ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-3/W4-2025
Latin American GRSS and ISPRS Remote Sensing Conference (LAGIRS) 2025, 10–13 November 2025, Foz do Iguaçu, Paraná, Brazil

HVAS: Detection of Vegetation Height near Electric Transmission Lines Using Deep Learning and Satellite Images

Jhon Majin Erazo[1], Jose Duenas Salazar[1], Richard Anthony Gomez[1] Jorge Benitez Caceres[1],
Flávio Grabiel Oliveira Barbosa[1], Santiago Duranti Piovesan[1], Mateus Andre Favretto[1], Paulo Alberto Violada[1],
Bruno Pereira da Costa[2], Marina de Siqueira[2],
Carlos Nascimento[3], Lucas Souza[3], Antonio Donadon[3]

[1] SENAI Institute of Innovation in Embedded Systems, Florianópolis, SC, Brazil - jhon.erazo@sc.senai.br
[2] SENAI Institute of Innovation in Renewable Energy, Natal, RN, Brazil - brunocosta@isi-er.com.br
[3] Minas Gerais Energy Company, Belo Horizonte, MG, Brazil - antonio.donadon@berkan.com.br

Keywords: Vegetation height, Deep Learning, Satellite images, Electric transmission lines, Semantic segmentation, Remote sensing.

Abstract

Vegetation encroachment near power transmission lines poses risks such as outages, fires, and increased maintenance costs. This study presents a method for estimating vegetation height using public satellite imagery and convolutional neural networks (CNNs). The approach involves segmenting dense vegetation areas, calculating height, and a new method for generating georeferenced alerts for risk analysis. Height estimates using Sentinel-2 and GEDI data showed a root mean square error (RMSE) of 7.7 meters and a mean absolute error (MAE) of 5 meters compared to high-resolution LiDAR data. The results demonstrate the originality of this study by identifying risk regions associated with the presence of vegetation near to power transmission lines, contributing to the improvement of vegetation management using public data in the electricity sector.

1. Introduction

Given the significant economic impact of energy losses in Brazil, totaling R$9.97 billion in 2023 according to (Agência Nacional de Energia Elétrica (ANEEL), 2025); the role of vegetation in causing short circuits and grid failures becomes a critical concern. In this context, efficient monitoring of tall vegetation near transmission lines is essential for preventive maintenance and operational safety.

To address this, vegetation mapping plays a key role by enabling the identification and characterization of risky areas. However, traditional methods such as field surveys or drone-based inspections, while accurate, are often limited by high operational costs and reduced spatial coverage. Similarly, LiDAR systems provide precise vegetation height estimation but are not viable for large-scale routine use due to cost constraints. Although commercial satellite imagery offers high-resolution data, its accessibility is often limited. Therefore, the use of free and publicly available satellite imagery—such as that provided by Sentinel-2A emerges as a scalable and economically feasible alternative, offering adequate spatial resolution and spectral range for large-scale vegetation monitoring in critical infrastructure environments.

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have significantly improved the ability to extract complex patterns from satellite data and automate vegetation classification and height estimation. Tree height is a critical indicator associated to biomass, carbon storage, and risk factors such as wildfire susceptibility and interference with human infrastructure, including power transmission lines. In that context, this study introduces HVAS (High Vegetation Alert System): a CNN-based framework for vegetation management that uses Sentinel-2A imagery to detect tall vegetation near high-voltage power lines in a region of the state of Minas Gerais in Brazil (see Figure 1).

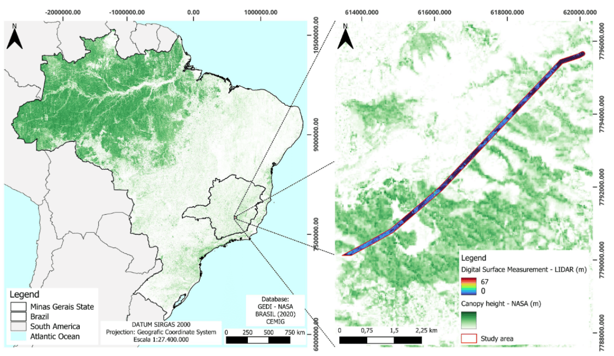

Figure 1. Study area in the western region of Minas Gerais, Brazil.

The system integrates vegetation density segmentation using the U-Net architecture (five different backbone architectures were tested), and height estimation through a U-Net model based on XceptionS2 and GEDI (Global Ecosystem Dynamics Investigation) data. By combining both outputs, HVAS identifies areas where vegetation exceeds 20 meters, providing a low-cost and scalable tool for infrastructure risk management. The proposed approach marks a significant advancement over existing methods, which typically address vegetation encroachment without explicitly estimating height or generating georeferenced alarms for vegetation management.

2. Related Work

Unmanned aerial vehicles (UAVs) equipped with RGB and hyperspectral sensors, often combined with LiDAR, have demonstrated high accuracy in vegetation analysis and canopy height estimation at local scales. For example, (Berra, 2020) and (Nezami et al., 2020) used UAVs to detect individual tree canopies and classify species with notable accuracy, while (Lim et al., 2015) and (De Almeida et al., 2021) estimated canopy height using UAV-based photogrammetry and LiDAR fusion. Despite their effectiveness, these methods face significant limitations in terms of weather conditions, operational cost, logistical complexity, and spatial coverage. Thus, frequent monitoring of this type of technology in large or remote areas become financially and practically unsustainable.

Commercial satellite imagery offers another alternative, with high-resolution sensors. Constellations such as GeoEye-1, planetScope and Maxar Vivid2 delivering detailed data suitable for canopy height modeling. For instance, in (Goldbergs, 2021) and (Tolan et al., 2024) the authors utilize imagery from these constellations, achieving height estimation errors of less than three meters when validated against LiDAR data. The high cost and limited geographic coverage of commercial satellite imagery pose significant constraints on its applicability for large-scale or continuous monitoring efforts (Sesnie et al., 2023).

In contrast, public satellite imagery such as Sentinel-2 present a low-cost alternative for large-scale vegetation monitoring. With 10–20 m spatial resolution and frequent revisit intervals (5 days), Sentinel-2 imagery is particularly suited for scalable applications. When integrated with deep learning models and GEDI data, these images can produce reliable canopy height estimates over extensive geographic areas. Studies such as (Perez et al., 2022), (Martins et al., 2024), and (Lang et al., 2023) have shown that U-Net-based models trained on Sentinel-2 and GEDI data can achieve strong performance in predicting tree height at regional and continental scales. Furthermore, (Torres de Almeida et al., 2022) successfully applied Sentinel imagery combined with LiDAR to detect tall vegetation near power transmission corridors highlighting its operational relevance for infrastructure safety.

In this paper, we propose the High Vegetation Alert System (HVAS), a novel framework that integrates two parallel U-Net-based models: one (model1, M1) for detecting regions of dense vegetation and another (model2, M2) for estimating vegetation height within power line corridors. By combining these outputs, HVAS enables accurate detection of high-risk areas where tall vegetation may interfere with infrastructure, offering a scalable and operational solution for preventive maintenance in the electrical power sector. To the best of our knowledge, this is among the first studies to perform large-scale vegetation height estimation specifically focused on power line rights-of-way.

3. Materials and method

This section presents the steps used for dataset processing and the procedures for training the CNNs models and proposed framework HVAS.

3.1 Dataset

The region of Minas Gerais, Brazil has a great diversity of vegetation due to its diversified geographical structure, composed of different types of land. This provides a favorable research context for the exploration of the height of tree groups. The study area included various types of vegetation, such as Cerrado forests, Atlantic Forest, pastures and agricultural cultivation areas. This study was carried out in the western region of Minas Gerais, as illustrated in Figure 1.

3.1.1 Training

The Sentinel-2 mission, part of the Copernicus program, provides multi-spectral images for monitoring vegetation, land use, and water quality. Composed of two satellites (Sentinel-2A and 2B), it captures data across 13 bands with spatial resolutions ranging from 10 to 60 meters, supporting environmental monitoring and disaster management.

For training the M1 model, three Sentinel-2A tiles were selected within the state of Minas Gerais, Brazil, covering distinct geographic areas and two different seasonal periods: the rainy season and summer. Only images with less than 50% cloud cover were used. Together, these tiles cover approximately $36,168.12\ \mathrm{km}^2$, ensuring a broader representation of vegetation types such as croplands, pasture, bare soil, and both high and low vegetation. From the three tiles, each of size (10980×10980) pixels at 10-m resolution, 1590 patches of $(512 \times 512 \times 3)$ were generated using the B4 (red), B3 (green), and B2 (blue) bands. The tile locations are shown in Figure 2.

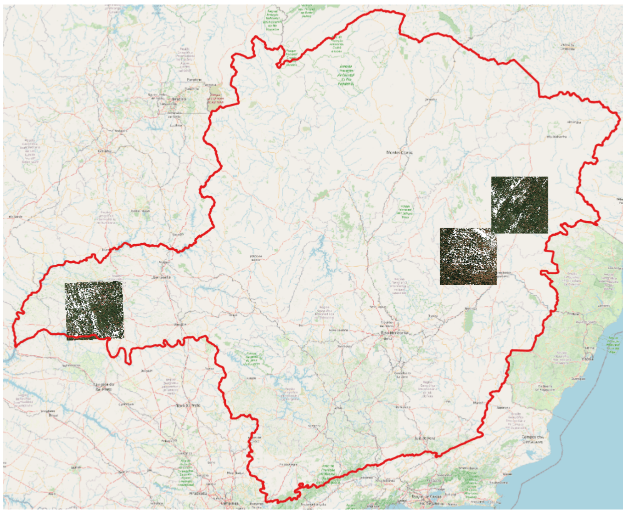

Figure 2. Detailed map of the image dataset used for delimitation/vectorization. Three tiles were used in different general mining regions.

To segment vegetation density, three classes were labeled: dense vegetation, clouds, and cloud shadows. This approach aimed to reduce the model's sensitivity to pixels lacking vegetation information, even after applying the Copernicus cloud mask. These categories were chosen because they occur most frequently in the selected images, with large areas covered by clouds and their shadows cast over regions with dense vegetation. Pixels corresponding to bare ground, bodies of water, or urban areas were excluded, as they do not contribute to the vegetation density estimate. Labeling was approached through the following strategies: 1) Labeling clouds and cloud shadows as separate classes to avoid misclassification as vegetation; 2) Using the Copernicus cloud mask as a pre-filter; 3) Manual review of suspicious samples or regions; 4) Exclusion of highly ambiguous regions between classes in the dataset.

To enhance vegetation texture in satellite images, a linear redistribution of pixel values between 2% and 88% of the tonal range was applied to each spectral band across all tiles. Additionally, Sentinel-2 reflectance data were normalized from their original 0–10,000 range to 0–1, focusing on typical forest reflectance values. These preprocessing steps improved the contrast and contributed to better segmentation performance by the M1 model.

https://doi.org/10.5194/isprs-annals-X-3-W4-2025-227-2026 | © Author(s) 2026. CC BY 4.0 License.

3.1.2 Evaluation dataset To compare the height predictions of our approach, we used a dataset collected in the field from different power transmission lines using a LiDAR. These data are composed of Digital Terrain Model (MDT) and Digital Surface Model (MDS) and orthophotos of the respective mapped areas, with a spatial resolution of 0.5 m per pixel. The LiDAR data were acquired in the months of 12/12/24 and 11/12/24 during winter and summer periods, where vegetation has changes in its shape and texture. These measurements include, among others, tree height and density. To compare the heights predicted by our models about to the LiDAR data, we used the mean value (centroid) of the $0.5m \times 0.5m$ pixels that intercept the values of the prediction with a resolution of 10 meters (satellite image used). For this purpose, operations on the MDS and MDT were used to identify the roughness and texture of the objects captured by the LiDAR, including trees, rocks, structures, houses, etc. Roughness is the degree of irregularity of the surface. It is calculated by the largest difference between a central pixel cell and its surrounding cell. The determination of roughness plays a role in the analysis of terrain elevation data, useful for calculations of morphology and physical geography in general. Finally, a filter is generated that aims to identify and/or separate areas with and without vegetation. These data are multiplied to focus only on vegetated regions.

3.2 Evaluation metrics

The metrics precision, accuracy, F1-Score, and Intersection over Union (IoU) were used to evaluate the performance of the semantic segmentation models. Precision and accuracy assess the correctness of pixel classification, while F1-Score measures the similarity between predicted and reference masks. IoU, a key metric, evaluates both the classification correctness and spatial alignment between the predicted and ground truth masks (Olczak et al., 2021).

To evaluate the predictions of the vegetation height model on the validation datasets, several metrics were used, including root mean square error ($RMSE$), mean error (MAE), coefficient of determination (R^2) and coefficient of concordance (d), which measures the correlation between predicted heights and actual height values (Chicco et al., 2021).

3.3 Proposed framework HVAS

Figure 3 shows the proposed HVAS for vegetation identification and height calculation near power transmission lines. There are three main stages in this HVAS: segmentation of regions with high vegetation density, calculation of vegetation height and generation of georeferenced alerts on possible risks.

3.3.1 Vegetation density segmentation In this work, regions with high vegetation density and tree height were segmented using the U-Net architecture (Ronneberger et al., 2015), known for its encoder-decoder "U"-shaped structure. A key feature of U-Net is the skip connections between encoding and decoding layers, which preserve spatial details during reconstruction. The encoder applies repeated blocks of 3×3 convolutions followed by ReLU activations and 2×2 max-pooling, doubling the feature channels at each step. The decoder performs bilinear upsampling, concatenates with corresponding encoder features, and applies 3×3 convolutions with ReLU, halving the feature channels at each stage. The final layer uses a 1×1 convolution to produce the segmentation output with the original spatial dimensions.

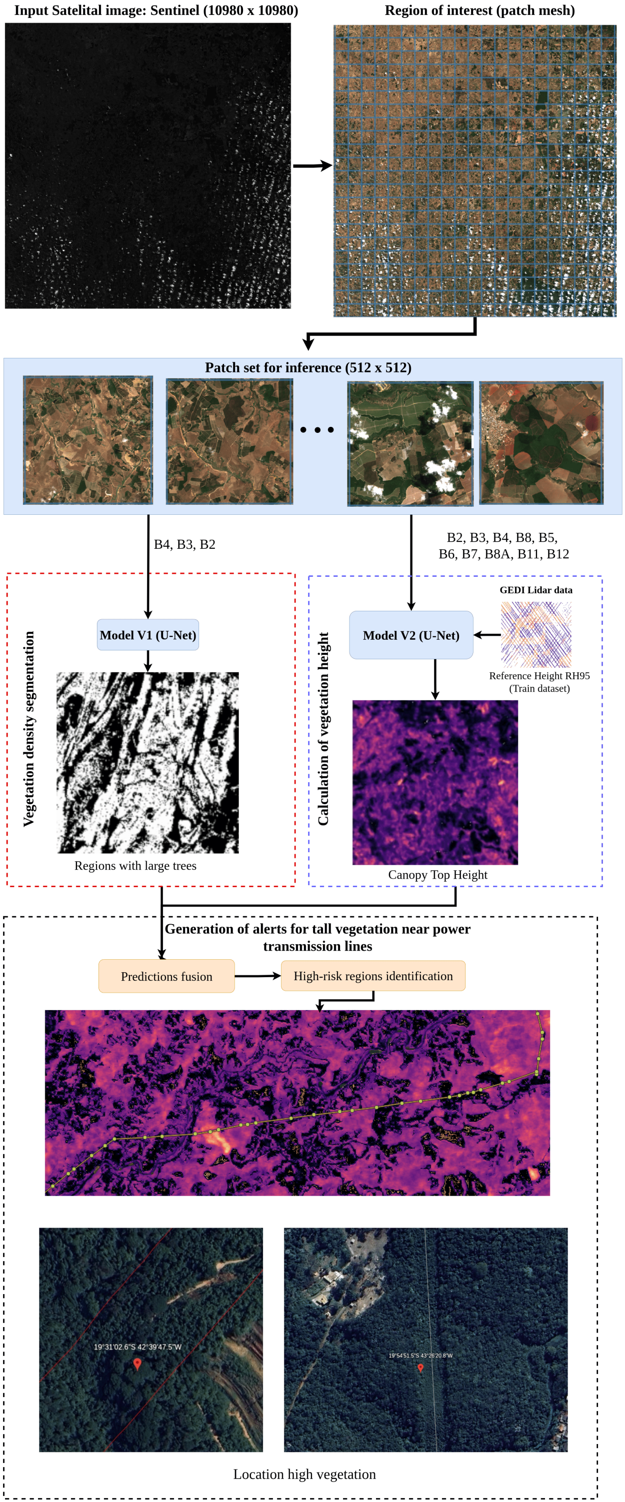

Figure 3. The proposed HVAS for calculation and identification of tall vegetation near electric transmission lines.

3.3.2 Calculation of vegetation height In this work, the U-Net model with the XceptionS2 architecture, previously developed and validated by (Lang et al., 2023), was utilized. This model was chosen for its proven effectiveness in the reference study, where various deep learning-based models were trained to produce a global canopy height map with a spatial resolution of 10 meters. The model training integrated data from dedicated space missions, such as height measurements provided by

GEDI, alongside optical satellite imagery from Sentinel-2.

To validate its applicability, the training configuration and key hyperparameters defined in the reference work—such as input image size, data normalization, and the order and number of spectral bands—were reused. However, in this study, the model was adapted to estimate vegetation height specifically in regions with high vegetation density, aiming to reduce the occurrence of false positives caused by instances not representing tall vegetation. A comprehensive visualization of the proposed model's complete pipeline is provided in Figure 4.

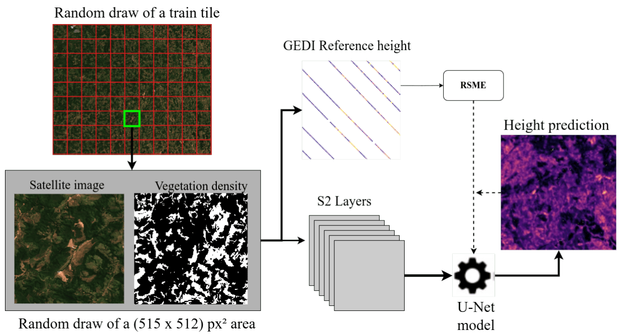

Figure 4. Overview of the vegetation height model using GEDI Lidar reference data and Sentinel2 L2A imagery over the state of Mina Gerais Brazil.

3.3.3 Generation of georeferenced alerts on possible risks

In this paper, we present an algorithm designed to automatically identify critical regions near power transmission lines, using predictions generated by models (heat map and semantic segmentation). The main objective is to determine the accurately the exact coordinates of these critical areas, allowing the personnel in charge to verify the alerts in the field and take the necessary corrective measures. Figure 5 shows, in general, the workflow of the proposed pipeline.

The pipeline begins with the processing of the tile with a structure of $[12 \times 10980 \times 10980]$, which is divided into patches of 512×512 pixels. Each of these patches is used as input for the two models, obtaining as outputs a prediction related to the density and another with the calculation of the height of the vegetation. Subsequently, these predictions are merged and the final prediction is reconstructed with the original size of the input tile. Next, the coordinates of the transmission lines are used to delimit the region of interest, analyzing exclusively the vegetation with heights greater than 20 meters within this area. Finally, alerts are generated based on the surface of the identified regions and their georeferenced location.

4. Experimental results

This section provides an analysis of the experiments conducted in the study. First, Section 4.1 discusses the results of the vegetation density segmentation model. Section 4.2 presents the outcomes of the vegetation height estimation. Finally, Section 4.3 analyzes the combined performance of the two cascaded models using in situ LiDAR data to detect tall vegetation within power transmission line corridors. All models were implemented and trained on an High Performance Computing (HPC) with a Linux Ubuntu 18.04 operating system, a 32 GB Nvidia A100 DGX video card and 1 TB of RAM.

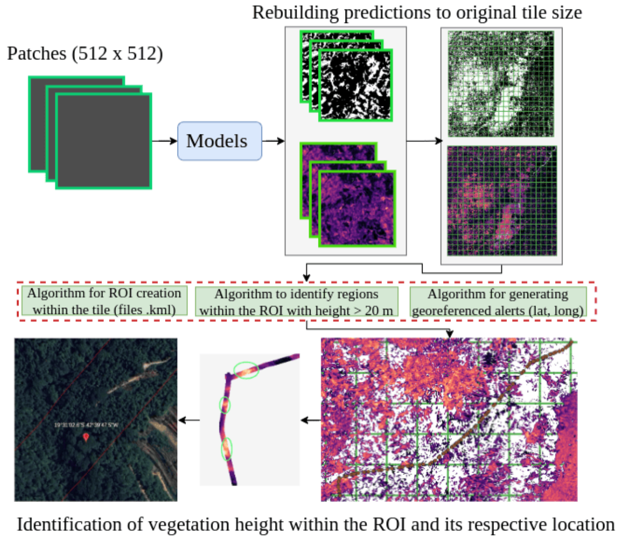

Figure 5. General pipeline for generating high vegetation alerts.

4.1 Model 1: vegetation density segmentation

CNN-based models typically require large datasets for effective generalization. To address this limitation, transfer learning with pre-trained ImageNet weights was employed, reducing training time and the risk of overfitting. Additionally, data augmentation techniques—such as flipping, scaling, and brightness adjustments—were applied to enhance the training set and improve model robustness.

Five backbone architectures ResNet34, ResNet50, EfficientNetV2, InceptionV3, and Seresnext50 were compared to evaluate their ability to learn vegetation characteristics from the reference dataset. The models were trained for 100 epochs with a batch size of 4, a fixed learning rate of 0.001, and a momentum of 0.9, using the Adam optimizer (Kingma, 2014). Training data were divided into 70% for training (1113 images), 20% for validation (318 images), and 10% for testing (159 images), with an input size of 512×512 pixels. Early stopping was applied to avoid overfitting. This configuration proved effective for the segmentation and interpretation of satellite images, supporting the analysis of high vegetation density regions. Table 1 presents the results of each segmentation model on the test dataset (159 images). All architectures demonstrated consistent performance, with Accuracy ranging from 0.87 to 0.92 and IoU from 0.82 to 0.85, highlighting their ability to segment dense vegetation in complex environments. The EfficientNetB7 model achieved the best results, with an IoU of 0.85 and Accuracy of 0.92, followed by ResNet-50 with 0.83 IoU and 0.89 Accuracy. Seresnext50 and InceptionV3 showed similar IoU values, although InceptionV3 performed slightly better in Accuracy. ResNet-34 obtained the lowest scores, likely due to limitations in capturing fine vegetation structures.

Backbones	Accuracy	Precision	F1-Score	Iou
Resnet-34	0.87	0.86	0.88	0.81
Resnet-50	0.89	0.89	0.90	0.83
Efficientnetb7	0.92	0.90	0.91	0.85
Inceptionv3	0.89	0.90	0.89	0.82
Seresnext50	0.88	0.88	0.87	0.83

Table 1. Vegetation density segmentation results on the test dataset.

From Table 1, it is possible to observe that Efficientnetb7 presented a high number of true positives and a reduced amount of false positives and negatives, achieving a good balance between precision (90%) and F1-Score (91%). These results are corroborated by Figure 6, which shows the loss curves during the training phase for all evaluated architectures. As illustrated, the Efficientnetb7 curve presents the low values (0.14).

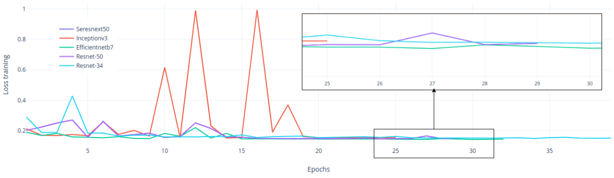

Figure 6. Evolution of IoU over training epochs.

Furthermore, Figure 7 illustrates the predictions obtained by EfficientNet-B7 on selected test images. The results show an average IoU exceeding 70%, even in challenging scenarios that include grass, crops, roads, clouds, and shadows. These findings confirm the model's robustness and capacity to produce segmentation masks closely aligned with the ground truth (GT).

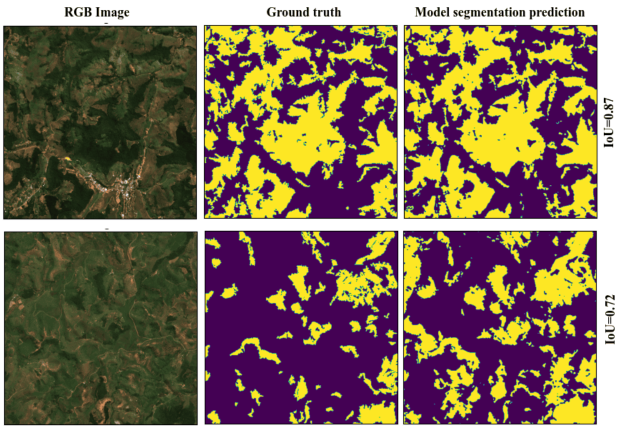

Figure 7. The V1 model's predictions for vegetation density segmentation were compared against expert-annotated ground truth, with IoU scores used to evaluate performance on the test images.

4.2 Model 2: Calculation of vegetation height

The model was retrained to focus on regions with high vegetation density by processing multichannel Sentinel-2 images and using expert-generated vegetation masks. Patches of 512×512 meters were sampled and paired with GEDI RH95 rasterized heights. The U-Net model was trained using the MAE metric as a loss function, considering only pixels with valid RH95 values. The training flow is detailed in Figure 4.

Satellite and LiDAR data were normalized following (Lang et al., 2023). The model was trained for 100 epochs with a batch size of 32, using the Adam optimizer and an initial learning rate of 0.0001, adjusted progressively to enhance convergence. Figure 8 presents a qualitative evaluation of the predictions from the vegetation density and height estimation models near power transmission lines. Despite the moderate resolution of satellite images (10 meters) and the complexity of the terrain, the models effectively segment high-density vegetation areas and estimate their heights, even in regions with diverse topography and vegetation.

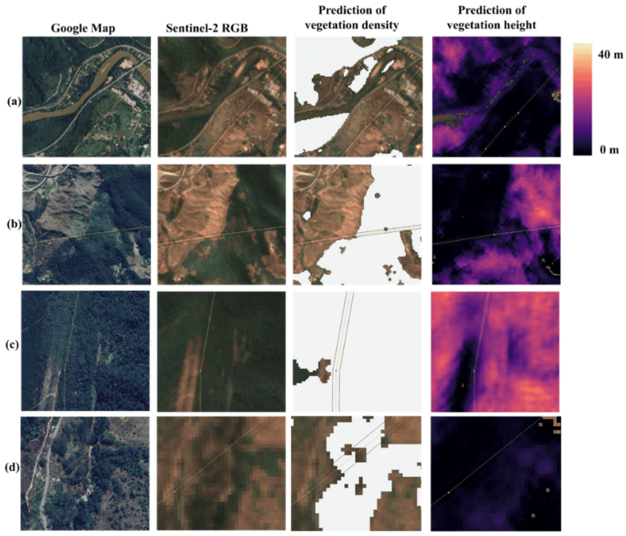

Figure 8. Model predictions for vegetation density and height near power transmission lines.

From Figure 8(c), it can also be noted that there are regions where the vegetation density model generates an erroneous mask in a region where there is exposed soil. However, due to the good training of the vegetation height model, this area is ignored. These results highlight the robustness of the approach used, evidencing the ability to identify and monitor critical vegetation zones in complex environments.

4.3 Evaluation with independent dataset captured in situ

In order to compare the results generated by HVAS for vegetation height estimation, a longitudinal profile of a power transmission line located in the municipality of Nova Lima, MG, was used. Based on this profile, graphs were generated representing the altitudes from the MDT, the MDS, and the predictions obtained by the models.

To correlate the LiDAR data (0.5 m) with the Sentinel-2 imagery (10 m), a grid representing each resolution was created. In Figure 9, the red cells (larger) correspond to the size of a Sentinel-2 pixel (10 m), while the black cells (smaller) correspond to the LiDAR pixels (0.5 m). To calculate the error, the following strategy was adopted:

- The centroid of each red cell (10 m) was determined—see Figure B, red dots;
- For each Sentinel-2 centroid, the corresponding pixel in the LiDAR grid (0.5 m) was identified—black dots;
- The height value was extracted at both resolutions at the same point (x, y);
- The difference between the heights obtained by LiDAR and Sentinel-2 was calculated, resulting in the error for each position.

Figure 10 shows a comparison between LiDAR data and the predictions generated by our HVAS models. The top panel displays a color-coded vegetation height map, highlighting height variations along the transect. The bottom panel shows the comparison of the height curves predicted by HVAS against the reference LiDAR data.

https://doi.org/10.5194/isprs-annals-X-3-W4-2025-227-2026

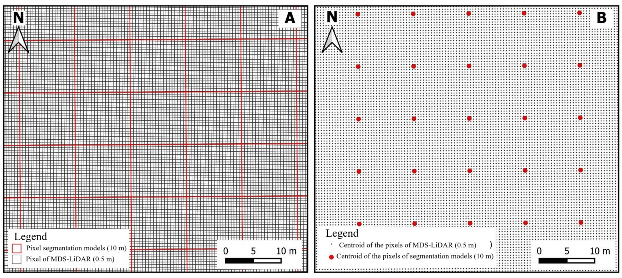

Figure 9. Visual details of the difference between the spatial resolution of two prediction dice of two models (10 m) and two LiDAR dice (0.5 m).

In general, the curves from the MDS and MDT show a similar behavior to the curves produced by the models of the HVAS, with a close correspondence in most sections of the terrain. This suggests that the models are able to capture general patterns of vegetation height in the areas corresponding to the power transmission lines. However, discrepancies are observed in specific areas, especially in areas with abrupt changes in topography or dense vegetation. In the longitudinal profiles presented, light green represents the prediction results of the HVAS models (10 m), while dark green corresponds to the MDS data with a resolution of 0.5 m.

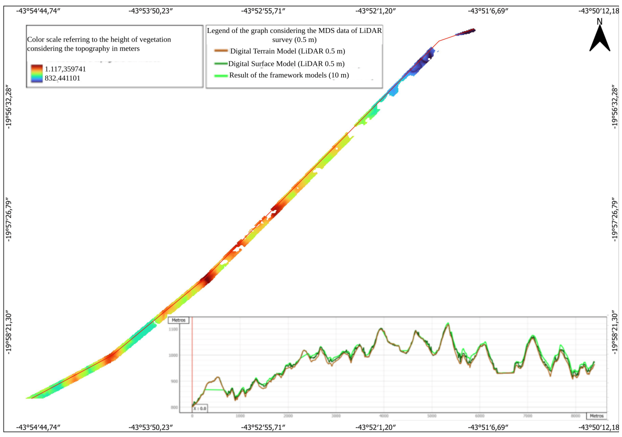

Figure 10. Geospatial comparison of HVAS vegetation height predictions and LiDAR data for the Barreiro–Taquaril transmission line section (Nova Lima, MG).

A comparison between the HVAS model predictions and the reference data reveals a similarity, although slight discrepancies are observed, with the predicted vegetation heights sometimes appearing relatively higher. Figure 11 shows an analysis of the similarity behavior between the HVAS predictions and the reference data: in (A), the vegetation height predicted by the model (light green line) shows a rectilinear trend relative to the topography (brown line). In (B) and (C), the profiles are very similar, reflecting lower vegetation heights in flat regions and the natural vertical development of vegetation on slopes and in valleys, respectively. In (D), the profiles are also similar, representing higher altitudes with lower vegetation growth.

Evaluating the average error values, we identified an overestimation of approximately 5 m in vegetation height presented by the HVAS models about to the MDS. This effect may be related to the difference in spatial resolution of the data used: while one pixel in the satellite image corresponds to 10 m (red box), each

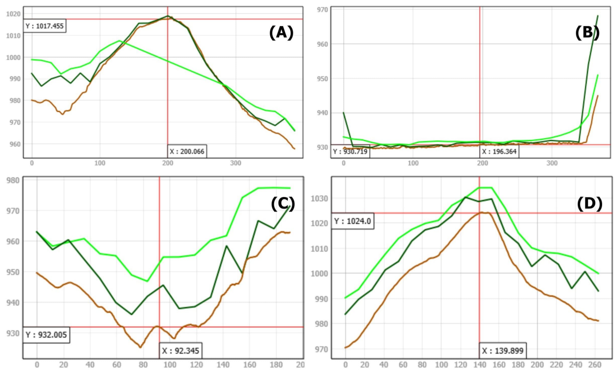

Figure 11. Graphical analysis of HVAS predictions and LiDAR data for the Barreiro–Taquaril transmission line section (Nova Lima, MG).

pixel in the LiDAR data represents an area of 0.5 m (black box). Table 2 presents the statistical analysis of the LiDAR dataset, showing a MAE of 5.04, indicating a slight overestimation of vegetation height by the model. The RMSE value of 7.71 suggests minimal deviation between the segmentation model data and the reference LiDAR values. This discrepancy is likely due to the difference in spatial resolution: satellite imagery pixels cover 10 m, while LiDAR data pixels cover 0.5 m.

Table 2. Statistical results of dataset.

Metrics	Results
(MAE) (m)	5.04
$RMSE$ (m)	7.71
Coefficient of concordance (d)	0.60
Coefficient of determination (R^2)	0.39

Willmott's coefficient of concordance (d) of 0.60 indicates moderate concordance between the datasets, suggesting significant variations. The R^2 value of 0.39 shows that 39% of the variation in the segmentation model data can be explained by LiDAR data. A Student's t-test, with a t-value of 47.934 and a p-value below $2.2e^{-16}$, revealed a statistically significant difference between the means of the predicted data and the LiDAR data ($p < 0.05$), supporting that the true difference is not zero. The 95% confidence interval (4.387 to 4.762) also confirms this result.

5. CONCLUSION

In this paper, a novel deep learning framework for detecting and estimating tree heights near power transmission lines was proposed. The HVAS integrates two sequential U-Net-based models: the first segments regions of dense vegetation, while the second estimates vegetation height only in the areas identified as high-risk. Based on the model predictions, a new georeferenced alarm generation pipeline is then applied to identify critical regions near transmission lines that require vegetation management. To support training and evaluation stages, a new image-based dataset for sentinel-L2A was developed from the region of Minas Gerais in Brazil; this dataset includes a wide variety of vegetation and diverse terrain conditions, including varying slopes and altitudes.

Among the five backbones architectures evaluated, the EfficientNetB7 demonstrated superior performance in vegetation density segmentation, achieving an Intersection over Union

(IoU) of 85% and an overall accuracy of 92%. For height estimation, the model's predictions were compared against high-resolution LiDAR data (0.5 m), resulting in a root mean square error ($RSME$) of 7.71 and a coefficient of determination (R^2) of 0.39. These results validate the potential of HVAS for real-world monitoring applications in the energy sector, particularly in identifying areas where vegetation growth may pose a threat to infrastructure. Moreover, by leveraging publicly available satellite data (sentinel-2 imagery and GEDI LIDAR data), the system provides a cost-effective and scalable solution, enabling vegetation height monitoring over large and diverse geographic areas without reliance on costly or proprietary high-resolution imagery.

The HVAS overcomes limitations observed in previous studies, focused only on detecting vegetation encroachment using high-resolution imagery, without estimating vegetation height. In contrast, our method integrates vegetation encroachment detection with height estimation, enabling the identification of critical zones in proximity to power transmission lines. By leveraging model-generated predictions, the approach facilitates precise geolocation of these high-risk areas, thereby enhancing the accuracy and utility of vegetation management strategies.

This study demonstrates the suitability of integrating deep learning models with publicly available satellite data for vegetation monitoring in critical infrastructure contexts. It points to two main directions for future research: (1) the use of transformer-based architectures, such as Vision Transformers and Swin Transformers, to enhance spatial and temporal feature modeling; and (2) the integration of multi-source satellite data, including Sentinel-1 SAR, Sentinel-2 imagery, and upcoming high-resolution missions. These advances aim to reduce measurement uncertainty, better capture vegetation dynamics, and improve predictive performance for risk assessment in the energy sector.

References

Agência Nacional de Energia Elétrica (ANEEL), 2025. Pesquisa sobre perdas - aneel. Acessado em: 16 maio 2025.

Berra, E. F., 2020. Individual tree crown detection and delineation across a woodland using leaf-on and leaf-off imagery from a UAV consumer-grade camera. *Journal of Applied Remote Sensing*, 14(3), 034501–034501.

Chicco, D., Warrens, M. J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science*, 7, e623.

De Almeida, D. R. A., Broadbent, E. N., Ferreira, M. P., Meli, P., Zambrano, A. M. A., Gorgens, E. B., Resende, A. F., de Almeida, C. T., Do Amaral, C. H., Dalla Corte, A. P. et al., 2021. Monitoring restored tropical forest diversity and structure through UAV-borne hyperspectral and lidar fusion. *Remote Sensing of Environment*, 264, 112582.

Goldbergs, G., 2021. Impact of Base-to-Height Ratio on Canopy Height Estimation Accuracy of Hemiboreal Forest Tree Species by Using Satellite and Airborne Stereo Imagery. *Remote Sensing*, 13(15), 2941.

Kingma, D. P., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.*

Lang, N., Jetz, W., Schindler, K., Wegner, J. D., 2023. A high-resolution canopy height model of the Earth. *Nature Ecology & Evolution*, 1–12.

Lim, Y. S., La, P. H., Park, J. S., Lee, M. H., Pyeon, M. W., Kim, J.-I., 2015. Calculation of tree height and canopy crown from drone images using segmentation. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 33(6), 605–614.

Martins, F., Godinho, S., Guiomar, N., Medinas, D., Rebelo, H., Segurado, P., Marques, J., 2024. Vegetation canopy height shapes bats' occupancy: a remote sensing approach. *GIScience & Remote Sensing*, 61(1), 2374150.

Nezami, S., Khoramshahi, E., Nevalainen, O., Pölönen, I., Honkavaara, E., 2020. Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks. *Remote Sensing*, 12(7), 1070.

Olczak, J., Pavlopoulos, J., Prijs, J., Ijpma, F. F., Doornberg, J. N., Lundström, C., Hedlund, J., Gordon, M., 2021. Presenting artificial intelligence, deep learning, and machine learning studies to clinicians and healthcare stakeholders: an introductory reference with a guideline and a Clinical AI Research (CAIR) checklist proposal. *Acta orthopaedica*, 92(5), 513–525.

Perez, G. G., Bourscheidt, V., Lopes, L. E., Takata, J. T., Ferreira, P. A., Boscolo, D., 2022. Use of Sentinel 2 imagery to estimate vegetation height in fragments of Atlantic Forest. *Ecological Informatics*, 69, 101680.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, Springer, 234–241.

Sesnie, S. E., Espinosa, C. I., Jara-Guerrero, A. K., Tapia-Armijos, M. F., 2023. Ensemble machine learning for mapping tree species alpha-diversity using multi-source satellite data in an Ecuadorian seasonally dry forest. *Remote Sensing*, 15(3), 583.

Tolan, J., Yang, H.-I., Nosarzewski, B., Couairon, G., Vo, H. V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., Vamaraju, J. et al., 2024. Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. *Remote Sensing of Environment*, 300, 113888.

Torres de Almeida, C., Gerente, J., Rodrigo dos Prazeres Campos, J., Caruso Gomes Junior, F., Providelo, L. A., Marchiori, G., Chen, X., 2022. Canopy Height Mapping by Sentinel 1 and 2 Satellite Images, Airborne LiDAR Data, and Machine Learning. *Remote Sensing*, 14(16), 4112.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-W4-2025-227-2026 | © Author(s) 2026. CC BY 4.0 License.