DEEP LEARNING AND REALITY CAPTURE FOR AUTOMATED RAILWAY BRIDGE INSPECTION WITH BIM ENRICHMENT

Rafael Cabral[1,2], Ricardo Santos[1,2], Diogo Ribeiro[1,2]
[1] iBuilt, School of Engineering, Polytechnic of Porto, Portugal
[2] CONSTRUCT, Faculty of Engineering, University of Porto, Porto, Portugal

Abstract

Manual inspection remains subjective and poorly integrated with digital workflows. Although computer vision enables automated damage detection, most approaches fail to translate observations into parametric Building Information Modelling (BIM). For bridges, this limitation persists within Bridge Information Modelling (BrIM), where geometry, semantics, and lifecycle information must align. This paper addresses the gap through a four-stage framework: optimal reality modelling, hybrid damage perception, automated Scan-to-BIM, and semantic enrichment via IFC objects. Validated on operational railway bridges, the method achieves a damage-detection mAP50 of 0.621 and geometric deviations under 23 mm, enabling BrIM generation directly from visual data.

Introduction

Railway bridges are critical infrastructure assets whose safety and serviceability depend on regular condition assessment. This requirement is becoming increasingly pressing as the European railway network ages; currently, more than 35% of railway bridges in the European Union are over 100 years old, with many approaching the end of their design service life (Tomor, 2013). Current inspection practice relies predominantly on manual visual surveys, which are inherently subjective, difficult to reproduce, and constrained by access, safety, and time limitations. Large-scale studies have demonstrated substantial inter-inspector variability in condition ratings, directly affecting the reliability of bridge management systems (BMS) and long-term maintenance planning (Moore et al., 2001).

To address these limitations, recent progress in reality capture technologies, specifically Unmanned Aerial Vehicles (UAVs), Terrestrial Laser Scanning (TLS), and Mobile Laser Scanning (MLS), has enabled the rapid acquisition of high-resolution geometric and visual data (Rakoczy et al., 2025, Olaszek et al., 2024). In parallel, deep learning techniques have achieved remarkable performance in detecting and segmenting structural damage from imagery (Ribeiro et al., 2025). However, a significant discontinuity remains between detection and digitalization.

Recent reviews on Digital Twin implementation (Nichols et al., 2025) indicate that while visual data is abundant, it is rarely used to update the engineering models serving as basis for structural analysis. Most inspection pipelines end at the level of 2D images or unstructured point clouds, failing to translate perception results into structured, parametric, and computable Building Information Modelling (BIM) representations (Yang et al., 2025). Within the bridge domain, this paradigm is referred to as Bridge Information Modelling (BrIM), extending BIM principles to infrastructure assets. This gap limits BrIM adoption for automated Operation and Maintenance (O&M), as current Scan-to-BIM approaches struggle to generalize across complex typologies without substantial manual intervention (Patil and Kalantari, 2025).

Unlike prior works that focus on isolated tasks, this work presents an end-to-end methodology for generating "as-damaged" BIM models directly from multimodal reality capture data. The framework is validated by a reinforced concrete railway bridge study located in Porto, Portugal.

Related works

Bridge inspection has traditionally relied on manual visual surveys that are time-consuming, subjective, and difficult to standardize (Kruachottikul et al., 2021). Recent research focuses on using reality capture, deep learning, and BIM/BrIM to improve efficiency and data quality along different parts of the inspection workflow.

UAV photogrammetry and terrestrial laser scanning (TLS) are widely used to generate accurate 3D as-is models of bridges, often by combining multiple sensors to overcome occlusions and access limitations (Mirzazade et al., 2023). For railway bridges, multimodal TLS–UAV fusion -has been applied to create precise 3D as-is models and compare them to as-designed models for geometric deviation assessment (Cabral et al., 2023). Other studies use 3D point clouds for automatic bridge component recognition to support Scan-to-BIM of existing bridges (Xu et al., 2026).

Numerous works propose CNN-based systems to detect, segment, and classify multiple defect types in bridge images, including cracks, corrosion, rust, spalling, and other surface damage . Advanced networks such as BridgeNet for multiclass damage segmentation (Huang et al., 2024), BGInet for bi-task bolt and rust detection from UAV imagery (Xu et al., 2025), and attention-enhanced YOLO variants for RC bridges (Ruggieri et al., 2025) achieve high detection accuracy and real-time performance. Several frameworks also address defect quantification, for example mapping image pixels to real-world dimensions or using orthophotos for crack measurements (Xu et al., 2025, Ayele et al., 2020).

A growing body of work links inspection data with BIM or digital twins to support asset management. Some studies integrate corrosion or surface defect measurements into BIM models to manage long-span or RC bridges more efficiently (Hattori et al., 2024, Cavieres-Lagos et al., 2024). Others develop BIM-based mixed-reality tools that overlay inspection information on 3D models to support on-site or remote bridge inspection and maintenance (Riedlinger et al., 2022, Zakaria et al., 2023).

A smaller set of studies tackles partial automation from vision-based damage detection to 3D localization and BIM linkage. Methods combining deep learning with Structure-from-Motion or reality capture integrate defect detections into 3D bridge models, reducing variability between images and enabling spatial recording of damage (Yamane et al., 2023). However, these workflows typically address either geometry reconstruction and component labelling (Kim et al., 2020) or damage detection and 3D localization (Yamane et al., 2024), rather than delivering a fully integrated, as-damaged BrIM that unifies multimodal capture, automated perception, Scan-to-BrIM, and semantic enrichment in a single, railway-bridge-specific framework.

Methodology

The proposed framework establishes a computational pipeline that connects raw sensor data to semantic BrIM models. As illustrated in *Figure 1*, the architecture is composed of four sequential modules: i) Optimal Reality Modelling; ii) Hybrid Damage Perception; iii) Automated Scan-to-BIM; and iv) Semantic BIM Enrichment via ray-casting and parametric object instantiation.

Optimal reality modelling

The first stage aims to generate high-fidelity 3D reality models suitable for geometric inference. Railway bridge environments are characterized by complex backgrounds, e.g. vegetation, catenary systems, and varying terrain, which introduce significant noise into both photogrammetric and LiDAR datasets.

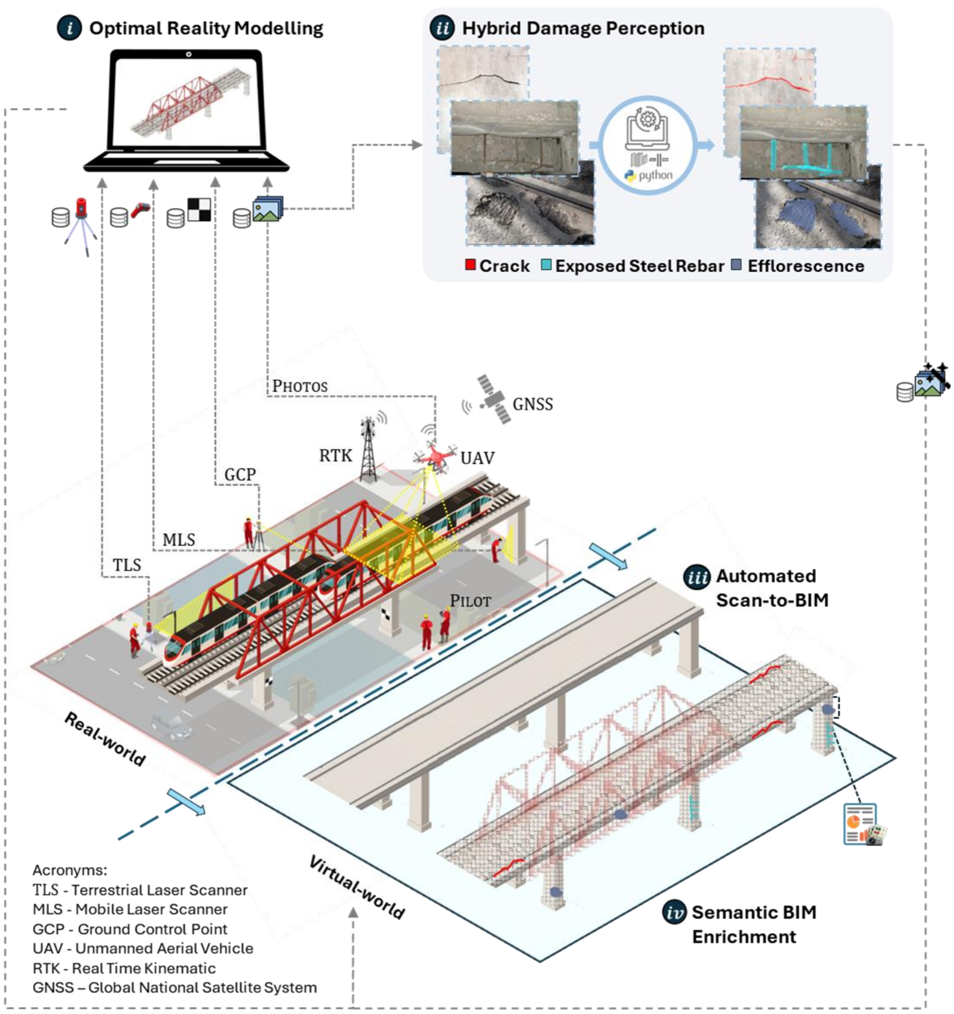


Figure 1: Overview of the proposed end-to-end framework, illustrating the four sequential stages: i) Optimal reality modelling, ii) hybrid damage perception, iii) Automated Scan-to-BrIM, and iv) semantic enrichment.

\

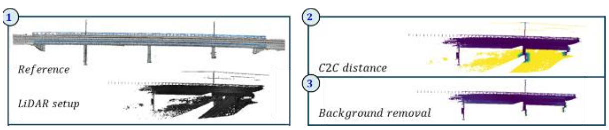
To mitigate these artefacts, a multi-modal fusion strategy was adopted that integrates background-aware filtering prior to reconstruction, illustrated in *Figure 2* (Cabral et al., 2025b). For photogrammetric data, a semantic segmentation network S processes each raw image $I \in R^{H \times W \times 3}$, generating a binary mask M where structural pixels are preserved ($M_{i,j} = 1$) and background pixels are suppressed ($M_{i,j} = 0$). This ensures that the subsequent Structure-from-Motion (SfM) pipeline reconstructs only relevant structural geometry. Simultaneously, LiDAR point clouds acquired via TLS or MLS undergo heuristic filtering based on multi-source geometric proximity (Cloud-to-Cloud distance).

Following data fusion, a Voxel Grid Filtering strategy addresses the non-uniform point densities inherent to multi-sensor capture. The 3D space is discretized into volumetric elements (voxels) V of dimension δ (e.g., 0.01 m). For a set of points $P = \{p_1, p_2, \ldots, p_n\}$ within a voxel V_k, the retained point centroid $p_{centroid}$ is computed in Equation (1).

$$p_{centroid} = \frac{1}{N_k} \sum_{i=1}^{N_k} p_i \tag{1}$$

where N_k is the number of points in voxel V_k. This downsampling preserves the underlying geometric topology while normalizing density, ensuring computational tractability for the subsequent deep learning stages without compromising the geometric fidelity required for defect quantification.

a)



b)

Figure 2: Optimal reality modelling strategy showing background removal for: a) imagery and b) LiDAR (adapted from Cabral et al., 2025b).

Hybrid damage perception

To achieve robust instance segmentation under operational conditions, the challenge of scale variation in high-resolution UAV imagery was addressed. Standard detectors often fail to identify fine cracks when images are downscaled. Thus, it is proposed a hybrid pipeline integrates YOLO11, Slicing Aided Hyper Inference (SAHI), and the Segment Anything Model (SAM) (Cabral et al., 2025a).

The detection phase utilises SAHI to systematically partition high-resolution images I_{HR} into overlapping patches P_k. Each patch is normalised to a 640×640 resolution, the native input size for the YOLO11 detector, ensuring that fine-grained features are preserved without the spatial degradation typically caused by global image downscaling. A YOLO11 detector, trained on railway-specific pathologies, processes each patch independently. The global detection set D_{global} is reconstructed by merging local predictions using Non-Maximum Suppression (NMS) to eliminate redundant bounding boxes.

To transition from coarse bounding boxes to pixel-perfect segmentation, the detected boxes B_{det} serve as geometric prompts for the Segment Anything Model (SAM), as depicted in *Figure 3*. Unlike traditional semantic segmentation networks trained on fixed classes, SAM leverages a Vision Transformer (ViT) backbone to generate zero-shot masks for each prompted instance. The segmentation objective is to minimize the boundary error between the predicted mask and the ground truth.

This hybrid approach ensures that fine-grained defects (e.g., hairline cracks) are detected at their native resolution while their boundaries are segmented with the precision of a foundation model.

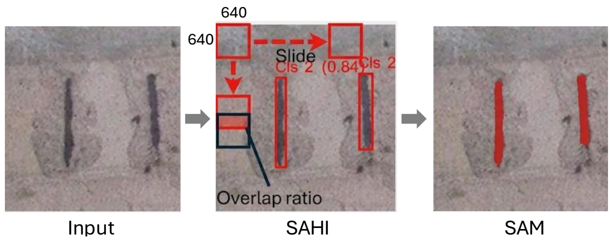


Figure 3: The hybrid perception pipeline: SAHI-assisted detection feeding into SAM-based segmentation.

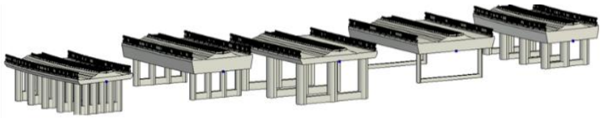
Automated Scan-to-BrIM

The third stage automates the conversion of unstructured point clouds into parametric Bridge Information Modelling (BrIM) elements. It replaces manual modelling with a deep learning-based keypoint regression network trained on synthetic data.

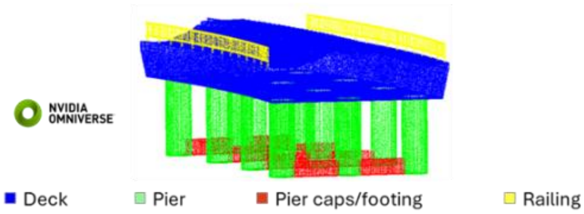
Given the scarcity of labelled bridge point clouds, it employs a procedural generation engine to create a synthetic dataset. This engine, based on Revit® and Dynamo©, is used to randomise bridge geometrical and structural parameters. Nvidia Omniverse© is then leveraged to simulate laser scanning technology and generate semantically labelled point clouds. *Figure 4* illustrates the overall data generation and perception workflow.

The reconstruction network based on Yolo-Pose accepts point cloud slices S_i along the bridge alignment and regresses a set of structural keypoints $K = \{k_1, k_2, \ldots, k_m\}$ defining the cross-sectional profile, as illustrated in *Figure 5*. The network is trained by minimising the Mean Squared Error (MSE) between predicted keypoints $\hat{K}$ and ground truth K_{GT} , as defined in Equation (2).

$$L_{keypoint} = \frac{1}{m}\sum_{j=1}^{m}\left|\left|\widehat{k_j} - k_{GT,j}\right|\right|^2 \qquad (2)$$



a)



b)

Figure 4. Synthetic bridge point cloud generation pipeline: a) bridge modelling and parameter randomisation using Revit® and Dynamo©; b) laser scanning simulation and semantic point cloud generation using NVIDIA Omniverse©.

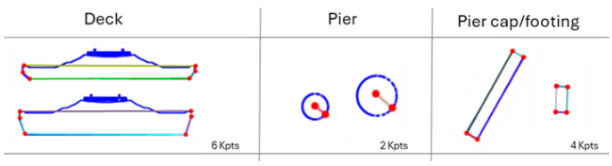


Figure 5: Labelled dataset for YOLO-Pose keypoints defining the cross-sectional profile.

Once inferred, these keypoints drive the instantiation of Industry Foundation Classes (IFC) entities. Structural elements are generated via lofting algorithms, ensuring the resulting model is fully parametric and interoperable, namely, IfcColumn, IfcBeam, IfCSlab and IfcWall (Amroune et al., 2022).

Semantic BIM Enrichment

The final stage transforms the geometric model into an "as-damaged" BrIM model by mapping 2D perception results into the 3D BIM environment.

Damage masks from Stage 2 are projected onto 3D space using a ray-casting algorithm. For a pixel u, v classified as damage, a ray $r(t)$ is cast from the camera center C through the pixel on the image plane. As defined in Equation (3).

$$r(t) = C + t \cdot K^{-1} \cdot [u, v, 1]^T \qquad (3)$$

where K is the camera intrinsic matrix. The intersection of $r(t)$ with the 3D mesh geometry determines the spatial location of the defect. To consolidate detections from multiple overlapping images, a spatial clustering algorithm refined by Otsu thresholding is applied to filter low-confidence projections. *Figure 6* illustrate this process

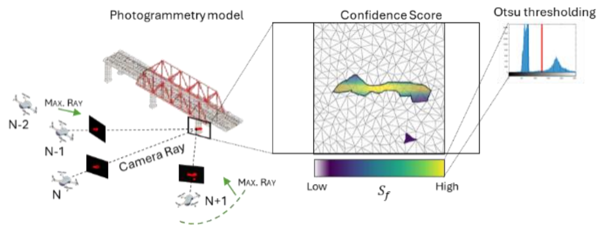


Figure 6: Multi-view ray-casting-based projection of 2D damage detections onto the 3D mesh for semantic enrichment.

The resulting 3D clusters are ortho-rectified to calculate engineering metrics (area, length, width). Finally, defects are instantiated as IfcBuildingElementProxy objects and hierarchically linked to their host structural element (e.g., IfcColumn) using the IfcRelContainedInSpatialStructure relationship.

This process establishes a direct link between visual inspection data and the BrIM model, enabling automated condition assessment and maintenance planning.

Validation and Results

Case Study Description

The proposed framework was experimentally validated on an operational reinforced concrete railway bridge located in Porto, Portugal, which is depicted in *Figure 7*. The structure consists of a single-track deck with a width of 11.33 m and a height of 0.80 m, supported by cylindrical piers (0.80 m in diameter and approximately 6.30 m in height), pier footing with a width of 0.40 m and a height of 0.80 m and integral abutments. The bridge exhibits heterogeneous degradation, including cracking, efflorescence, and exposed steel rebars.

This combination of structural complexity and varied damage mechanisms provides a case study for evaluating the robustness and reliability of the proposed automated inspection pipeline. All data acquisition activities were conducted in full compliance with national aviation regulations and railway safety requirements. The UAV flights were carried out with the necessary authorisations and under the supervision and support of the relevant infrastructure and railway authorities.

Figure 7: Aerial view of the bridge located in Porto (Portugal) used for experimental validation.

\

Reality Modelling Quality Assessment

The accuracy of the reality model serves as the foundational constraint for all downstream perception and reconstruction tasks. For the case study, the effectiveness of the proposed multi-modal fusion strategy, i.e. combining semantic background removal for photogrammetry with heuristic filtering for LiDAR, was evaluated based on geometric completeness.

Qualitative inspection of the fused point cloud confirms that the upstream filtering successfully suppressed high-frequency noise sources typical of railway corridors, as illustrated in *Figure 8*. This pre-processing yielded a structurally focused dataset in which the primary load-bearing elements (piers, deck slab, pier footings, and abutments) maintained surface continuity. Quantitatively, the voxel grid filtering strategy significantly reduced the data volume while homogenising point density.

a)

b)

Figure 8: geometric completeness: a) Raw reality modelling containing noise; b) Optimal reality model after filtering.

Damage Perception Performance

The damage identification performance was quantitatively evaluated using a hybrid deep learning pipeline designed to detect and segment three critical railway bridge pathologies, namely, cracks, efflorescence, and exposed rebar.

The model was trained using a multi-source transfer learning strategy to address the scarcity of labelled infrastructure data. The foundational weights were initialised from a model pre-trained on the MS COCO dataset, followed by fine-tuning on a curated subset of the public DACL10K bridge damage dataset (3,737 images). The final domain adaptation was performed on a custom high-resolution UAV dataset comprising 645 training images and 161 validation images (5280×3956 pixels), annotated with pixel-precise masks.

The proposed hybrid pipeline, which synergises a YOLO11 detector with Slicing Aided Hyper Inference (SAHI) and the Segment Anything Model (SAM), achieved a Mean Average Precision (mAP50) of 0.621 for detection across all classes on the validation set.

This step proved critical for the Scan-to-BIM stage by removing non-structural outliers, the effective geometric topology was preserved, allowing the keypoint regression network to operate on clean cross-sectional profiles.

In total, 170 distinct damage instances were identified and segmented, comprising 75 cracks, 13 efflorescence regions, and 82 areas of exposed reinforcement. Figure 9 qualitatively confirmed that the pipeline effectively identified the damage from the UAV image. A class-specific performance analysis highlighted the benefits of the hybrid architectural choices:

Cracks. The native YOLO11 output, enhanced by SAHI's high-resolution inference tiling, proved instrumental for linear defects. It achieved an Intersection over Union (IoU) of 0.495, successfully recalling narrow cracks that typically vanish during standard image downsampling.

Area Defects. For pathologies with complex morphologies but clearer boundaries, the two-stage refinement using SAM (prompted by YOLO bounding boxes) outperformed native segmentation. This approach yielded IoU scores of 0.331 for efflorescence and 0.205 for exposed steel rebars.

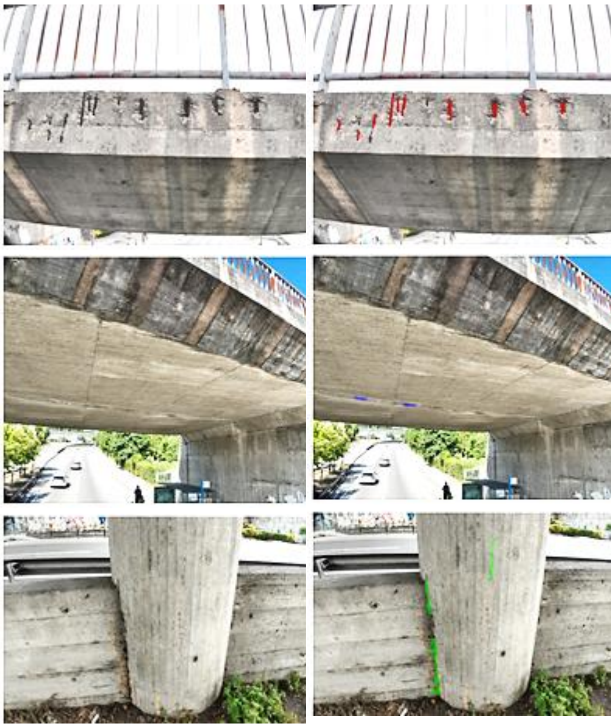
By producing well-defined masks, the framework ensures that metric quantification is derived from precise morphological geometry rather than coarse approximations.

Geometric Accuracy of Scan-to-BIM

The geometric validation of the generated Bridge Information Model (BrIM) was assessed by computing the signed Euclidean deviations between the generated IFC elements and the ground-truth reality capture point cloud. *Figure 11* depicts the BrIM model. The global mean geometric deviation for the bridge was approximately 14.46 mm. It is important to note that the upper surface of the deck was not modelled due to occlusion by the ballast and track superstructure; thus, this does not reflect in the deviation calculation. This accuracy aligns with the requirements for Level of Development (LOD) 300 (BIMForum, 2024), which requires elements be graphically represented as specific systems with precise quantities, sizes, shapes, locations, and orientations.

A component-level error analysis reveals distinct performance characteristics across structural elements. The pier caps and footings exhibited the lowest mean deviation (5.32 mm), followed closely by the abutments (9.29 mm). The cylindrical piers showed a mean deviation of 19.6 mm, while the deck slab presented the highest deviation at 22.89 mm. This confirms that synthetic domain randomisation is a viable strategy for overcoming the scarcity of labelled infrastructure datasets.

\



a) b)

Figure 9: Qualitative results of the hybrid perception engine: a) original image; b) prediction damage, i.e., cracks (green), efflorescence (blue), and exposed rebar (red).

Semantic Enrichment and Damage Quantification

The final phase of validation focused on the semantic enrichment capability, the transformation of the BrIM model from a simple structural component into an "as-damaged" BrIM model. The ray-casting and clustering algorithms successfully mapped the 2D perception results onto the 3D IFC entities, establishing a coherent spatial database of defects.

Figure 11: the automated generated parametric BrIM geometry (LOD 300).

The enriched BIM environment enabled the automatic extraction of aggregate damage metrics. *Figure 10* illustrates the damage mapped to the georeferenced BrIM model, with corresponding dimensional attributes and reference image as well as the corresponding orthophoto. For the case study bridge, the primary contribution is the instance-level granularity; each defect is instantiated as an IfcBuildingElementProxy and hierarchically linked to its host element (e.g., a specific pier). Beyond local statistics, the system quantified a total crack length of 13.82 m, an efflorescence surface area of 0.112 m², and an exposed reinforcement area of 0.241 m². This structured data format allows asset managers to query the model for element-specific condition ratings, bridging the gap between raw visual data and actionable maintenance decision-making.

Discussion

This research resolves the historical discontinuity between automated damage detection and structural digitalization by integrating multimodal reality capture with Bridge Information Modelling (BrIM). While conventional pipelines typically terminate at 2D imagery or unstructured point clouds, the proposed framework embeds quantified defects as *IfcBuildingElementProxy* objects. This approach transforms a static BrIM into an interoperable "as-damaged" model, supplying asset

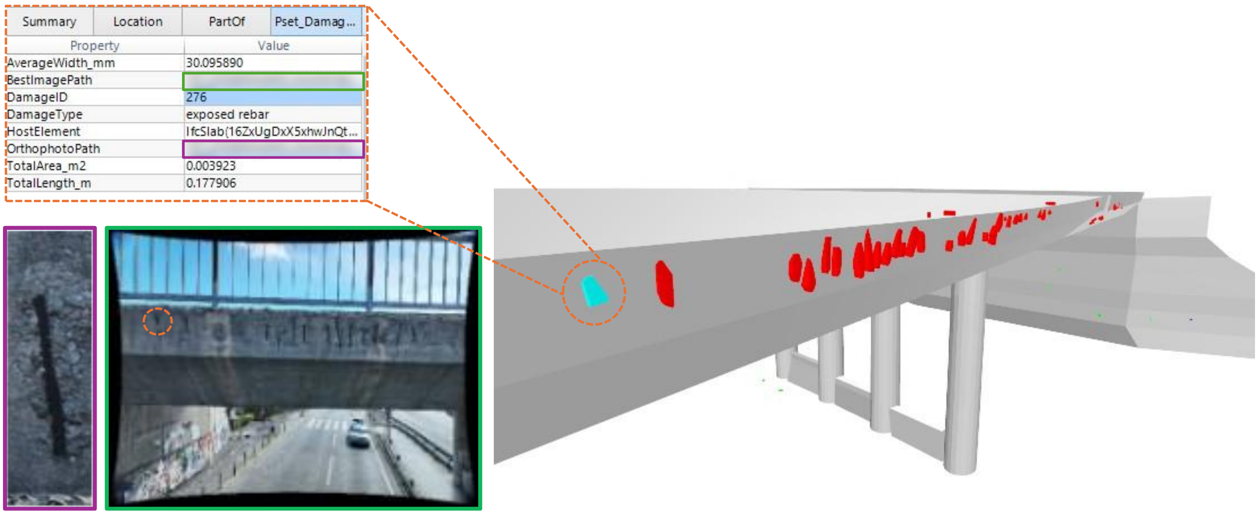


Figure 10: the "as-damaged" BrIM model visualizing mapped defects (light blue) alongside the associated property window displaying quantified dimensional attributes (e.g., area, length) and links to the source reference image and generated orthophoto.

\

managers with actionable, element-specific condition metrics for automated maintenance decision-making.

Experimental validation on an operational railway bridge demonstrated that high-fidelity reality modelling is a strict prerequisite for automated inspection in complex environments. By employing a multi-modal fusion strategy—combining semantic background removal for photogrammetry with heuristic LiDAR filtering—the framework successfully isolated structural geometry from background noise (e.g., vegetation, catenary systems). Consequently, the automated reconstruction achieved a global mean geometric deviation of 14.46 mm, satisfying Level of Development (LOD) 300 requirements.

The findings also highlight the inadequacy of standard object detectors for capturing fine-grained infrastructure pathologies, as standard image downscaling often eliminates hairline cracks. Integrating Slicing Aided Hyper Inference (SAHI) with YOLO11 preserved these linear defects at their native high resolution. Subsequent processing with the Segment Anything Model (SAM) generated zero-shot, pixel-perfect masks, ensuring damage quantification was driven by precise morphology rather than coarse bounding box approximations.

To address the scarcity of annotated 3D point clouds for Scan-to-BIM, the framework successfully implemented synthetic domain randomization. Generating simulated data via Revit®, Dynamo©, and NVIDIA Omniverse© proved highly effective for training accurate keypoint regression networks, yielding component-level deviations as low as 5.32 mm.

Despite robust performance, current limitations include the inability to model the bridge deck's upper surface due to track superstructure occlusions, alongside the methodology's exclusive focus on surface-level manifestations of damage.

Conclusions

This study presented and validated an end-to-end computational framework that effectively brings together multimodal reality capture and semantic Bridge Information Modelling (BrIM). By integrating AI-driven background removal, a hybrid YOLO11–SAHI–SAM perception engine, and synthetic-data-trained Scan-to-BIM reconstruction, the methodology overcomes the scalability limitations inherent to traditional manual inspection.

The experimental validation on the bridge located at Porto (Portugal) confirms that high-fidelity reality modelling is a prerequisite for automated inspection. The proposed multi-modal background filtering strategy proved essential for reducing geometric noise, enabling automated reconstruction algorithms to converge with millimetre-level accuracy (mean deviations < 23 mm).

Furthermore, the results demonstrate that standard object detectors are insufficient for the complexity of infrastructure pathologies. The integration of Slicing Aided Hyper Inference (SAHI) and the Segment Anything Model (SAM) was critical for decoupling detection from segmentation, enabling precise quantification of fine-grained defects, such as cracks, that are typically lost (due to resolution) in standard down sampling processes. Ultimately, the true engineering value of the framework lies in the semantic enrichment stage, which successfully transforms static a BrIM model into an "as-damaged" BrIM model. By embedding quantified defects as intelligent IfcBuildingElementProxy objects, the system delivers interoperable, structured data directly usable for predictive maintenance and asset management.

Future research will prioritise the comparative validation of this automated framework against traditional manual inspection benchmarks for the same asset. Additionally, the synthetic training strategy will be extended to encompass complex masonry and steel truss typologies, thereby enhancing model generalisation. Finally, efforts will be directed towards integrating subsurface non-destructive testing (NDT) data into the BrIM environment, providing a holistic view of structural health that extends beyond surface-level defects.

Acknowledgments

This work was financially supported by Funding-UID/04708 of the CONSTRUCT-Instituto de I&D em Estruturas e Construções-funded by Fundação para a Ciência e a Tecnologia, I.P./MCTES through the national funds, and co-financed by COMPETE 2030, by Portugal 2030, as well as, by the NORTE 2030 Programme through the European Regional Development Fund (ERDF), under project FireBIM, reference NORTE2030-FEDER-00358400. It was additionally supported by the grant UI/BD/150970/2021 Portuguese Science Foundation, FCT/MCTES.

References

Amroune, A., Cuillière, J. C. & François, V. (2022). Automated Lofting-Based Reconstruction of CAD Models from 3D Topology Optimization Results. Computer-Aided Design, 145, 103183.

Ayele, Y. Z., Aliyari, M., Griffiths, D. & Droguett, E. L. (2020). Automatic Crack Segmentation for UAV-Assisted Bridge Inspection. Energies, 13, 6250.

Bimforum. (2024). Level of development (LOD) specification [Online]. Available: https://bimforum.org/resource/lod-level-of-development-lod-specification/ [Accessed January 2 2025].

Cabral, R., Oliveira, R., Ribeiro, D., Rakoczy, A. M., Santos, R., Azenha, M. & Correia, J. (2023). Railway Bridge Geometry Assessment Supported by Cutting-Edge Reality Capture Technologies and 3D As-Designed Models. Infrastructures, 8, 114.

Cabral, R., Santos, R., Correia, J. & Ribeiro, D. (2025a). A Hybrid YOLO and Segment Anything Model Pipeline for Multi-Damage Segmentation in UAV Inspection Imagery. Sensors (Basel), 25.

Cabral, R., Santos, R., Correia, J. & Ribeiro, D. (2025b). Optimal reconstruction of railway bridges using a machine learning framework based on UAV photogrammetry and LiDAR. Structure and Infrastructure Engineering, 173, 1-21.

Cavieres-Lagos, S., Muñoz La Rivera, F., Atencio, E. & Herrera, R. F. (2024). Integration of BIM tools for the facility management of railway bridges. Applied Sciences, 14, 6209.

Hattori, K., Oki, K., Sugita, A., Sugiyama, T. & Chun, P.-J. (2024). Deep learning-based corrosion inspection of long-span bridges with BIM integration. Heliyon, 10.

Huang, L., Fan, G., Li, J. & Hao, H. (2024). Deep learning for automated multiclass surface damage detection in bridge inspections. Automation in Construction, 166, 105601.

Kim, H., Yoon, J. & Sim, S. H. (2020). Automated bridge component recognition from point clouds using deep learning. Structural Control and Health Monitoring, 27, e2591.

Kruachottikul, P., Cooharojananone, N., Phanomchoeng, G., Chavarnakul, T., Kovitanggoon, K. & Trakulwaranont, D. (2021). Deep learning-based visual defect-inspection system for reinforced concrete bridge substructure: a case of Thailand's department of highways. Journal of Civil Structural Health Monitoring, 11, 949-965.

Mirzazade, A., Popescu, C., Gonzalez-Libreros, J., Blanksvärd, T., Täljsten, B. & Sas, G. (2023). Semi-autonomous inspection for concrete structures using digital models and a hybrid approach based on deep learning and photogrammetry. Journal of Civil Structural Health Monitoring, 13, 1633-1652.

Moore, M., Phares, B. M., Graybeal, B., Rolander, D., Washer, G. & Wiss, J. (2001). Reliability of visual inspection for highway bridges, volume I. Turner-Fairbank Highway Research Center.

Nichols, L., Ashmawi, A. & Nguyen, P. H. D. (2025). Digitalizing Bridge Inspection Processes Using Building Information Modeling (BIM) and Business Intelligence (BI). Applied Sciences-Basel, 15, 10927.

Olaszek, P., Maciejewski, E., Rakoczy, A., Cabral, R., Santos, R. & Ribeiro, D. (2024). Remote Inspection of Bridges with the Integration of Scanning Total Station and Unmanned Aerial Vehicle Data. Remote Sensing, 16, 4176.

Patil, J. & Kalantari, M. (2025). Automatic Scan-to-BIM-The Impact of Semantic Segmentation Accuracy. Buildings, 15, 1126.

Rakoczy, A. M., Ribeiro, D., Hoskere, V., Narazaki, Y., Olaszek, P., Karwowski, W., Cabral, R., Guo, Y., Futai, M. M. & Milillo, P. (2025). Technologies and platforms for remote and autonomous bridge inspection–review. Structural Engineering International, 35, 354-376.

Ribeiro, D., Rakoczy, A. M., Cabral, R., Hoskere, V., Narazaki, Y., Santos, R., Tondo, G., Gonzalez, L., Matos, J. C., Massao Futai, M., Guo, Y., Trias, A., Tinoco, J., Samec, V., Minh, T. Q., Moreu, F., Popescu, C., Mirzazade, A., Jorge, T., Magalhaes, J., Schmidt, F., Ventura, J. & Fonseca, J. (2025). Methodologies for Remote Bridge Inspection-Review. Sensors (Basel), 25, 5708.

Riedlinger, U., Klein, F., Hill, M., Lambracht, C., Nieborowski, S., Holst, R., Bahlau, S. & Oppermann, L. (2022). Evaluation of mixed reality support for bridge inspectors using BIM data: Digital prototype for a manual task with a long-lasting tradition. i-com, 21, 253-267.

Ruggieri, S., Cardellicchio, A., Nettis, A., Renò, V. & Uva, G. (2025). Using attention for improving defect detection in existing RC bridges. IEEE Access, 13, 18994-19015.

Tomor, A. (2013). Life-cycle assessment and deterioration models for masonry arch bridges. WIT Transactions on Ecology and the Environment, 179, 535-546.

Xu, J., Shi, M., Cabral, R., Ribeiro, D., Yu, W., Wu, H. & Narazaki, Y. (2026). Domain-generalizable point cloud instance segmentation of bridge components using class-balanced dynamic thresholding. Automation in Construction, 181, 106631.

Xu, T., Wu, Y., Qin, Y., Long, S., Yang, Z. & Guo, F. (2025). Automatic steel girder inspection system for high-speed railway bridge using hybrid learning framework. Computer-Aided Civil and Infrastructure Engineering, 40, 1508-1527.

Yamane, T., Chun, P.-J. & Honda, R. (2024). Detecting and localising damage based on image recognition and structure from motion, and reflecting it in a 3D bridge model. Structure and Infrastructure Engineering, 20, 594-606.

Yamane, T., Chun, P. J., Dang, J. & Honda, R. (2023). Recording of bridge damage areas by 3D integration of multiple images and reduction of the variability in detected results. Computer-Aided Civil and Infrastructure Engineering, 38, 2391-2407.

Yang, X. F., Castillo, E. D., Zou, Y., Wotherspoon, L., Yang, J. X. & Li, H. (2025). Automated Concrete Bridge Damage Detection Using an Efficient Vision Transformer-Enhanced Anchor-Free YOLO. Engineering, 51, 311-326.

Zakaria, M., Karaaslan, E. & Necati Catbas, F. Real-time ai-based bridge inspection using mixed reality platform. Structures Congress 2023, 2023. 120-131.

\