ISSN: 2447-3359

REVISTA DE GEOCIÊNCIAS DO NORDESTE

Northeast Geosciences Journal

v. 11, nº 2 (2025)

https://doi.org/10.21680/2447-3359.2025v11n2ID41141

Quantifying Forest Loss in Legal Amazon Settlements through AI-Driven Remote Sensing

Quantificação da Perda Florestal em Assentamentos da Amazônia Legal por Meio de Sensoriamento Remoto com Inteligência Artificial

Sabrina do Carmo Alves[1]; Laura Coelho de Andrade[2]; Arthur Amaral e Silva[3] ; Helóisa Sâmela Borges[4]; Izabela Farias Oliveira[5]; Darlan Miranda Nunes[6]; Maria Lucia Calijuri[7]

[1] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: sabrinacarmoalves@yahoo.com ORCID: https://orcid.org/0000-0001-8618-5840
[2] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: laura.andrade@ufv.br ORCID: https://orcid.org/0000-0003-3693-2208
[3] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: arthuramaral.e.a@gmail.com ORCID: https://orcid.org/0000-0001-5548-459X
[4] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: heloisa.borges@ufv.br ORCID: https://orcid.org/0009-0002-0220-3154
[5] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: izabela.fo@gmail.com ORCID: https://orcid.org/0009-0009-7399-9592
[6] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: darlan.nunes@ufv.br ORCID: https://orcid.org/0000-0001-5566-7919
[7] Federal University of Viçosa, Department of Civil Engineering, Viçosa/MG, Brazil. Email: calijuri@ufv.br ORCID: https://orcid.org/0000-0002-0918-2475

Abstract: The Amazon rainforest, a critical component of the Earth's climate system, faces increasing deforestation, particularly in settlement areas established through agrarian reform programs. This study investigates forest loss in the Alcobaça Settlement and Juma Tract in the Legal Amazon from 2018 to 2022 using remote sensing and artificial intelligence (AI) techniques. Sentinel-2 satellite imagery was analyzed using machine learning algorithms, including Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and Convolutional Neural Networks (CNNs), to classify land cover and quantify deforestation trends. The results demonstrate that CNN outperformed the other classifiers, achieving the highest accuracy and better identifying deforestation patterns over time. The trained model was then applied to the Juma Tract to assess its generalization capability. Although the CNN approach proved effective, it overestimated deforestation by 8.32% in 2022 compared to manual classification, highlighting challenges in transferring machine learning models to different regions without additional calibration. The findings emphasize the potential of AI-driven remote sensing for large-scale environmental monitoring while underscoring the necessity of localized training and validation to improve classification accuracy. This research contributes to the development of automated methods for forest loss assessment, providing valuable insights for environmental management and policy-making in the Amazon.
Keywords: Remote Sensing; Machine Learning; Deforestation; Amazon.

Resumo: A floresta amazônica, um componente crucial do sistema climático da Terra, enfrenta um aumento do desmatamento, especialmente em áreas de assentamentos estabelecidas por meio de programas de reforma agrária. Este estudo investiga a perda florestal no Assentamento Alcobaça e no Trecho Juma, localizados na Amazônia Legal, no período de 2018 a 2022, utilizando técnicas de sensoriamento remoto e inteligência artificial (IA). Imagens do satélite Sentinel-2 foram analisadas por meio de algoritmos de aprendizado de máquina, incluindo Máquinas de Vetores de Suporte (SVM), Florestas Aleatórias (RF), Árvores de Decisão (DT) e Redes Neurais Convolucionais (CNNs), para classificar a cobertura da terra e quantificar as tendências de desmatamento. Os resultados demonstraram que as CNNs superaram os demais classificadores, atingindo a maior acurácia e melhor identificando os padrões de desmatamento ao longo do tempo. O modelo treinado foi então aplicado ao Trecho Juma para avaliar sua capacidade de generalização. Embora a abordagem com CNN tenha se mostrado eficaz, ela superestimou o desmatamento em 8,32% em 2022 em comparação com a classificação manual, evidenciando os desafios na transferência de modelos de aprendizado de máquina para diferentes regiões sem calibração adicional. Os resultados destacam também o potencial do sensoriamento remoto com IA para o monitoramento ambiental em larga escala, ao mesmo tempo em que reforçam a necessidade de treinamento e validação localizados para melhorar a precisão das classificações. Assim, esta pesquisa contribui para o desenvolvimento de métodos automatizados de avaliação da perda florestal, fornecendo informações valiosas para a gestão ambiental e a formulação de políticas públicas na Amazônia.
Palavras-chave: Sensoriamento Remoto; Aprendizado de Máquina; Desmatamento; Amazônia.

Received: 14/08/2025; Accepted: 22/10/2025; Published: 28/10/2025.

1. Introduction

The Amazon, the largest tropical forest on the planet, spans around 5,2 km² and plays a fundamental role in climate regulation, oxygen production, and the preservation of water resources and biodiversity (SALISBURY et al., 2012; FEARNSIDE, 2017; GATTI et al., 2021). In addition to storing 25% of the Earth's carbon, the region is vital for local communities, including Indigenous and traditional populations, who depend on the forest for their livelihoods (NEPSTAD et al., 2008). However, land conflicts involving rural workers, Indigenous peoples, and agribusiness — exacerbated by illegal activities such as deforestation and land grabbing — pose significant challenges to conservation and the implementation of effective public policies (BECKER, 2013; MIRANDA; PERES; CARVALHO, 2019). Recent data from INPE (2024) show that over 13% of deforestation in the Legal Amazon occurs within rural settlements, where traditional monitoring remains limited, reinforcing the need for innovative analytical approaches.

Given its limited staff and infrastructure, INCRA has signed Decentralized Execution Agreements (TEDs) with universities, which have been crucial in expanding its technical capacity. These agreements enable academic institutions to support INCRA in research, surveys, and analyses, contributing to land regularization and territorial planning in the Legal Amazon. TEDs foster knowledge exchange between managers and researchers, promoting the adoption of new technologies and methodologies. Moreover, these partnerships strengthen land governance, increase policy efficiency, and optimize financial resources, with universities ensuring transparency and credibility (SILVA et al., 2022).

In 2017, INCRA partnered with the Federal University of Viçosa (UFV) through a TED, resulting in the RADIS Project, which, upon its completion in late 2023, supported the occupational review for land regularization of 308 rural settlements, benefiting over 30,000 families. Additionally, since 2021, INCRA and UFV have been developing the AMARIS Project, focused on improving the registration of federal public lands for land regularization purposes, applying innovative methodologies to enhance occupation mapping and ensure greater legal security in the region.

Federal Tracts represent state-owned lands for various uses, whereas settlement projects are planned units for land redistribution and family farming, ensuring legal security and access to public policies for sustainable rural development (SANTOS; OLIVEIRA, 2021). Managing these vast areas in the Legal Amazon remains challenging, given their socio-environmental complexity and the limitations of in situ monitoring, which makes remote technologies indispensable.

Remote sensing combined with artificial intelligence (AI) algorithms such as Machine Learning and Deep Learning has become a key tool for environmental and territorial management, enabling real-time forest loss detection and land use monitoring (COSTA; SOUZA; SILVA, 2020; RUDORFF et al., 2020). While previous studies have mapped large-scale deforestation, few have focused on rural settlements and federal tracts, representing a gap in applied territorial governance research. AI models can handle large datasets, detect subtle changes, and anticipate degradation trends (PEREIRA et al., 2023), enhancing monitoring precision and public policy effectiveness (SOARES et al., 2021).

In this context, this study analyzes and quantifies forest loss in a Settlement Project (SP) located in Pará, within the Legal Amazon, from 2018 to 2022, using satellite imagery and AI techniques. The best-performing model was also applied to a Federal Tract under INCRA’s management, demonstrating the method’s transferability and potential to support territorial management. This study seeks to answer: (1) What is the accuracy of the proposed AI-based method for forest loss detection? and (2) How does it compare to traditional remote sensing techniques in identifying deforestation within Federal Tracts?

2. Material and Methods

2.1 Study Area

For this study, the selected study areas were the Alcobaça Settlement Project (SP), located in the state of Pará, in the municipality of Breu Branco (Figure 1), and the region of Apuí, in the state of Amazonas, where is situated the Federal Juma Tract (Figure 2).

[1] https://dataspace.copernicus.eu/

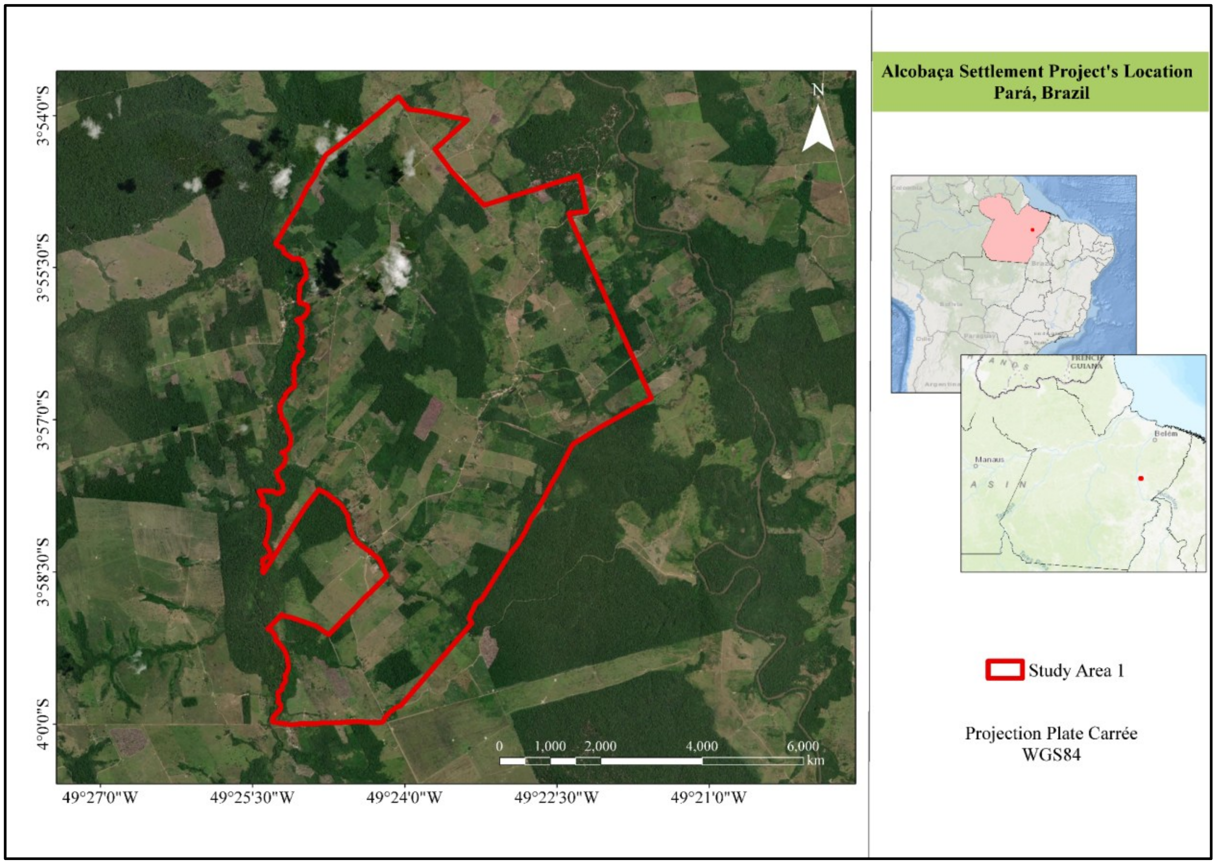


Figure 1 – Location of the Alcobaça SP – Pará, Brazil.
Source: Authors (2025).

The state of Pará is the second-largest federative unit in Brazil in terms of territorial extension, with an area of 1,245,870.704 km², of which 1,745.77 km² is urbanized (IBGE, 2022). The Alcobaça Settlement Project (PA) covers an area of approximately 5,000 hectares.

[1]https://dataspace.copernicus.eu/

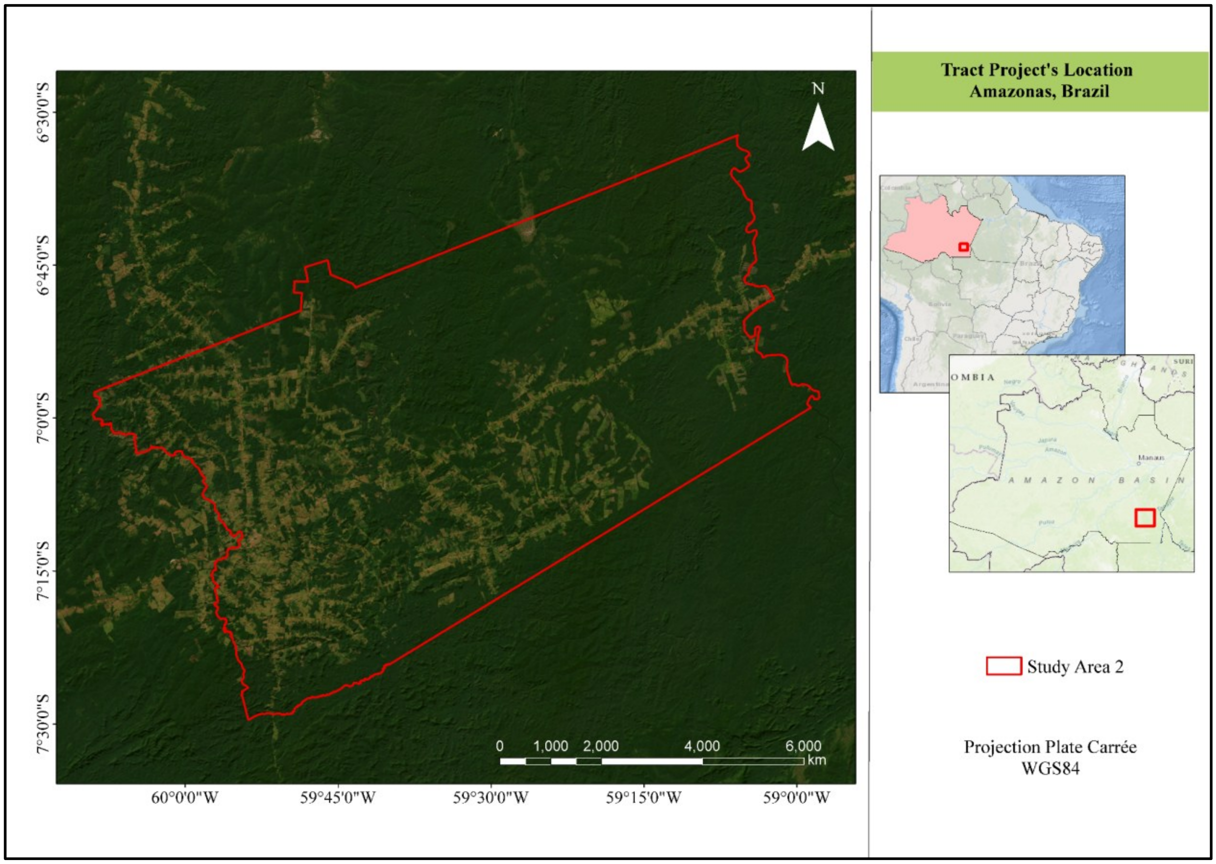


Figure 2 – Location of the Juma Tract, Amazonas, Brazil.
Source: Authors (2025).

The Juma Land, located in the municipality of Apuí, Amazonas, was originally established as the Directed Settlement Project (DSP) "Rio Juma" in 1982, covering approximately 689,000 hectares, within the framework of INCRA's colonization policies for the Amazon (GALUCH; DA COSTA, 2023). Initially intended for family farming, the area underwent an intense process of land speculation, land concentration, and livestock expansion, leading to its loss of settlement status. This process resulted in the reclassification of the territory as public land, consolidating its integration into the formal land market. The change favored the expansion of agribusiness and contributed to the increase in irregular occupation and deforestation in the region (SANTOS; PONTES FILHO, 2024).

2.2 Methods

The methodology adopted in this study follows a structured sequence of steps, as illustrated in the flowchart in Figure 3. The main steps include database preparation, processing, result acquisition, and analysis, as detailed below.

[1]https://dataspace.copernicus.eu/

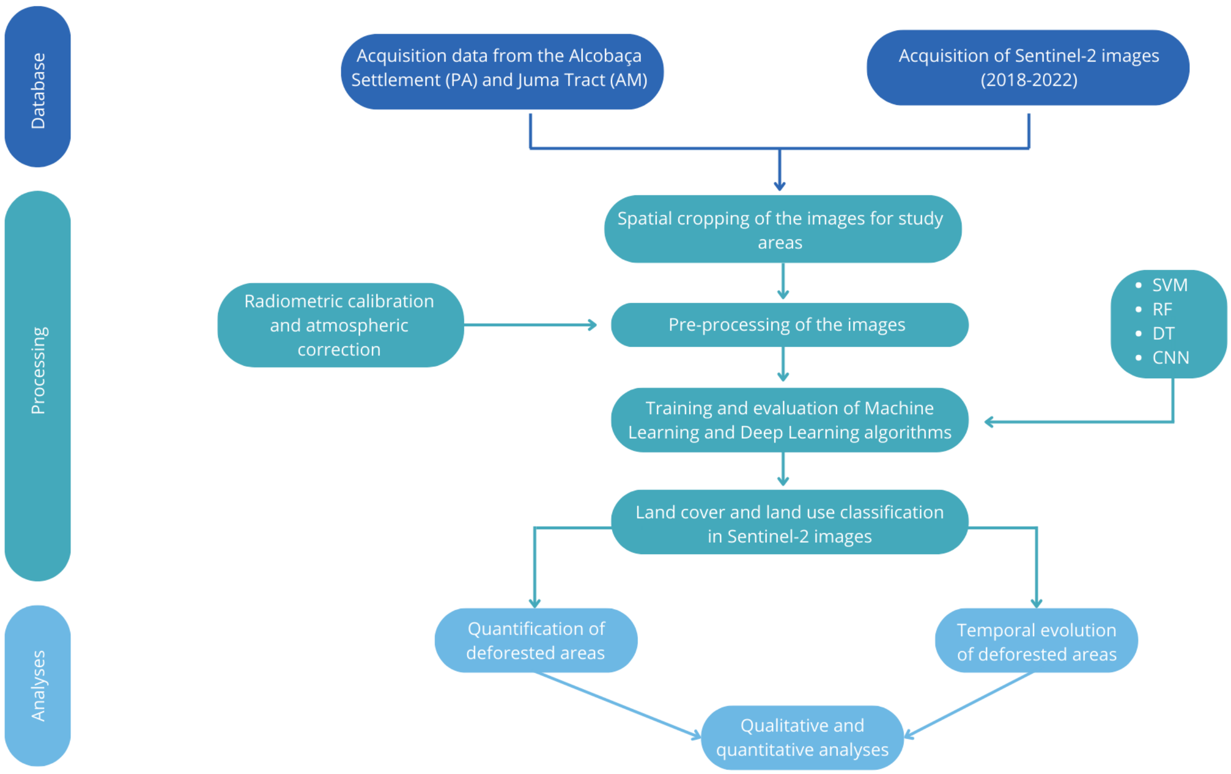


Figure 3 – Methodology flowchart.
Source: Authors (2025).

2.2.1 Data source

Initially, the study area was defined, corresponding to the Alcobaça Settlement Project (SP) in Pará, which is included in the scope of the RADIS-UFV Project, as well as the region of Juma Tract, located in the state of Amazonas, which is the study area of the Amaris Project.

It is important to note that for the Alcobaça SP, the study area was limited to the settlement itself. In contrast, for the Juma Tract, the entire portion available in the orbital image was analyzed, with the aim of classifying a "worst" scenario, with a larger area and greater diversity of spectral responses.

Once the areas of interest were established, orbital images from the Sentinel-2 satellite were obtained through the Copernicus portal[1] for the years 2018, 2019, 2020, 2021, and 2022 for the Alcobaça SP, and for the years 2018 and 2022 for the Juma Tract, aiming to quantify the deforested area in the latter over four years.

For this study, only spectral bands with a spatial resolution of 10 meters (blue, green, red, and near-infrared) were considered. Additionally, efforts were made to select images with the lowest possible cloud cover for both locations.

2.2.2 Processing

The next step consisted of preprocessing the selected Sentinel-2 images for both locations. First, radiometric calibration was performed using the image metadata, adjusting the radiance values to correct the bands' discrepant gains and offsets.

Next, atmospheric correction was applied using the Dark Object Subtraction method. This process allowed for the removal of atmospheric scattering effects by subtracting the lowest pixel value as the background signature of the image.

With the preprocessed images, a spatial clipping was performed for the Alcobaça SP, using a bounding rectangle around the study area, resulting in five (05) images corresponding to each analyzed year. Subsequently, samples were prepared for training the machine learning models, considering the specific characteristics of each scene. A set of spatially

[1]https://dataspace.copernicus.eu/

distributed points was defined across the study area, representing the six land cover and land use classes considered: Forest, Deforestation, Bare Soil, Clouds, Shadows, and Hydrography (Figure 4).

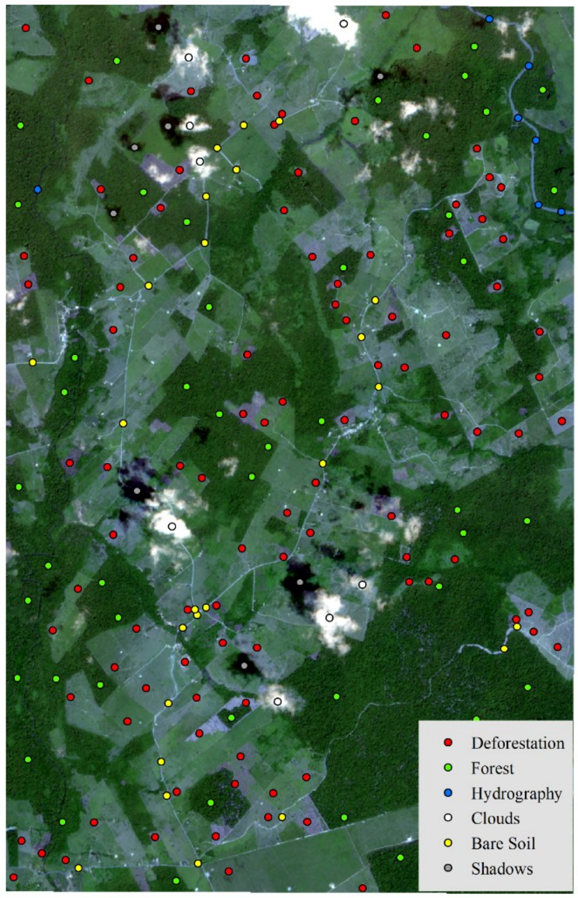


Figure 4 – Distribution of points in Alcobaça SP.
Source: Authors (2025).

2.2.3 Artificial Intelligence Algorithms

Among the Artificial Intelligence methods (Machine Learning and Deep Learning), Support Vector Machine, Random Forest, Decision Tree, and Convolutional Neural Network were tested, as these methods are widely used in orbital image classification (PAL; MATHER, 2005; BELGIU; DRĂGUŢ, 2016; MAGGIORI et al., 2017; SHARMA; KUMAR, 2016; ZHU et al., 2017).

2.2.3.1 Support Vector Machine (SVM)

The SVM algorithm is a widely applied tool in regression and classification tasks. Its main approach consists of finding a hyperplane that functions as a decision surface, separating data classes in a high-dimensional space (GRAF et al., 2004; SOUSA, 2009). The core characteristic of SVM is its ability to minimize the margin between the classes and the separating hyperplane, ensuring robustness and a high generalization capacity in the final results (BONESSO, 2013).

2.2.3.2. Random Forest (RF)

[1]https://dataspace.copernicus.eu/

One of the main advantages of this algorithm is the diversity introduced by the randomness in the selection of data subsets and features, allowing Random Forest to be resilient to outliers and noise in the data, ensuring more reliable generalization. The method stands out for its accuracy, especially in situations where data are scarce or there is high dimensionality in the feature space, demonstrating great potential for application in complex real-world systems.

It is noteworthy that in this study, the RandomForestClassifier function was employed with controlled authority, where the majority vote of the trees is used to classify the data.

2.2.3.3. Decision Tree (DT)

The Decision Tree method builds models using simple rules derived from data, offering ease of use, support for diverse data types, and transparency, thus being considered a "white-box" model (AKAR; GÜNGÖR, 2012). However, it may produce overly complex trees, leading to overfitting and reduced generalization. Techniques like pruning and depth limitation help address this. The model can also be unstable and biased with unbalanced data. Optimal tree construction often relies on heuristics that may not ensure globally best solutions (LEE; CHEANG; MOSLEHPOUR, 2022; NOWOZIN et al., 2011).

2.2.3.4.Convolutional Neural Network

Convolutional Neural Networks (CNNs) are deep learning models suited for grid-like data such as satellite images. Their layered architecture learns spatial hierarchies from input data, enhancing performance in remote sensing tasks (MAGGIORI et al., 2017; ZHU et al., 2017). Convolutional layers extract features like edges and textures, learned via backpropagation, while pooling layers reduce spatial dimensions and computational cost (CIRESAN et al., 2011; MAGGIORI et al., 2017). Unlike traditional methods that require handcrafted features, CNNs learn relevant patterns automatically. Yet, they depend on large, well-labeled datasets, which can be a limitation in remote regions (ZHU et al., 2017).

2.2.4. Analyses - Metrics for Assessing Model Performance

The evaluation of the models for the SP Alcobaça was carried out using the parameters of accuracy, precision, recall, F1-score, Kappa index, Standard Deviation and RMSE (Root Mean Squared Error), providing a comprehensive assessment of their quality. Additionally, the areas of vegetation loss occurrence were quantified according to the algorithms, and a temporal analysis of its evolution was also performed.

For the Juma Tract, where the best model was applied for the years 2018 and 2022, the methodology evaluation was conducted by comparing the forest loss area classified manually by a specialist and automatically for the aforementioned years, also with the parameters accuracy, precision, recall, F1-score and Kappa index. This analysis could also be assessed qualitatively.

3. Results

3.1 Alcobaça Settlement

The Sentinel-2 images covering SP Alcobaça, corresponding to the years 2018 to 2022, were classified using algorithms trained based on the provided samples. The performance of each algorithm was analyzed using performance metrics applied to each of the generated models. Subsequently, the trained models were applied to classify the images. Based on the resulting land cover and land use classes, the deforested areas in the SP were quantified over the years.

Table 1 presents the accuracy values obtained for each algorithm for the images from the five analyzed years.

Table 1 – Accuracy after training each algorithm.

Algorithm	2018	2019	2020	2021	2022
	Accuracy of the trained algorithm (%)				
SVM	92.31	92.31	95.06	93.75	95.18

[1]https://dataspace.copernicus.eu/

RF	96.61	90.91	96.69	97.92	96.77
DT	93.22	93.51	96.69	92.71	95.16
CNN	94.87	98.08	96.30	95.31	95.18

Source: Authors (2025).

The values presented in Table 1 above were extracted from the model evaluation reports after training and testing.

To assess the performance of the different algorithms applied to satellite image classification, Table 2 presents the average accuracy, mean deviation, root mean square error and mean kappa index.

Table 2 – Performance Evaluation of the Algorithms.

GENERAL STATISTICS	ALGORITHM			
	SVM	**RF**	**DT**	**CNN**
Mean Accuracy (%)	93.72	95.78	94.26	95.95
Mean Deviation	1.13	1.95	1.33	0.99
Root Mean Squared Error	2.83	17.18	8.01	3.62
Mean Kappa Index	0.89	0.90	0.88	0.92

Source: Authors (2025).

It is noteworthy that the average Kappa index values for the tested algorithms ranged between 0.88 and 0.92, with CNN achieving the highest value (0.92), followed by Random Forest (0.90), SVM (0.89), and Decision Tree (0.88). CNN also had the lowest mean deviation (0.99), whereas Random Forest had the highest mean deviation (1.95), suggesting greater variability in model performance over the analyzed years.

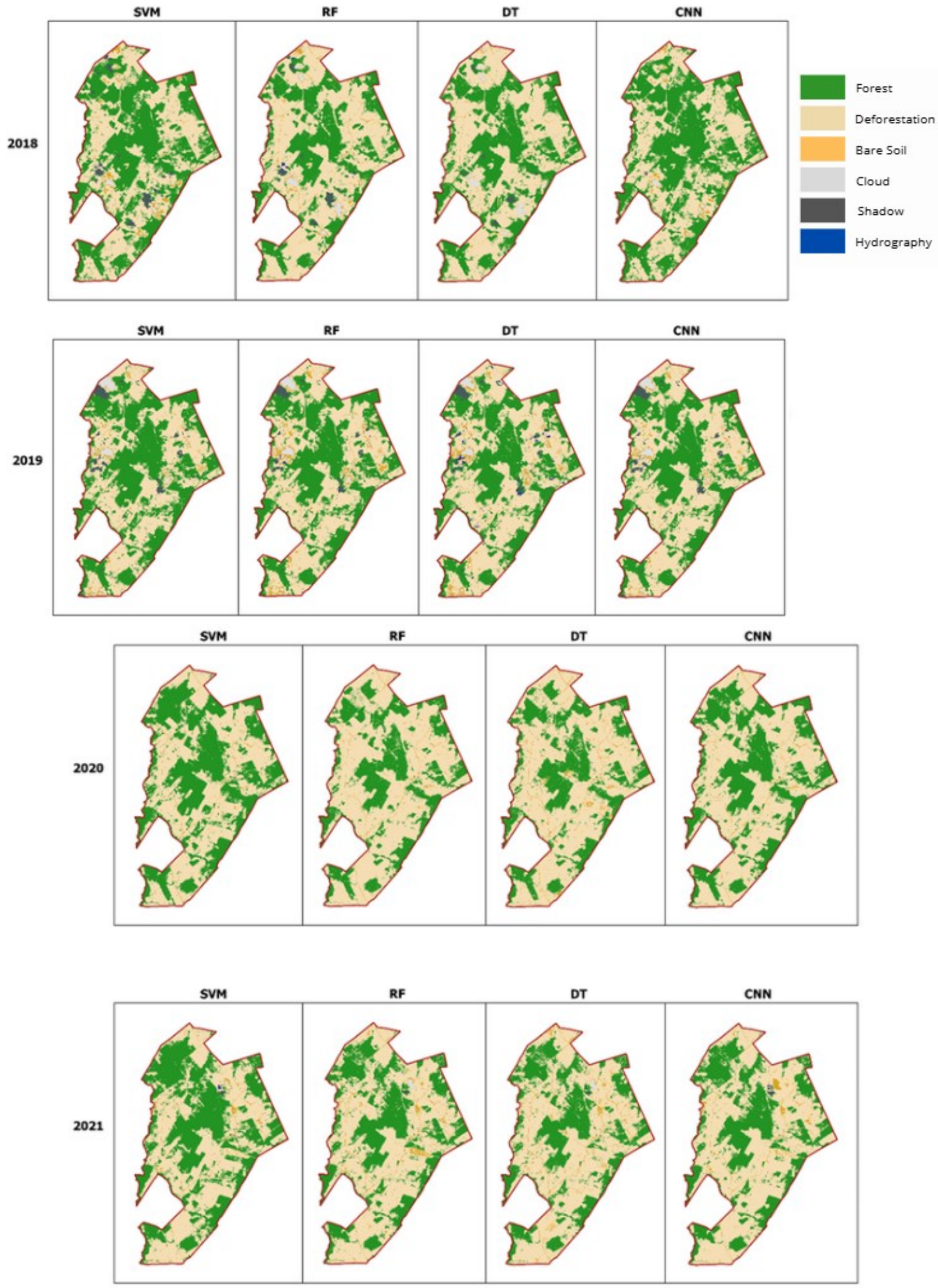
Thus, the use of CNNs stood out due to their ability to detect small deforestation fragments that were not identified by other algorithms. This capability stems from the CNNs' structure, which allows them to capture subtle features in the images. CNNs exhibited the highest average accuracy and the lowest mean deviation, and in 2019, they reached 98.08% accuracy, the best among all algorithms.

The SVM algorithm also performed well in terms of mean deviation, showing a lower RMSE but a slightly lower average accuracy compared to the other models. On the other hand, the Random Forest algorithm demonstrated effectiveness in generalizing classification patterns across the analyzed years, but it had the highest recorded mean deviation. Inconsistencies were observed, particularly in 2019, which may be attributed to the algorithm's sensitivity to specific variables, such as variability in climatic conditions and the heterogeneity of input data.

Nevertheless, the choice of the algorithm should be carefully considered based on several factors, including the spatial resolution of the images, the representativeness of the samples used, the prevailing environmental conditions during image acquisition, and the availability of high-quality training data. Each of these factors can significantly influence model performance, making algorithm selection a process that should be tailored to the specific application context (CIHLAR, 2000; FOODY, 2002; YU et al., 2014).

Regarding land cover and land use in the SP area during the analyzed period (2018 to 2022), a trend of decreasing forest cover was observed, indicating an increase in deforestation in the SP area (Figure 5).

[1]https://dataspace.copernicus.eu/



[1]https://dataspace.copernicus.eu/

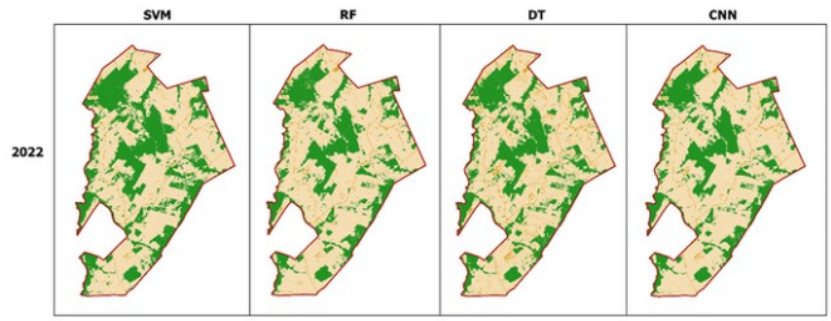


Figure 5 – Land Cover and Land Use in SP Alcobaça, resulting from the classification of images over the years by each algorithm.
Source: Authors (2025).

During the five-year period analyzed in this study, a comparative assessment was conducted of the areas classified as forest by the different algorithms studied. Table 3 presents the extent of these areas in hectares, highlighting the reduction in forest cover within the boundaries of the Settlement Project (SP). These data indicate the magnitude of deforestation and the relative efficiency of each algorithm in capturing changes in vegetation cover over time, contributing to a better understanding of the environmental impacts in settlement areas.

Table 3 – Forest Cover Loss in SP Alcobaça Over a 5-Year Period.

Algorithm	**2018**	**2022**	**Total Difference in Forest Areas in SP Alcobaça Over 5 Years (ha)**
	Area (ha)		
SVM	2210.93	1652.32	558.6
RF	1676.98	1337.410	339.6
DT	1644.97	1424.56	220.4
CNN	1848.50	1351.42	497.1

Source: Authors (2025).

3.2 Juma Tract

Based on the results obtained for SP Alcobaça, where CNN demonstrated the best performance and highest agreement, this algorithm was selected for the automatic classification of Juma Tract, in the Amazon region, using the calibrated model for the year 2018.

The decision to apply the CNN model trained on SP Alcobaça directly to Juma Tract region was based on the similarity in vegetation cover and land occupation patterns between the two areas. However, it is important to emphasize that, despite sharing ecological characteristics, differences in land use dynamics and environmental conditions may affect the model's accuracy when transferred to a new location. Furthermore, since no new training data were used for Juma, the model's generalization may lead to variations in the results.

Figure 6 illustrates the classification performed using the convolutional neural network and the manual classification for the image containing Juma Tract.

[1]https://dataspace.copernicus.eu/

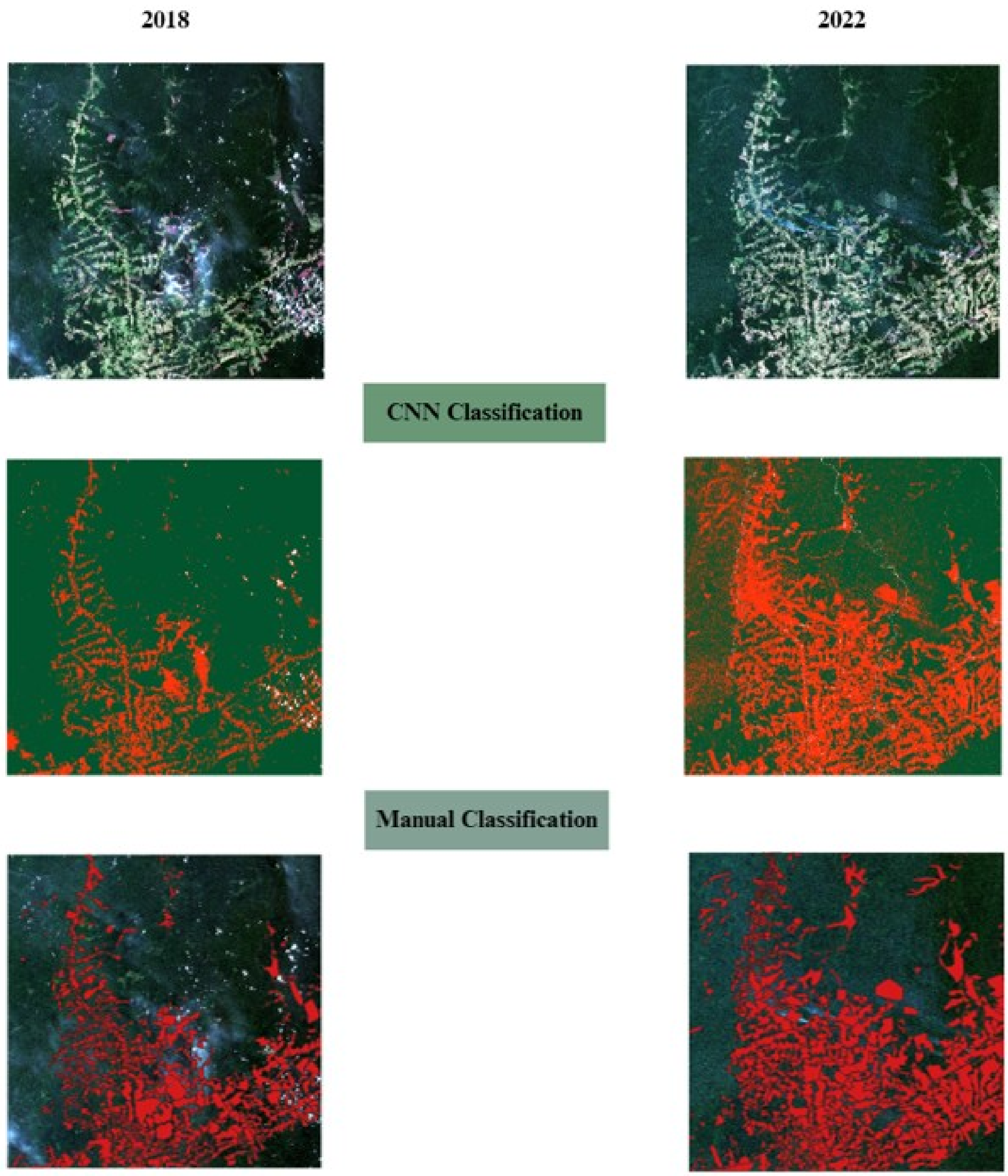


Figure 6 – CNN and Manual Classification for the region of Juma Tract.
Source: Authors (2025).

The quantitative analysis revealed that between 2018 and 2022, the CNN estimated a forest cover loss of approximately 50,492.97 ha. This discrepancy suggests that the CNN overestimated deforestation in 2022, identifying about 8.32% more lost area compared to the manual classification. In contrast, for the year 2018, the CNN underestimated forest loss by 5.23%, indicating that the model may have exhibited bias over time.

Table 4 highlights the percentage of equivalent values calculated using the CNN and manual classification.

[1]https://dataspace.copernicus.eu/

Table 4 – Forest Loss Areas According to CNN and Manual Classification.

	CNN (ha)	**Manual (ha)**	**Absolute Difference (ha)**	**Difference (%)**
2018	273,086.03	288,150.07	15,064.04	5.23
2022	461,413.68	425,984.75	35,428.93	8.32
Forest Cover Loss (5 years)	188,327.65	137,834.68	50,492.97	36.63

Source: Authors (2025).

The results for the average metrics calculated for Juma Tract are presented in Table 5, highlighting the classification performance based on accuracy, Kappa index, recall, precision, and F1-score.

Table 5 – Metrics for Juma Tract Classification.

Average Metrics	**CNN**
Accuracy (%)	82.29
Precision (%)	72.69
Recall (%)	79.57
F1-Score (%)	75.97
Kappa	0.62

Source: Authors (2025).

The classification results indicate a satisfactory performance, with an average accuracy of 82.29% and a Kappa index of 0.62, suggesting a moderate to substantial agreement between the automatic classification and the manual reference. The average precision of 72.69% and recall of 79.57% show that the model successfully identified most areas of interest, although there is still a considerable rate of false positives. The F1-Score of 75.97%, which balances precision and recall, reinforces that the automatic classification exhibits a consistent performance, although it is not free from errors.

Furthermore, it is important to mention that the complete image covers an area of 1,197,538.08 ha (11,975.38 km²), demonstrating that the differences between the manual and CNN classifications are relatively small when compared to the total area.

For the validation of the manual classification, a topology with the non-overlapping rule for polygons was applied to correct possible human errors during execution. Additionally, it is worth noting that the vectorization process took approximately 80 hours to complete for both images (2018 and 2022).

4. Discussions

4.1 Alcobaça Settlement

The analysis of supervised classification models for SP Alcobaça revealed a high level of accuracy and agreement with the reference classification. Authors such as Fleiss, Levin, and Paik (2003) classify a Kappa value above 0.75 as excellent, while Landis and Koch (1977) suggest that a Kappa value above 0.81 indicates near-perfect agreement, aligning with the results presented in this study.

Among the tested models, CNN exhibited the best overall performance, achieving the highest accuracy and lowest mean deviation. This superior performance can be attributed to its ability to capture complex spatial and spectral patterns

[1]https://dataspace.copernicus.eu/

in the image, as evidenced by its higher Kappa index and detection of small deforestation fragments (Table 2). CNNs have gained prominence in land use and land cover classification due to their capacity to extract intricate patterns, reducing errors commonly observed in traditional methods such as Random Forest (RF) and Support Vector Machines (SVM) (Maggiori et al., 2016). However, despite their high accuracy, CNNs pose challenges such as high computational cost and the need for large volumes of labeled data, which may limit their application in cases with limited training samples. In this context, authors like Maggiori et al. (2016) and Zhu et al. (2017) highlight that strategies such as semi-supervised learning and data augmentation are valuable for mitigating these limitations, allowing for better model generalization.

The SVM model demonstrated an intermediate performance, achieving an average accuracy of 93.72% and the lowest root mean square error (2.83) among the tested models. However, its slightly lower performance may be attributed to difficulties in defining an optimal hyperplane for class separation in areas with complex spectral patterns, as pointed out by Pal and Mather (2005).

Finally, the Decision Tree (DT) model had the lowest average Kappa index (0.88) and a higher tendency toward overfitting since it generates deep, highly specialized trees based on training samples. This behavior reduces the model's ability to generalize to unseen data, leading to inferior performance (Nowozin et al., 2011). On the other hand, Sharma and Kumar (2016) argue that applying attribute selection techniques and pruning can reduce overfitting while maintaining greater interpretability compared to other decision tree-based methods, depending on the dataset.

Beyond the statistical analysis of the models, the annual variation in forest cover in SP Alcobaça demonstrated a progressive increase in deforestation between 2018 and 2022, as shown in Table 5. These findings reinforce the importance of continuous monitoring in the region and the application of AI-based approaches to provide rapid and accurate information on vegetation cover changes.

4.2 Juma Tract

For Juma Tract, a different approach was applied, in which the CNN model trained in SP Alcobaça was used without introducing new training samples, characterizing an unsupervised classification. This strategy was adopted to evaluate the model's generalization in an area with relatively similar characteristics of the Amazon biome but without specific refinements for its environmental conditions and land use patterns.

The results indicated that the model overestimated deforestation in Juma in 2022 when compared to the manual classification, with a difference of approximately 35,000 hectares (8.32%). This discrepancy can be explained by the fact that the model was trained with data from SP Alcobaça for 2018, where the spectral characteristics of the classes are slightly different from those of Juma in 2022. Another determining factor for this difference may be the confusion between exposed soil and deforestation, as the CNN mistakenly classified open areas as deforested due to spectral similarities.

The classification metric results indicate satisfactory performance, with an accuracy of 82.29% and a Kappa index of 0.62, suggesting a moderate to substantial agreement between the automatic classification and the manual reference (CONGALTON; GREEN, 2019). However, the relationship between precision (72.69%) and recall (79.57%) reveals an imbalance in commission and omission errors, also evident in the area calculations. The lower precision compared to recall indicates that the model is more prone to commission errors, incorrectly classifying certain areas as deforested when they are not, which contributes to overestimation of deforestation (FOODY, 2002). Conversely, the recall value suggests that omission errors also occur, where some areas belonging to a class were not correctly identified, potentially leading to an underestimation of certain phenomena (OLOFSSON et al., 2014). The F1-Score of 75.97% confirms that the classification maintains a balanced performance between these errors but still exhibits a tendency for the model to incorrectly include pixels in the target class.

Another relevant point is that the spatial resolution of the images (10 meters per pixel) amplified the impacts of classification errors. Since each error corresponds to 100 m² of misclassified area, small prediction failures can result in considerable differences in the total deforestation estimate. This issue has already been reported in studies analyzing the influence of spatial resolution in AI-based classification models (CIHLAR, 2000; YU et al., 2014).

Additionally, the manual classification approach required approximately 80 hours of vectorization, reinforcing the advantage of AI in automating environmental monitoring processes. However, the results indicate that applying a model trained in a different area requires careful consideration and specific adjustments to improve accuracy. Potential improvements include model recalibration with local samples, transfer learning techniques to fine-tune weights based on the new region, modifications in the neural network architecture, and ensuring the training data align temporally with the classification period (ZHU et al., 2017; SILVEIRA et al., 2020). These refinements could help mitigate spectral and spatial discrepancies, reducing classification errors.

[1]https://dataspace.copernicus.eu/

Finally, the analysis of Juma Tract highlights the importance of testing the adaptability of trained models before applying them to new areas, reinforcing that the accuracy of a supervised model can decrease significantly when applied without refinement in a new context.

5. Conclusions

The analysis of forest cover loss in the Settlement Project (SP) Alcobaça through supervised classification demonstrated the effectiveness of AI applied to remote sensing for quantifying deforestation. The comparison between different algorithms revealed that CNN achieved the best performance, capturing complex spatial patterns and ensuring high accuracy in detecting deforested areas. Based on these results, the model trained for 2018 in SP Alcobaça was applied to Juma Tract region to assess the feasibility of transferring the learning process to another recently occupied area in the Amazon. However, the results indicated that, despite a good initial match, there was an overestimation of forest loss in 2022, highlighting the need for specific adjustments for the new region.

Juma Tract has been one of the main areas of agricultural frontier expansion in the Amazon, with high rates of land speculation, deforestation, and conversion of forest areas into pastures. The transition from a Settlement Project to Public Land intensified these dynamics, consolidating its integration into the formal land market and increasing challenges for environmental and land governance. In this context, the adoption of advanced monitoring techniques, such as AI-based unsupervised classification, becomes a fundamental strategy for the rapid and accurate detection of land use changes, enabling more effective enforcement and conservation actions.

Thus, this study reinforces the importance of continuous monitoring of deforestation in the Amazon, both to support conservation policies and to prevent illegal forest conversion practices. The strategic use of AI applied to remote sensing can revolutionize biome monitoring, providing fast, accurate, and scalable responses for territorial management in the context of settlement projects and public lands. However, its implementation must be carried out responsibly, ensuring that analyses accurately reflect field reality, preventing distortions that could compromise environmental and land-use decisions in the region. Future research should consider applying transfer learning or fine-tuning to adapt models to new regions, with a better resolution satellite systms, as well as conducting sensitivity analyses to better understand the influence of environmental variables on classification outcomes.

References

AKAR, Ö.; GÜNGÖR, O. Classification of multispectral images using Random Forest algorithm. Journal of Geodesy and Geoinformation, v. 1, n. 2, p. 105–112, 2012.

AKAR, Ö.; GÜNGÖR, O. Classification of multispectral images using Random Forest algorithm. Journal of Geodesy and Geoinformation, v. 1, n. 2, p. 105–112, 2012.

BECKER, B. (2013). Amazônia: economia, ecologia e desafios do desenvolvimento sustentável. São Paulo: Editora XYZ.

BECKER, B. K. (2013). Revisão das políticas de ocupação da Amazônia: é possível identificar modelos para projetar cenários? Estudos Avançados, 19(53), 71-86. Disponível em: https://docs.ufpr.br/~adilar/GEOPOL%C3%8DTICA2019/Geopolitica%20da%20Amazonia/Amaz%C3%B4nia_Pol%C3%ADtica%20de%20ocupa%C3%A7%C3%A3o.pdf

BELGIU, M.; DRĂGUŢ, L. Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, v. 114, p. 24-31, 2016.

BEZERRA, J. F., FERREIRA, M. C., & OLIVEIRA, R. S. (2021). Desafios da regularização fundiária na Amazônia Legal: aspectos jurídicos e institucionais. Revista Brasileira de Estudos Rurais e Urbanos, 6(1), 85-102. Disponível em: https://www.rberu.org.br/desafios-fundiarios-amazonia

BÖHM, S. M.; SOBIERANSKI, A. C. Análise de Performance de um Algoritmo de Reconhecimento Facial por Visão Computacional Aplicado a Sistemas Embarcados, 2021.

BONESSO, D. Estimação dos Parâmetros do Kernel em um Classificador SVM na Classificação de Imagens Hiperespectrais em uma Abordagem Multiclasse, 2013.

[1]https://dataspace.copernicus.eu/

Brazil. (2018). Decree No. 9,311, of March 15, 2018. Regulates provisions of Law No. 8,629/1993 related to agrarian reform. Official Gazette of the Union. Retrieved from http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/d9311.htm

BREIMAN, L. Random Forests. Random Forests, 2001. Disponível em: https://link.springer.com/article/10.1023/a:1010933404324. Acesso em: 25 ago. 2024.

CARVALHO, W. D. S., FILHO, F. J. C. M., & SANTOS, T. L. D. (2021). Uso e cobertura do solo utilizando a plataforma google earth engine (gee): estudo de caso em uma unidade de conservação / land use and land cover using the google earth engine platform (gee): case study in a conservation unit. Brazilian Journal of Development, 7(2), 15280-15300.

CIHLAR, J. Land cover mapping of large areas from satellites: status and research priorities. International journal of remote sensing, v. 21, n. 6-7, p. 1093-1114, 2000.

CIRESAN, D. C. et al. Flexible, High Performance Convolutional Neural Networks for Image Classification, 2011.

CONGALTON, R. G., & GREEN, K. (2019). Assessing the accuracy of remotely sensed data: principles and practices. CRC press.

COSTA, L. R., SOUZA, A. F., & SILVA, T. P. (2020). Aplicações de sensoriamento remoto para a gestão territorial na Amazônia: Desafios e perspectivas. Revista Brasileira de Geografia, 39(4), 281-299.

FEARNSIDE, P. M. (2017). Desmatamento na Amazônia: impactos e soluções. Brasília: Editora Universidade.

FERREIRA, A. M. M., & SALATI, E. (2005). Forças de transformação do ecossistema amazônico. Estudos Avançados, 19(53), 25-40. Disponível em: https://www.scielo.br/j/ea/a/Y77GvYY7DkMnkRZynzQY5Qx/

FLEISS,J.L;LEVIN, B. AND PAIK, M.C. Statistical Methods for Rates and Proportions. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, sep 2003.

FRANKLIN, S. E., & WULDER, M. A. (2002). Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. Progress in Physical Geography, 26(2), 173-205.

GADELHA, R. M. A. F. (2002). Conquista e ocupação da Amazônia: a fronteira Norte do Brasil. Estudos Avançados, 16(45), 31-50. Disponível em: https://www.scielo.br/j/ea/a/8Y9Y8Y9Y8Y9Y8Y9Y8Y9Y8Y9Y/

GALUCH, M. V.; DA COSTA, A. M. Trajetórias e (des) ocupação: migrações, conflitos e expansão da fronteira agropecuária no sul do amazonas. Amazônica-Revista de Antropologia, v. 15, n. 1, p. 72-95, 2023.

GATTI, L. V., ET AL. (2021). O papel da Amazônia na regulação climática global. Journal of Climate Science, 12(4), 350-364.

GRAF, H. P. et al. Parallel Support Vector Machines: The Cascade SVM, 2004.

HOSSAIN, M. I. Support Vector Machine, 2022.

INTRODUÇÃO ÀS REDES NEURAIS CONVOLUCIONAIS. In: DEEP LEARNING BOOK. 2022. Disponível em: https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais. Acesso em: 25 ago. 2024.

LEE, C. S.; CHEANG, P. Y. S.; MOSLEHPOUR, M. Predictive Analytics in Business Analytics: Decision Tree. Advances in Decision Sciences, , v. 26, n. 1, p. 1–30, 2022.

LOPES, M. V. R. Tratamento de Imprecisão na Geração de Árvores de Decisão, 2016.

MAGGIORI, E., TARABALKA, Y., CHARPIAT, G., & ALLIEZ, P. (2017). Convolutional Neural Networks for large-scale remote-sensing image classification. IEEE Transactions on Geoscience and Remote Sensing, 55(2), 645-657. https://doi.org/10.1109/TGRS.2016.2612821

[1]https://dataspace.copernicus.eu/

MEDEIROS, I. N.; TSEN, J. M.; FAVAN, J. R. Deep Learning Para Identificação e Classificação em Imagens de Manguezais, 2021.

MIRANDA, C., PERES, C. A., & CARVALHO, L. (2019). Governança fundiária e sustentabilidade na Amazônia: dilemas e perspectivas. Estudos Avançados, 33(96), 175-195. Disponível em: https://www.scielo.br/j/ea/a/fundiaria-amazonia

MIRANDA, E., ET AL. (2019). Conflitos fundiários e desenvolvimento na Amazônia Legal. Estudos Sociais, 22(2), 202-221.

NEPSTAD, D., ET AL. (2008). Amazônia: proteção e desenvolvimento sustentável. Rio de Janeiro: Editora UFRJ.

NOWOZIN, S. et al. Decision tree fields. Barcelona, Spain. International Conference on Computer Vision, 2011. Disponível em: http://ieeexplore.ieee.org/document/6126429/. Acesso em: 28 ago. 2024.

NOWOZIN, S., VEZHNEVETS, A. S., & LAMPERT, C. H. (2011). On parameter learning in CRF-based approaches to object class image segmentation. Proceedings of the 2011 International Conference on Machine Learning (ICML), Bellevue, WA.

OLOFSSON, P., FOODY, G. M., HEROLD, M., STEHMAN, S. V., WOODCOCK, C. E., & WULDER, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. Remote sensing of Environment, 148, 42-57.

PAL, M., & MATHER, P. M. (2005). Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 26(5), 1007–1011. https://doi.org/10.1080/01431160512331314083

PEIXOTO, F. (2009). Linha do tempo: Entenda como ocorreu a ocupação da Amazônia. Imazon. Disponível em: https://imazon.org.br/imprensa/linha-do-tempo-entenda-como-ocorreu-a-ocupacao-da-amazonia/

QUINLAN, J. R. Induction of decision trees. Machine Learning, v. 1, n. 1, p. 81–106, 1986.

RIGATTI, S. J. Random Forest, p. 9, 2016.

Salisbury, J., et al. (2012). Ecologia da Amazônia: uma visão geral. São Paulo: Editora ABC.

SANTOS, A. L., & OLIVEIRA, M. R. (2021). Regularização fundiária e reforma agrária no Brasil: desafios e perspectivas. Revista de Direito Agrário e Desenvolvimento Rural, 15(1), 45-67.

SENS, D. F. Rede Neural Convolucional para classificação de imagens SAR de desflorestamento na Amazônia, 2021.

SHARMA, H.; KUMAR, S. A Survey on Decision Tree Algorithms of Classification in Data Mining. International Journal of Science and Research (IJSR), v. 5, n. 4, 2016.

SILVA, J. C. D. Rede Neural Convolucional Para o Diagnóstico de Rolamentos em Baixa Rotação, 2024.

[1]https://dataspace.copernicus.eu/