The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-G-2025
ISPRS Geospatial Week 2025 "Photogrammetry & Remote Sensing for a Better Tomorrow…", 6–11 April 2025, Dubai, UAE

Evaluation of a UAS-based Bridge Inspection Framework with Automated Damage Candidate Suggestion and Human-in-the-loop Damage Assessment

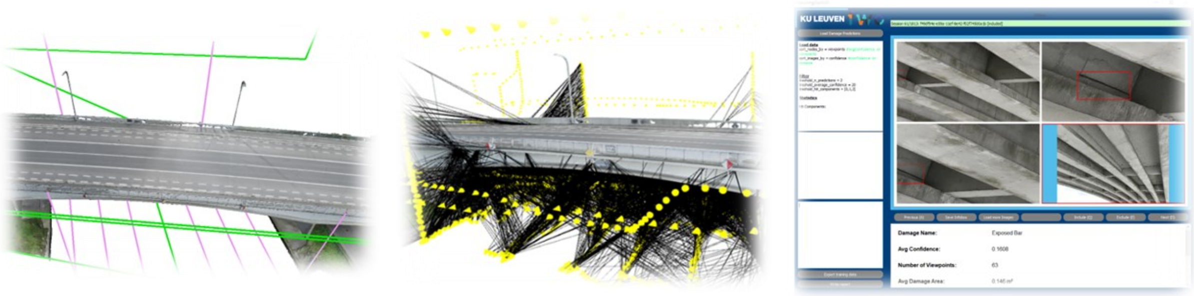
Erkki T. Bartczak[1*], Maarten Bassier[1] and Maarten Vergauwen[1]

[1] Department of Civil Engineering, Faculty of Engineering Technology, Geomatics Research Group, KU Leuven, Gebroeders De Smetstraat 1, B-9000 Gent, Belgium.
*erkkitobias.bartczak@kuleuven.be

Keywords: UAS, bridge inspection, automated damage detection, infrastructure monitoring, photogrammetry

ABSTRACT

Unmanned aerial systems have shown considerable potential to improve bridge inspection procedures by enhancing safety, efficiency, and data quality. However, developing a comprehensive system that integrates safe flight planning, data acquisition, automated damage detection, and reporting while addressing practical challenges remains complex. In this paper, we present a UAS-based bridge inspection pipeline that has been validated through a real-world case study. The system generates safe and efficient flight routes, which were closely followed by the UAS, achieving an RMSE of 0.67 m and ensuring successful camera alignment and precise photogrammetric 3D models with a mean distance error of 3.2 cm compared to terrestrial laser scans. The damage detection system predicts potential damages, maps them onto the 3D model, computes various characteristics, and aggregates the predictions into a manageable set of damage candidates. A human-in-the-loop validation via a graphical user interface refines these results, producing a verified damage report in less than 4.25 hours, providing accurate, actionable data for effective bridge maintenance. Our experimental results indicate that the proposed approach is both practical and effective for comprehensive bridge inspections. This paper systematically evaluates the system, highlights key strengths and limitations, and provides critical insights for future improvements.



a) b) c)

Figure 1: Visualization of the UAS-based bridge inspection framework. A) Generation of semi-automated flight paths, b) Ray-casting of damage detection bounding boxes onto the model and spatial filtering, c) Human-in-the-loop damage assessment in a graphical user interface.

1. Introduction

Regular inspections are crucial for ensuring the safety and longevity of bridge infrastructure. Over time, bridges face numerous challenges, including increased traffic loads, environmental exposure, and material degradation. Early detection of structural damages is essential to prevent catastrophic failures and to plan timely maintenance procedures.

Traditional manual inspection methods involve visual assessments by inspectors using equipment like scaffolding, under-bridge inspection vehicles, or rope access techniques. These methods present several challenges. They are time-consuming and labour-intensive, often requiring lane closures or traffic disruptions, which can lead to traffic impairment and increased safety risks for inspectors (Ichi und Dorafshan 2024). Accessibility to certain parts of the bridge, such as underdeck areas or high elevations, can be limited, potentially causing some damages to go unnoticed. Additionally, manual inspections are subjective and may lack consistency in damage detection and reporting.

To address these limitations, the use of Unmanned Aerial Systems (UAS) has been proposed to capture high-resolution images and videos of bridge structures, including hard-to-reach areas, without the need for extensive equipment setups or traffic disruptions. They enhance safety by keeping inspectors off the structure and reduce inspection times while providing comprehensive data for analysis (Rachmawati und Kim 2022).

Despite these benefits, current UAS-based inspection methods face challenges that restrict their widespread adoption. A major issue is the lack of integration across different stages of the inspection process, from flight planning over data analysis to damage assessment. Furthermore, automated damage detection models exhibit parameter-dependent performance, making their reliability inconsistent during inference, particularly in real-world conditions. Moreover, a practical alternative must accommodate for various bridge damage classes, analogue to conventional inspections. This study presents an integrated UAS-based bridge inspection pipeline that combines:

- Semi-automated data acquisition
- Comprehensive suggestion of damage candidates using deep learning object detection
- Damage mapping and spatial filtering
- Human-in-the-loop data assessment for final damage classification and reporting

While various studies have proposed components of UAS-based inspection, e.g., most prominently machine learning architectures for damage detection (Guo et al. 2024), few have demonstrated a cohesive system that can handle all stages from data acquisition to damage assessment in one integrated pipeline, for instance Lin et al. (2021). The workflow proposed in this paper is designed to overcome the fragmented nature of existing methods, offering a unified approach for obtaining, processing, and analysing bridge inspection data.

This pipeline addresses multiple challenges commonly encountered in UAS-based inspections. Flight path generation is optimized to support automated flight routes in GNSS-denied areas, while ensuring sufficient image overlap for subsequent photogrammetric reconstruction. Automated damage detection uses a trained deep learning model that is carefully validated to manage domain shifts between conventional inspection photographs and UAS-captured images. Additionally, by embedding spatial information into the damage predictions, inspectors gain immediate contextual understanding. The end-to-end approach is validated and thoroughly evaluated on a representative multi-span concrete bridge under realistic operational conditions.

The remainder of this paper is organized as follows**: Section 2** reviews relevant literature on UAS-based inspections, **section 3** details our proposed methodology and **section 4** presents the case study. **Section 5** presents the results of the study, including metrics for flight accuracy, photogrammetric model precision, and damage detection performance. **Section 6** provides a discussion of the implications and limitations of the findings, and **section 7** concludes with a summary of contributions and directions for future research.

2. Related work

The field of UAS-based bridge inspection has seen significant advancements in recent years, spanning multiple domains including flight path planning, photogrammetric reconstruction and automated damage detection frameworks.

The first critical component for realizing UAS assisted bridge inspections concerns the generation of safe, efficient and reliable UAS flight paths. Wang et al. (2022a) developed flight planning strategies optimized for efficient data acquisition and optimal image results during photogrammetric processing. However, the practical execution of such flight routes remains challenging, as this involves navigating in GNSS-denied environments, such under the deck. Various solutions have been proposed to address this issue, including visual odometry (Lu et al. 2024), SLAM, fiducial markers (Wang et al. 2023a), or the use of ultrasonic receivers (Kang und Cha 2018). Despite these advancements, the accuracy of UAS flight execution, particularly in these GNSS-denied environments, is rarely reported, leaving a gap in understanding how accurately these flights can be performed in such conditions.

Photogrammetric reconstruction using UAS imagery enables the creation of detailed 3D models of bridge structures, relying on Structure-from-Motion (SfM) algorithms to process overlapping images captured from various viewpoints. However, achieving accurate and comprehensive 3D reconstructions presents several challenges. Tang et al. (2024) evaluated the accuracy of a UAS-based photogrammetric model for a railroad bridge, using 20 manually placed Check Points (CPs), and reported a root mean square error (RMSE) of 1-2 cm. Similarly, Chen et al. (2019) reported a deviation of 0.2 - 3.2 cm when analysing UAS-derived 3D models to Terrestrial Laser Scans (TLS) in a cloud-to-cloud (C2C) comparison. Despite these results, reliance on Ground Control Points (GCPs) for georeferencing remains a significant limitation, as this process is time-consuming and labour-intensive, particularly for large or complex bridges (Tang et al. 2024). This effort can be drastically reduced by replacing the GCPs with the Real-Time Kinematic (RTK) positioning data from the UAS for georeferencing (Štroner et al. 2021). However, this approach has yet to be explored in bridge applications. Additionally, misalignment issues often occur during SfM processing, especially in underdeck areas, due to insufficient image overlap and inconsistent viewpoints (Wang et al. 2023b). Ensuring sufficient overlap between flight routes is critical for mitigating these issues.

Automated damage detection is essential for managing the large image and video data obtained by UAS-assisted bridge inspections. Deep learning models, such as Faster Region-Convolutional Neural Networks (Faster R-CNN), Single Shot Detector (SSD), and You Only Look Once (YOLO) variants, have been widely applied to bridge inspection tasks and have demonstrated high accuracy in controlled environments (Mittal et al. 2020; Liang et al. 2023). The CODEBRIM dataset includes common concrete damage classes such as exposed reinforcement bars, cracks, corrosion stains, efflorescence, and spalling, and has been used to train classification models with reported accuracies ranging from 0.8 to 0.9 for each damage class (Mundt et al. 2019). Similarly, Wang et al. (2022b) trained an object detection model to identify cracks and exposed bars with a mean average precision (mAP) of 0.832. However, these controlled settings often do not reflect the complexity of real-world bridge inspection imagery. In contrast, real-world UAS bridge inspection imagery typically contains smaller damages and highly variable backgrounds, such as trees, varying lighting conditions, and other environmental factors, all of which complicate the detection process. Sliding Window modules have been proposed to improve model performance, addressing the small object problem (Akyon et al. 2022). The discrepancy between training data real-world UAS inspection imagery is commonly referred to as the "domain gap," which typically leads to decreased performance when applying the same model (Fang et al. 2025). A thorough evaluation of the domain gap in bridge damage detection remains unaddressed.

Additionally, detection alone is insufficient for providing actionable insights to bridge managers. The ability to accurately map detected damages to a 3D model of the bridge is crucial for effective maintenance planning. In one of the most comprehensive studies to date, Lin et al. (2021) demonstrate that incorporating a damage mapping step significantly increases model performance, improving both the accuracy and usability of detection results. Despite advances in damage detection, further evaluation is needed to determine how integrating damage mapping and spatial filtering can improve performance and usability within a comprehensive bridge inspection framework.

Lastly, damage assessment and reporting require efficient methods to interpret and interact with the data. Several studies have advanced interactive tools for damage assessment to bridge this gap. For example, Tang et al. (2024) demonstrated how detailed, textured photogrammetric models can be used to identify missing bolts and corrosion in steel bridges. Similarly, Seo et al. (2018) used the photogrammetric model and compared the results to a conventional inspection report, finding that the UAS approach showed comparable overall quality and even identified additional previously missed damages. However, this process was carried out manually, pointing to a limitation of current methods that require further automation. In contrast, Lin et al. (2021) presented an interactive web-based viewer for bridge inspection that allows inspectors to verify automated damage detections using both 2D and 3D data. The system includes advanced features such as measurement tools and overlay views of different inspection epochs, allowing inspectors to observe damage progression over time to assess severity. These advancements highlight the need to evaluate and refine interactive damage assessment tools to determine the optimal toolset for large-scale bridge inspections.

3. Methodology

In this study, we propose a comprehensive methodology for UAS-based bridge inspection that begins with flight path generation, data acquisition, photogrammetric reconstruction, a complex damage detection and mapping system and final human-in-the-loop data assessment.

The UAS flight routes are generated based on the known 3D trajectory of the bridge and the height of the starting point. First, an overview flight is conducted to capture the bridge from both sides. The underdeck inspection flight path includes vertical maneuvers at the start and endpoints of each underdeck route, aligning these positions with those captured in the overview flight. This alignment ensures that images from both the underdeck pass, and the overview flight can be accurately matched for reliable 3D model construction. Because GNSS signals are often obstructed underneath the bridge, the UAS relies on its inertial measurement unit (IMU) and 360° obstacle avoidance system to maintain stable flight in these areas. Upon exiting the underdeck, the UAS reestablishes its RTK-GNSS connection and readjusts its trajectory accordingly. To verify the accuracy of this approach, the camera positions obtained in the subsequent image alignment step are compared against the planned path.

After data acquisition, Agisoft Metashape v2.0 is used to process all collected images and generate a 3D model of the bridge. Underdeck images are aligned through common visual features only, as their RTK data are discarded due to signal obstruction, whereas the georeferenced overview flight images retain their RTK coordinates for precise georeferencing. Combining these datasets enables successful and accurate camera alignment. The final mesh is simplified to 200,000 vertices, balancing detail with computational efficiency and textured using ten 8192×8192 texture maps. Photogrammetric accuracy is assessed via a C2C comparison against TLS data, offering an objective measure of the model's accuracy. The model is manually segmented into components such as the road surface, installations (e.g., streetlights, traffic signs), ground surface, and underdeck sections. Additionally, the camera positions and undistorted images are exported from the 3D model to facilitate subsequent damage detection.

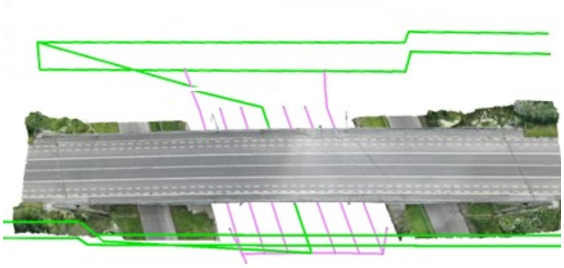


Figure 2: Flight path generation. The underdeck inspection flight (pink) passes underneath the bridge and connects to the overview flight containing RTK-GNSS data.

Because damage detection in real-world environments can be challenging due to variations in illumination, scale, and structural configuration s, detecting exposed reinforcement bars is prioritized in this work. These damages are typically easier to identify than fine cracks, making them well-suited for developing and testing an automated detection pipeline. A YOLOv8 (Jocher et al. 2023) object detection model was thus trained on a curated dataset of 2,000 documentation images from conventional bridge inspections (0.4 MP to 48 MP resolution), each manually annotated for exposed rebars. Various metrics, including bounding box frequencies and sizes, guided the selection of 600 images for validation to ensure the validation images represent the training dataset. Transfer learning was applied using the weights from pretraining on the COCO dataset and **Table 1** summarizes the hyperparameters used during training.

To further analyse the performance of the trained model and to quantify potential domain gaps, additional experiments were conducted on external sets of images. A subset of 43 images was specifically curated from the CODEBRIM dataset regarding exposed bars, excluding cases with excessively large or centrally located damages. A subset of 18 UAS-based inspection images was also annotated to refine inference parameters for complex under-bridge scenes. By comparing the results from these small but diverse datasets, we evaluate the capabilities of the model in various scenarios (**Figure 3**).

Following object detection, each bounding box prediction is projected onto the 3D model via a raytracing procedure. This process captures both geometric attributes (e.g., 3D location, approximate damage area, confidence) and semantic data (e.g., damaged bridge component). The projected points on the mesh are used to calculate damage areas and to create 3D bounding boxes, while averaged measurements (e.g., damage center location, surface normal) are derived from multiple images that capture the same damage from different viewpoints. These aggregated data structures, combined with essential image information (e.g., file names, path, intrinsic matrix), are stored in a Resource Description Framework (RDF) g raph to maintain consistency and allow complex querying (**Figure 4**).

Table 1: Training hyperparameters

Parameter	Value	Description
epochs	400	Maximum number of training epochs
batch	Auto	Batch size automatically adjusted based on GPU memory availability (trained on NVIDIA GeForce RTX4070 Ti)
imgsz	1024	Image size (width and height); images are resized/padded to 1024×1024 for training
optimizer	Auto	Automatically selects between SGD and Adam based on dataset characteristics
momentum	0.937	Momentum value for the optimizer to stabilize and accelerate training
weight_decay	0.0005	Weight decay (L2 regularization) used to mitigate overfitting
iou	0.7	Intersection-over-Union threshold for non-maximum suppression

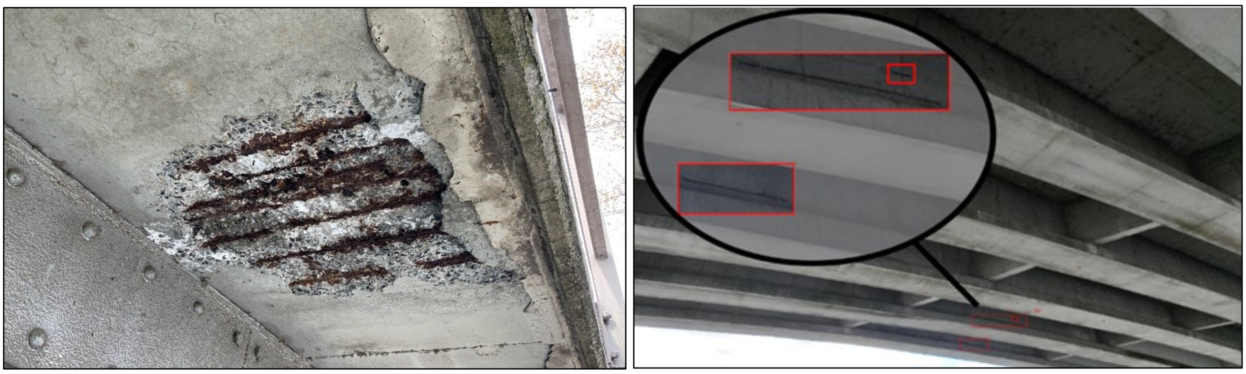


Figure 3: Domain gap example images. The quality of the typical training images (left) and the UAS inspection images (right) differs drastically.

In the final damage assessment step, the RDF graphs provide the basis the human-in-the-loop evaluation. Detected damages are visualized in a user interface where each damage node is visualized through multiple images including bounding boxes at the same time, offering a comprehensive understanding (**Figure 1c**). Additionally, the interface shows further information about the damage candidate in a text box which permits manual edits, such as adding notes or updating the review status. By coupling quantitative model outputs with expert oversight, the approach ensures robust data validation and reliable final assessments suitable for practical bridge maintenance applications.

4. Case study

We demonstrate the feasibility of the proposed pipeline through a real-world inspection of the Merendreebrug, a three-span concrete bridge near Ghent, Belgium (**Figure 5**). Measuring approximately 130 m in total length with a 60 m main span, the structure features seven I-girders under a 20 m-wide roadway. Multiple traffic participants, including cars on the deck and ships navigating underneath, introduce additional operational complexity, in addition to the car and bicycle traffic at the abutments. This environment underscores the need for careful flight planning to ensure safety, maintain compliance with EU regulations, and capture comprehensive imagery of critical underdeck areas. In this specific case study, no special flight permit was required, as the operation was conducted under the Open category’s A2 subcategory, in accordance with *Commission Implementing Regulation (EU) 2019/947.* The UAS maintained a minimum horizontal distance of 5 meters from uninvolved persons by operating in low-speed mode. However, the deployment of UAS-based inspection pipelines in other locations may require compliance with regional aviation and infrastructure safety regulations. Additionally, the chosen bridge involves previously documented issues, such as exposed bars, spalling, corrosion stains, delamination, and corroded drainage pipes, highlighting the importance of capturing high-resolution imagery across difficult-to-reach underdeck areas.

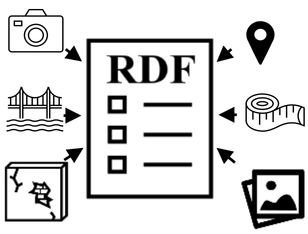


Figure 4: Content of final RDF graph. Various data types are linked and stored in one file.

A DJI Mavic 3 Enterprise UAS equipped with a 4/3″ CMOS camera, RTK-GNSS module, and omnidirectional obstacle avoidance was selected to perform the inspection flights. The underdeck flight route included 7 passes beneath the deck at approximately 1 m/s with a clearance of approx. 5 m to the structure and 2 m above the water surface. When ship traffic appeared, the flight was paused, and the UAS was maneuvered upward until the waterway was clear. To ensure flight safety, the pilot should be positioned in a location that provides a clear line of sight to monitor approaching traffic and assess potential hazards. Underneath the main span, the loss of GNSS signal caused notable flight instabilities especially in altitude, which reduced once the UAS exited the underdeck area and reestablished RTK connectivity.

The camera was controlled manually to allow for maximum control and to capture the faces of the I-girders at slight lateral angles, with a maximum pitch of 30° upward (**Figure 7**). Under sunny conditions, the overview flight was configured with automatic exposure at a target shutter speed of approximately 1/1000 s at a flight speed of 3 m/s. By contrast, a manual setting of 1/640 s, ISO 800, and f/4 was used during the underdeck route to achieve bright and sharp imagery at 1 m/s; although parts of the exterior became overexposed, these areas were sufficiently covered by the overview flight. To maximize overlap and coverage consistency, images were recorded at 0.7-second intervals. In addition, handheld images were collected around the abutments where vehicular and pedestrian activity prohibited close-in flight operations; the operator walked a path analogous to the underdeck trajectory, using similar camera settings to maintain consistent exposure and data quality. **Table 2** summarizes the total flight time and number of images acquired for each route.

Figure 5: Field work at the Merendreebrug. The image illustrates the complexity of the I-girder cross-section.

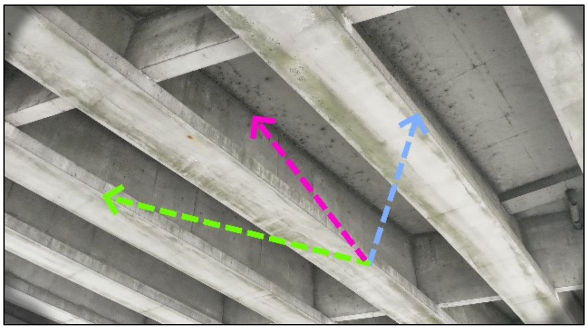


Figure 7: Perspective of inspection images. To improve coverage, the camera angle (pink) was directed in an angle from the flight direction (green).

Parallel to the aerial data collection, a TLS survey was performed using a Leica P30 scanner at 22 setups, taking about 4.5 hours on site.

5. Results

This section presents the results of the experiments regarding the flight path accuracy, geometric accuracy of the resulting photogrammetric point cloud and the performance of the automated damage detection and mapping system.

The comparison of the camera positions against the planned flight route (**Figure 6**) shows for the overview flight a mean deviation of 0.16 m horizontally (XY) and 0.05 m vertically (Z), corresponding to an overall mean deviation (XYZ) of about 0.17 m and an RMSE of 0.21 m, which is higher than the expected 0.03 m accuracy of the RTK module. In contrast, the underdeck flight showed mean deviations of roughly 0.38 m in XY and 0.25 m in Z, resulting in a 0.51 m mean XYZ deviation and an overall RMSE of 0.67 m. Maximum deviations in the underdeck route extended beyond 2.1 m horizontally and 1.6 m vertically, particularly when RTK signals were lost beneath the bridge and toward the end of the flight route. Moreover, the manually entered starting point altitude for flight route generation may have introduced a baseline offset in the vertical dimension.

The photogrammetric reconstruction was performed using medium-quality settings for image alignment and dense point cloud generation. **Table 3** summarizes the principal processing steps, with the creation of depth maps requiring 10.5 hours of the total 28.6 hours. To focus on the main structure, regions not captured in the TLS survey, such as road surface as well as the ground area, were removed manually. Subsequently, the SfM model was coarsely aligned to the TLS dataset using known global coordinates from a set of Leica targets. This was followed by an iterative fine-alignment process, during which corresponding points in both datasets were identified and matched within cross section sub-clouds. Once aligned, the photogrammetric cloud was compared to the TLS reference, yielding a mean deviation of approx. 3.2 cm. Only a few localized regions, particularly smaller compartments in the underdeck, exceeded 6 cm. **Figure 9** illustrates the distribution of these deviations, indicating that most points remain well below this threshold.

The trained YOLO models were initially evaluated on the CODEBRIM validation subset to identify the best-performing configuration under optimal conditions. A series of Bayesian hyperparameter sweeps determined that a model setup (YOLOv8n at 1000 px input size) achieved an F1-score of 0.70 at a confidence threshold of 0.11 (**Figure 10**). Building on these findings, the selected model was then applied to 18

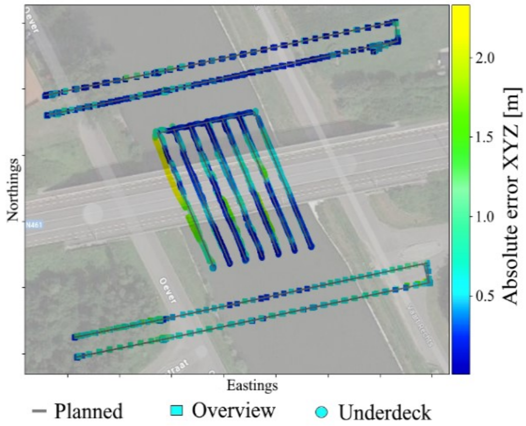


Figure 6: Accuracy of executed flight routes. The color code shows deviations from the planned flight route.

Table 2: UAS data acquisition

Flight route	Images	Flight time
Overview	346	13
Underdeck	1282	19
Abutment 1	1247	12*
Abutment 2	1263	14*

*handheld

Table 3: Photogrammetric processing

Processing step	Setting	Time[h]
Camera alignment	Medium	6
Depth map generation	Medium, Aggressive filtering	5
Dense point cloud	120 M points including depth maps	10.5
Meshing	Medium, 13 M faces, based on depth maps	1
Simplify model	200 000 faces	0.1
Texture	10 x 8192 x 8192	6
	=	**28.6**

Table 4: Fine-tuned inference parameters on UAS inspection images

Model	Parameters	Recall
Yolov8n_1000px	Image Size: 3000	0.422
Yolov8n_1000px + SAHI	slice_width = 1000 slice_height = 480	0.375

Figure 8: Inference on inspection images.

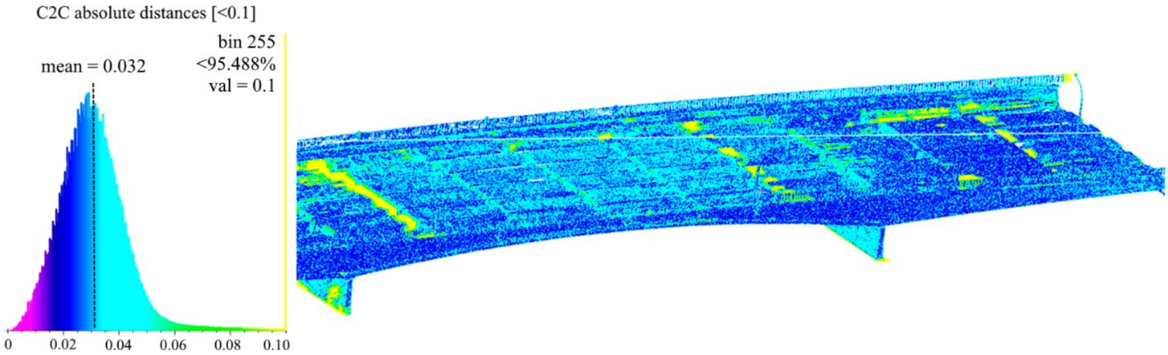


Figure 9: SfM-to-TLS comparison. The mean distance of the photogrammetric point cloud to the TLS is 3.2 cm. The histogram shows that only few areas, especially close to the cross-bracings show a higher deviation than 6 cm.

representative images from the UAS inspection flight to fine-tune parameters for optimizing Recall at a confidence threshold of 0.1. Several image input sizes were tested between 224 and 4000 px, including variants aided by the SAHI module. The best results from this experiment are summarized in **Table 4**. **Figure 8** illustrates that optimizing for high Recall resulted in a substantial number of False Positives (FP).

From the 1,282 underdeck images, the detection model produced a total of 16,822 bounding box predictions. After projecting onto the 3D model, 2,194 did not intersect with the mesh. Additionally, 1,972 predictions were mapped onto the road surface, 1,594 onto the ground, and 252 onto various installations. Excluding these irrelevant associations resulted in a set of damage candidates on the underdeck and girder regions. The remaining detections were consolidated by their centre point location within a 0.3 m radius, to 1,813 combined damage candidate nodes.

In the final assessment step, the damage candidate nodes were loaded into the graphical user interface, where they could be quickly labelled as “included” or “excluded” by updating each node’s revision status attribute through a key-binding shortcut. The final confirmation step took approximately 4.0 hours and reduced the set to 66 nodes. In a subsequent pass, the viewer was configured to display only these included nodes, allowing inspectors to refine their classification and add notes where relevant. This second step required an additional 0.25 hours and produced 24 exposed bars, 16 corrosion spots, 9 likely cases of concrete delamination, 8 cases of spalling without visible reinforcement, 5 cracks and 4 instances of drainage damage (two flagged as ‘uncertain’ for further investigation) as illustrated in **Figure 11**. By consolidating the damage candidates and validating each damage node, the pipeline produced a concise, actionable list of critical defects that can inform subsequent maintenance decisions.

6. Discussion

In the following discussion, we discuss our key findings and reflect on their implications and limitations.

The flight route analysis indicates that the results may suffer from a constant offset due to an inaccurately selected starting point height during route generation. This systematic error suggests that the UAS followed the generated route more precisely than the data implies. To address the maximum observed deviations, a minimum safety clearance of 2 m from the bridge is recommended for the underdeck flights. Moreover, while the overall RMSE of 0.67 m for the underdeck flight may seem high, it compares reasonably with recent findings. For example, Wang et al. (2023a) achieve an RMSE of 0.3 - 0.4 m using fiducial markers for stereo visual-inertial localization, which requires customized UAS and the setup of fiducial markers. Although the flight inaccuracy did not compromise subsequent tasks in this study, scenarios that demand higher positional accuracy, such as attaining a lower GSD for detecting fine concrete cracks, may require more precise flight maneuvers.

The C2C comparison between the SfM dense point cloud and the TLS reference yielded a mean distance of 3.2 cm, consistent with centimetre-level accuracies reported in the literature (Mohammadi et al. 2021; Abdel-Maksoud 2024). However, the limited literature and varying measurement methods indicate a need for further research. Due to a safety distance to the pillars, the flight routes resulted in a lack of coverage in the first compartment between the cross-bracings. It is advisable to

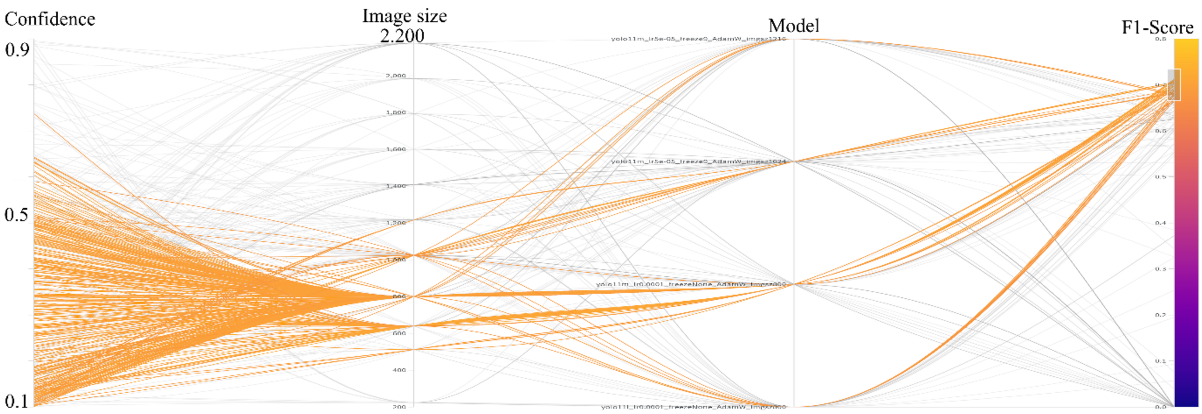


Figure 10: Sweep in Weights and Biases. Evaluating multiple model variations on the CODEBRIM subset using various input parameters during inference (Weights & Biases, 2022).

A) Exposed bars | B) Corrosion spots | C) Delamination | D) Spalling | E) Concrete Cracks | F) Drainage | G) FPs

Figure 11: Final damage assessment results. The images are cropped out directly from the bounding box predictions and illustrate various relevant bridge damage classes.

acquire additional handheld images in these regions to address potential coverage gaps. Finally, replacing the approximate camera calibration with a fully calibrated model could further enhance geometric accuracy while reducing computational costs.

The modest improvement in F1-score (+0.1) of the selected YOLOv8n model over its closest variants on the CODEBRIM validation subset may appear marginal under controlled conditions. However, its superior generalization became evident when applied to high-resolution UAS inspection images, where other models exhibited significant drops in Recall. This underscores the need to optimize not only the training hyperparameters but also to align inference parameters to overcome the domain gap problem for bridge defects detection. Using a low confidence threshold resulted in a large number of FPs, but also enabled the detection of damages the model was not explicitly trained on. The subsequent aggregation step drastically reduced the candidate nodes for review, though further filtering based on factors such as the number of viewpoints or implausible estimated damage areas could further reduce manual effort. Moreover, the RDF structure permits targeted searches for damages in specific image regions. Selecting image with optimal viewing angles and using known working distances to optimize inference parameters during a second prediction step could further enhance prediction results. Similarly, an additional classification stage on the bounding box predictions appears promising for future work. Such a classifier could differentiate between damage classes and reduce FPs by identifying healthy bridge areas, a particularly valuable feature given that concrete joints, dirt, and spider webs were frequently misclassified as exposed bars in this study.

The results of the damage assessment step demonstrate the value of combining automated detection with targeted human supervision. Although the automated pipeline efficiently scanned thousands of images and generated numerous candidate damage nodes, human review proved effective for filtering FPs and refining the detections. This approach aligns with the *EU AI Act's*[1] emphasis on maintaining human responsibility in critical decision-making processes. Notably, the entire review process required only 4.25 hours, indicating that an initial broad-scale automated screening followed by focused manual validation can be both efficient and sufficiently rigorous for practical bridge maintenance. However, the current study did not compare the results of the system to the latest conventional inspection report yet, which is an important extension for future investigations. This study also revealed opportunities to enhance the user interface. Future iterations should consider integrating a high-resolution 3D model to help identify damages that are difficult for the detection model to generalize, e.g. large spalling. Allowing users to add new annotations during the review process and enabling manual combination of damage nodes regardless their 3D location could further enhance the workflow and improve the final damage inventory.

7. Conclusion

This study presents an integrated UAS-based bridge inspection pipeline that combines flight path planning, semi-automated flight execution, photogrammetric reconstruction, automated damage detection, and human- in-the-loop assessment. Demonstrated on the Merendreebrug near Ghent, Belgium, the system achieved a photogrammetric accuracy within 3.2 cm relative to TLS data and an RMSE of 0.67 m in flight path tracking under challenging underdeck conditions. Although optimizing for recall in damage detection led to a high rate of false positive predictions, subsequent spatial grouping and expert review enabled the rapid identification of critical defects within 4.25 hours. These results suggest that UAS-based methods can enhance the efficiency, safety, and quality of bridge inspections.

While these findings are promising, they derive from a single case study and indicate the need for further testing across diverse bridge types and environments. Future work should focus on refining flight planning and hyperparameter settings, incorporating secondary classification techniques to further reduce false positive predictions, and exploring adaptive strategies to improve detection accuracy. Overall, this study contributes to ongoing efforts in civil engineering to leverage technology-driven approaches for infrastructure monitoring, highlighting the importance of integrating automated processes with expert oversight to achieve robust, actionable outcomes.

Author Contributions: E.T.B. is the main author of the work and conceived the method. M.B. is the direct supervisor. M.V. is the supervisor. All authors have read and agreed to the published version of the manuscript.
Conflicts of Interest: There are no conflicts of interest.

[1] European Commission. (2021, April 21). *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*

References

Abdel-Maksoud, Hany (2024): Combining UAV-LiDAR and UAV-photogrammetry for bridge assessment and infrastructure monitoring. *Arab. J. Geosci*. DOI: 10.1007/s12517-024-11897-5

Agisoft LLC. (2022): Agisoft Metashape Professional, Version 2.0. https://www.agisoft.com

Akyon, Fatih Cagatay; Altinuc, Sinan Onur; Temizel, Alptekin (2022): Slicing Aided Hyper Inference and Fine-tuning for Small Object Detection. *Computer Vision and Pattern Recognition*. DOI: 10.1109/ICIP46576.2022.9897990

Chen, S; Laefer, D. F.; Mangina, E.; Zolanvari, S. M. I.; Byrne, J. (2019): UAV Bridge Inspection through Evaluated 3D Reconstructions. *Journal of bridge engineering*. DOI: 10.1061/(ASCE)BE.1943-5592.0001343)

Fang, Xiang; Easwaran, Arvind; Genest, Blaise; Suganthan, Ponnuthurai Nagaratnam (2025): Your data is not perfect: Towards cross-domain out-of-distribution detection in class-imbalanced data. *Expert Systems with Applications*. DOI: 10.1016/j.eswa.2024.126031

Guo, Jingjing; Liu, Pengkun; Xiao, Bo; Deng, Lu; Wang, Qian (2024): Surface defect detection of civil structures using images: Review from data perspective. *Automation in Construction*. DOI: 10.1016/j.autcon.2023.105186

Ichi, Eberechi; Dorafshan, Sattar (2024): Evaluation of Infrared Thermography Dataset for Delamination Detection in Reinforced Concrete Bridge Decks. *Applied Sciences*. DOI: 10.3390/app14062455

Jocher, G.; Chaurasia, A.; Qiu, J. (2023): YOLOv8: Ultralytics Next-Generation Object Detection and Segmentation Model. https://github.com/ultralytics

Kang, Dongho; Cha, Young-Jin (2018): Autonomous UAVs for Structural Health Monitoring Using Deep Learning and an Ultrasonic Beacon System with Geo-Tagging. *Computer-Aided Civil and Infrastructure Engineering*. DOI: 10.1111/mice.12375

Liang, Han; Lee, Seong-Cheol; Seo, Suyoung (2023): UAV-Based Low Altitude Remote Sensing for Concrete Bridge Multi-Category Damage Automatic Detection System. *Drones*. DOI: 10.3390/drones7060386

Lin, Jacob J.; Ibrahim, Amir; Sarwade, Shubham; Golparvar-Fard, Mani (2021): Bridge Inspection with Aerial Robots: Automating the Entire Pipeline of Visual Data Capture, 3D Mapping, Defect Detection, Analysis, and Reporting. *Journal of Computing in Civil Engineering* DOI: 10.1061/(ASCE)CP.1943-5487.0000954

Lu, Bing-Xian; Tsai, Yu-Chung; Tseng, Kuo-Shih (2024): GRVINS: Tightly Coupled GNSS-Range-Visual-Inertial System. *Journal of Intelligent & Robotic Systems*. DOI: 10.1007/s10846-023-02033-8

Martin Mundt; Sagnik Majumder; Sreenivas Murali; Panagiotis Panetsos; Visvanathan Ramesh (2019): CODEBRIM: COncrete DEfect BRidge IMage Dataset. *Computer Vision and Pattern Recognition*. DOI: arxiv.org/abs/1904.08486

Mittal, Payal; Singh, Raman; Sharma, Akashdeep (2020): Deep learning-based object detection in low-altitude UAV datasets: A survey. *Image and Vision Computing*. DOI: 10.1016/j.imavis.2020.104046

Mohammadi, Masoud; Rashidi, Maria; Mousavi, Vahid; Karami, Ali; Yu, Yang; Samali, Bijan (2021): Quality Evaluation of Digital Twins Generated Based on UAV Photogrammetry and TLS: Bridge Case Study. *Remote Sensing*. DOI: 10.3390/rs13173499

Rachmawati, T.S.N.; Kim, S. (2022): Unmanned Aerial Vehicles (UAV) Integration with Digital Technologies toward Construction 4.0: A Systematic Literature Review. *Sustainability*. DOI: 10.3390/su14095708

Seo, Junwon; Duque, Luis; Wacker, Jim (2018): Drone-enabled bridge inspection methodology and application. *Automation in Construction*. DOI: 10.1016/j.autcon.2018.06.006

Štroner, Martin; Urban, Rudolf; Seidl, Jan; Reindl, Tomáš; Brouček, Josef (2021): Photogrammetry Using UAV-Mounted GNSS RTK: Georeferencing Strategies without GCPs. *Remote Sensing*. DOI: 10.3390/rs13071336

Tang, Zhiyuan; Peng, Yipu; Li, Jian; Li, Zichao (2024): UAV 3D Modeling and Application Based on Railroad Bridge Inspection. *Buildings*. DOI: 10.3390/buildings14010026

Wang, Feng; Zou, Yang; Del Rey Castillo, Enrique; Ding, Youliang; Xu, Zhao; Zhao, Hanwei; Lim, James B.P. (2022a): Automated UAV path-planning for high-quality photogrammetric 3D bridge reconstruction. *Structure and. Infrastructure Engineering*. DOI: 10.1080/15732479.2022.2152840

Wang, Feng; Zou, Yang; Zhang, Cheng; Buzzatto, Joao; Liarokapis, Minas; Del Rey Castillo, Enrique; Lim, James B.P. (2023a): UAV navigation in large-scale GPS-denied bridge environments using fiducial marker-corrected stereo visual-inertial localisation. *Automation in Construction*. DOI: 10.1016/j.autcon.2023.105139

Wang, Wenjun; Su, Chao; Fu, Dong (2022b): Automatic detection of defects in concrete structures based on deep learning. *Structures*. DOI: 10.1016/j.istruc.2022.06.042

Wang, Xuguang; Demartino, Cristoforo; Narazaki, Yasutaka; Monti, Giorgio; Spencer, Billie F. (2023b): Rapid seismic risk assessment of bridges using UAV aerial photogrammetry. *Engineering Structures*. DOI: 10.1016/j.engstruct.2023.115589

Weights and Biases (2020). Experiment Tracking with Weights and Biases. Available from https://wandb.com