INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH AND INNOVATIVE STUDIES
ISSN: 2820-7157
www.ijsrisjournal.com

October 2025 Volume 4 Number 5 12-21

Received Date: August 21, 2025 Accepted Date: September 13, 2025 Published Date: October 01, 2025

Drone Photogrammetry and AI for Monitoring Paved Roads Distresses in Cameroon – A Case Study of Elig-Effa-Melen-EMIA Road Section in Yaoundé

MEKANDA YAKAN Eugène Salomon [1], LIMALEBA Roger Blaise [2], BIKIE Gerald Anicet [3]

1. Department of Land Survey, National Advanced School of Public Works (NASPW), Yaoundé, Cameroon.
2. Département de Topographie-Cadastre, École Nationale Supérieure des Travaux Publics (ENSTP), Yaoundé, Cameroun

Abstract

This study presents an innovative approach for automated pavement degradation monitoring in Cameroon, specifically focusing on the city of Yaoundé. We integrate drone photogrammetry for high-resolution images with artificial intelligence, leveraging a YOLOv11 model for precise detection and segmentation of various road surface degradations. A comprehensive Python workflow was developed to automate the entire processing chain, from orthomosaic tiling and AI model inference to georeferencing predictions and generating GIS-ready layers. The methodology was applied to a 2.5 km road section (Elig-Effa–Melen– Carrefour EMIA) in Yaoundé. The trained model achieved a F1-score of 77% with a precision of 86% and recall of 70% for the detection and F1-score of 81% with a precision of 86% and recall of 78% for the segmentation. The automated workflow processed the road section in 5 minutes 47 seconds, detecting 438 instances of degradation with a total affected surface area of 662.96 m², demonstrating significant time savings compared to days or weeks required for manual inspection. The generated degradation maps and an interactive dashboard provide objective, actionable insights for road managers, enabling proactive maintenance planning and optimized resource allocation. While acknowledging limitations such as the need for further validation with field surveys and the current focus on 2D surface data, this research highlights the scalability and cost-effectiveness of the proposed solution. It offers a promising pathway toward intelligent and efficient road network management in developing countries, by providing timely and accurate data to support data-driven maintenance strategies.

Keywords: Photogrammetry, Drone, Artificial Intelligence, Paved Roads, Degradations, YOLOv11.

1. Introduction

Road infrastructure is a cornerstone of economic development and social connectivity, particularly in developing countries. In Cameroon, the road network spans over 121,873 km as of 2024, yet only about 8.4% of these roads are paved (Woof M. J., 2025). This vital network faces continuous deterioration due to increasing traffic loads, demanding tropical climates, and natural aging of materials. The progressive degradation of road surfaces—manifesting as cracks, potholes, rutting, and other distresses—threatens infrastructure durability and user safety, with severe economic and social consequences if left unaddressed. Traditionally, road condition assessment in

Cameroon and elsewhere has relied on labor-intensive visual inspections or on specialized survey vehicles equipped with cameras, laser scanners, or other sensors. While such methods provide valuable data, they are often costly, time-consuming, subjective, and limited in spatial coverage (Tan and Li, 2019). Manual surveys are difficult to replicate consistently across large urban areas, and their infrequent periodicity, constrained by limited resources, impedes timely maintenance. In cities like Yaoundé, where the road network extends hundreds of kilometers, traditional approaches have become insufficient for regular, standardized, data-driven maintenance planning.

In this context, a promising alternative emerges: the combination of unmanned aerial vehicles (UAVs or drones) for photogrammetry and modern deep learning-based artificial intelligence (AI). Drones have proven to be versatile tools for infrastructure inspection, offering substantial advantages in speed, safety, and coverage. Equipped with high-resolution cameras, UAVs can capture detailed aerial imagery of road surfaces from various angles and altitudes, enabling comprehensive mapping of pavement conditions through photogrammetry. When coupled with AI, this approach allows for faster, more precise, and systematic pavement inspections than traditional methods. High-resolution orthoimages derived from drone flights can be processed with computer vision models to automatically detect and quantify road distresses, providing objective assessments in a fraction of the time required by manual surveys.

Recent studies worldwide have explored the integration of drone imagery and AI for automated road damage detection. In Spain, Silva et al. (2023) demonstrated the effectiveness of deep learning models (YOLO family) in identifying cracks and potholes from UAV images. Their work achieved up to 73% mAP (mean Average Precision) at 0.5 IoU using YOLOv7, highlighting the potential of UAV-based surveys for road condition monitoring. However, the authors noted challenges regarding model generalizability: models trained on one region or lighting condition often underperform when applied to different environments. This underscores the need for context-specific datasets and locally tailored training, as variations in pavement materials and environmental conditions can significantly affect detection performance. In another study, Tan and Li (2019) in China used UAV photogrammetry to reconstruct 3D road models and detect pavement distress. Their method achieved high precision, with discrepancies of around 1 cm in distress depth measurement compared to field surveys, demonstrating that drone-based surveys can meet engineering accuracy requirements for road defect characterization.

In Cameroon, pioneering research has begun to apply drone photogrammetry to road monitoring. Nguimbock (2020), for example, conducted one of the first local studies using drone images to map pavement degradations on a Yaoundé roadway. That study involved manually digitizing road defects (cracks and potholes) on high-resolution orthomosaics and comparing the results to field measurements. The findings showed a strong correlation between drone-derived measurements and ground truth, with an average area deviation of only 0.07 m² between GNSS-measured and photo-derived defect areas. This confirmed the reliability of drone orthomosaics (when properly georeferenced with ground control points) for capturing the extent of road surface damage. However, Nguimbock's approach did not incorporate machine learning for automation; the identification of defects was essentially manual and focused on the pavement surface layer alone. The literature thus highlights two critical gaps that our work aims to address: (1) the lack of an integrated, automated workflow that can process drone imagery end-to-end for road defect detection (from image acquisition to GIS mapping), and (2) the need to consider the road infrastructure holistically as a system (including not just the pavement lanes, but also related features like sidewalks, shoulders, and drainage elements). By building upon these prior studies, our research seeks to develop a scalable UAV+AI based monitoring system tailored to the context of Yaoundé, and to demonstrate how such a system can improve upon traditional road maintenance practices through more frequent, accurate, and cost-effective assessments

2. Methodology

Our methodology integrates UAV-based photogrammetry, deep learning segmentation, and GIS-based post-processing to detect and map degradations on paved roads. The overall approach consists of four main steps: (i)data acquisition via drone flights, (ii) dataset preparation and AI model training, (iii) automated inference and georeferencing of detections, and (iv) integration of results into a Geographic Information System (GIS) for analysis and visualization. Our approach is based on the following diagram:

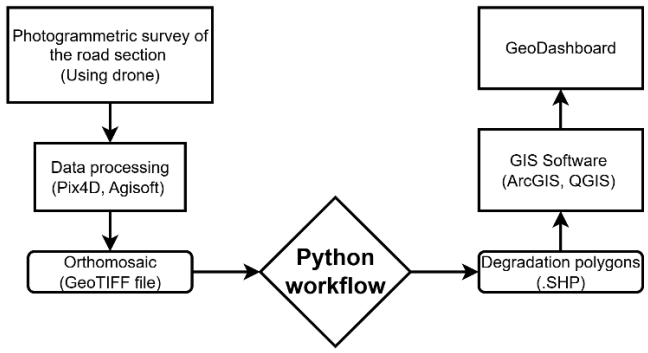


Figure 1 : Processing approach

2.1 Study area

The case study is conducted in Yaoundé, Cameroon's capital, which provides a representative urban environment with diverse road conditions and common degradation types. Within Yaoundé, a specific road section known as Elig-Effa–Melen–Emia (approximately 2.5 km long) (Figure 2), was selected for detailed analysis. This section was chosen due to its varied pavement conditions and observable distresses, making it an ideal testbed for the proposed monitoring system.

The selected road segment traverses residential and commercial districts, thus experiencing significant traffic. Prior to data collection, necessary flight authorizations were obtained from the Cameroon Civil Aviation Authority, and survey operations were coordinated to ensure safety and compliance with local regulations.

Figure 2 : Elig-Effa-Melen-Emia Road section

2.2 Data acquisition

Aerial data were acquired using a *DJI Mavic 3 Enterprise (M3E)* drone (Figure 3). The M3E is a compact quadcopter equipped with an RTK (Real-Time Kinematic) module for centimeter-level geolocation accuracy of images. It carries an integrated 20 MP wide-angle camera (4/3″ CMOS sensor) capable of high-resolution capture even in lower light conditions. We planned flight missions to achieve a ground sampling distance (GSD) of 2-3 cm/pixel on the ground, which is sufficient to resolve fine pavement details such as small cracks. The drone was flown at an altitude and overlap configuration optimized for photogrammetry: typically, 80% forward overlap and 70% side overlap between images, ensuring robust 3D reconstruction and full coverage of the road surface.

Figure 3 : DJI Mavic 3 Enterprise

To enhance absolute positioning accuracy of the orthomosaic, we placed Ground Control Points (GCPs) on and around the road prior to flying. These GCPs were marked on the pavement (painted targets at easily identifiable locations) and precisely surveyed with dual-frequency GNSS receivers in Network RTK mode (Figure 4).

An average horizontal accuracy of ±1 cm and vertical accuracy of ±2 cm was achieved for the GCP coordinates. The spatial arrangement of GCPs covered the extremities and key points along the road section to support an accurate georeferencing of the photogrammetric models. The data acquisition campaign took place over a one-month period, accounting for suitable weather windows and coordinating around traffic to minimize disruptions.

Figure 4: GCP surveying

2.3 Photogrammetric processing

All collected images were processed using Pix4Dmapper to generate the high-resolution orthomosaics of several roads in Yaoundé. In Pix4D, a new project was created for each road subsection flight, using the WGS84-UTM Zone 32N coordinate system (EPSG:32632) appropriate for Yaoundé. We incorporated the local vertical datum (EGM2008 geoid model) to convert elevations to orthometric heights consistent with local surveying practice. For each road section, the processing workflow included: (1) *Aerial triangulation*, (2) *Dense point cloud generation*, (3) *Digital Surface Model (DSM)* and *Orthomosaic*. Figure 5 shows an extract of the generated orthomosaics.

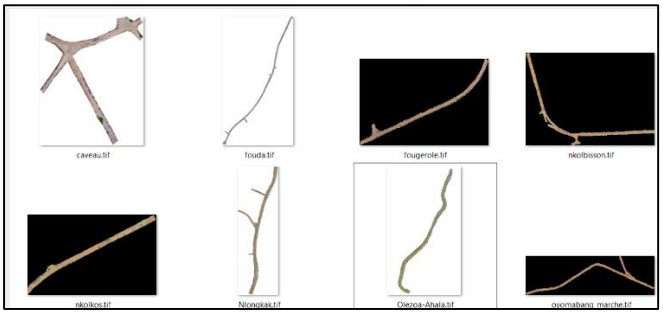


Figure 5 : An extract of the orthomosaics dataset

2.4 Dataset preparation and annotation

From the orthomosaics acquired from various road sections in Yaoundé, we generated a dataset of image jpeg *tiles* for use in training the AI model. The orthomosaics were divided into smaller patches (640×640 pixels), which is a convenient size for modern deep learning frameworks and avoids memory issues when processing high-resolution images. We extracted around 1000 image tiles and each tile was labeled for the presence of road surface degradations. Four damage classes were used as shown in Table 1:

Table 1 : Selected degradation classes

Class	*C1*	*C2*	*C3*	*C4*
Degradation	Cracks	Potholes	Raveling/stripping	Edge failures

Using a combination of manual annotation and semi-automated tools based on SAM (Segment Anything Model), we marked instances of these defects on the images. Annotation was performed with the aid of the Roboflow platform, which allowed labeling each defect with a polygon (mask) or bounding box and an associated class label (Figure 6). To increase the robustness of the model, we applied data augmentation techniques on the training images: rotations, flips, brightness/contrast adjustments, and scaling were used to synthetically vary the dataset. This helps the model generalize better to different orientations and lighting conditions of defects. The prepared dataset was then split into training, validation, and test sets (approximately 70% for training, 20% validation, 10% testing) ensuring that images from the same area did not leak into multiple sets.

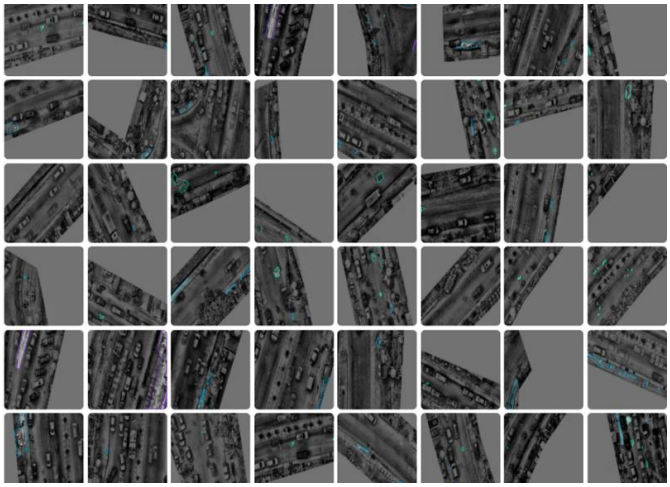


Figure 6 : An extract of the annotated dataset

2.5 YOLOv11 model development

We employed a custom variant in the YOLO (You Only Look Once) family of object detectors, referred to here as YOLOv11, for pavement damage detection and segmentation. The YOLO architecture is known for its efficiency in real-time object detection, using a single neural network to predict bounding boxes and class probabilities in one evaluation. Our version extends this concept by also predicting segmentation masks for each detected object, effectively performing instance segmentation of road defects. The model training was carried out on the Google Colab cloud infrastructure, leveraging an NVIDIA Tesla T4 GPU with 16 GB of video memory (VRAM), 12 GB of RAM, and 112 GB of storage.

The training process was configured with a set of carefully selected hyperparameters to ensure robust model convergence and effective transfer learning. The model was trained over 750 epochs with a batch size of 16 and an input resolution of 640 × 640 px. Optimization used SGD with momentum (0.937) and an initial learning rate of 0.01 with cosine decay. Training was initialized with YOLOv11 pre-trained COCO weights. This configuration enabled efficient execution of the training and optimization processes while ensuring a reasonable computation time for the dataset size considered.

The performance of the model was evaluated using standard metrics commonly employed in object detection and instance segmentation tasks. These include precision (1), recall (2), and F1-score (3), which provide insight into the model's accuracy and ability to detect true distress instances while minimizing false detections. /

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (1)$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (2)$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

- *TP*: True Positives
- *FP*: False Positives
- *FN*: False Negatives

Additionally, we report the mean Average Precision (mAP) at an Intersection-over-Union (IoU) threshold of 0.5 (mAP@0.5), and over a range of thresholds from 0.5 to 0.95 (mAP@[0.5:0.95]), following the COCO evaluation protocol, given by (3) and (4). These metrics offer a robust framework for assessing both the localization and classification performance of the model.

$$\text{mAP@0.5} = \frac{1}{N} \times \sum_1^n AP_i \quad (3)$$

$$\text{mAP@[0.5:0.95]} = \frac{1}{10N} \times \sum_1^N \sum_{t=0.5}^{0.95} AP_i(t) \quad (4)$$

- *N*: number of classes (N=4 in our case)
- *t*: IoU threshold
- $AP_i(t)$*:* average precision for class iii at threshold ***t***

2.6 Integrated processing workflow

A major innovation of this study is the development of an integrated Python-based workflow (Figure 7) to automate the inference and mapping process end-to-end. This workflow was implemented as a sequence of modules that take the orthomosaic as input and output georeferenced degradation polygons in shapefile format for GIS analysis, without human intervention.

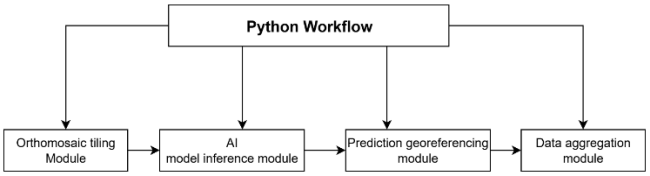


Figure 7 : Python workflow

This workflow was designed as a modular and automated solution, orchestrating the entire process from orthomosaics tiling to the generation of final GIS layers. Its architecture aimed to minimize human intervention and maximize processing fluidity. The key components of the workflow are:

- **The orthomosaic tiling module:** This module programmatically slices the large orthomosaic into the same tile size that the AI model expects (640×640 px), using a sliding window approach. Along with each tile, it generates a corresponding world file (a text file containing the georeferencing coordinates for that tile (*.jgw*)). This ensures that the pixel coordinates of segmentation can later be translated back to real-world coordinates.
- **The AI model inference module:** This module loads the trained YOLOv11 model and runs it on each image tile in turn. The model outputs, for each tile, a set of predicted bounding boxes, segmentation masks, confidence scores, and class labels (C1–C4). We capture these predictions and initially store them in a JSON format (for each tile), which contains the pixel coordinates of the mask polygons and the class and confidence for each detected object. The automation here allows batch processing of hundreds of tiles rapidly, far faster than manually running a detector on each image patch.
- **Prediction georeferencing module:** In this step, the local pixel coordinates of each detected polygon are converted to global map coordinates (latitude/longitude or easting/northing). Using the world file info for each tile and libraries such as GeoPandas and Shapely in Python, we transform each mask polygon from the tile's coordinate space into the orthomosaic's coordinate system using a first-order affine transformation, as defined in the GeoTIFF specifications (Ritter & Ruth, 1997) and implemented in the GDAL library, given by (5):

$$\begin{cases} X = scale_x \cdot x + x_{\text{top left}} \\ Y = scale_y \cdot y + y_{\text{top left}} \end{cases} \quad (5)$$

  - *x, y*: pixel coordinates of the point in the tile
  - *X, Y*: geographic coordinates (usually in meters in the projected CRS)
  - $scale_x$: pixel resolution along the X-axis
  - $scale_y$: pixel resolution along the Y-axis
  - $x_{\text{top left}}, y_{\text{top left}}$: geographic coordinates of the top-left corner of the tile.

The results for all tiles are then merged. If a defect spans multiple tiles (rare due to tile size and overlaps), the polygons can be merged or handled to avoid duplication. We output a georeferenced vector file (shapefile) containing all detected road defects, each with attributes like class, area, and confidence score. This module effectively bridges AI output with GIS input automatically.

- **Data aggregation module:** Finally, the workflow compiles all georeferenced detections into a single layer and performs some summary calculations. It computes the total count of defects and the sum of their areas per class for the entire road section.

2.7 GIS integration

The aggregated geospatial dataset is saved as a shapefile feature class, ready to be loaded into GIS software. We then import this layer into ArcGIS Pro for visualization and further spatial analysis. An interactive dashboard was also developed (using ArcGIS Dashboard) to visualize the results dynamically. The dashboard links the map of detected defects with charts and statistics (e.g., number of potholes, total cracked area, density of defects per kilometer), and allows filtering by defect type or road segment. This helps road managers interpret the data and identify priority areas.

Overall, the methodology emphasizes reproducibility and automation. By using this integrated approach, once the drone imagery is collected and the model is trained, the entire workflow from raw images to actionable maps can be executed quickly without manual intervention at each stage. This is crucial for scaling up the solution to city-wide road networks on a regular basis.

3. Results and discussions

3.1 Model Performance

3.1.1 Global quantitative performance metrics

After training, the YOLOv11 model showed strong performance in detecting and segmenting the four defined classes of road degradation. Table 2 reports the global metrics at the end of training for detection (B, bounding boxes) and segmentation (M, masks).

Table 2*:* Model performance metrics values

Task	Precision	Recall	mAP@0.50	mAP@0.50–0.95
Detection (B)	0.86	0.70	0.75	0.48
Segmentation (M)	0.86	0.78	0.73	0.38

These values indicate a well-balanced model that favors relatively high recall while maintaining good precision, thereby limiting false positives. Class-wise precision shows high values for potholes and raveling, consistent with Table 3.

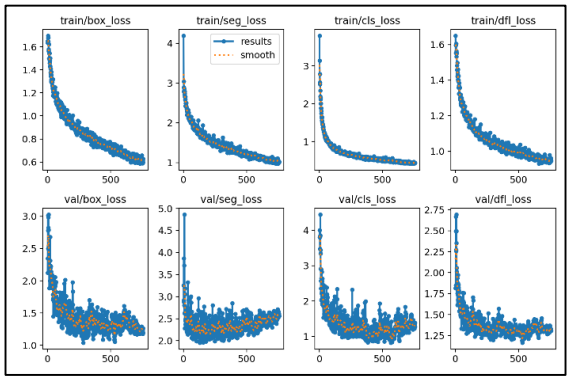
Table 3: Class-wise model performance (precision)

Class	Potholes	Raveling	Cracks	Spalling / Edge failures
Precision (%)	92	80	65	67

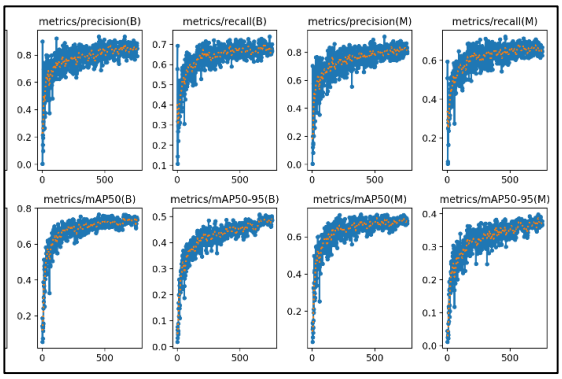
3.1.2 Training curves

The analysis of the training curves (Figure 8) was essential to understand the model's learning process and to identify potential issues such as overfitting or underfitting. The training losses (train/box_loss, train/seg_loss, train/cls_loss, train/dfl_loss) decrease rapidly and then stabilize, indicating proper convergence. The validation losses (val/box_loss, val/seg_loss, val/cls_loss, val/dfl_loss) follow the same trend with normal, mild fluctuations and no marked divergence from training.

Regarding the metrics, precision reaches about 0.86 for both detection and segmentation, while recall is around 0.70 for detection and 0.78 for segmentation. The mAP at IoU 0.50 is approximately 0.75 (B) and 0.73 (M), whereas mAP@0.50–0.95 is lower—0.48 (B) and 0.38 (M)—reflecting the increasing strictness of higher IoU thresholds.



(A)



(B)

Figure 8: Training loss curves (A) and metrics curves (B)

The YOLOv11 model's performance metrics (Table 3.2) and the training curves indicate a robust ability to detect and segment road distresses:

- **High precision (0.86):** Most reported detections are correct. The damage mapping is reliable and limits false positives. This is crucial for operational planning, as resources will not be deployed unnecessarily on undamaged segments.

- **Recall—moderate for detection (0.70) and higher for segmentation (0.78):** This indicates that some defects may still be missed, notably very fine cracks or areas with low contrast/shadows. In practice, this implies that targeted field checks remain relevant in sensitive segments, while segmentation helps recover part of the cases missed by detection alone.

- **mAP@0.50 (0.75 for B, 0.73 for M):** These values are compatible with network-scale screening and prioritization. Conversely, mAP@0.50–0.95 evaluates localization accuracy at stricter IoU thresholds: the lower scores (0.48 for B; 0.38 for M) suggest that fine adjustment of boxes and mask contours remains improvable, especially for irregular shapes (raveling, alligator cracking).

Regarding the training process, the decrease followed by stabilization of train/val losses points to stable learning, with no marked overfitting.

In summary, the combination of high precision and solid mAP@0.50 enables the production of actionable damage masks for maintenance. The main avenues for improvement concern recall on difficult classes and fine localization accuracy (mAP@0.50–0.95), which can be enhanced by enriching the dataset (rare classes) and refining the annotations.

3.2 Application to the Elig-Effa-Melen-Emia Road section

3.2.1 Detection and segmentation of road damage

We applied the fully trained model and the automated workflow to the Elig-Effa–Melen–Carrefour EMIA Road section orthomosaic (Figure 9). The final orthomosaic had a GSD of 2.3 cm/px and covered the entire Elig-Effa–Melen–Emia corridor. The Pix4D's quality report, showed an RMSE (root mean square error) of 1 cm on GCPs, confirming the precision of the mapping.

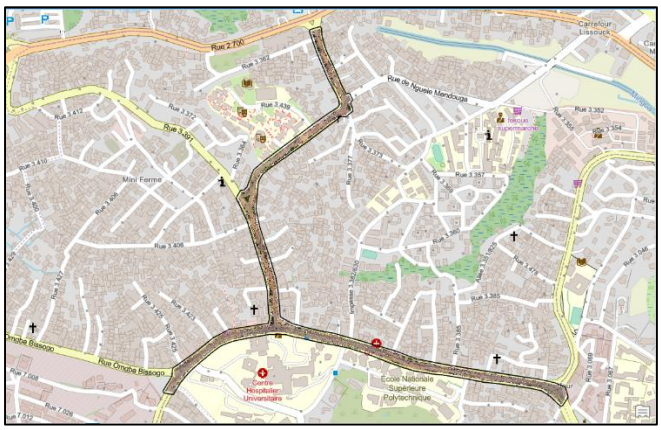


Figure 9: Elig-effa-Melen-EMIA road section orthomosaic

The model's output on this section provides a detailed inventory of detected pavement distresses. In total, the AI detected 438 instances of degradation along the 2.5 km section. These include various sizes of cracking, potholes, edge spalls, and areas of raveling. The detection mask centroids are georeferenced and mapped in Figure 10, which displays the spatial distribution of all defects across the section. Table 4 summarizes the count and total area for each class of defect identified on this road.

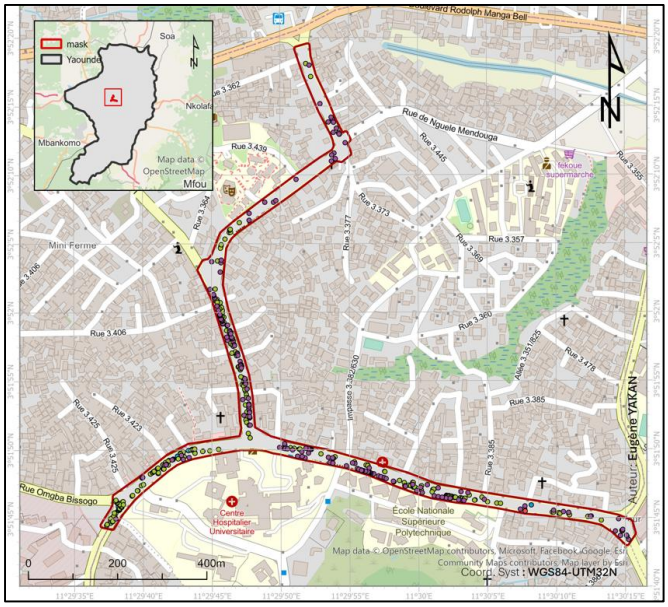


Figure 10: spatial distribution of distress across the whole road section

Table 4: Detection statistics

Classes	*Cracks (C1)*	*Potholes (C2)*	*Raveling (C3)*	*Edge Failures (C4)*	*Total*
Instances	7	225	204	2	**438**
Sum of Area (m²)	102.236	248.166	298.244	14.316	**662.964**

These figures indicate that raveling (surface material loss) and cracking are the most prevalent types of distress by area on this section, whereas potholes, though fewer in number, contribute significantly to the degraded area. The total degraded surface area is approximately 662.96 m², which is the sum of all detected polygons areas. This corresponds to a certain percentage of the road section's pavement area. This level of quantification is extremely useful for maintenance planning, it provides an objective measure of how much of the pavement needs repair, and of what type.

Each colored polygon in Figure 11 represents a detected defect: cracks in blue, potholes in red, raveling in green, and edge failures in orange. The background is the orthomosaic.

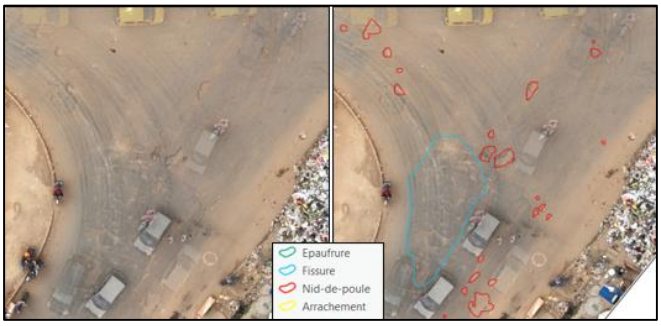


Figure 11: Example of detected road damages.

Notably, the processing efficiency of the system was very high. The entire inference and mapping workflow (tiling the orthomosaic, running the model on all tiles, georeferencing results, and aggregating to GIS) completed in just 5 minutes 47 seconds on a PC with 16Go of RAM and an NVIDIA GPU with 8Go of VRAM, for the 2.5 km road. In contrast, a traditional manual road survey of the same length, even a simplified windshield survey, might take a field team multiple days to cover and record all distresses, and still would lack the spatial precision of our method. This highlights the enormous time-saving potential of the approach. Analyses that used to require weeks of field work and data transcription can now be obtained in near real-time once the aerial images are captured. The rapid turnaround enables more responsive maintenance – for example, authorities could survey an entire city's roads in a few days and quickly identify emerging problem areas for intervention, rather than waiting for citizen complaints or periodic inspections.

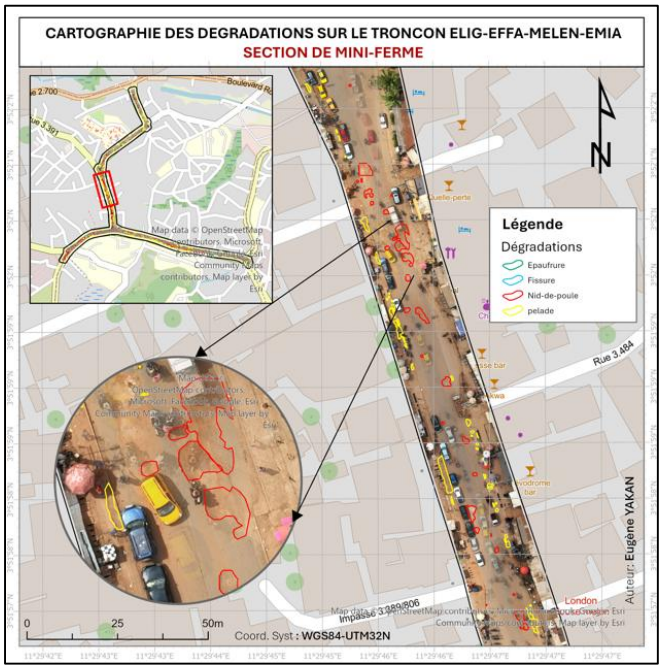


Figure 12: Detected Road damages map of the Mini-ferme road section

3.2.2 Spatial analysis and visualization

Beyond raw detection counts, we performed spatial analyses on the output to derive additional insights. We created damage density maps of the defects by calculating the centroid of each detected polygon and then mapping these points using each area field as weight. We generated the maps (Figure 13 and Figure 14), in which road stretches are color-coded (blue for low density, red for moderate, yellow for high). Figure 13 clearly highlighted that the Mini-Ferme stretch of the road was the most deteriorated, corroborating observations from the field that this area has long-standing maintenance issues. Such areas stand out on a density plot and can be prioritized for more detailed engineering evaluation or urgent repair.

The distress severity of road sections (6) might be evaluated by computing the percentage of surface area distressed in 100 m road intervals according to guidelines from pavement management systems like FHWA's Distress Identification Manual (2013). We can classify segments as low, medium, or high severity based on thresholds of area damaged.

$$\text{Severity Index (\%)} = \left(\frac{A_d}{A_t}\right) \times 100 \qquad (6)$$

- A_d: Total surface area of detected distresses in the segment (m²),
- A_t: Total surface area of the segment (m²)

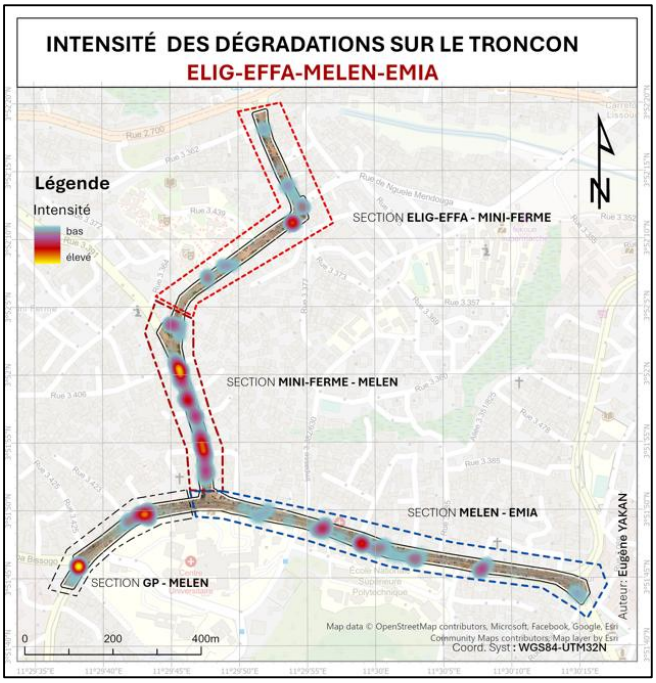


Figure 13: Damage density map of the whole road section

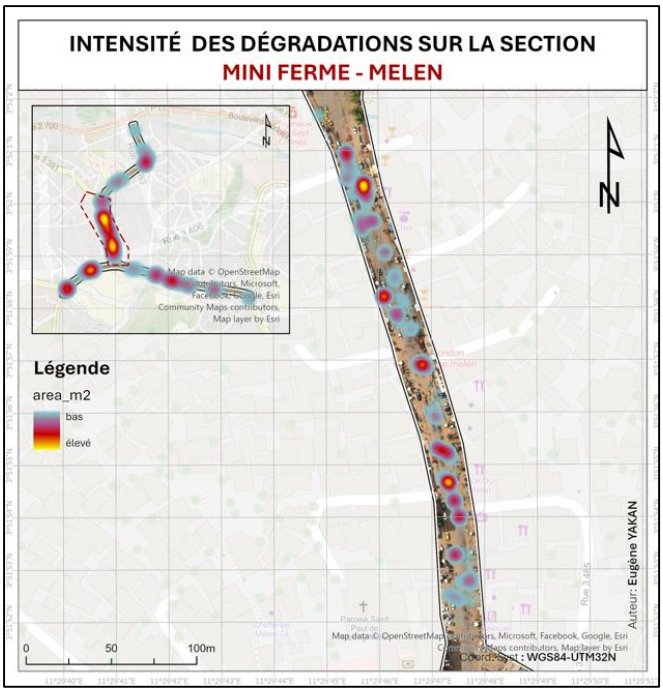


Figure 14: Damage density map Mini ferme -Melen road section

The georeferenced outputs were also loaded into an interactive dashboard (GeoDashboard). This interactive dashboard (Figure 15) allows users to pan/zoom the map of Yaoundé and see all detected defects. It includes filters like show only potholes, or only a particular survey date if multiple surveys are loaded and displays summary statistics like "Total crack area in view" or "Number of defects by type in this neighborhood." This kind of tool provides a dynamic way for decision-makers to query the data. After repairs are done, a new drone survey's results can be compared to the previous one to verify that defects were properly fixed or to see if new ones appeared.

Figure 15: GeoDashboard

3.2.3 Validation of results

As the Elig-effa-Melen-Carrefour EMIA section was rehabilitated during the study period, direct field surveys for validation were not feasible. Therefore, an alternative, scientifically justified validation strategy was adopted, based on comparing the results of our Python workflow with manual digitalization on the orthomosaic (Figure 16). This approach was supported by:

- **Error propagation theory**: Mathematically, if a surface $A = L \times W$, then differentiation gives $dA = LdW + WdL$. The error on the area ΔA thus propagates from errors on linear dimensions (ΔL, ΔW). Conversely, a small surface deviation necessarily implies that linear, and thus planimetric (X, Y), errors are even smaller. This property justifies the use of digitized areas as an indirect indicator of planimetric precision.

- **Data quality**: The orthomosaic used had a GSD of 2.3 cm/px and an RMSE of 0.01 m. Although slightly lower than Nguimbock's data (GSD = 1.83 cm/px; RMSE = 1.01 cm), these precision levels remain largely sufficient for digitizing road degradations, in accordance with literature standards.

Figure 16: Manually Digitized polygons and Python workflow generated polygons

While the visual detection and manual digitalization method revealed more degradations (457 degradations, 19 more) than those automatically detected and segmented, the comparison primarily focused on a subset of degradations detected by both

methods, totaling 25 degradations. Table 6 shows that the surface area deviations between the two methods remained generally small. In most cases, the absolute deviation did not exceed 10 cm², and the relative deviation remained below 3.66%. These results confirm that automated AI detection faithfully reproduces the surface areas of manually identified degradations, consistent with Nguimbock's (2020) observations.

Table 2: Difference in area between AI and Manual digitized polygons

Damages	**$(\Delta Area)_{mean}(m^2)$**
Cracks	0.06
Raveling	0.033
Potholes	0.035
Edge failures	0.057
Mean	**0.046**

It is worth noting that the manual mapping did find approximately 19 very small cracks that the AI missed (hence 457 vs 438 count). Many of these were hairline cracks at the limit of image resolution or in areas of heavy shadow where even the human mapper was uncertain. Some could be false positives on the human side. Nonetheless, from a maintenance perspective, missing a few tiny low-severity cracks is not critical, whereas identifying the major distresses correctly is more important. We are satisfied that the AI caught the vast majority of significant defects and that its quantification of those defects (areas) is reliable.

4. Conclusion

This study successfully validated an integrated approach combining drone photogrammetry and deep learning-based artificial intelligence for automated monitoring of paved road conditions in Cameroon. Focusing on a case study in Yaoundé, we demonstrated that high-resolution drone imagery, when processed through a tailored YOLOv11 model and automated GIS workflow, can rapidly and accurately detect common pavement degradations such as cracks, potholes, raveling, and edge failures. The approach transformed what is traditionally a manual, subjective process into a fast, objective, and reproducible one. Key achievements include the precise quantification of over 400 defects on a 2.5 km road segment and the generation of detailed maps and statistics within minutes of data processing. These results underscore a significant advancement in efficiency – effectively turning days of fieldwork into minutes of automated analysis – and highlight the scalability of the solution for city-wide deployment. While the outcomes are promising, further work is recommended to enhance and generalize the approach. First, integrating 3D analysis (using the dense point cloud or digital elevation models from drone data) could allow assessment of pothole depth or rut depth, providing a more comprehensive measure of degradation severity. This could involve coupling the current 2D model with a depth estimation module or using emerging techniques like volumetric change detection. Second, expanding the training dataset with more varied examples of road distresses (including those from different cities or regions) will improve the model's robustness across diverse conditions. Continuous learning could be implemented, where the model is periodically retrained with new data from subsequent surveys to capture new types or appearances of defects.

5. References

[1] Federal Highway Administration (FHWA). (2013). *Distress Identification Manual for the Long-Term Pavement Performance Program* (FHWA-HRT-13-092). McLean, VA: U.S. Department of Transportation.

[2] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788. https://doi.org/10.1109/CVPR.2016.91

[3] Silva, L. A., Leithardt, V. R. Q., López Batista, V. F., Villarrubia, G., & de Paz, J. F. (2023). Automated Road Damage Detection Using UAV Images and Deep Learning Techniques. *IEEE Access*, *11*, 62918–62931. https://doi.org/10.1109/ACCESS.2023.3287770

[4] Tan, Y., & Li, Y. (2019). UAV Photogrammetry-Based 3D Road Distress Detection. *ISPRS International Journal of Geo-Information*, *8*(9), 409. https://doi.org/10.3390/ijgi8090409

[5] Ritter, N., & Ruth, M. (1997). *The GeoTIFF data interchange standard for raster geographic images*. Open Geospatial Consortium (OGC).

[6] Woof, M. J. (2025, August 4). Cameroon is developing its road network. *Global Highways*. Retrieved from https://www.globalhighways.com/news/cameroon-developing-its-road-network.

[7] Nguimbock, D. S. (2020). *Utilisation de la photogrammétrie par drone pour le relevé des dégradations de la chaussée : cas du tronçon Nkolbisson–Leboudi à Yaoundé*. Engineer's thesis, École Nationale Supérieure des Travaux Publics, Yaoundé, Cameroon.