A Framework for Vision-Based 3D Inspections for Maintenance Activities and Digital Twin Integration

Panagiotis Vrachnos
Information Technologies Institute
CERTH
Thessaloniki, Greece
panosvrahnos@iti.gr

Carlos Ramonell
Universitat Politecnica de Catalunya
UPC
Barcelona, Spain
carlos.ramonell@upc.edu

Ilias Koulalis
Information Technologies Institute
CERTH
Thessaloniki, Greece
iliask@iti.gr

Konstantinos Ioannidis
Information Technologies Institute
CERTH
Thessaloniki, Greece
kioannid@iti.gr

Irina Stipanovic
Department of Construction Management and Engineering, University of Twente
Enschede, The Netherlands
i.stipanovic@utwente.nl

Stefanos Vrochidis
Information Technologies Institute
CERTH
Thessaloniki, Greece
stefanos@iti.gr

2024 International Conference on Content-Based Multimedia Indexing (CBMI) | 979-8-3503-7844-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/CBMI62980.2024.10859253

***Abstract*—Vision-based monitoring methods have been actively studied in the construction industry as they can automatically generate information related to progress, productivity, and safety. 3D reconstruction is key in such monitoring techniques, allowing the inference of job-site context, the creation of digital counterparts of physical spaces, and the comparisons between as-designed and as-built conditions. However, 3D applications in construction currently produce large volumes of unstructured data and unusable point clouds, which are time-consuming to convert into an interactive environment for Building Information Modelling (BIM) or Digital Twins. While radiance field rendering methods are increasingly gaining traction, the adoption of generated Neural Radiance Fields or Gaussian Splatting models by digital construction technology is still tentative. This study introduces a framework that uses Neural Radiance Fields (NeRFs) to improve 3D inspection and maintenance on construction sites. It merges NeRF's high-resolution, real-time 3D modelling with an interactive platform, facilitating detailed remote site analysis and defect detection. The framework incorporates a custom version of TurboNeRF, tailored specifically for construction site inspections. Through this paper, we aim to highlight the potential of combining 3D imaging technology, the use of drone imagery and ontological models to improve construction site management practices.**

***Index Terms*—Digital Twin, Ontology, Remote Inspection, Defect Detection, Neural Radiance Field**

I. Introduction

The construction industry is inextricably linked with progress monitoring and site inspection. Construction progress monitoring is highly significant, and its absence from a construction project can lead to delays in project completion as well as additional costs. Traditional methods of progress monitoring require manual data entry, involving laborious human effort, consuming considerable time and prone to human error [1]. An accurate report can be beneficial, whether produced during the course of the construction for its progress or upon completion for maintenance purposes. In this way, stakeholders stay informed about the project's status, enabling them to make crucial decisions [2].

Over the past few years, digital twins were introduced into the construction field [3], showing their potential and gaining significant recognition among the research community. Digital twins are virtual replicas of physical objects that dynamically exchange real-time data, enabling continuous interaction between the physical and digital environment. Thus, they stand as one of the most empowering technologies for achieving asset management, facilitating the ability to understand, monitor and optimise the operations of real-world entities. The connection between the physical and digital worlds enables data analysis, progress monitoring and inspection to proactively identify potential issues, prevent delays and efficiently manage projects through simulations of current and future work. Digital twins empower their potential through the interdisciplinary field of computer vision, which serves as the means for computers to understand visual information.

Computer Vision is closely connected to the engineering and construction sector. In the past decade the rapid evolution of Computer Vision based techniques has revolutionised the construction progress monitoring and inspection processes, offering invaluable support for strategic decision-making by management [4]. So far, computer vision methods include procedures such as data acquisition, processing, 3D reconstruction and defect/damage detection, among others [5]. Traditional methods for data acquisition often involve expensive hardware, such as LiDAR devices [6], and time consuming tasks . Moreover, these methods typically result in unstructured point clouds lacking precise geometry representation. Large real-world complex environments such as construction sites are difficult to replicate accurately. A multitude of factors including variations in lighting, shading, and weather conditions, can negatively impact 3D reconstruction methods that rely on Structure from Motion (SfM) techniques [7]. As a result, engineers today require inspection methods focusing on effective maintenance activities to optimise the life-cycle of infrastructures while minimising operational time and costs.

Accurate representation of complex scenes in 3D space through remote inspection techniques is crucial for maintenance activities. Recently, Neural Radiance Fields (NeRFs) [8] have attracted significant interest for their ability to represent volumetric scenes as a continuous function through a neural network fed with 2D images and their corresponding camera poses. The outcomes offer high-detail photo-realistic renderings, enabling the synthesis of novel views. Consequently, the current research on NeRF interactive systems and applications is limited, with scientists directing their focus toward addressing the lack of methods allowing stakeholders to interact with the digital counterpart.

In this paper, we propose an interactive framework that leverages NeRF capabilities to enhance 3D inspection and maintenance activities on construction sites. Specifically, we develop an interactive framework that combines the benefits of UAV-based data acquisition, high detailed realistic 3D rendering produced by NeRF and defect detection to facilitate detailed remote site analysis and damage identification, improving construction site management practices. Our interactive framework integrates an automated UAV visual data collection, specifically designed for construction projects. In this manner, a dataset suitable for an accurate 3D reconstruction is provided. Subsequently, these data are fed into a custom implementation of TurboNeRF [9] tailored for construction infrastructures inspections, generating high-detailed real-time 3D renders within an interactive platform. We further extend the framework's applicability by enabling automated defect detection, thereby empowering stakeholders to conveniently and immediately inspect the project's status. We demonstrate our framework on a real highway bridge construction project.

II. Related Work

The construction industry has long been struggling by the lack of autonomous systems for management, monitoring, maintenance and safety purposes. In the context of maintenance processes, traditional methods of civil infrastructure status assessment typically require skilled inspectors for visual inspection, combined with appropriate decision-making criteria. In this manner, the inspection turns into a tedious and time-consuming task, posing even risks for inspectors' safety. Thus, the scientific community has turned its attention to the integration of computer vision techniques for remote inspection within 3D reconstruction and defect detection, in an attempt to evaluate the current infrastructure status, mitigating risks, human effort and saving considerable time.

NeRFs have revolutionised the field of 3D computer graphics since their emergence. Traditional methods involve storing triangles or colour values for the complete voxel grid on disk, leading to memory-related challenges. Contrary, NeRF enables high precision 3D representation of intricate real-world scenes, solely through neural network training. In comparison with photogrammetric techniques which lead to sparse, unstructured point clouds, NeRF methods learn to represent the scene as a continuous function. Their advantage in representing a colour's point depends on the viewpoint, enabling for the capture of diverse lighting effects such as reflections and transparencies. These effects are commonly observed in the construction sector, highlighting the importance of integrating NeRF methods into construction activities [10].

In the study by Zhexiong Shang and Zhigang Shen, [11], the authors explore the integration of Visual Simultaneous Localisation and Mapping (SLAM) with Unmanned Aerial Vehicles (UAVs) to facilitate real-time 3D mapping of construction sites.The paper introduces a system architecture and experimental setup designed to overcome accessibility issues on construction sites through the use of UAVs, enabling the capture of inaccessible areas from the air. Furthermore, these near real-time capture methods often result in 3D point clouds that face challenges in terms of data scale and structure. Furthermore, information within the captured images is often underused, thus representing a missed opportunity for semantically enriched results.

Computer vision techniques are undoubtedly transforming the construction industry by enabling the automatic generation, analysis, and application of visual data throughout the lifecycle of construction projects. Beyond 3D reconstruction and digitisation of physical spaces, these technologies are emerging for inspection purposes [12]. Currently, in the field of computer vision, many researchers have dedicated their effort to developing image-based automatic non-destructive testing (NDT) methods for contactless or even remote detection systems. The majority of the initial approaches in literature rely upon mature image processing techniques to detect depicted cracks and discriminate them from the background [13]. Deep learning has greatly expanded the efficiency and robustness of traditional vision-based defect detection for a wide variety of visual defects, from cracks and delaminations to corrosion [14]. A collection of studies has explored the use of UAV-based visual data combined with detection techniques. Decision-making algorithm that employs CNNs optimised for the detection of cracks documented in [15]–[17].

Computer vision damage detection capabilities are improving fast, continuously enriching the semantic descriptions of visualised scenes. In the construction domain, these semantic descriptions need to be adequately interconnected with existing models of constructed assets, and combined with many other information sources. For these reasons, Digital Twin models are required to evolve flexibly, and allow domain-agnostic enrichment. Thus, establishing digital twin models driven by a combination of ontologies from different domains is being increasingly considered, providing the backbone of what are called 'Semantic Digital Twins' [18]. An ontology is a computer-readable conceptualisation of some domain-specific knowledge [19]. They are managed through Semantic Web standards such as RDF [20] and OWL [21] , providing graph-like data (knowledge graphs [22]) that are interoperable and self-described to a computer. Semantic digital twins have been already implemented in different projects at different scales. The World Avatar project [23] carried out by the Cambridge CARES group in Singapore, is an example of implementation of a Semantic Digital Twin to improve city management, while the Ashvin project [24] exemplifies their use in platforms to manage individual assets [25].

Accelerating damage detection and annotation workflows through computer vision tools, and connecting the results to vaster systems (Digital Twins) is a meaningful topic that requires exploration.

III. Framework Description

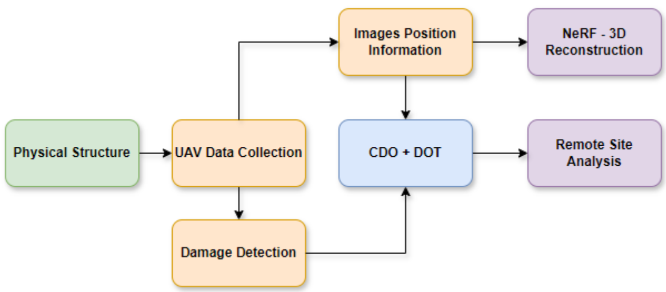


Fig. 1. Visualisation of the proposed visual-based framework for 3D inspection in maintenance activities. The process describes the transformation of physical structure into digital replica. Collected data are processed and registered into the ontology, where they are linked with damage detection data, resulting in a detailed 3D site analysis.

The methodology for developing an operational image-based Digital Twin for infrastructural structures, as proposed in this study, is shown in Figure 1. The process begins with a description of the physical structure in terms of geometry and position in a georeferenced coordinate system. Drones scans are used to obtain a large number of RGB images of the physical structure, a volumetric rendering of the structure's radiance field, achieved through the employment of a custom NeRF implementation based on Instant-NGP [26]. During this step, a detailed CAD model of the structure can be registered at the same georeferenced coordinate system that is used during the drone flights for easier comparison with the reconstructed 3D model. In parallel, a deep learning algorithm analyses the RGB images to identify and categorise construction defects. For the task of classifying defects and damages, the You Only Look Once (YOLO) version 8 algorithm, a real-time object detection system, was selected based on the team's prior research work. The type, position, and size of the detected defects/damages are extracted from the AI algorithms and mapped to the radiance field rendering in a Blender interactive environment, thereby creating the first version of the Digital Twin. Information on the position, type, and size of defects/damages from the Digital Twin can be used to determine if maintenance is needed on the physical structure.To enable the integration and interoperability of the detection results with the semantic digital twins, it is necessary to create an ontology for storing the damaged areas patterns, damage classification, related images, linking the damages to their geometrical representation to the digital object. In our study, data follow the Concrete Damage Ontology [27]

The Digital Twin can be updated, based on the new images. In the future, the proposed Digital Twin can be used to facilitate immersive inspection of the area even with VR glasses using the VR plugin of Blender application to evaluate its structural integrity.

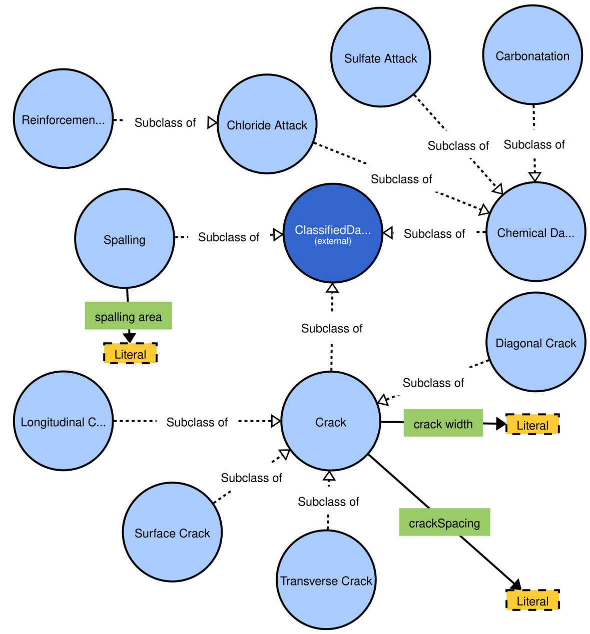


Fig. 2. Concrete Damage Ontology (CDO). Visualisation using WebVOWL tool [27]

A. Data collection

Data collection is a significant process for effective inspection of infrastructures. However, there are scenarios where infrastructures are located on a large scale or at high-altitude, making accessibility challenging. For this purpose, a UAV operation was carefully designed with the intention of thoroughly covering the area. The integration of a high-resolution camera into the UAV provided footage of the scene from multiple perspectives, encompassing otherwise inaccessible areas and offering a more unobstructed view of the site, while minimising the requirement for human labour. The primary objective is to create a comprehensive dataset consisting of high-resolution images, detailing the target area from varying viewing angles and altitudes. In this way, sufficient overlap of images is ensured, which significantly enhances the accuracy of 3D reconstruction.

The density of images is selected to ensure the necessary overlap between the images, both vertically and horizontally. The amount of overlap necessary for effective NeRF reconstructions is not specified as a straightforward percentage or metric; therefore, we adhered to the widely accepted guideline for photogrammetry, which defines the minimum overlap between the photos as 60%.

B. NeRF-based 3D Reconstruction

In photogrammetry involving a single camera, an object in a real-world coordinate system is projected to a 2D image plane. To restore the 3D information from 2D images, relationships

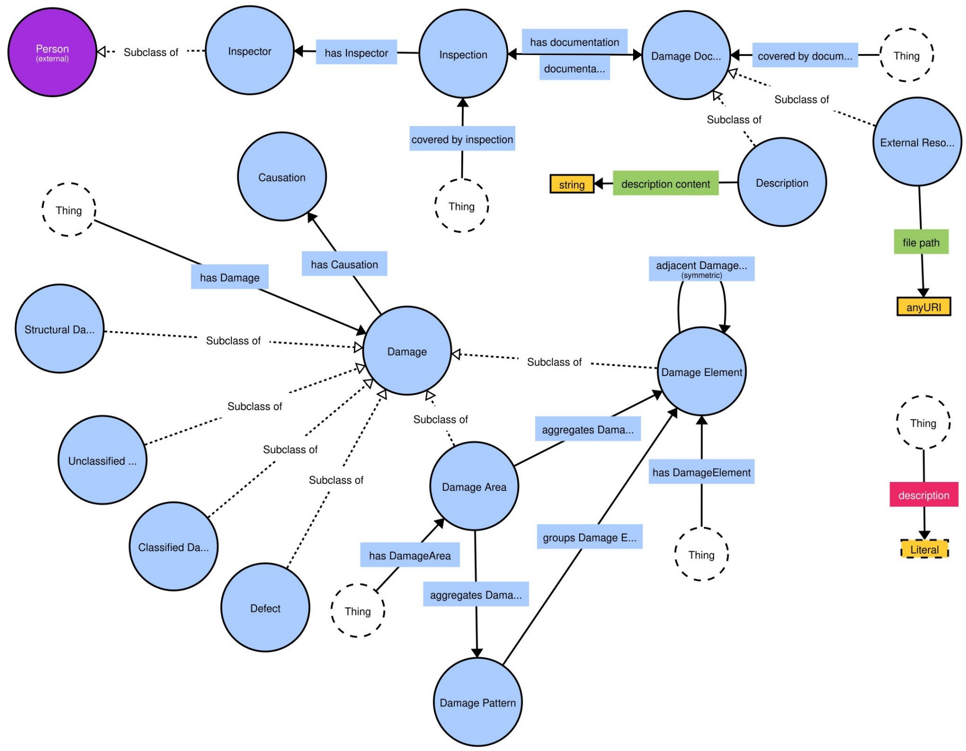


Fig. 3. Damage Topology Ontology (DOT) [28]. Visualisation using WebVOWL tool

between the image plane and real locations should be estimated. Traditional methods employed voxel grids and polygon meshes to store scenes, which were deemed inefficient. In contrast, NeRFs utilise a neural network which is trained to accurately represent the intricate details of a scene. To this end, the creation of a comprehensive and information-rich dataset is imperative. In the initial stage, the raw footage is processed to extract sequential RGB frames. Emphasis is directed towards determining the sampling frequency for frame extraction, thereby ensuring the required amount of overlap. Subsequently, the position and the orientation of each image in 3D space are estimated via a process known as camera pose estimation. In detail, through this process, the spatial location (x, y, z) and viewing direction (θ, ϕ) of each point are defined, constituting continuous 5D coordinate vectors. Then, these vectors are mapped to a higher-dimensional space, enabling the model to accurately represent high-frequency scene content. The transformed vectors are fed as input into the neural network, which learns to predict through training process the corresponding view-dependent RGB colour (c) and volume density (σ). This approach, enables each point to be distinctly represented when observed from various perspectives, thereby enhancing the scene's overall representation. Once the network is trained, high-detailed photo-realistic 3D rendering is provided.

In order to incorporate NeRF capabilities with a semantic digital twin, TurboNeRF was applied, a custom NeRF implementation integrated into the interactive Blender platform. TurboNeRF is based on Instant-NGP [26], leveraging multi-resolution position hash encoding to achieve fast training of large scenes.

C. Defect Detection

The scope of defect/damage detection on construction surfaces is covered either by the category of semantic segmentation or by the category of object detection. Given the recent progress in YOLO-based object detection algorithms, which include benefits such as high classification accuracy and real-time throughput , we applied object detection to the task at hand. Incoming images originated from drone footage using a high-resolution camera, reinforcing the requirement for highly efficient neural networks. A YOLO-based architecture was chosen, for simplicity and due to the high network throughput and the availability of open-source implementations in common machine-learning libraries.

In order to enable the detector's output interoperability and integration into semantic digital twins, an ontology that explains the output data should be created and published, or the output data should conform to an existing ontology. In this study, data are mapped to the existing Concrete Damage Ontology (CDO) [27], as depicted in Figure 2. The ontology allows modelling the three classes currently identified by the detector, i.e. cracks, spalling and corrosion stains due to steel reinforcement degradation.

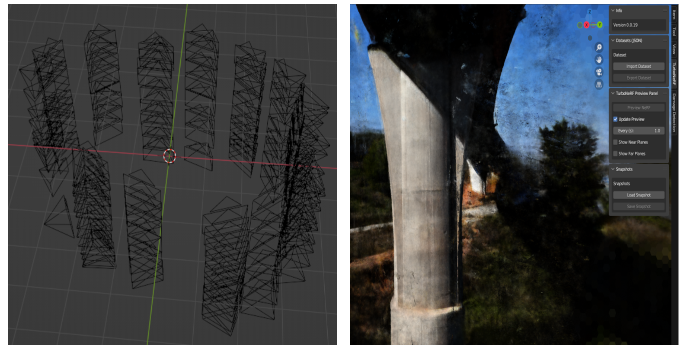
D. Integration of the defect detection tool with Semantic Digital Twins

The process for integrating the defect detection into our Blender-incorporated interactive Digital Twin tool unfolds in the following steps. Initially, damages are detected via the automated AI network, allowing for the estimation of the localisation of the damage in pixel coordinates as long as the camera's position in the 3D space. This information then is combined to registered the identified damages within the 3D virtual space. The real world images are combined with the 3D digital space, creating a cohesive digital representation. The key element in this process is calculating the camera position information using the images captured from a camera and synthesising 3D virtual volumetric spaces using the position information. As a final step, this synthesised information is encoded according to the ontology. The CDO ontology is directly linked to another ontology, the Damage Topology Ontology (DOT) [28]. The DOT ontology (see Fig. 3) allows modelling damaged areas, patterns, as well as to relate them to inspection processes and inspection-related data. It also allows linking the damages to their detected geometrical representation, as well as to pictures and built element models. This integration allows for a structured representation of detected defects, their exact locations, and dimensions within the digital twin framework. By mapping the spatial and visual data to the ontology, we enable seamless information flow between data processing tools and semantic digital twins, which, through RDF graphs, can link damage graphs to models of processes and built products, and many other models representing the asset's context.

IV. Visual Interface

In this section, the deployment of the proposed framework for vision-based 3D inspection for maintenance activities is presented in detail. Specifically, the developed framework was tested over a highway overpass located in the Metropolitan Area of Barcelona. The main object was to generate a 3D representation of one of the bridge pillars while simultaneously identifying any potential damages. Given the peculiarity of the bridge, spanning over railway lines and rivers, with an estimated length of 846 meters, manual inspection or data acquisition using hand-held devices would substantially diminish the effectiveness and efficiency of inspection process. This would lead to increased scanning time, human effort and costs. Considering the aforementioned challenges, we deem that the data acquisition and defect detection in an automated manner, coupled with the 3D representation of the construction site for remote inspection and maintenance activities, constitute a semantic digital twin that facilitates efficient and meaningful analysis of the infrastructure.

In detail, data collection is conducted with the use of a DJI Mavic Pro UAV model. Through a meticulously planned operation, the UAV's trajectory was optimised to cover the scene from diverse perspectives, distances, and altitudes. Consequently, a high-resolution video capturing the target from multiple viewing angles was created, with a resolution of 8000×6000. In order to compose a dataset conducive to subsequent 3D reconstruction process, 135 RGB frames were extracted and then resized to 1500×1000, striking on balance



(a) (b)

Fig. 4. (a) Initial framework visualisation: Results of the camera poses estimation in 3D space. (b) Results of 3D render after 65k training steps

between dataset size and overlap. Next, camera poses and intrinsic camera parameters were estimated, utilising COLMAP [29] Structure from Motion tool.

Once the dataset was created, then is fed as input to the TurboNeRF model. The first part of the framework visible to the user presents the camera poses in 3D space (see Fig 4a). When $start\ training$ button is executed, each pixel's colour and density within the bounding scene are optimised. Also framework enable pausing the training at any point to evaluate the ongoing results. Figure 4b illustrates a comprehensive training after $65k$ steps.

Then, the framework is extended with a custom add-on, developed inside Blender's interface, named Defect Detection, as shown in Figure 5. This addition offers an automated method for defect detection and also interacts with the 3D render. Upon selecting the Defect Detection custom add-on, a menu with various operations is presented, providing users with flexibility and control. Initially, user can navigate through the 3D space and select any camera to view the scene from the corresponding perspective. Within the first option, called $visualize\ image$, the scene is split, with the left half expanding to display a window containing the corresponding real image, while the right half continues to display the 3D render, showing the perspective of the selected camera in the 3D scene. In the second panel, named Damage Detection, two additional buttons are available. The first one, called $Detect\ Damage$, activates the damage detection algorithm. Upon execution, the real-time object detection method YOLOv8 is applied, identifying the position, size, and type of defects or damages. In our specific case, the types of defects/damages include cracks, delamination, and corrosion stains. After the completion of the detection process, the specific type and pixel-wise localisation of the damage are encoded using the CDO and DOT ontologies, linking them with the spatial location of the cameras in 3D virtual space. At the same time, within Blender's interface, a new collection is dynamically generated, called Damage Detection Cameras, encompassing all the cameras capturing one or more detected defects or damages. Each defect's type is visually represented to the

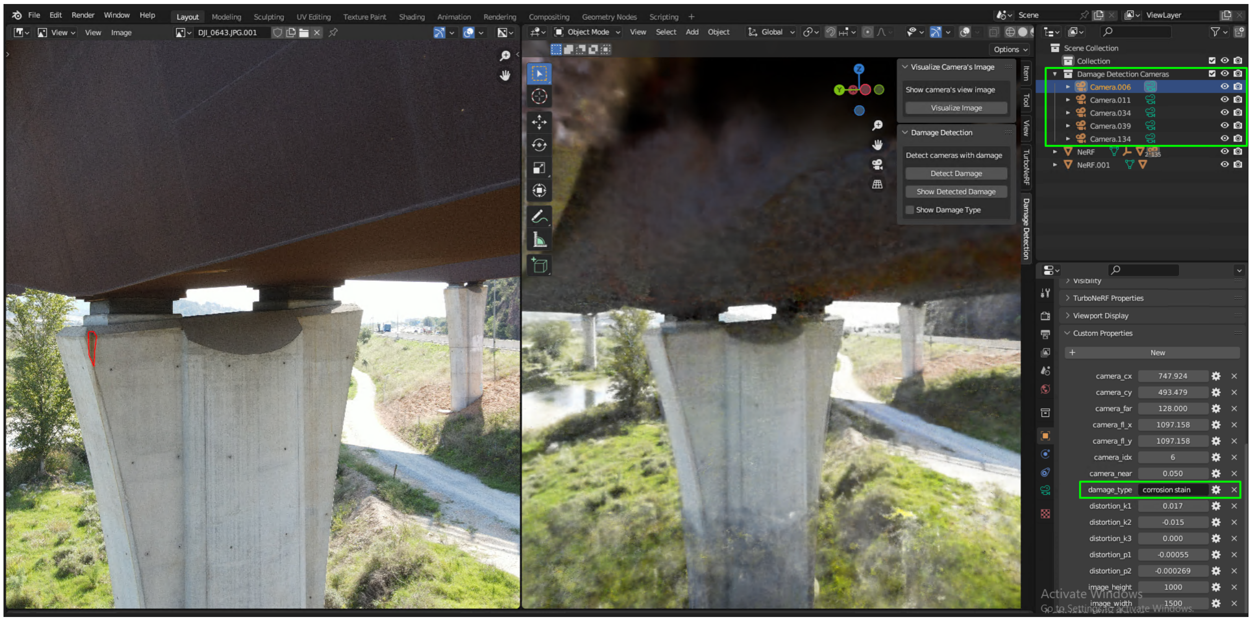


Fig. 5. Illustration of damage detection extension. Top right: Newly generated collection of cameras capturing detected defects. Bottom right: Type of damage highlighted. Left half: Exact location of defect/damage outlined in red polygon. Right half: Corresponding camera view in 3D virtual space.

user through the corresponding camera's properties, in the field named $damage\_type$. Finally, when selecting a camera from the newly generated collection, $Show\ Detected\ Damage$ operation activates this linkage, illustrating the precise location of the defect/damage. Specifically, in the left half is depicted the real image, with the affected area outlined by a red polygon for easy identification. Simultaneously, the 3D render automatically adjusts its viewpoint, providing a corresponding view of the detected area within the immersive 3D virtual environment.

Overall, our framework provides a continuous exchange of data and information facilitated through the mapping of visual and spatial data within the ontology.

V. Conclusion

In this paper, we propose a framework for a tool for detecting and localising defects through an interactive 3D model of a construction site built from 2D images taken by a DJI commercial UAV. The framework also provides a method to integrate the detected defects into semantic Digital Twins. The Instant-NeRF algorithm in Blender trains the 3D rendering, estimating camera parameters and building a radiance field representation. Additionally, image data processed by a semantic segmentation Unet detector provides semantic and geometrical defect descriptions. Defects detected conform to a defect-specific domain ontology, the Concrete Damage Ontology, which enables relating damages with data from other knowledge domains, making it easier for managers to understand the condition of built assets and make better-informed decisions. Moreover, the automated framework emphasises on human-centric features, avoiding time-consuming and risky manual inspections, highlighting worker security and satisfaction. We examine how this framework is used to manage defect detection processes carried out in the field, as well as how it helps automating the inspection information flow using computer vision-based automatic defect detection methods in a real bridge located in Barcelona. As this research is at a conceptual level, we need to validate the proposed system and prototype with more construction domain data. Our machine vision-based Defect Inspection System overcomes significant limitations for use on construction sites by integrating with existing BIM tools to simplify creating defect management information and providing markers for defect localisation and classification. Future research will assess the practical applicability of our system in the construction field, focusing on efficiency and effectiveness. We also aim to expand the scope of our research by incorporating quantitative results, evaluating the detected defects in the virtual environment with ground truth data, thus given a more comparative analysis with existing methods. This work, showcasing the workflow and functionality, aims to shift from current reactive defect management practices in the construction industry to a more proactive approach.

Acknowledgments: This work was supported by the EC-funded research and innovation programme H2020 ASHVIN: "Digitising and transforming the European construction industry" under the grant agreement No.958161, SUN: "Social and hUman ceNtered XR" under the grant agreement No.101092612 and SEISMEC: "Supporting European Industry Success Maximization through Empowerment Centred development under the grant agreement No.101135884".

References

[1] J. Teizer, "Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites," *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 225–238, 2015.

[2] M. Sami Ur Rehman, M. T. Shafiq, and F. Ullah, "Automated computer vision-based construction progress monitoring: a systematic review," *Buildings*, vol. 12, no. 7, p. 1037, 2022.

[3] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," *Transdisciplinary perspectives on complex systems: New findings and approaches*, pp. 85–113, 2017.

[4] S. Xu, J. Wang, W. Shou, T. Ngo, A.-M. Sadick, and X. Wang, "Computer vision techniques in construction: a critical review," *Archives of Computational Methods in Engineering*, vol. 28, pp. 3383–3397, 2021.

[5] V. K. Reja, K. Varghese, and Q. P. Ha, "Computer vision-based construction progress monitoring," *Automation in Construction*, vol. 138, p. 104245, 2022.

[6] J. Wang, W. Sun, W. Shou, X. Wang, C. Wu, H.-Y. Chong, Y. Liu, and C. Sun, "Integrating bim and lidar for real-time construction quality control," *Journal of Intelligent & Robotic Systems*, vol. 79, pp. 417–432, 2015.

[7] V. Hoskere, Y. Narazaki, T. A. Hoang, and B. F. Spencer Jr, "Towards automated post-earthquake inspections with deep learning-based condition-aware models," *arXiv preprint arXiv:1809.09195*, 2018.

[8] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[9] J. Perlman, "Turbonerf-blender," https://github.com/JamesPerlman/TurboNeRF-Blender, Year you accessed the repository.

[10] P. Vrachnos, M. Krestenitis, I. Koulalis, K. Ioannidis, and S. Vrochidis, "A framework for 3d modeling of construction sites using aerial imagery and semantic nerfs," in *International Conference on Multimedia Modeling*. Springer, 2024, pp. 175–187.

[11] Z. Shang and Z. Shen, "Real-time 3d reconstruction on construction site using visual slam and uav," in *Construction Research Congress 2018*, 2018, pp. 305–315.

[12] I. Koulalis, N. Dourvas, T. Triantafyllidis, K. Ioannidis, S. Vrochidis, and I. Kompatsiaris, "A survey for image based methods in construction: from images to digital twins," in *Proceedings of the 19th International Conference on Content-based Multimedia Indexing*, 2022, pp. 103–110.

[13] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, "Computer vision algorithms and hardware implementations: A survey," *Integration*, vol. 69, pp. 309–320, 2019.

[14] R. Ali, J. H. Chuah, M. S. A. Talip, N. Mokhtar, and M. A. Shoaib, "Structural crack detection using deep convolutional neural networks," *Automation in Construction*, vol. 133, p. 103989, 2022.

[15] F. Kucuksubasi and A. Sorguc, "Transfer learning-based crack detection by autonomous uavs," *arXiv preprint arXiv:1807.11785*, 2018.

[16] D. Choi, W. Bell, D. Kim, and J. Kim, "Uav-driven structural crack detection and location determination using convolutional neural networks," *Sensors*, vol. 21, no. 8, p. 2650, 2021.

[17] J. Casas *et al.*, "D5.1 shm digital twin requirements for residential, industrial buildings and bridges," Unknown, Technical Report, 2022.

[18] C. Boje, A. Guerriero, S. Kubicki, and Y. Rezgui, "Towards a semantic construction digital twin: Directions for future research," *Automation in construction*, vol. 114, p. 103179, 2020.

[19] B. Chandrasekaran, J. R. Josephson, and V. R. Benjamins, "What are ontologies, and why do we need them?" *IEEE Intelligent Systems and their applications*, vol. 14, no. 1, pp. 20–26, 1999.

[20] "Rdf 1.1 primer (w3c working group note)," Web page, accessed on April 3, 2024. [Online]. Available: https://www.w3.org/TR/rdf11-primer/

[21] P. Hitzler, M. Krötzsch, B. Parsia, P. F. Patel-Schneider, and S. Rudolph, "Owl 2 web ontology language primer . w3c recommendation (11 december 2012)," *Online unter: https://www. w3. org/TR/2012/REC-owl2-primer-20121211*, 2012.

[22] A. Hogan, E. Blomqvist, M. Cochez, C. d'Amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier *et al.*, "Knowledge graphs," *ACM Computing Surveys (Csur)*, vol. 54, no. 4, pp. 1–37, 2021.

[23] J. Akroyd, S. Mosbach, A. Bhave, and M. Kraft, "Universal digital twin-a dynamic knowledge graph," *Data-Centric Engineering*, vol. 2, p. e14, 2021.

[24] "Ashvin – digitising and transforming the european construction industry," Web page, accessed on April 3, 2024. [Online]. Available: https://www.ashvin.eu/

[25] C. Ramonell, R. Chacón, and H. Posada, "Knowledge graph-based data integration system for digital twins of built assets," *Automation in Construction*, vol. 156, p. 105109, 2023.

[26] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM transactions on graphics (TOG)*, vol. 41, no. 4, pp. 1–15, 2022.

[27] "Concrete damage ontology," Web page, accessed on April 3, 2024. [Online]. Available: https://github.com/Alhakam/dot/blob/master/Extension/cdo.ttl

[28] "Damage topology ontology," Web page, accessed on April 3, 2024. [Online]. Available: https://w3id.org/dot/0.8

[29] J. L. Schonberger and J.-M. Frahm, "Structure-from-motion revisited," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4104–4113.