Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Machine Learning-Based Deforestation Detection in the Amazon Rainforest: A Data-Driven Comparative Study

Júlia R. M. do Nascimento[1], Wallace Casaca[2]

Instituto de Biociências, Letras e Ciências Exatas (IBILCE)/UNESP, São José do Rio Preto, SP

The preservation of the Amazon Rainforest, a crucial ecosystem for global climate regulation, requires effective and systematic monitoring of deforestation activities. In this context, machine learning (ML) algorithms offer robust, well-established tools for detecting and mapping deforestation occurrences from remote sensing (RS) imagery. In fact, by combining advanced ML-based techniques with RS technologies, one can monitor and analyze vast amounts of forest areas, identifying potential changes in land cover over time while still facilitating conservation measures and sustainable land management practices.

Considering the above context, in this work we conduct a comparative analysis of three machine learning-based frameworks: *Isolation Forest* (IF), *One-Class Support Vector Machine* (OC-SVM) and *Convolutional Neural Network* (CNN). More precisely, by integrating remote sensing data collections, image processing techniques, and machine learning algorithms, we adapt and refine the aforementioned machine intelligent approaches for the task of deforestation detection in the municipality of São Félix do Xingu, Pará, Brazil: one of the largest municipalities in the world in terms of native land area.

In order to computationally detect deforested areas, our image classification methodology relies on data from the *Google Earth Engine* (GEE) [2], an immersive and functional platform renowned for its extensive geospatial data catalog and cloud-based capabilities. More specifically, we collect aerial images from the GEE Landsat-8 catalog, which provides high-resolution imagery across various spectra. Images from three different years (2014, 2020 and 2021) were taken and then processed by applying masks and spectral indices, including the *Normalized Difference Vegetation Index* (NDVI) [1], to enhance the image quality assessments.

Regarding the specific classification methods taken in our comparative evaluation, we employ the following approaches: (a) *Isolation Forest* (IF) [1], which detects anomalies in high-dimensional data by applying binary trees to isolate data points deviating from normal patterns; (b) *One-Class Support Vector Machine* (OC-SVM) [4], which classifies specific samples while rejecting others and; (c) *Convolutional Neural Network* (CNN) [3], which is a neuron-driven architecture that utilizes convolution and pooling layers to extract relevant features while reducing dimensionality. Hyperparameter tuning was also carried out to improve the performance of all machine learning algorithms, ensuring optimal settings for each model's parameters.

Aiming to identify the most suitable machine intelligent approach for detecting deforestation in the studied portion of the Brazilian Amazon rainforest, we first train the classification models using a Remote Sensing dataset of 3.100 images, and subsequently test their performance on a separate subset of 310 images. In terms of labeling, these images include both deforested and non-deforested areas.

[1]julia.rm.nascimento@unesp.br
[2]wallace.casaca@unesp.br

Tabela 1: Quantitative evaluation metrics for the machine learning methods.

Method	Accuracy	Specificity	Sensitivity	Precision	F1-score
IF	0.8161	0.6896	0.8292	0.9628	0.8910
OC-SVM	0.8710	1.0000	0.8576	1.0000	0.9233
CNN	0.9935	0.9655	0.9964	0.9964	0.9964

Table 1 provides an overview of the quantitative evaluation measures for all the machine learning methods when they are employed on the testing subset. The tabulated scores reveal that the CNN-based framework outperformed both the IF and OC-SVM methods in terms of accuracy, achieving a remarkable score of 0.9935. Although IF demonstrated competitive performance, especially in precision (0.9628) and F1-score (0.8910), it fell short of CNN. Similarly, OC-SVM showed a satisfactory accuracy (0.8710) and the best specificity, but its performance in other metrics was relatively lower compared to CNN. In summary, the results highlight the superiority of CNN in deforestation detection, showcasing its potential as a robust and reliable approach in environmental monitoring applications.

In this work, a comparative analysis of three ML-based frameworks was performed by implementing and adapting them for the deforestation detection application in the São Félix do Xingu region: an Amazon rainforest portion located in the state of Pará, Brazil. The obtained results indicated that the CNN-based framework outperformed the others in terms of accuracy, sensitivity, and F1-score, emerging as the most effective approach for deforestation detection in the examined region. Overall, these findings underscore the potential of machine learning techniques in enhancing deforestation monitoring efforts, thereby contributing to the conservation and sustainable management of the Amazon Rainforest ecosystem.

Acknowledgment

The authors would like to thank the São Paulo Research Foundation (FAPESP - #2013/07375-0, #2022/13665-0 and #2023/14427-8) and the Brazilian National Council for Scientific and Technological Development (CNPq – #316228/2021-4) for providing funding for this research.

Referências

[1] V. L. S. Gino, R. G. Negri, F. N. Souza, E. A. Silva, A. Bressane, T. S. G. Mendes e W. Casaca. “Integrating unsupervised machine intelligence and anomaly detection for spatio-temporal dynamic mapping using remote sensing image series”. Em: **Sustainability** 15.6 (2023), p. 4725. DOI: 10.3390/su15064725.

[2] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau e R. Moore. “Google Earth Engine: Planetary-scale geospatial analysis for everyone”. Em: **Remote sensing of Environment** 202 (2017), pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

[3] T. Kattenborn, J. Leitloff, F. Schiefer e S. Hinz. “Review on Convolutional Neural Networks (CNN) in vegetation remote sensing”. Em: **ISPRS Journal of Photogrammetry and Remote Sensing** 173 (2021), pp. 24–49. DOI: 10.1016/j.isprsjprs.2020.12.010.

[4] A. E. O. Luz, R. G. Negri, K. G. Massi, M. Colnago, E. A. Silva e W. Casaca. “Mapping fire susceptibility in the Brazilian Amazon forests using multitemporal remote sensing and time-varying unsupervised anomaly detection”. Em: **Remote Sensing** 14.10 (2022), p. 2429. DOI: 10.3390/rs14102429.