studies emphasize the need for deep learning networks specifically designed for deforestation segmentation, leveraging tailored datasets to address the challenge of limited sample availability. Therefore, this study compares well-established architectures, such as U-Net and ResNet, with newer DL-based detection models, including YOLO versions 8, 9, and 11.

3. Methodology

3.1. Datasets

In order to assess the DL approaches selected in this study, we utilize two publicly available datasets, provided by [Bragagnolo et al. 2021]. Both datasets consist of images captured by the Sentinel-2 Level 2A satellite, containing four spectral bands (RGB and NIR) at a resolution of 512×512 pixels. The images were acquired from the Amazon and Atlantic Forest biomes, as illustrated in Figure 1. The Amazon dataset comprises 499 training images, 100 validation images, and 20 test images, while the Atlantic Forest image collection includes 485 training images, 100 validation images, and 20 test images. The ground-truth data for each dataset was manually generated using the GRASS GIS 7.6.1 software suite, assigning binary values to the classes: class 0 (black) corresponds to forested regions, while class 1 (white) represents deforested areas.

To ensure consistent training conditions across all networks, only the RGB bands were used, as YOLO-type networks natively accept three-channel inputs. A preprocessing step was also performed to adapt the labels for YOLO-based models to the expected format, which represents each class using polygon coordinates instead of ground-truth images, resulting in saved polygons for two classes: Deforestation and Forest.

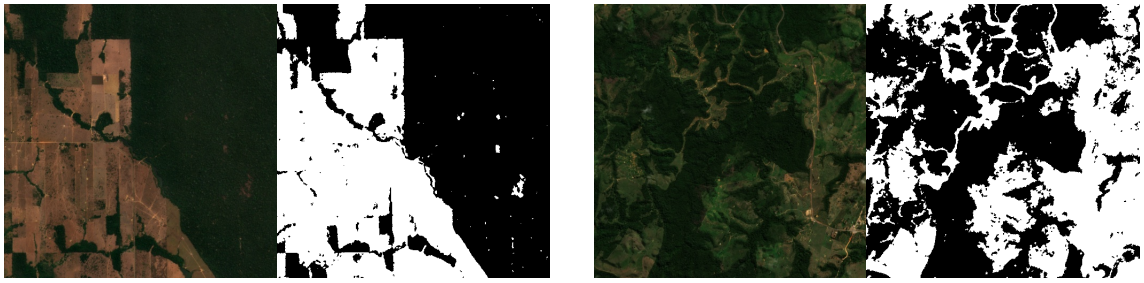

Figure 1. Sample images from the two datasets: (a) and (b) represent an image and its corresponding ground-truth from the Amazon dataset, while (c) and (d) illustrate the same for the Atlantic dataset.

3.2. Deep neural networks for automated deforestation segmentation

We implement, adjust and compare a total of ten DL architectures for the semantic segmentation of deforested regions in both the Amazon and Atlantic Forest biomes. Among the evaluated deep neural networks, three are well-established: U-Net, ResNet50, and FCN32, while the remaining seven are YOLO-based models with different versions and number of parameters. A detailed description of each DL network is provided below.

U-Net: The U-Net network is a semantic segmentation framework proposed by Ronneberger et al. [Ronneberger et al. 2015], initially designed for biomedical image analysis. The architecture consists of two main phases: an encoder and a decoder. The encoder progressively downsamples the input, reducing its spatial dimensions while extracting relevant features. The decoder then upsamples the feature maps, reconstructing the segmented image and forming the characteristic U-shaped structure of the model. U-Net is a widely recognized segmentation architecture that has been successfully applied in various domains, including medical imaging [Li et al. 2024], satellite imagery [Pathak et al. 2024], and other fields requiring precise object delineation.

ResNet50: ResNet50 is a deep residual network that incorporates shortcut connections to mitigate the vanishing gradient problem when training very deep architectures. Originally proposed by He et al. [He et al. 2016] for general image classification tasks, ResNet50 has also been widely adopted as a backbone in segmentation models due to its strong feature extraction capabilities.

FCN32: Fully Convolutional Networks (FCNs), introduced by Long et al. [Long et al. 2015], are segmentation frameworks that replace the fully connected layers of traditional neural networks with convolutional layers, enabling end-to-end pixel-wise predictions. This architectural modification enables the generation of dense segmentation maps instead of single-label image classifications. FCNs have been trained on general-purpose datasets containing multiple object classes, demonstrating their versatility in various segmentation tasks.

YOLOv8: The You Only Look Once (YOLO) networks are widely recognized for their real-time object detection capabilities. Since version 8, released by Ultralytics [Jocher et al. 2023], the developers have expanded the architecture's scope to address other computer vision tasks, including semantic segmentation. YOLOv8 incorporates techniques such as bottlenecks and specialized layers to extract multiscale features for precise object delineation. This network is available in five different sizes, each offering a trade-off between accuracy and computational efficiency. In this study, we evaluate the small, medium, and large variants of YOLOv8.

YOLOv9: The ninth version of YOLO, developed by [Wang and Liao 2024], builds upon the Ultralytics YOLOv5 framework and introduces two significant innovations: Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN). PGI ensures that complete input information is utilized in the computation of the objective function, leading to more reliable gradient updates for optimizing network weights. GELAN, on the other hand, is a network architecture designed to enhance parameter efficiency. YOLOv9 demonstrated superior performance on the COCO dataset; however, it requires greater computational resources and longer training times compared to its YOLOv8 counterparts.

YOLOv11: Also developed by Ultralytics [Jocher and Qiu 2024], YOLOv11 is optimized for high accuracy with fewer parameters than its predecessors. Available in five different sizes (number of parameters), YOLOv11, like YOLOv8, supports multiple tasks beyond object detection and semantic segmentation. In our comparative analysis, we evaluate three of its variants: small, medium, and large.

3.3. Experimental design and assessment criteria

All networks assessed in our experiments were implemented using the PyTorch library. An extensive hyperparameter tuning process was conducted to determine the optimal configuration for each network during training. This process was performed using the Optuna library, exploring parameters such as learning rate, dropout rate, number of filters, and batch size, according to each model's specifications. A total of 50 trials were executed, each comprising the initial training epochs, for both datasets. The YOLO models, however, already incorporate advanced optimization mechanisms such as warmup strategies and an Auto Learning Rate Finder in their standard training process, reducing the impact of tuning. At the end of the tuning process, all models were analyzed using the best set of parameters identified for the segmentation tasks. The networks were trained with the ADAM optimizer. The BCEWithLogitsLoss function was employed for U-Net, ResNet50, and FCN32, while the YOLO models utilized their standard loss function.

The models were trained in the Google Colaboratory environment utilizing an NVIDIA Tesla T4 GPU, which features 2560 CUDA cores, 320 Tensor Cores, 40 RT Cores, and 16 GB of GDDR6 memory. Each network was trained for 50 epochs, with a batch size of 16 images. This number of epochs was chosen based on the available computational resources, being sufficient for the models to converge while avoiding overfitting, thus balancing training time and overall model performance.

To quantitatively assess the performance of the trained models, four widely adopted similarity metrics were selected from the literature: F1-Score, Overall Accuracy, Precision, and Recall [de Bem et al. 2020, Ortega Adarme et al. 2020,

Torres et al. 2021]. These metrics, commonly used in the segmentation task, are defined in terms of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), as can be seen in Table 1. Precision quantifies the proportion of correctly predicted positive instances among all predicted positives, while Recall measures the proportion of correctly identified positive instances relative to all actual positives. The F1-Score represents the harmonic mean of Precision and Recall, providing a balanced measure of model performance. Finally, Overall Accuracy (OA) indicates the proportion of correctly classified samples relative to the total number of samples evaluated.

Table 1. Evaluation metrics and their respective formulas.

Precision	Recall	F1-score	Overall Accuracy
$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	$\frac{TP+TN}{TP+TN+FP+FN}$

To complement our comparative analysis, the computational complexity of each model was assessed, reporting the total number of parameters. In general, models with a higher number of parameters exhibit greater complexity, which can result in increased computational burden as well as slower inference and training times. Lastly, the average inference time on the test set for both datasets was also collected.

4. Results and Discussion

In this section, the results obtained from each model are presented and discussed, organized into quantitative and qualitative analyses, with separate evaluations for each analyzed dataset from the Brazilian biomes.

4.1. Amazon forest dataset

Table 2 summarizes the overall performance of each DL model on the Amazon test dataset, with best scores in bold. The second column reports the total number of parameters (in millions), with the small versions of YOLOv11 (10.10M) and YOLOv8 (11.80M) being the lightest models, while FCN32 (134.26M) is the heaviest. The third column displays the average inference time (in seconds) for each model. YOLOv8s (0.0142s) emerges as the fastest model, closely followed by YOLOv11s (0.0143s), which is expected given their optimization for real-time detection. The fourth column lists the mean and standard deviation of the F1-Score, where U-Net (0.9588 ± 0.0315) and ResNet50 (0.9497 ± 0.0347) achieve the highest scores, highlighting their robustness in accuracy.

Overall, the tabulated scores indicate that U-Net and ResNet50 achieve the highest accuracies, with U-Net leading across all metrics. However, this model is also the slowest, with a mean inference time of 0.0701s. In contrast, the smaller YOLO variants, particularly YOLOv11s, demonstrate competitive performance despite their lightweight architectures. YOLOv11s achieves an F1-Score of 0.9336 and an OA of 0.9491, while YOLOv8m maintains an F1-Score of 0.9319 and an OA of 0.9467. The YOLO-based models balance performance and efficiency, making them well-suited for real-time applications. One can also observe that YOLOv8s and YOLOv9c exhibit higher standard deviations, particularly in Recall, suggesting potential instability across different runs.

Table 2. Comparison of DL models in the Amazon forest dataset.

Models	Params (M)	Time (s)	F1-Score	OA	Precision	Recall
U-Net	31.03	0.0741	**0.9588 ± 0.0315**	**0.9696 ± 0.0188**	**0.9558 ± 0.0354**	**0.9639 ± 0.0500**
ResNet50	18.07	0.0334	*0.9497 ± 0.0347*	*0.9606 ± 0.0286*	*0.9470 ± 0.0367*	*0.9544 ± 0.0535*
FCN32	134.26	0.0528	0.8920 ± 0.0692	0.9217 ± 0.0319	0.8900 ± 0.0616	0.8971 ± 0.0907
YOLOv8s	*11.80*	**0.0142**	0.8781 ± 0.2118	0.8975 ± 0.2124	0.8782 ± 0.2179	0.8815 ± 0.2123
YOLOv8m	27.30	0.0312	0.9319 ± 0.0364	0.9467 ± 0.0259	0.9352 ± 0.0460	0.9299 ± 0.0407
YOLOv8l	46.00	0.0435	0.9356 ± 0.0341	0.9498 ± 0.0241	0.9369 ± 0.0460	0.9350 ± 0.0306
YOLOv9c	27.90	0.0397	0.8891 ± 0.1693	0.9185 ± 0.1114	0.9305 ± 0.0656	0.8830 ± 0.1863
YOLOv11s	**10.10**	*0.0143*	0.9336 ± 0.0350	0.9491 ± 0.0234	0.9416 ± 0.0418	0.9266 ± 0.0401
YOLOv11m	22.40	0.0341	0.9354 ± 0.0374	0.9513 ± 0.0233	*0.9399 ± 0.0432*	0.9320 ± 0.0437
YOLOv11l	27.60	0.0387	0.9109 ± 0.1101	0.9323 ± 0.0846	0.9356 ± 0.0424	0.9022 ± 0.1431

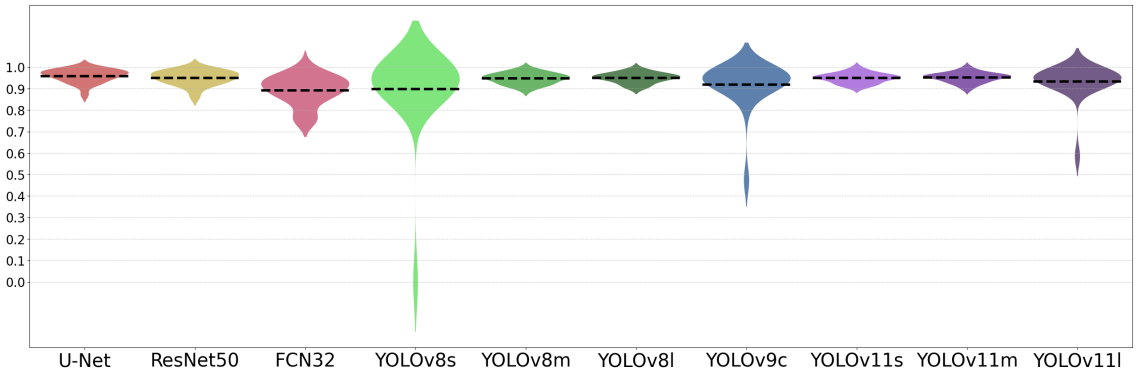

Figure 2. F1-Score distribution for DL models in the Amazon forest dataset.

For a clearer understanding, Figure 2 shows the distribution of prediction accuracy using the F1-Score metric, indicating that YOLOv8s, YOLOv9c, and YOLOv11l exhibit outliers in the test set. Upon closer examination of the affected sample, one can check that it corresponds to an image with significant cloud cover, while its ground-truth annotation classifies it as entirely deforested. This misclassification in the dataset directly impacts both the training and evaluation processes, leading to a reduction in the overall performance metrics. Despite this issue, the remaining YOLO models demonstrate a more consistent distribution of results, concentrating the results between 0.9 and 1.0.

Figure 3 shows the overlap between the ground-truth and the predictions generated by each deep neural model. In this comparison, ground-truth is marked in green, predicted areas in magenta, and overlapping regions appear in white. From the analyzed image-sample, one can inspect that all models detect large deforested areas, with edge refinement being the primary differentiating factor among them. U-Net and ResNet50 provide the best coverage of ground-truth areas, minimizing the presence of uncovered green pixels, however, the ResNet50 tends to produce more false positives, as indicated by the magenta edges around detected regions. FCN32 also struggles with precise boundary delineation. Meanwhile, the YOLO networks demonstrate a balanced trade-off between contour definition and accuracy, but face challenges in detecting smaller deforested areas.

By analyzing the general performance of the networks across all test images, Figure 4 presents the cumulative confusion matrix, illustrating the percentage distribution of pixels classified as TP, TN, FP, and FN. First, the distribution suggests that the Amazon dataset has well-balanced classes. Second, the U-Net and ResNet50 models demonstrated

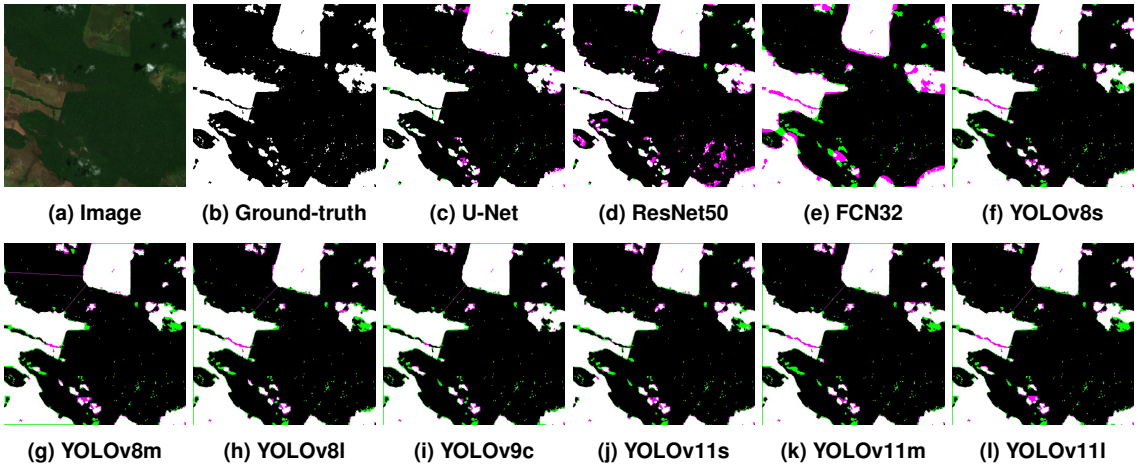

Figure 3. Qualitative comparison of a deforestation sample from Amazon dataset.

the most favorable distributions, whereas FCN32 exhibits the weakest performance due to a higher FP rate. Among the YOLO-derived models, all achieve high TP prediction rates, however, YOLOv8s, YOLOv9c, and YOLOv11l presented higher FN errors, likely influenced by the misclassification of outlier samples.

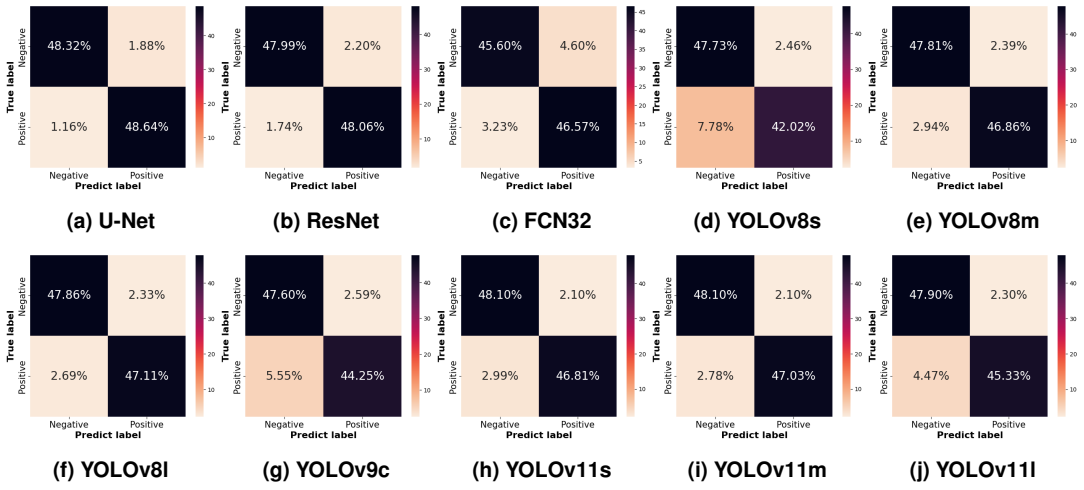

Figure 4. Cumulative confusion matrix for the Amazon forest dataset.

4.2. Atlantic forest dataset

Having examined the first dataset, our evaluation was extended to the second dataset using the same approach. Table 3 presents the evaluation metrics w.r.t. the Atlantic Forest dataset, emphasizing key performance differences among the DL models. In this case, the neural networks show lower accuracy than on the Amazon dataset, indicating that this dataset poses a greater challenge. While U-Net and ResNet50 continue to demonstrate superior accuracy, the trained models YOLOv8s, and YOLOv11s emerge as more suitable alternatives for applications requiring computational efficiency.

Concerning the distribution of scores for the test set, a similar pattern is observed, as illustrated in Figure 5. While the overall average is slightly lower compared to the

Table 3. Comparison of DL models in the Atlantic forest dataset.

Models	Params (M)	Time (s)	F1-Score	OA	Precision	Recall
U-Net	31.03	0.0679	**0.9524 ± 0.0243**	**0.9363 ± 0.0356**	**0.9544 ± 0.0273**	**0.9523 ± 0.0461**
ResNet50	72.26	0.1005	*0.9332 ± 0.0327*	*0.9105 ± 0.0477*	0.9358 ± 0.0365	*0.9346 ± 0.0650*
FCN32	134.26	0.0483	0.8704 ± 0.0551	0.8241 ± 0.0574	0.8190 ± 0.0642	0.9308 ± 0.0595
YOLOv8s	*11.80*	**0.0142**	0.9188 ± 0.0334	0.8700 ± 0.0534	0.9328 ± 0.0351	0.9060 ± 0.0399
YOLOv8m	27.30	0.0312	0.9082 ± 0.0635	0.8847 ± 0.0685	0.9192 ± 0.0937	0.9024 ± 0.0375
YOLOv8l	46.00	0.0440	0.9156 ± 0.0389	0.8969 ± 0.0300	0.9374 ± 0.0316	0.8955 ± 0.0505
YOLOv9c	27.90	0.0387	0.9145 ± 0.0413	0.8964 ± 0.0312	0.9361 ± 0.0301	0.8948 ± 0.0570
YOLOv11s	**10.10**	*0.0143*	0.8999 ± 0.0849	0.8860 ± 0.0539	0.9314 ± 0.0379	0.8764 ± 0.1135
YOLOv11m	22.40	0.0336	0.9194 ± 0.0341	0.9001 ± 0.0279	0.9360 ± 0.0305	0.9039 ± 0.0441
YOLOv11l	27.60	0.0383	0.9178 ± 0.0377	0.8991 ± 0.0315	*0.9392 ± 0.0299*	0.8983 ± 0.0512

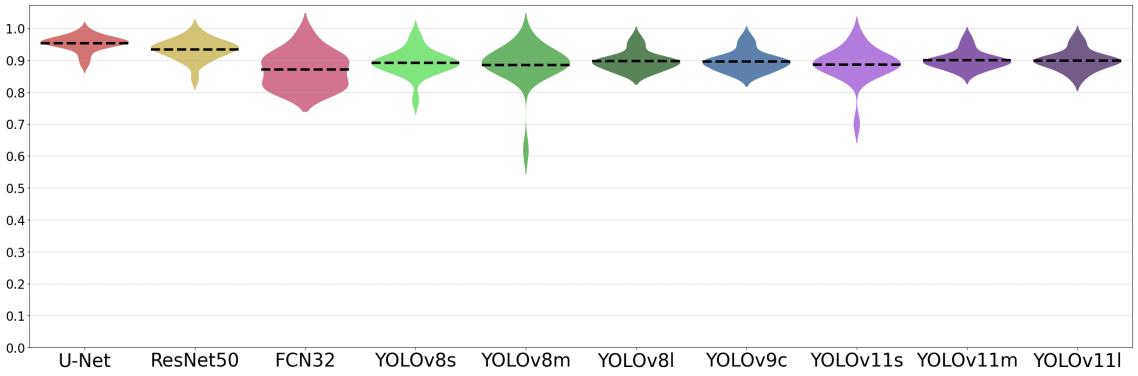

Figure 5. F1-Score distribution for DL models in the Atlantic forest dataset.

Amazon dataset, the key characteristics remain consistent. The YOLO networks achieve a more compact result distribution, although YOLOv8m and YOLOv11s present outliers.

The qualitative analysis of Figure 6, which depicts a large deforested portion of the Atlantic biome, reveals that the networks maintain similar behaviors across different datasets, despite the higher difficulty level in delineating the targets in this dataset. U-Net and ResNet accurately capture small regions of the ground-truth, but struggle to classify certain areas marked as deforested. This observation is further supported by the confusion matrices in Figure 7, which highlight an imbalance in the Atlantic dataset, as it contains a higher proportion of positive samples. This imbalance directly affects network performance, contributing to the differences observed between the two datasets analyzed. Further examining Figure 7, one may conclude that YOLO networks achieve higher TN rates, with YOLOv11l reaching 27.72%. This shows that these models are more conservative in classifying deforested areas, potentially reducing false positive detections.

5. Conclusion

This study implemented, systematically tuned, and compared several deep learning architectures for deforestation segmentation, focusing on the Amazon and Atlantic Forest biomes. These include both well-established segmentation networks, such as U-Net, ResNet50, and FCN32, as well as modern detection frameworks, including YOLOv8, YOLOv9c, and YOLOv11 with various parameter settings. Our quantitative and visual analysis revealed that U-Net and ResNet50 achieved the highest segmentation accuracy. However, lighter YOLO variants, particularly YOLOv8 and YOLOv11, offered a favorable trade-off between accuracy and computational efficiency, making them well-suited

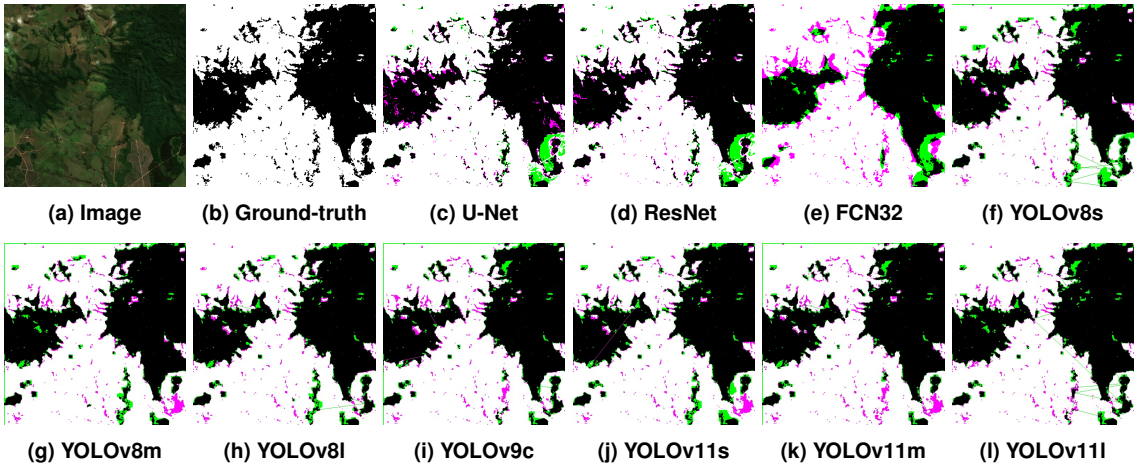

Figure 6. Qualitative comparison of a deforestation sample from Atlantic dataset.

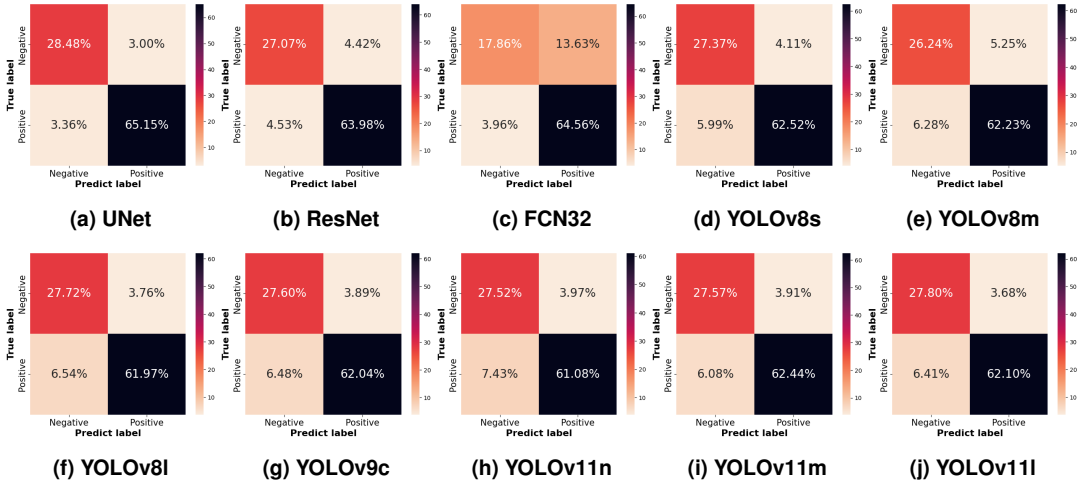

Figure 7. Cumulative confusion matrix for the Atlantic forest dataset.

for real-time monitoring applications. These findings reinforce the potential of deep learning for large-scale deforestation detection. By evaluating both the accuracy and computational efficiency of the deep neural models, our analysis provides practical insights for real-time monitoring applications, where computational resources may be limited.

This study also contributes by addressing deforestation in the Atlantic Forest biome through deep learning techniques. Although previous research on deforestation in this region has been limited, this work serves as an important starting point, providing practical guidelines and benchmarks for future studies.

Given the critical importance of image quality and accurate ground-truth annotations, future research should prioritize refining segmentation techniques and incorporating additional spectral data to enhance both accuracy and real-world applicability. Another potential approach is to expand the datasets, both in size and diversity [Ribas et al. 2025], to improve the generalization capabilities of the models.

Acknowledgments

This research was funded by the São Paulo Research Foundation (FAPESP), under grants 2023/14427-8, 2024/04718-8, and 2024/01610-1; the National Council for Scientific and Technological Development (CNPq), under grants 316228/2021-4, 305206/2024-9 and 305220/2022-5; and the Coordination for the Improvement of Higher Education Personnel (CAPES), Finance Code 001.

References

Bragagnolo, L., da Silva, R. V., and Grzybowski, J. M. V. (2021). Amazon and atlantic forest image datasets for semantic segmentation - 10.5281/zenodo.4498086.

Cabral, B. F., Yanai, A. M., de Alencastro Graça, P. M. L., Escada, M. I. S., de Almeida, C. M., and Fearnside, P. M. (2024). Amazon deforestation: A dangerous future indicated by patterns and trajectories in a hotspot of forest destruction in brazil. *Journal of Environmental Management*, 354:120354.

Careli, A., Vilas Boas, E., Teixeira, E., Silva, E., Aquino, G., and Pereira de Figueiredo, F. (2024). Deep learning segmentation models evaluation for deforestation monitoring embedded systems. In *2024 International Conference on Intelligent Cybernetics Technology & Applications*, pages 274–278.

de Bem, P. P., de Carvalho Junior, O. A., Fontes Guimarães, R., and Trancoso Gomes, R. A. (2020). Change detection of deforestation in the brazilian amazon using landsat data and convolutional neural networks. *Remote Sensing*, 12(6).

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.

Jocher, G., Chaurasia, A., and Qiu, J. (2023). Ultralytics yolov8.

Jocher, G. and Qiu, J. (2024). Ultralytics yolo11.

Kovačovič, P., Pirnik, R., Kafková, J., Michálik, M., Kanalikova, A., and Kuchár, P. (2025). Satellite-based forest stand detection using artificial intelligence. *IEEE Access*, 13:10898–10917.

Li, X., Fu, C., Wang, Q., Zhang, W., Sham, C.-W., and Chen, J. (2024). Dmsa-unet: Dual multi-scale attention makes unet more strong for medical image segmentation. *Knowledge-Based Systems*, 299:112050.

Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440.

Maurenza, D., Crouzeilles, R., Prevedello, J. A., and Almeida-Gomes, e. a. (2024). Effects of deforestation on multitaxa community similarity in the brazilian atlantic forest. *Conservation Biology*, n/a(n/a):e14419.

Md Jelas, I., Zulkifley, M. A., Abdullah, M., and Spraggon, M. (2024). Deforestation detection using deep learning-based semantic segmentation techniques: a systematic review. *Frontiers in Forests and Global Change*, 7:1300060.

Ortega Adarme, M., Queiroz Feitosa, R., Nigri Happ, P., Aparecido De Almeida, C., and Rodrigues Gomes, A. (2020). Evaluation of deep learning techniques for deforestation detection in the brazilian amazon and cerrado biomes from remote sensing imagery. *Remote Sensing*, 12(6).

Pathak, A., Kamani, M., and Priyanka, R. (2024). Satellite image segmentation via image quality enhancement and modified unet architecture. In *2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, pages 646–651.

Ribas, L. C., Casaca, W., and Fares, R. T. (2025). Conditional generative adversarial networks and deep learning data augmentation: A multi-perspective data-driven survey across multiple application fields and classification architectures. *AI*, 6(2):32.

Rodrigues, J., Dias, M. A., Negri, R., Hussain, S. M., and Casaca, W. (2024). A robust dual-mode machine learning framework for classifying deforestation patterns in amazon native lands. *Land*, 13(9):1427.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241. Springer.

Silpalatha, G. and Jayadeva, T. (2025). Accelerating fast and accurate instantaneous segmentation with yolo-v8 for remote sensing image analysis. *Remote Sensing Applications: Society and Environment*, 37:101502.

Torres, D. L., Turnes, J. N., Soto Vega, P. J., Feitosa, R. Q., Silva, D. E., Marcato Junior, J., and Almeida, C. (2021). Deforestation detection with fully convolutional networks in the amazon forest from landsat-8 and sentinel-2 images. *Remote Sensing*, 13(24).

Wang, C.-Y. and Liao, H.-Y. M. (2024). Yolov9: Learning what you want to learn using programmable gradient information.