Image-based and GIS-assisted Approach for Mapping and Scoring Vegetation Encroachment in Electric Power Distribution Systems

Nelson R. Silva, Walmir Freitas,
Fernanda C. L. Trindade
School of Electrical and Computer Engineering
Universidade Estadual de Campinas
Campinas-SP, Brazil
{n241415, ferfeec}@unicamp.br

Luiz F. Ferro
ETH Zürich
Department of Information Technology and Electrical Engineering
Zürich, Switzerland
lferro@ethz.ch

***Abstract*— In electric power distribution systems, vegetation encroachment is a leading cause of power outages that can be mitigated by continuous monitoring and pruning of trees. This paper proposes an image-based methodology to detect vegetation in urban zones that may pose risks to energy distribution, using freely available satellite and street-level images. First, segmentation and filtering techniques are applied to satellite imagery to extract vegetation contours. These are then georeferenced and cross-referenced with the distribution utility's geographic information system (GIS) database to identify intersections with low-voltage and medium-voltage power lines and assign scores. Finally, street-level images are processed with a deep learning model to detect individual trees and estimate their height to refine classification. The result enables the identification of vegetation posing threats to the distribution network, enabling the implementation of preventive maintenance strategies, reducing financial losses and enhancing operational safety.**

***Index Terms*— Distribution system, image segmentation, satellite imagery, vegetation encroachment, predictive maintenance**

I. Introduction

Vegetation is commonly identified as a major factor contributing to power outages and electrical incidents. Contact between trees and electrical lines or equipment may result in short circuits, posing safety risks, causing financial losses for utilities, and interrupting vital services for customers. In 2024, vegetation accounted for 16.2% of primary causes of outages, and it can be indirectly associated with weather-related outages, which represented 19.3% of the causes [1]. Moreover, faults in distribution lines pose safety risks to the public, exposing them to potential arcing, electrical fires, and electrocution. Pruning limits the encroachment of tree branches onto energized conductors and, when done effectively, can reduce storm-related tree damage by as much as 63% [2]. However, developing strategies for continuous monitoring and preventive maintenance represents challenges due to high costs and the complexity of identifying the locations that require the highest priority in urban centers. According to the American Public Power Association, 52.9% of surveyed utilities reported that budgets for tree management were insufficient [3], reinforcing the complexity of guaranteeing reliability while keeping operational costs within budget constraints.

In this context, this work presents a proof-of-concept of an image-based method to identify vegetation encroachment in urban centers and assign a per-tree score. Although UAV- or LiDAR-based inspections may be highly accurate, they are typically harder to deploy at scale due to higher costs, operational logistics, regulatory constraints, and limited revisit frequency. To the authors' knowledge, the proposed approach is original [4], relying primarily on openly available satellite imagery to identify vegetation and assign scores, enabling scalable monitoring. Street-level images are used as an optional complementary source to estimate vegetation height and support the assessment of potential contact with distribution lines. The primary objective of this study is to develop a scalable, cost-effective screening approach for mapping and scoring vegetation encroachment, thereby supporting utility decision-making for targeted inspection and pruning.

This paper is organized as follows. In Section II, an overview of the proposed methodology is presented, followed by a detailed description of each step. The results of the methodology applied to real data are presented in Section III. The discussions and conclusions are in Section IV and Section V, respectively.

II. Methodology

The methodology starts by preprocessing satellite images and identifying vegetation contours using color segmentation. These contours are then cross-referenced with the power utility's geographic database to identify intersections with low-voltage (LV) and medium-voltage (MV) power lines. Each

This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), in part by the São Paulo Research Foundation (FAPESP) under grants #2021/11380-5 and # 2025/01669-9, and in part by the Brazilian National Council for Scientific and Technological Development (CNPq) under grant #310099/2023-4.

979-8-3315-8138-1

contour is scored according to its dimensions and the characteristics of the intersecting circuits, such as cable geometry and voltage level. To improve the scoring, the YOLOv8 (You Only Look Once, version 8.3.221) deep learning model is applied to street-level imagery to detect individual trees and estimate their height [5]. This height information refines the ranking by confirming which trees physically intersect with power lines. Satellite imagery is retrieved using the Mapbox Raster Tiles API and street-level images from Mapillary and KartaView. The proposed methodology employs a geographic database containing both geospatial and electrical information on distribution network assets, including MV and LV lines and protection devices. Although the case studies presented focus on Brazilian cities, the approach is generalizable and can be applied in any region where geographic network data are available. The results can support vegetation management strategies to prevent outages and enhance network reliability.

The following subsections present the proposed methodology applied to the feeder shown in Fig. 1, encompassing all stages from satellite image acquisition to vegetation segmentation, georeferencing, network integration, and risk assessment.

A. Tiles Retrieval

Map tiles are graphical units of a map grid served at different zoom levels, each defining a specific geographic scale. This study uses Mapbox's Raster Tiles API at zoom level 18, which offers 256×256-pixel tiles with a resolution of about 0.597 m/pixel. Tiles are requested in 512×512 format for better visualization, but keep the same effective spatial resolution. The coordinate system used is EPSG:4326 [6]. Given geographic coordinates and a zoom level, the API retrieves the corresponding map tile from a global grid that divides the Earth's surface into square tiles, each representing a specific region at the selected zoom level. The grid is indexed from the top-left corner of the map, which serves as the origin (0, 0), and each tile's position is referenced by its own upper-left corner. Based on the latitude, longitude, and zoom level, the API computes the indices of the tile that contains the requested location within this global grid.

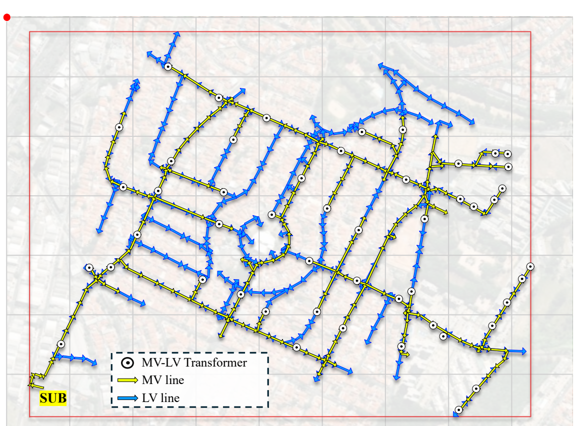

Figure 1. Study region built by merging the tiles, with the gray grid indicating their boundaries and the bounding box highlighted in red. Basemap: © Mapbox, © OpenStreetMap, © Maxar.

All LV and MV segments of each feeder are used to extract its smallest axis-aligned bounding box (AABB). The AABB coordinates are converted into tile coordinates (x, y) to determine the number of tiles needed to cover the feeder, enabling high-resolution image acquisition for the entire area. A total of 63 tiles is retrieved and merged into a single 4608x3584 pixel image covering the entire feeder (Fig. 1).

B. Vegetation Segmentation

This study adopts color-based segmentation for vegetation extraction instead of deep learning. Although YOLOv8 was trained on 300 images, with numerous trees per image and over 100 epochs, it failed to accurately detect and separate individual trees (Fig. 2). Deep learning models could improve accuracy but require extensive annotated data and high computational resources. In contrast, color-based segmentation offers efficient, consistent results with low computational cost and no annotation effort. Non-tree elements are later removed by cross-referencing with power line data. Thus, color segmentation provides a practical solution when imagery is captured during predominantly green canopy periods. Similar challenges in tree detection using deep learning models applied to high-resolution satellite imagery can also be observed in [7].

To segment the vegetation in Fig. 2, the image is first converted from RGB to HSV, which separates color from brightness and improves the distinction of green tones under variable lighting. A bilateral filter is then applied to smooth textures while preserving edges, facilitating clearer contour detection. Morphological operations (opening and closing) are applied to the extracted contours to remove small artifacts and smooth boundaries, improving vegetation delineation. The enhancement effect is illustrated in Fig. 3, which compares the contour mask before and after the operation.

The final processing step involves filtering out small contours below a minimum perimeter threshold, as these typically correspond to noise or minor vegetation elements that do not represent a risk to power lines. In contrast, large contiguous vegetation areas, where multiple trees are merged,

Figure 2. Detection of individual trees in satellite imagery using the YOLOv8 model (confidence score ≥ 0.5). Basemap: © Mapbox, © OpenStreetMap, © Maxar.

979-8-3315-8138-1

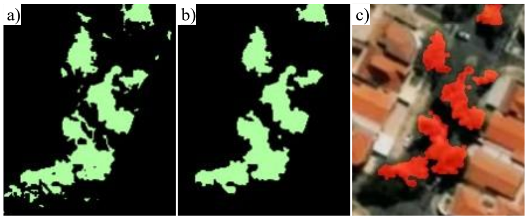

Figure 3. Vegetation mask before and after morphological processing: (a) original mask, (b) and (c) after applying morphological opening and closing. Basemap: © Mapbox, © OpenStreetMap, © Maxar.

are divided into smaller circular regions (Fig. 4). This segmentation allows each circle to be associated with height information and assigned a specific score according to the electrical circuit it intersects. The resulting mask, showing both the filtered and subdivided vegetation regions, is presented in Fig. 4 overlaid on the study area, with vegetation boundaries highlighted in green for clarity.

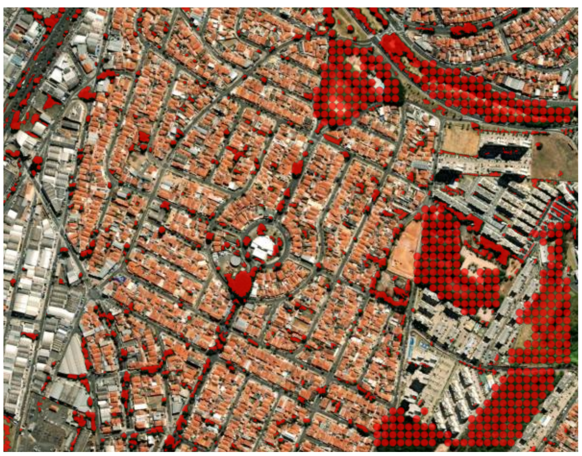

Figure 4. Segmentation results, showing the segmented vegetation areas highlighted in red. Basemap: © Mapbox, © OpenStreetMap, © Maxar.

C. Georeferencing the Contours

With the segmentation results (Fig. 4), the contours are georeferenced. For the tiles shown in Fig. 1, only their grid positions (x, y) are initially known. The geographic coordinates of each tile's top-left pixel, which also serves as its reference point, are calculated using the relationships previously defined for latitude and longitude (1) and (2). Based on the zoom level and tile size (512px), the spatial resolution S_{pixel} is determined, allowing each contour pixel to be converted from image coordinates to geographic coordinates. Each contour is then associated with the tile it lies within, enabling full georeferencing within the distribution network map.

$$S_{pixel} = 2\pi \cdot 6{,}378{,}137 \cdot \frac{\cos(latitude)}{2^{zoom} \cdot 512}$$

$$latitude_{new} \approx latitude_{ref} + (y_{pixel\_ref} - y_{pixel}) \cdot S_{pixel} \quad (1)$$

$$longitude_{new} \approx longitude_{ref} - (x_{pixel\_ref} - x_{pixel}) \cdot S_{pixel} \quad (2)$$

where, $(latitude_{ref}, longitude_{ref})$ represent the geographic coordinates of the reference point of the closest tile (e.g., red dot in Fig. 1), $(y_{pixel\_ref}, x_{pixel\_ref})$ refers to the pixels of the closest tile reference point, and (y_{pixel}, x_{pixel}) indicate the pixels coordinates to be converted.

D. Building the Oriented Network Graph

Using the geospatial data of the power lines, a directed graph is constructed to maintain the radial structure of the network (Fig. 1). In this configuration, MV circuits extend from the substation to the MV–LV transformers, while LV circuits branch from these transformers to end customers. Although only the geographic coordinates are required for the score assignment step, constructing a directed graph is essential for the subsequent analysis that estimates how many customers would be affected by an interruption caused by each identified vegetation region.

E. Cross-referencing and Tree Scoring

With the directed graph (Fig. 1) and the georeferenced vegetation contours (Fig. 4), the datasets are cross-referenced to remove segmented elements that do not intersect the network. Scores are then assigned to the remaining vegetation regions based on the voltage level of the intersected circuits and the perimeter of each contour, which reflects the extent of the detected vegetation.

To compute the outage risk score for each tree, we define a heuristic prioritization index on a 0-1 scale to highlight urgent cases for inspection and pruning. Two factors are weighted: (i) contour perimeter (proxy for vegetation size), contributing up to 0.40 as a threshold to avoid oversized contours dominating the score; and (ii) the intersected circuit type, set to 0.25 for LV, 0.35 for MV, or 0.60 when both are intersected to reflect higher potential impact. Vegetation regions intersecting only multiplexed cables are excluded due to minimal outage risk. The resulting score controls contour opacity to visually represent relative risk levels, as shown in Fig. 5.

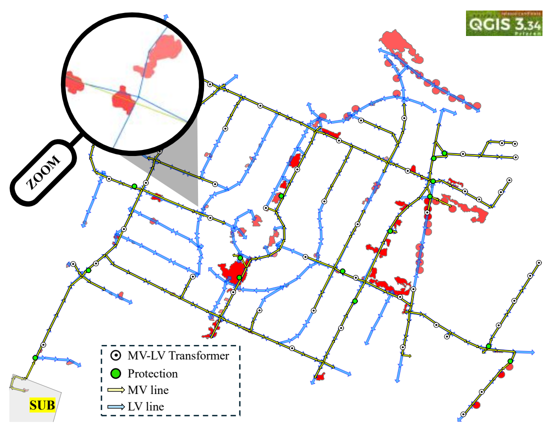

Figure 5. Scores assigned to vegetation contours (red intensity), based on perimeter and intersections with MV (yellow) and LV (blue) segments from the directed graph.

F. Number of customers possibly affected by an outage

In addition to the scores assigned to each vegetation region, a complementary analysis is conducted to estimate the number of customers affected by a fault caused by each vegetation. The process begins by calculating the number of customers

979-8-3315-8138-1

connected downstream of each protection device node (highlighted in green in Fig. 5), distinguishing four categories: LV customer units (LVCU), MV customer units (MVCU), LV generating units (LVGU), and MV generating units (MVGU). For each vegetation contour intersecting the network, the directed graph is then traversed to identify the nearest upstream protection device, which defines the network segment that would be isolated in the event of a fault. By identifying these disconnected branches, the number of affected customers can be estimated, providing both the vegetation's score and its potential outage impact.

G. *Estimating Tree Heights Through Street-level Imagery*

A YOLOv8 model is trained to detect individual trees using a dataset of 100 manually annotated street-level images from Mapillary and KartaView. To enhance variability, the dataset is expanded with data augmentation (rotations, flips, distortions, brightness/contrast adjustments, and coarse dropout to simulate occlusions), while YOLOv8's transfer-learning backbone and multi-scale training further support generalization across viewpoint, scale, illumination, and partial occlusions common in street-level scenes. Precise segmentation is unnecessary because the objective is to estimate tree height directly from detected bounding boxes. Unlike satellite imagery, where trees appear as small, dense clusters that often require pixel-level segmentation, street-level images provide object-scale cues that align with YOLOv8's detection-based design.

For each previously extracted contours centroid (Fig. 5), the Mapillary and KartaView APIs are queried to retrieve up to six nearby images within a square bounding box of approximately 30 meters. Collecting multiple images per location increases the likelihood of capturing the entire tree, enabling more accurate height estimation. Since both are public, crowdsourced platforms, street-level imagery is not available for all vegetation contours; therefore, height estimation is performed only when at least one corresponding image exists. The selected image is the one with the smallest angular difference between the camera orientation and the centroid azimuth, excluding images with partial tree visibility.

As illustrated in Fig. 6, *a* represents the tree centroid's azimuth, *c* the camera orientation relative to north, *P* the photo's viewing direction, and V the ideal viewing direction for a clear observation of the vegetation. The objective is to minimize *d*, the angular deviation between *V* and *P*, thereby selecting the image that provides the best view of the tree.

After selecting the most suitable street-level image for each vegetation contour, the trained model detects tree bounding

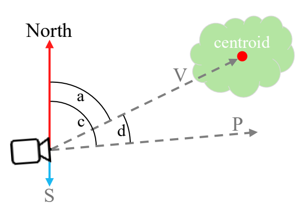

Figure 6. Illustration of how the angular difference *d* is determined.

boxes. The final detection is chosen based on both confidence score and estimated height, excluding unrealistic values that indicate incorrect distance. This combined filtering, supported by the prior image selection process, ensures more reliable height estimation. The vertical field of view (FOV) used in the calculations is obtained directly from each image's metadata, with values averaging around 80° across the dataset. As illustrated in Fig. 7, the tree height (*h*) represents a fraction of the total view height (*T*), which is determined by the camera's FOV (*r*) at a specific distance (*g*), where g corresponds to the distance between the camera location (centroid) provided by the API and the centroid of the previously segmented contour.

The vegetation height is estimated using (4), which links the pixel ratio to the actual heights between *T* and *h* in meters.

$$T_{metres} = \tan\left(\frac{r}{2}\right) \cdot g \cdot 2 \tag{3}$$

$$\frac{T_{pixels}}{h_{pixels}} = \frac{T_{meters}}{h_{meters}} \quad \therefore \quad h_{meters} = \frac{T_{meters} \cdot h_{pixels}}{T_{pixels}} \tag{4}$$

where, *g* represents the distance between the camera and the vegetation contour centroid, *r* refers to the camera's FOV. T_{pixels} denotes the picture height in pixels, and h_{pixels} is the selected bounding box height in pixels.

After estimating each tree's height, the results are compared with the standard circuit heights defined in the local power utility's technical specifications. An indicator is added to the dataset to identify whether a tree reaches the corresponding circuit height. As illustrated in Fig. 8, the opacity of the gray contours varies with their score, representing vegetation without available height information (Fig. 5). For contours with known height, opacity remains constant, as these allow verification of whether trees are in contact with the network.

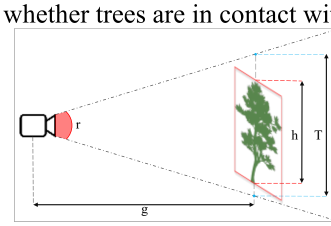

Figure 7. Estimation of tree height from street-level camera.

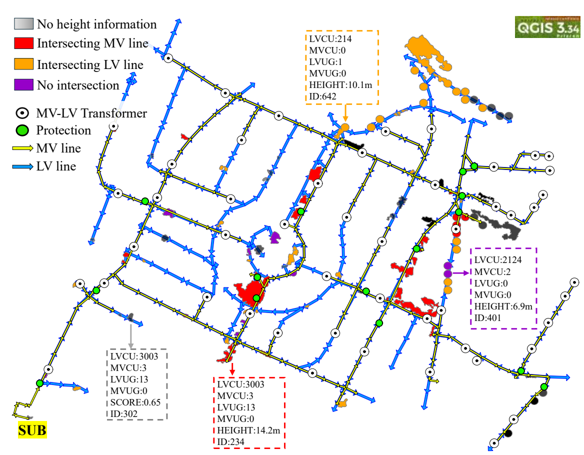

Figure 8. Results of the developed methodology.

979-8-3315-8138-1

III. RESULTS

Fig. 9 presents an additional test conducted to evaluate the performance of the proposed methodology. In this example, one contour (highlighted in pink in Fig. 9) from the analyzed feeder was selected to evaluate the segmentation performance, assigned vegetation score, number of affected units, and vegetation height. Although this vegetation was confirmed to be tall enough to reach both circuits, it received a score of 0.74, of which 0.35 corresponds to the intersection with MV segments and 0.39 to the weight assigned to its contour perimeter. In this case, the score was not increased due to contact with the LV network, as the LV segment is multiplexed. By traversing the directed graph, it was determined that a potential fault caused by this tree would interrupt service to 2 MV customers (MVCU) and 265 LV customers (LVCU). The number of affected units is explained by the presence of a protection device located immediately upstream of the vegetation site.

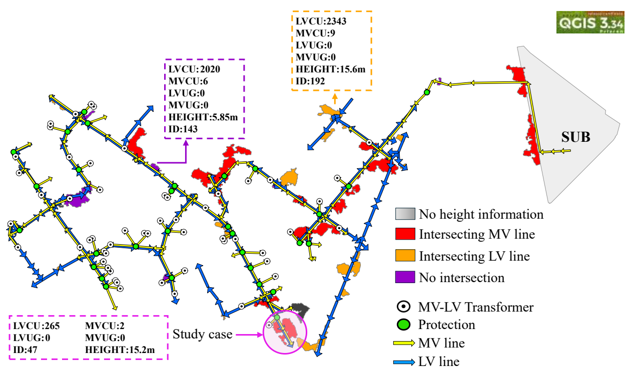

Figure 9. Results of the proposed methodology applied to an entire feeder.

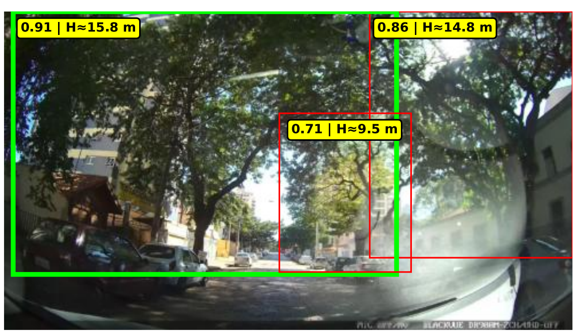

Figure 10. Tree detection bounding boxes generated by the YOLOv8 model using a Mapillary street-level image. Image source: © Mapillary, by marciosampaio, licensed under CC-BY-SA.

Using the contour's coordinates, a corresponding street-level image was retrieved from Mapillary to generate a visual representation and perform model inference. As illustrated in Fig. 10, the selected bounding box presents a high confidence score and a realistic height estimation, accurately representing the full tree structure – from the trunk base to the canopy. The estimation method determined that the targeted tree in Fig. 10 occupies 82.5% of the image's vertical pixels. Given the measured distance of 12m, and the camera's FOV of 77º (camera BlackVue DR900M-2CH), the trigonometric relation illustrated in Fig. 7 and (4) yields a tree height of 15.8m. The procedure classified this tree as in contact with both circuits, since it exceeds the 13m height of the supporting pole, according to the utility's database. Moreover, the tree shown in Fig. 10 matches the pink contour identified in Fig. 9, reinforcing the consistency of the proposed methodology.

IV. CONCLUSION

The proposed image-based and GIS-assisted methodology effectively integrates remote sensing, spatial analysis, and image processing to assess vegetation encroachment in power distribution networks. It enables the identification of vegetation position and the estimation of vegetation height, producing results consistent with visual validation. Despite inherent limitations in publicly available imagery and temporal mismatches between data sources, the methodology demonstrates scalability and adaptability across different geographic contexts. Overall, this practical framework approach advances vegetation management by fostering preventive maintenance, strengthening grid resilience, and promoting the sustainable integration of remote sensing technologies into decision-making.

Although the proposal depends on images captured during periods of active foliage, the color-based segmentation proved efficient, accurately detecting vegetation under power lines with minimal computational cost and no annotation effort, even when tree geometry was irregular or canopies overlapped. The YOLO-based height estimation also delivered reliable outcomes but could benefit from more diverse training data to better handle complex scenarios.

Future work will consist of unifying the scores into a single index, providing quantitative performance metrics, and evaluating the gains from using higher-resolution images and other contour detection models. Also, the methodology can benefit from integrating meteorological risk factors into the existing scoring framework.

REFERENCES

[1] American Public Power Association. (2024 Sample Annual Benchmarking Reliability Report). Washington, D.C., USA, Apr. 2024. [Online]. Available: https://www.publicpower.org/reliability-tracking/annual-reliability-reports

[2] M. Nazari, S. Hosseinian, E. Shakeri, and S. Nazari, "Vegetation management in power distribution networks: A review of challenges and opportunities," Sustainability, vol. 14, no. 2, Art. no. 904, Jan. 2022. [Online]. Available: https://www.mdpi.com/2071-1050/14/2/904

[3] University of Wisconsin–Stevens Point and CNUC, *Utility Vegetation Management in North America: Results from a 2019 Utility Forestry Census of Tree Activities and Operations*, 2020. [Online]. Available: https://www.waa-isa.org/wp-content/uploads/2020/12/Utilities Vegetation-Management-Miller-and-Hauer-1.pdf

[4] F. Bahreini, A. Hammad, "Detection of Vegetation Proximity to Power Lines: Critical Review and Research Roadmap," *Forests*, 16, 1658, 2025. Available: https://doi.org/10.3390/f16111658

[5] Ultralytics, *YOLO version 8.3.221 Documentation*. Ultralytics, 2023. Available: https://docs.ultralytics.com/models/yolov8/

[6] "EPSG:4326 — WGS 84", EPSG Geodetic Parameter Dataset, International Association of Oil & Gas Producers, 2025. [Online]. Available: https://epsg.io/4326

[7] A. S. Santos et al., "Estimation of vulnerable areas to faults caused by tree vegetation in power distribution systems," *IEEE Access*, v. 12, p. 48472-48484, 2024.

979-8-3315-8138-1