# System Integration of Multi-Modal Sensor for Robotic Inspection of Power Lines

Gustavo Fardo Armênio, João Henrique Campos Soares, Maria Luiza Cenci Stedile,
Oswaldo Ramos Neto, Ronnier Frates Rohrich and Andre Schneider de Oliveira
Graduate Program in Electrical and Computer Engineering
Program in Computer Science
Universidade Tecnológica Federal do Paraná
Curitiba, Brazil
e-mail: {gustavofardoarmenio | jsoares.2021 | mstedile | oswaldo}@alunos.utfpr.edu.br
{rohrich | andreoliveira}@utfpr.edu.br

***Abstract*—This study aimed to predict current and future issues in high-voltage-transmission lines using an integrated, specially designed multimodal-robotic sensor system for inspection. The system comprises several distinct sensors employed for the analysis of specific spectrums, such as, thermal, acoustic, spatial, visual, spectroradiometric, and referencing. Information obtained from different viewpoints and interfaces at different times are standardized and correlated to obtain composite-inspection data. This sensor (coupled to a cable-driven robotic platform) is intended to execute autonomous inspection of transmission elements by working over the power lines.**

***Keywords-inspection; multimodal; robot*.**

## I. INTRODUCTION

Power-grid functionality is reliant on electric-transmission-line integrity and reliability. Transmission lines are the backbone of electricity-distribution networks and are susceptible to threats ranging from environmental to human-induced disruptions. Weather-related events alone account for a significant proportion of transmission-line failures, underscoring the need for robust inspection protocols.

Proactive, inspection strategies can enhance the reliability of the power-transmission infrastructure and contribute to cost savings and operational efficiency. Investments in preventive maintenance (including routine transmission-line inspections) yield substantial returns by reducing outage durations, averting system failures, and minimizing associated economic losses. Advanced inspection technologies, such as Unmanned Aerial Vehicle (UAVs) [1]; light detection and ranging (LiDAR) [2]; and thermal imaging [3], facilitate comprehensive assessments of line components and prompt identification of defects and vulnerabilities [4], [5].

Traditionally, these inspections have relied on manual, visual assessments and single-sensor technologies such as infrared cameras or optical sensors. However, recent advancements in sensor technology have facilitated the development of multimodal sensors that can integrate multiple sensing capabilities into a single system—enhancing the effectiveness, accuracy, and efficiency of power-line inspections.

This integration allows for comprehensive data collection from different perspectives, allowing detection of potential issues, with greater accuracy [6]–[8]. For instance, in electrical systems, thermal imaging can detect hotspots (indicating potential overheating or electrical faults), whereas optical cameras provide high-resolution images for the visual inspection of physical damage or anomalies. LiDAR generates three-dimensional (3D) models of power lines and the surrounding vegetation, helping to identify encroachments and structural issues. Acoustic sensors detect partial discharge and other such signals, indicative of electrical malfunctions. By leveraging these diverse, sensing capabilities, multimodal sensors can identify a greater range of defects and conditions relative to single-sensor systems.

The integration of multiple sensing modalities enhances the accuracy and reliability of inspections [9], [10]. Each sensor type has its own strengths and limitations, and combining them mitigates the individual weaknesses. For example, optical cameras may be impeded by poor lighting conditions; however, thermal imaging can still detect issues under low-light conditions. Similarly, LiDAR can penetrate foliage to some extent, providing a clearer view of the power-line surroundings than optical cameras alone. Moreover, the fusion of data from different sensors allows for the cross-verification of findings, reducing false positives and negatives [11]. This redundancy ensures that the detected anomalies are genuine, enabling more reliable, maintenance decisions and actions.

With a multimodal sensor-equipped drone or vehicle, a single pass can gather comprehensive data, reduce inspection times, and reduce labor costs [12], eliminating the need for multiple passes. The high level of detail and accuracy provided by multimodal sensors leads to earlier detection of potential issues, preventing minor problems from escalating into major failures. Proactive-maintenance reduces downtime and repair costs, contributing to cost-effective, power-line management.

Power-line inspection is hazardous and often requires personnel to work at significant heights or close to high-voltage equipment. The use of multimodal sensors that are mounted on drones or robotic systems, reduces the need for human inspectors to operate in dangerous environments [13]–[15] and permits inspections in inaccessible and hazardous areas. The diverse data collected by multimodal sensors are ideal for integration with advanced analytics and machine-learning algorithms. By analyzing both historical and real-time data, these systems can predict potential failures and proactively recommend preventive-maintenance actions, thereby enhancing the overall reliability and resilience of the power grid.

This paper presents a novel, multimodal sensor coupled to an autonomous inspection robot (moving over an electric cable) for transmission-line inspection. The multispectral sensor integrates several perception sources to produce a unique inspection map.

The paper is organized into five sections. Section 2 discusses the concept of *LaRa* autonomous inspection robot. Section 3 discusses the proposed approach for *MultiSpectrum* sensor integration. Section 4 explains the experimentation and evaluation. Finally, Section 5 shows the conclusions.

## II. *LaRa*: Autonomous Robot for Multi-Modal Predictive Inspection of High-Voltage Transmission Lines

The mobile robot autonomously performs inspections by traveling directly over the electrical cables. The autonomous robot for the multimodal predictive inspection of high-voltage transmission lines (*LaRa*) is designed to attach to the cable and move with precision, carrying the multimodal inspection system, as shown in Figure 1.

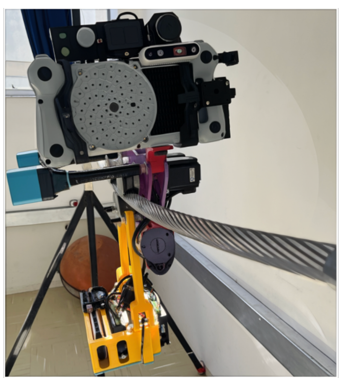

Figure 1. The *LaRa* robot.

Two wheels are used to ensure support on the electrical cable: one wheel is free, and the other is driven by a servomotor. The third wheel is part of a connecting rod–crank system that moves the non-actuated wheel toward the cable, maintaining a clamping pressure similar to that of a robotic claw. This wheel can also move linearly away from the cable, allowing the robot to be removed and perform obstacle suppression maneuvers.

The cable-gripper system is mounted on a structure consisting of two parallel plates separated by fixed spacers, as shown in Figure 2. Between these plates, a connecting rod–crank system moves the fixing wheel at the bottom of the cable. The motors are fixed to the front part of the claw, which interferes with the stabilization of the system on the cable, leading to rotation around the cable and potential falls.

The *LaRa* robot features a lower luggage rack fixed with two articulated arms to ensure that the weight is always directed toward the gravitational force at the center of the cable gripper. The luggage rack houses the electronic control system, motor power, control system, and battery of the robot.

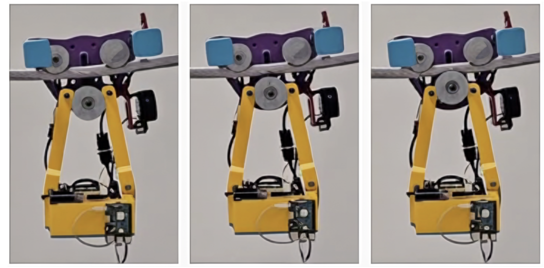

Figure 2. Cable-gripper system in action.

The center of mass of the system is aligned with the cable center, which is achieved by introducing two counterweight arms. One of these arms also serves as a support for the attachment of the multimodal inspection sensor.

## III. The Architecture of *MultiSpectrum* Sensor

The *MultiSpectrum* sensor comprised several sensors, specially designed to evaluate electric faults (Figure 3). All the sensors were integrated into a stacked inspection map. This approach was detailed further in an earlier study [16].

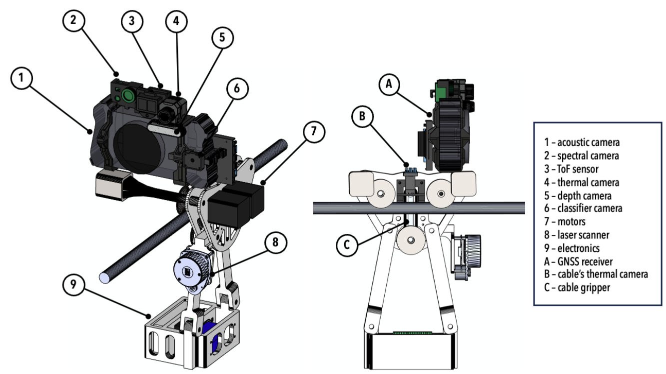


Figure 3. Modules of *LaRa* robot.

Figure 4 illustrates the integration of various sensor modules within the multispectral system designed for robotic inspection of transmission lines. This multimodal-sensor suite comprises several interconnected components, each serving a distinct function to ensure comprehensive monitoring and analysis of transmission line conditions.

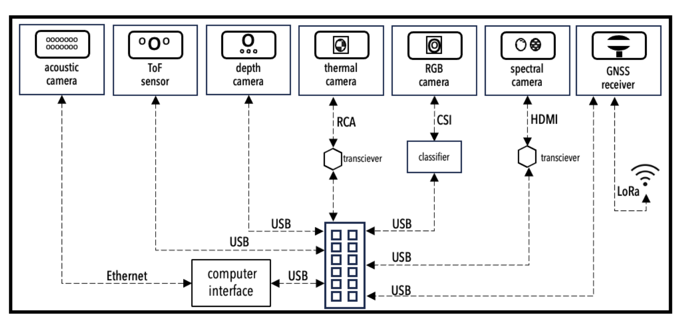


Figure 4. Integration of sensor modules.

The Real-Time Kinematic (RTK) Global Navigation Satellite System (GNSS) receiver provides geospatial data, enabling precise location tracking by robot inspectors. It connects via a Universal Serial Bus (USB) for data transmission and supports LoRa communication for long-range, low-power wireless connectivity aimed at RTK accuracy.

The spectral camera, equipped to capture a wide range of wavelengths, offers a detailed analysis of the material properties of the transmission lines and communicates with the central system through a high-definition multimedia interface to ensure high-quality data transfer; the aim was to analyze the proximity of the electric elements to vegetation.

The Red-Green-Blue (RGB) camera captures standard color images essential for visual inspection. It interfaces with the system using a Camera Serial Interface (CSI), that feeds directly into the Graphics Processing Unit (GPU) classifier for real-time image processing.

The thermal camera detects heat signatures and hotspots and identifies potential overheating issues or faults. It uses an RCA connection coupled with a transceiver to convert and transmit data through USB.

The depth camera provides 3D data, crucial for assessing the spatial relationships and physical conditions of transmission lines and their surrounding environment. It connects using a USB.

A Time-of-Flight (ToF) sensor measures the time required for a light signal to reflect from the object. It provides precise distance measurements and communicates with the system via USB. Sensor calibration and global referencing are crucial.

The acoustic camera captures sound waves to detect anomalies that may not be visible or detectable through other sensors, and is integrated into the system using an Ethernet connection for reliable data transfer.

### A. Multi-modal sensing

Integrating multimodal sensors involves combining data from various sensors to understand the environment or system comprehensively and accurately. The integration process leverages the strengths of each sensor type, compensating for individual sensor weaknesses and providing a richer dataset. The key to successful multimodal sensor integration lies in effective data fusion. Data-fusion algorithms combine information from different sensors to produce more accurate, reliable, and coherent information. This process often involves synchronizing data streams, spatially and temporally aligning data, and filtering noise.

The challenges in multimodal sensor integration include ensuring interoperability between different sensor types, managing large volumes of data, and maintaining real-time processing capabilities. Ensuring interoperability involves addressing various technical and operational issues because different sensors often have distinct communication protocols, data formats, and sampling rates. To integrate these sensors seamlessly, a common framework or middleware is required to translate and standardize the data from each sensor type.

The Petri net flow for the multispectral sensor system illustrates the comprehensive workflow involved in the multimodal inspection of transmission lines, as shown in Figure 5. The process begins with the system in a ready state (p1), which is initialized and prepared for inspection. Upon starting the inspection (t1), the system waits for inputs from various sensors, including spectral (p2), depth (p3), RGB (p4), thermal (p5), distance (p6), ToF (p7), GNSS (p8), and acoustic (p13) data.

Each type of sensor input underwent specific acquisition and processing steps. The spectral images were filtered (t9) and registered (t15) to align them accurately, resulting in a filtered spectral image (p9) and registered spectral image (p16). The depth images were resized (t10) and warped (t16) to correct any distortions, producing a resized depth image (p10) and depth layer (p17). The RGB images were classified (t11) to identify relevant features, resulting in a classified RGB image (p11). The thermal images were resized (t12), decomposed into component parts (t17), and registered (t20) to align with the other sensor data, resulting in a resized thermal image (p12), decomposed thermal image (p18), and registered thermal image (p22). The acoustic images were resized (t18), filtered to remove noise (t21), and registered (t23) for accurate alignment, resulting in an adjusted acoustic image (p19) and a filtered acoustic image (p23).

After initial processing, the system combined and adjusted the data layers. Spectral images were warped (t19) and integrated into a spectral layer (p21), thermal images that underwent thermal warping (t22) were integrated into a thermal layer (p24), and acoustic images were warped (t24) and integrated into an acoustic layer (t26). These processed layers were stacked to form a comprehensive inspection map (p25).

The inspection map was further refined through georeferencing (p27) to ensure that the data were accurately mapped to real-world coordinates. The final outputs of this process included detailed inspection maps (p28) that provided a thorough overview of the inspection results and geospatially contextualized data, indicating the precise locations of the inspected areas (p29).

### B. Global localization of multi-modal inspection

The multispectral sensor employs an RTK-GNSS to ensure the correct localization of electric components and to allow correlation between different inspections. The RTK GNSS employs a stationary base station and *LaRa* robot (i.e., rover) to obtain highly accurate positioning data with centimeter-level accuracy. The base station measures signals from the GNSS satellites and calculates the errors caused by atmospheric conditions, satellite-orbit inaccuracies, and other factors. These corrections are sent to the *LaRa* robot in real-time, through a communication link (LoRa), allowing it to adjust its calculations and achieve higher accuracy. The *LaRa* robot then applies these corrections to improve positional accuracy, as illustrated in Figure 6.

Figure 7 illustrates the process of integrating sensor data for locating the nearest power pole on a transmission line. The process begins by measuring distances using ToF and depth sensors to accurately measure nearby objects. The system then computes the nearest point relative to the sensor. The

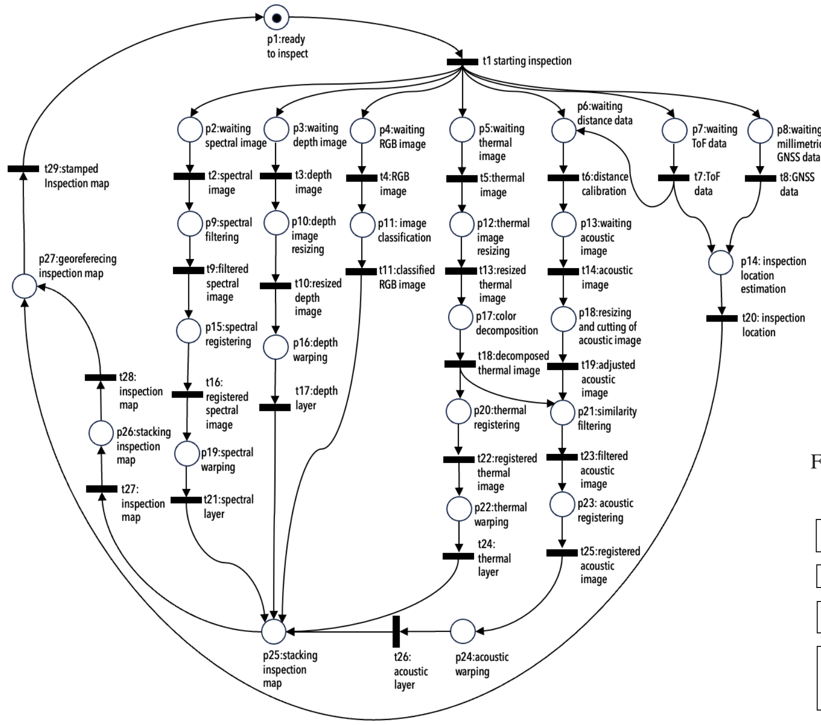


Figure 5. Petri-Net of of Multi-modal sensing of MultiSpectrum sensor.

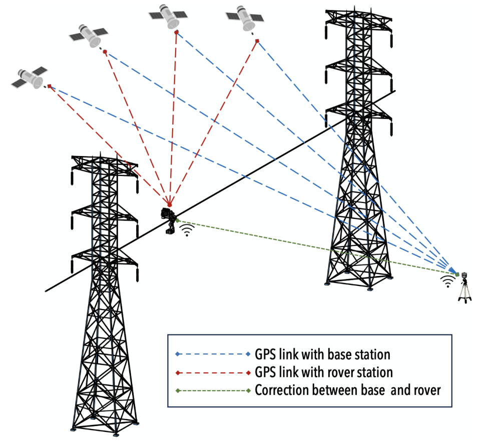


Figure 6. Scheme of global localization through RTK GNSS.

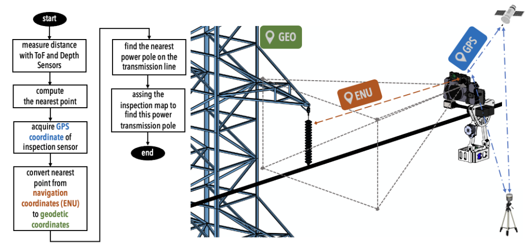


Figure 7. Transformation from local to global coordinates.

RTK-GNSS coordinates of the inspection sensor are acquired, providing its geographic position.

The nearest point, initially in the East-North-Up (ENU) coordinates (a local Cartesian coordinate system), was converted to geodetic coordinates (latitude, longitude, and altitude). The system identified the nearest power pole on the transmission line by matching it to a map or a database of pole locations. Finally, an inspection map was assigned to the identified power-transmission pole, which linked the sensor data to a specific location in the transmission infrastructure.

This process effectively integrates local measurements and global positioning to precisely locate power poles for inspection and correlates them with previous inspections, allowing for the prediction of future behaviors.

### *C. Object classification and recognition*

Here, the objective was to develop a device capable of detecting key elements (such as insulators, transmission towers, and dampers) along transmission lines, in real-time, utilizing local processing with energy consumption compatible with battery-usage. Initially, the primary component of the device was defined as a tool capable of detecting objects in an image with high reliability. The eighth (state-of-the-art) version of the YOLO (You Only Look Once) neural-network architecture (YOLOv8) was selected owing to its optimization ease and flexibility of application. YOLOv8 is provided through an SDK maintained in the Ultralytics library and features a simple Python interface that facilitates the configuration of network parameters and training procedures [17].

To enable the near-real-time processing of a neural network such as YOLO, it is necessary to have hardware, capable of supporting the parallel processing of the network layers. An NVIDIA Jetson Nano B01 with 4GB of random-access memory was selected because of its compact size, low-energy consumption, and graphical-acceleration capabilities. Their applications are further supported by multiple interfaces with other devices and peripherals. By utilizing the MIPI CSI input, it is possible to attach a Raspberry Pi V2 camera designed for embedded systems (with reduced energy consumption and weight) for environmental image capture. Additionally, the UART TTL serial-interface pins enabled communication between the detection device and other computers using an FT232 Serial-USB converter. For training, approximately 540 images were selected from photos and videos of transmission-line, drone inspections. The images were labeled with rectangular annotations. The classes were named after the key elements: Transmission Tower, Insulator, Damper, and Transformer. The training was performed using 200 epochs, a batch size of 16 samples, and an image size of 640 × 640 pixels.

The training results were visualized in a confusion matrix (shown in Figure 8). A high accuracy for transmission towers and isolators, with true-positive rates of 90% and 88%, respectively, can be seen. The Damper class had a lower true positive rate of 64%, and the Transformers were correctly classified at 85%.

The training was also evaluated using a bar chart (Figure 9), which illustrates the distribution of instances for four classes: Damper, Isolator, Transmission Tower, and Transformer. The Damper class had the fewest instances, with fewer



 

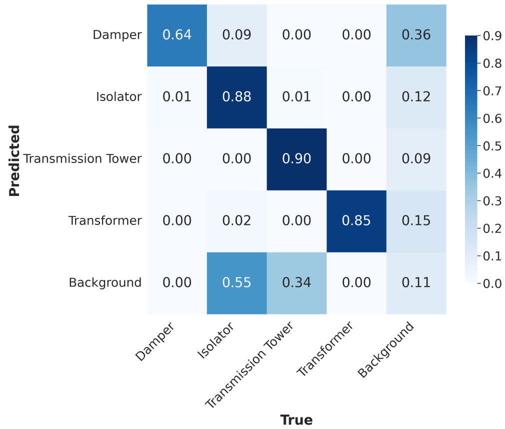


Figure 8. Confusion matrix of training.

than 500 examples, indicating that this class was relatively underrepresented in the dataset. In contrast, the Isolator class had the highest number of instances, with approximately 3000 examples, suggesting that the model had more data to learn from for this class, potentially leading to higher prediction accuracy. The Transmission Tower class had a moderate number of instances, approximately 1500, providing a balanced amount of data for model training compared with the others. The Transformer class had the fewest instances after the damper, with fewer than 500 examples, which, like the Damper class, might have affected the ability of the model to accurately predict this class.

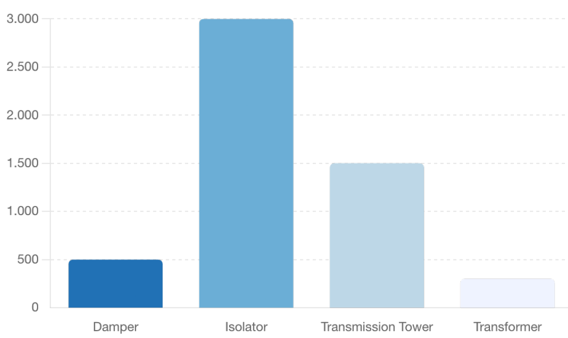


Figure 9. Instances of training.

Figure 10 presents two graphs tracking the mean average precision at IoU = 0.50 (mAP50) and mAP at IoU = 0.95 (mAP95) metrics over 200 training epochs. The mAP50 graph demonstrated rapid initial improvement from around 0.30 to approximately 0.83, indicating that the model quickly learns to detect objects with moderate IoU thresholds. The curve then showed a more gradual increase as the training progressed, stabilizing at approximately 0.83. This suggests that the model achieved high precision for easier detection and maintained consistent performance towards the end of the training period.

The mean Average Precision (mAP) in the range 0.50 < Intersection over Union (IoU) < 0.95 (mAP50-95) graph starts lower, around 0.20, but steadily increases throughout the training process, reaching approximately 0.56. This reflected

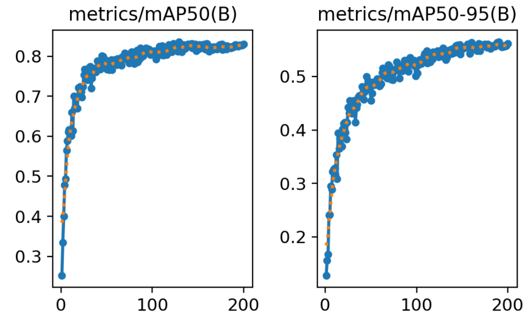


Figure 10. Training accuracy analysis.

the growing ability of the model to handle more challenging detection scenarios, although with a slower improvement compared with the mAP50 metric. The gradual rise and final values indicated that while the model performed well, its precision decreased as the IoU threshold increased. The output of evaluation of object classification can be seen in Figure 11.

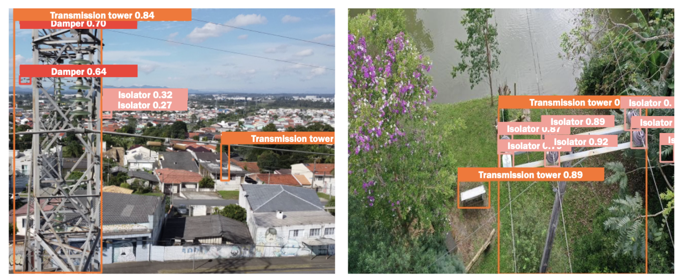


Figure 11. Evaluation of object classification.

## IV. Multi-Modal Inspection

Multi-modal inspection is consolidated into a comprehensive multi-layer inspection map with global referencing for a specific transmission tower. Each layer of the map represents a distinct spectrum of analysis for the transmission line elements, as illustrated in Figure 12.

The first layer utilizes visual analysis to identify visible faults in high resolution and recognize power line elements. This identification is performed using object classification and recognition methods and serves as a foundation for subsequent layers. The second layer employs depth spectrum analysis, enabling volumetric inspection of elements and spatial correlation. The third layer is dedicated to thermal analysis, detecting anomalies in thermal profiles and identifying hotspots. The fourth layer focuses on acoustic spectrum analysis, examining distortions in the acoustic response of elements to diagnose malfunctions such as breaks, wear, and the corona effect. The fifth layer analyzes the vegetation spectrum, evaluating the proximity of vegetation to the elements and its potential to cause electrical arcs.

## V. Conclusions

This paper presents a multispectral-sensor system for the multimodal-robotic inspection of high-voltage-transmission

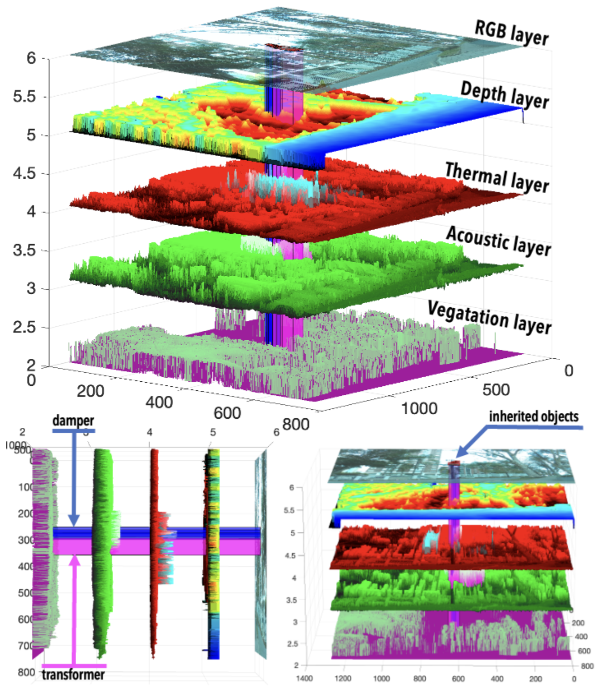

Figure 12. Multi-Modal inspection map.

lines. The system integrates various sensors — thermal, acoustic, spatial, visual, spectroradiometer, and referencing — to enable the accurate prediction of current and future issues. Standardizing and correlating data from these sensors provides comprehensive inspection results, enhancing the accuracy and reliability of power-line maintenance.

A key feature of the multispectral sensor is the RTK-GNSS, which ensures precise localization with centimeter-level accuracy and is crucial for correlating data from different inspections. The system employs a stationary base station and *LaRa* robot to provide real-time corrections, thereby improving the positional accuracy of the electric components along the transmission lines. Additionally, the device uses the YOLOv8 neural network for the real-time detection of elements such as insulators, transmission towers, and dampers, chosen for its high reliability and ease of application. The training and evaluation of the YOLOv8 model highlighted potential accuracy variations based on the class representation. Overall, the multispectral-sensor system, with its advanced integration of RTK-GNSS and YOLOv8, offers a state-of-the-art solution for the autonomous and efficient predictive inspection of power lines.

## ACKNOWLEDGMENT

The project is supported by the National Council for Scientific and Technological Development (CNPq) under grant number 407984/2022-4; the Fund for Scientific and Technological Development (FNDCT); the Ministry of Science, Technology and Innovations (MCTI) of Brazil; the Araucaria Foundation; the General Superintendence of Science, Technology and Higher Education (SETI); and NAPI Robotics.

## REFERENCES

[1] Z. Wang, Q. Gao, J. Xu, and D. Li, "A review of uav power line inspection", in *Advances in Guidance, Navigation and Control: Proceedings of 2020 International Conference on Guidance, Navigation and Control, ICGNC 2020, Tianjin, China, October 23–25, 2020*, Springer, 2022, pp. 3147–3159.

[2] N. Munir, M. Awrangjeb, and B. Stantic, "Power line extraction and reconstruction methods from laser scanning data: A literature review", *Remote Sensing*, vol. 15, no. 4, p. 973, 2023.

[3] F. Shams *et al.*, "Thermal imaging of utility power lines: A review", in *2022 International Conference on Engineering and Emerging Technologies (ICEET)*, IEEE, 2022, pp. 1–4.

[4] L. Yang *et al.*, "A review on state-of-the-art power line inspection techniques", *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 12, pp. 9350–9365, 2020.

[5] A. Sriram and T. Sudhakar, "Technology revolution in the inspection of power transmission lines-a literature review", in *2021 7th International Conference on Electrical Energy Systems (ICEES)*, IEEE, 2021, pp. 256–262.

[6] O. Kullu and E. Cinar, "A deep-learning-based multi-modal sensor fusion approach for detection of equipment faults", *Machines*, vol. 10, no. 11, p. 1105, 2022.

[7] O. Kähler, S. Hochstöger, G. Kemper, and J. Birchbauer, "Automating powerline inspection: A novel multisensor system for data analysis using deep learning", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 747–754, 2020.

[8] F. Najar, M. Ghommem, S. Kocer, A. Elhady, and E. M. Abdel-Rahman, "Detection methods for multi-modal inertial gas sensors", *Sensors*, vol. 22, no. 24, p. 9688, 2022.

[9] J. Petrich, Z. Snow, D. Corbin, and E. W. Reutzel, "Multi-modal sensor fusion with machine learning for data-driven process monitoring for additive manufacturing", *Additive Manufacturing*, vol. 48, p. 102 364, 2021.

[10] C. Martin, E. Schaffernicht, A. Scheidig, and H.-M. Gross, "Multi-modal sensor fusion using a probabilistic aggregation scheme for people detection and tracking", *Robotics and Autonomous Systems*, vol. 54, no. 9, pp. 721–728, 2006.

[11] P. Karle, F. Fent, S. Huch, F. Sauerbeck, and M. Lienkamp, "Multi-modal sensor fusion and object tracking for autonomous racing", *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 7, pp. 3871–3883, 2023.

[12] P. Nooralishahi, F. López, and X. P. Maldague, "Drone-enabled multimodal platform for inspection of industrial components", *IEEE Access*, vol. 10, pp. 41 429–41 443, 2022.

[13] M. Jeon, J. Moon, S. Jeong, and K.-Y. Oh, "Autonomous flight strategy of an unmanned aerial vehicle with multimodal information for autonomous inspection of overhead transmission facilities", *Computer-Aided Civil and Infrastructure Engineering*, 2024.

[14] B. Jalil, D. Moroni, M. Pascali, and O. Salvetti, "Multimodal image analysis for power line inspection", in *Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence, Montreal, QC, Canada*, 2018, pp. 13–17.

[15] S. Cao *et al.*, "Multi-sensor fusion and data analysis for operating conditions of low power transmission lines", *Measurement*, vol. 190, p. 110 586, 2022.

[16] R. F. Rohrich and A. S. de Oliveira, "Multispectrum inspection of overhead power lines.", in *20th International Conference on Informatics in Control, Automation and Robotics*, 2023, pp. 119–126.

[17] G. Jocher and Ultralytics, *Yolo by ultralytics*, https://github.com/ultralytics/yolov8, Accessed: 2024-07-09, 2023.

 