The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W12-2026
11th Intl. Workshop 3D-ARCH "3D Virtual Reconstruction and Visualization of Complex Architectures", 10–12 February 2026, Ancona, Italy

From Survey to Action: AI-Driven Severe Damage Mapping

Kai Zhang [1], Asli Tekin[1], Chiara Mea[1], Younan Stephanie[2], Michel Chalhoub[2], Francesco Fassi [1]

[1]3D Survey Group, ABC Department, Politecnico di Milano, Via Ponzio 31, 20133 Milano, Italy
– {kai.zhang, chiara.mea, francesco.fassi}@polimi.it, asli.tekin@mail.polimi.it
[2] DISTRUCT Solutions Sarl, Deek bldg, Hadara street, Zalka, Lebanon
– {stephanie.younan, chalhoub}@distruct.com

Keywords: AI, Segmentation, Heritage in Risk, Survey, Post-disaster, Damage Assessment.

Abstract

Post-disaster rapid response relies on the timely acquisition of information. Specifically, assessing structural integrity of building requires fast and reliable analysis, which can be supported by quantitative damage assessment based on the complete 3D geometry knowledge of structures. Such information can be derived either from quickly acquired imagery (e.g., video frames or panoramas) or directly from 3D data. Today, data collection is relatively straightforward thanks to UAV surveys and mobile mapping systems; however, extracting actionable information remains time-consuming when performed manually. This highlights the need for automated methods that can localize damage, identify critical issues, and support interpretation and decision-making.
In recent years, artificial intelligence (AI) has attracted substantial attention in this domain, driven by the emergence of models that deliver fast and effective performance across a range of perception tasks. Yet the success of these approaches remains strongly conditioned by data availability and quality. In many application settings—especially those involving rare or highly specific post-disaster scenarios—representative training samples are scarce, underscoring the need for dedicated datasets construction, together with methods capable of learning from limited data. In this research, a dataset dedicated to the detection of cracks was collected. The aim is to have quick and straightforward identification of decay related to the structural stability of the building. Trained YOLOv11 object detection and segmentation model were tested on two case studies collected in Lebanon. These two case studies feature structural damage under severe external force, representative for the post-disaster scenarios analysis. The research evaluated the proposed hybrid approach (involving deep learning and machine learning methods and integrated with photogrammetric workflow), retrieving architectural severe damage location from 2D and localizing them effectively in 3D to assess global analysis, by comparing data with the ground truth in both 2D and 3D data.

1. Introduction

The recent development of survey and data processing technology has proven the possibility of quick survey and analysis of many types of physical environments, providing comprehensive solutions for both extensive and complex environmental, urban, architectural and even industrial scenarios. The characteristics of the survey tools can serve as a basis for the next step of real-time application, which can be found useful for 3D data interpretation, increasing the knowledge of structural characteristics of the object. The information extraction process must be rapid, even if not fully real-time, as time efficiency is critical for enabling prompt intervention. This is the case for high-risk post-disaster or post-conflict scenarios, where immediate reaction and real-time decision-making are vital. This paper examines the applicability of AI to survey methodologies in these scenarios, focusing on post-event responses, particularly the structural threats caused by cracks. The study evaluates the expected capabilities of AI techniques and contemporary survey tools, identifies areas for methodological improvement to reduce future adverse impacts. The long-term idea is organizing the decay phenomena into emergency levels, using AI to allow classification and quantification of the decay, together with the 3D mapping to localize the issues with the aim to improve the level of understanding of the damage.
Nowadays, modern survey methods are diverse and increasingly mature. Cameras and laser scanners can acquire swift and accurate 2D and 3D data; drones provide flexibility of digitalization activities. Furthermore, mobile mapping and Simultaneous Locationing and Mapping (SLAM) allow for autonomous mapping activities. These advances in technology make it feasible to conduct a fast survey in the severely damaged area, as the complicated physical condition greatly limits the time availability of data acquisition.
Post-disaster response critically depends on accurate decisions made under severe time constraints. A central challenge is the effective use and analysis of the large volumes of data acquired in the immediate aftermath of an event, when the speed of identifying priorities can directly influence outcomes. Because available time is limited, comprehensive assessment of these datasets quickly becomes difficult, especially when carried out manually. In this context, AI-based automatic processing offers a promising solution. Deep learning methods for 2D data analysis have already demonstrated strong performance across a wide range of applications, suggesting clear potential for accelerating post-disaster assessment workflows. However, research on automated post-disaster data processing remains limited, and, importantly, there is still a lack of standardized, publicly available datasets specifically collected in post-disaster scenarios. Therefore, effort is required in this field.
This research integrates object detection and segmentation of 2D images into photogrammetric processing pipelines. A dataset is elaborated for the proposed methods, allowing the practical tests on the data collected in Lebanon, evaluating the feasibility and potentiality of using AI for post-disaster survey activities.

2. Related Works

Post disaster analysis has become a central research area in disaster risk engineering, with numerous studies emphasizing the need for rapid and reliable information to support emergency response and early recovery (Shafian et al., 2024). Review papers typically organize the work into several thematic circles centered

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-2-W12-2026-519-2026

on data collection modalities, such as satellite and aerial imagery, unmanned aerial vehicles (UAV) acquired data, ground-based imaging, synthetic aperture radar (SAR) observations, and crowd sourced visual information across both weather related and seismic events.

These modalities have been widely applied in real post disaster scenarios: satellite and ultra-high resolution aerial imagery combined with transformer based architectures have enabled multi class building damage recognition after hurricanes, achieving 88% accuracy following Hurricane Michael (Singh et al., 2023), while crowd AI frameworks integrating street level and satellite imagery reached 83% accuracy for residential damage classification in hurricane affected areas (Cheng et al., 2023). UAV based mapping has similarly been used to generate rapid, high resolution damage assessments in earthquake affected regions (Gu et al., 2024), and SAR based models have demonstrated strong performance in detecting collapsed buildings under cloud covered or night-time conditions (Macchiarulo et al., 2025; Rao et al., 2023). More recent multi-modal approaches that fuse imagery with non-imagery data (such as wind speed, building characteristics, and proximity to the hazard footprint) have further improved performance, with the Multi Modal Swin Transformer achieving 92.67% accuracy after Hurricane Ian (Xue et al., 2024).

Recent studies also highlight a shift toward advanced data collection technologies and computational methods, reflecting the field's move toward scalable, data driven assessment workflows rather than traditional manual inspections (Shafian et al., 2024). However, despite this methodological evolution, the development of image-based damage detection systems remains constrained by the limited availability of extensive, well labelled datasets, which continues to hinder the training of robust and generalizable models (Dondi et al., 2025).

Across these studies, disaster-related damage is conceptualized as a multi-mechanism phenomenon that can arise from both seismic actions and pre-existing deterioration processes (Dogan et al., 2023). Early work shows that the intricate and unique architectural configurations of heritage structures complicate systematic inspections, making the identification and interpretation of post-disaster damage more challenging (Wang et al., 2020). More recent investigations focusing on reinforced concrete buildings highlight that earthquake-induced deterioration commonly manifests through excessive deformation, cracking, concrete spalling, and, in severe cases, partial or total collapse, all of which pose significant risks to structural safety and economic resilience (Feng et al., 2024).

Among these, cracks are considered one of the most revealing forms of damage, prompting the development of parametric annotation schemes capable of capturing their variability and generating realistic crack instances under expert defined constraints to better represent the diversity of severe damage scenarios (Dondi et al., 2025).

For certain disaster that could have effects on regional scale, for example earthquakes, have received considerable attention in AI-based assessment studies. At the regional scale, researchers employ damage detection, damage classification, and multi-source data-fusion approaches using CNN-based models applied to satellite imagery, ground-motion records, and social-media content, with satellite-derived data being the most widely used (Jia and Ye, 2023).

For example, a multisource machine learning framework was applied to the 2015 Nepal earthquake to perform pixel level damage classification and micro-zone damage ratio estimation across five grades (Chen et al., 2022). Extending image-based post-earthquake analysis to a broader international context, post-earthquake images from Haiti, Nepal, Taiwan, Ecuador, and Pohang are evaluated using a four level damage dataset (no damage, light, moderate, and severe) to perform structural damage level recognition (Zhuang et al., 2025). In Türkiye, a dataset of 1,040 damaged RC elements collected after the 2019 Silivri and 2020 Elazig earthquakes enabled the development of a model distinguishing earthquake-induced damage from reinforcement corrosion with 90.62% accuracy (Dogan et al., 2023). Similarly, a four level damage dataset (undamaged, slightly damaged, moderately damaged, and heavily damaged) collected from buildings affected by the 2023 Türkiye earthquake is used to classify crack-related damage severity, with the DenseNet201–KNN hybrid model achieving the highest accuracy of 94.62% among the evaluated architectures (Reis et al., 2025). Additional studies detect spalling and crack patterns on plastered masonry walls using ground-based and UAV imagery, though performance remains sensitive to lighting, background variation, and irregular geometries (Ling et al., 2025).

At the individual building scale, post-disaster AI research spans a wide range of tasks, beginning with the detection and segmentation of surface level defects such as cracks, spalling, and material loss, which serve as the earliest and most direct indicators of structural distress (Bai et al., 2021; Reis et al., 2025). Deep learning models are also employed to segment crack patterns on earthquake damaged masonry buildings, preserving their geometry and continuity to support reliable diagnostic interpretation (Pantoja Rosero et al., 2022). Moving beyond element level defects, building level approaches classify overall damage severity into multiple grades by analysing visual cues such as cracking, spalling, deformation, and partial collapse (Saquella et al., 2025). In parallel, post-earthquake image datasets have been used to classify structural damage across a full spectrum (from light hairline fractures to complete collapse) with MobileNet achieving the highest accuracy among seven tested models (Zhuang et al., 2025). Recent advances in lightweight object detection architectures have further strengthened crack level assessment, with improved YOLOv8 based models achieving real-time and high accuracy detection of concrete surface cracks through attention enhanced modules and optimized convolutional structures (Dong et al., 2024).

3. Methodology

3.1 General workflow

The presented research starts from the initiative of establishing a generalizable workflow that can be used for general post-disaster decision making. It aims to generate a compact dataset that provides representative information to train the AI model. 2D data derived from various sources is used for learning phase, associating semantic to pixels. These semantics can, later, contribute to 3D analysis thanks to inference performed on the same 2D data used to reconstruct the asset thanks to 3D survey. The proposed methods are tested on manual annotated ground truth in both 2D and 3D.

The first step involves a process of data preparation, including data collection, data processing and data annotation process, with the purpose of training a deep learning architecture.

As to detect the decay phenomena, this research tested two methods: the first one use YOLO v11(Jocher and Qiu, 2024) , specifically YOLO11m[1] for object detection task. It is trained to

[1] In YOLO11m, the suffix m stands for medium, indicating a mid-sized model variant of YOLO11 that offers a balance between accuracy and computational cost.

https://doi.org/10.5194/isprs-archives-XLVIII-2-W12-2026-519-2026 | © Author(s) 2026. CC BY 4.0 License.

detect the wanted category and output the associated bounding boxes. Then, the detection box results are used to prompt the pretrained model Segment Anything Model (SAM2) (Kirillov et al., 2023; Ravi et al., 2024), to segment the object within the boxes. The method is later addressed as YOLO-SAM. The second method is to directly train YOLO segmentation model (YOLO11m-seg) to segment the crack in images. Both methods will generate binary masks of detected objects.
Second step is to apply the trained model to unseen dataset, projecting the 2D outcome to 3D. To do this, the 2D masks are used to recreate the model texture, the normalized weighted score of confidence ($\mathrm{W(p)} \in [0,1]$, see below) will be mapped on the 3D model, acquiring the grayscale texture ($\mathrm{G(u,v)} \in [0,255]$) in 3D after its rasterized, interpolated, and quantized realization in the texture coordinate domain.

$$W(p) = \frac{\sum_{i \in \mathcal{V}(p)} w_i(p)}{\max_q \left(\sum_{i \in \mathcal{V}(q)} w_i(q) \right)}$$

where p = a surface point / texture element
$\mathcal{V}(\mathrm{p})$ = set of images where p is visible
$\mathrm{w_i(p)}$ = contribution weight of image i to point p
$\mathrm{W(p)} \in [0,1]$ = final grayscale value

3.2 Training Dataset

Based on a previous study (Zhang et al., 2024), this research is intended to limit the detection categories to the specific post-damage scenarios, in sense of level of emergency, and to support the decision making. Given the nature of post-disaster damage, cracks frequently constitute the most immediate and morphologically distinct indicators of structural distress. Their formation is typically associated with mechanisms such as stress concentration, environmental loading, and the progressive degradation of material properties in systems ranging from concrete to asphalt and metal (Jiya et al., 2016). In concrete structures, they often represent the earliest visible sign of deterioration, and if left unaddressed, their propagation can significantly reduce structural integrity and escalate into severe failures with substantial human and economic consequences (Nyathi et al., 2024). As cracks extend, they compromise load-bearing components and may lead to irreversible damage, making timely identification critical for preventing further deterioration (Kaveh and Alhajj, 2024; Ranieri et al., 2025). These characteristics underline the need for a representative dataset that captures the variability and complexity of crack patterns observed in real post-disaster scenarios.
To address this, and to ensure methodological robustness and contextual representativeness, a dataset was assembled from a diverse range of open-access and publicly available sources. Unlike conventional crack datasets -typically captured under controlled conditions, e.g. nadiral perspective and light environment, and limited to uniform materials or lighting- our collection reflects the complexity of real-world post-disaster environments. The incorporated datasets include the Damage Detection Dataset for Concrete Structures with Multi-Feature Backgrounds (Raushan et al., 2024), the Post-earthquake Concrete Crack Dataset (PECCD) (Mayya et al., 2025) collected along the Syrian Coast following the 2023 Turkey-Syria earthquake, and the ImageOP-Crack Dataset (Luiz Carvalho Ottoni and Toledo Cordeiro Ottoni, 2025), which documents 24 religious buildings across four historic Brazilian cities (Ouro Preto, São João del-Rei, Tiradentes, and Itapecerica). These sources collectively span a wide range of resolutions (416 × 416 pixels to 4608 × 3456 pixels), camera types, and compression qualities, capturing surfaces composed of concrete and plaster with deterioration levels ranging from micro-cracking to severe structural failure.
In addition, the dataset was enriched with post-disaster and post-conflict documentation sourced from international news media. These images, captured under highly variable illumination conditions (including harsh sunlight, deep shadow, and mixed artificial lighting) further contribute to the dataset's heterogeneity. Images of the damaged town hall in Sant'Agostino di Ferrara were collected from coverage of the 2012 Emilia-Romagna earthquake, while visual records of missile-induced damage to the Spaso-Preobrazhensky Cathedral in Odesa were compiled from global news outlets following its inclusion in the UNESCO List of World Heritage in Danger in 2023 (UNESCO World Heritage Centre, 2023).
Furthermore, photogrammetric imagery from multiple case studies was provided by DISTRUCT Solutions Sarl. The images were collected from conflict-affected sites in Tripoli and Beirut, as well as from earthquake-affected areas in Turkey, and were used to represent post-conflict and post-earthquake damage typologies.
Altogether, the image sets incorporated documentation from Italy (2012 Emilia-Romagna earthquake), Ukraine (2023 Odesa missile strikes), Lebanon (conflict-related damage in Tripoli and Beirut), Turkey and Syria (2023 earthquake), and Brazil (heritage records from four historic cities) (Figure 1). It consists of, in total, 744 images.

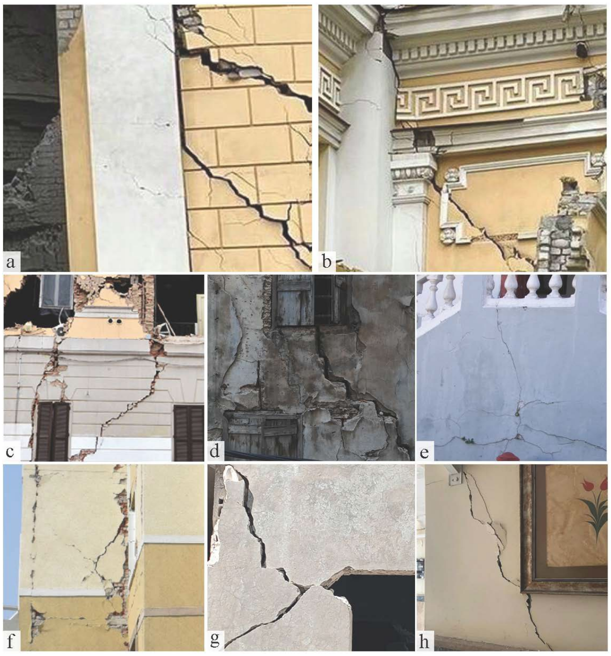

Figure 1. Image samples from the dataset: (a-b) Ukraine (2023 missile strike), (c) Italy (2012 earthquake), (d) Lebanon (conflict-related damage), (e) Brazil (heritage sites), (f) Mixed, (g) Syria (2023 earthquake), (h) Turkey (2023 earthquake).

To optimize the dataset for training, a structured preprocessing pipeline was implemented. The raw imagery was first divided into 2 × 2 patches, a step designed to reduce intra-image variability and enhance the detection of localized crack features. Subsequently, all images underwent resolution normalization: samples with a shorter edge exceeding 640 pixels were rescaled so that the shorter edge equaled 640 pixels, while preserving the original aspect ratio. This operation standardized the dataset to a uniform resolution of 640 × n pixels, thereby ensuring consistency across samples.

Following cropping and normalization, the dataset expanded to 2,976 images. A curation stage was then applied, in which non-relevant samples lacking visible crack patterns were eliminated. The final dataset is comprised of 2,034 images, representing a balanced and domain-specific collection of concrete crack scenarios (**Figure 2**). In total, there annotated by human 13344 crack instances, with commonly 6.56 instances per image, with area taking up 11.86%. Accordingly, the dataset was partitioned into training, validation, and test subsets following a 70–20–10 ratio, yielding 1,423 images for training, 407 images for validation, and 204 images for testing.

An YOLO11m model is trained based on this annotated dataset. It arrived convergence at 150 epoch with performances on evaluation set at 46.5% precision, 0.46 F1 score at confidence of 0.217, with 41% mean Average Precision (mAP) at Intersection over Union (IoU) 0.5.

For the segmentation task, an identical preprocessing workflow to that used in the detection pipeline was employed to maintain methodological consistency across experiments. The segmentation stage utilized the pre-processed dataset, in which images had already been divided into 2 × 2 patches and rescaled to a shorter-edge resolution of 640 pixels during the earlier detection pipeline. From this processed pool, 132 images were selected to represent a broad range of crack morphologies and surface conditions across the building envelope (Figure 2). There are annotated by human 839 crack instances, with commonly 5.05 instances per image, with box area taking up 45.59%, mask area taking up 7.84%. Additionally, the dataset also leverages an open source dataset on crack segmentation using online annotation tool Roboflow (University, 2022). This dataset contains in total 4029 samples with mostly the size of 416 × 416 pixels (Figure 2). There are present 5290 crack instances, with commonly 1.31 instances per image, with annotated box area taking up 24.47%, mask area taking up 1.89%.

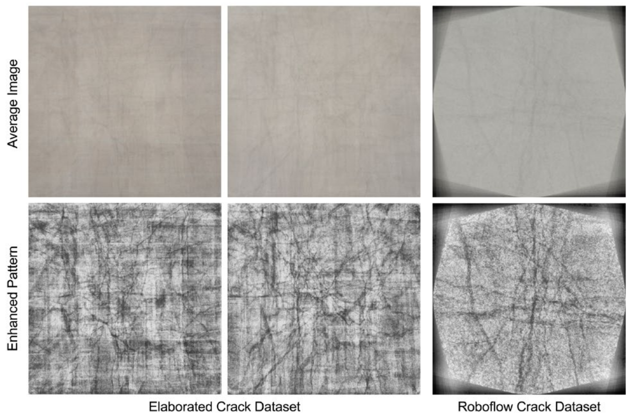

Figure 2. Image data analysis of crack: creating average image (mean of colour values) of random 50 samples (up) and corresponding CLAHE enhanced pattern (bottom)

In the final configuration, the manual annotated dataset followed a 70–15–15 split: 70% images were allocated to the training set, while the validation and test set each contained 15% images. Prior to training, the new annotated training subset was duplicated 3 times to better focus on the post-disaster scenario, enhancing the model's ability to learn fine-grained crack boundaries.

An YOLO11m-seg model is trained based on this annotated dataset. The model was trained with images compressed to 640 pixels by side, arrived convergence at 45 epoch with performances on evaluation set at 0.45 F1 score at confidence of 0.316.

4. Case studies

This research tests AI trained models on survey data collected from Tripoli, Lebanon. Two case studies were selected, both located near the Citadel of St. Gilles, within the historic urban fabric of the old city. They represent the common features of the involved sites with significant cultural heritage value. At the same time, they have all suffered severe architectural damage (including cracks, wall displacements, ceiling collapses, and loss of architectural elements), necessitating a prompt analysis of next structural consolidations. The survey and analysis in this condition are vital for accessing the structural threats.

4.1 The first case study

First case study is a three-story traditional residential building constructed with masonry load-bearing walls and finished with lime-based plaster, Baghdadi timber ceilings, and a roof covered with red clay tiles. The building reflects construction techniques and material typologies characteristic of late Ottoman-period urban housing in Tripoli, both in its structural system and architectural expression, and represents a typical example of traditional residential architecture in the region (Figure 3).

Figure 3. Image of the case study

This building presents multiple forms of structural and material deterioration that compromise both its architectural integrity and structural performance. In-plane cracking is widely observed in the load-bearing masonry walls, occurring mainly as vertical, diagonal, and stepped cracks. These patterns indicate stress redistribution within the masonry, likely related to long-term material degradation, incompatibility between original materials and later interventions, and a lack of maintenance. In several areas, cracking follows mortar joints and masonry interfaces, suggesting a progressive reduction in cohesion and shear capacity. In addition, the building exhibits significant out-of-plane deformation of masonry walls, including bulging, tilting, and localized displacement from their original vertical alignment. Partial collapse of ceilings represents another major form of degradation.

Overall, the combined effects of cracking, deformation, and localized collapse have led to a substantial loss of original materials, including masonry units, lime mortar, plaster layers, and timber elements. These damage mechanisms have significantly reduced the building's structural continuity and its ability to function as an integrated load-bearing system. Without timely stabilization and conservation measures, the ongoing deterioration poses a serious risk to both the structural safety and long-term preservation of the building

The survey work of this residential building was conducted in 2025, which involved photogrammetric survey and laser scanning. The overall dataset with 237 images includes 160 aerial images using a DJI Mini 4 Pro at 4032*2268 resolution, and 77 camera images at 5772*3648 resolution for 3D reconstruction purpose. All the images are directly collected from the site.

4.1.1 The ground truth preparation

To evaluate the results, a subset of the dataset was manually annotated, and cracks were identified by hand both on 2D images and the 3D point cloud. The 2D images annotation were used to identify smaller cracks not directly clearly visible on 3D and were projected to the 3D in a second phase creating a map of confidence.

In the end, over 600+ instances were labelled. Bounding-box annotations of cracks were used to prompt segmentation by the SAM2 model, resulting in masks representing the decay phenomenon (Figure 4). In most cases, the target objects are correctly delineated; however, adjacent surrounding regions are often included in the segmentation results.

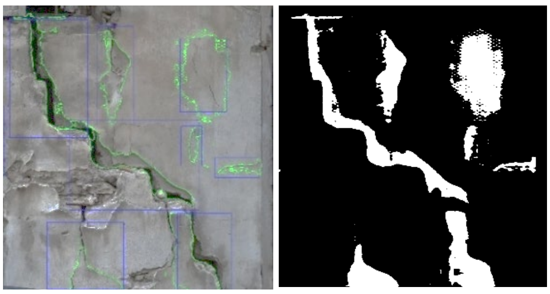

Figure 4. Using annotated boxes to prompt SAM2 for generating the masks of crack.

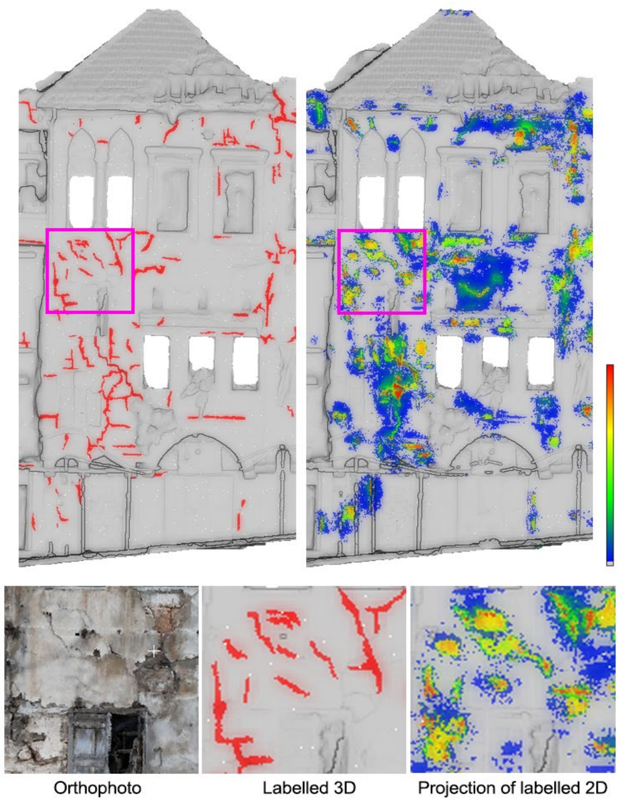

Figure 5. Annotated crack directly on 3D (top left) comparing with the recreated the 3D texture (top right) using the SAM2 masks.

These masks were subsequently used to regenerate the model texture using a mosaic blending mode. During this process, target regions in the 3D model are colorized according to the corresponding masks, with priority given to optimal viewing geometry, resolution, and image sharpness. As a result, the final value reflects the accumulation of overlapping detections within the same areas. Interpretated the value from blue to yellow and red, as a 3D heat map, the visualization represents the probability of decay occurrence, as shown in Figure 5: higher values are rendered in red, indicating an higher likelihood of crack presence; the blue indicates this surface region is less supported by image evidence.

4.1.2 The cracks detection methodology

As mentioned before; all the data were elaborated using the two methods: YOLO-SAM and YOLO11m-seg. The detection model and segmentation models were evaluated on 2D data with manual annotated ground truth. While YOLO11m detection arrives at box precision 58%, 0.318 mAP at IoU 0.5, the recall reaches 6.2% and 16.9% at 0.4 and 0.2 confidence threshold correspondingly. The YOLO11m-seg detection arrives at 15.5% mAP at IoU 50%, while box F1 reaches 0.12, the mask F1 reaches only 0.03 at confidence 0.156. Inspecting the plotted results, the YOLO11m object detection model demonstrates strong performance in identifying cracks, but it tends to emphasize minor cracks. That means that, for crack delineation, the segmentation model produces more accurate results and preserves the spatial continuity of the crack geometry. But YOLO11m-seg is prone to misclassifying material gaps (e.g., between the window and the wall) as cracks and being easily distracted by the presence of wires. (Figure 6)

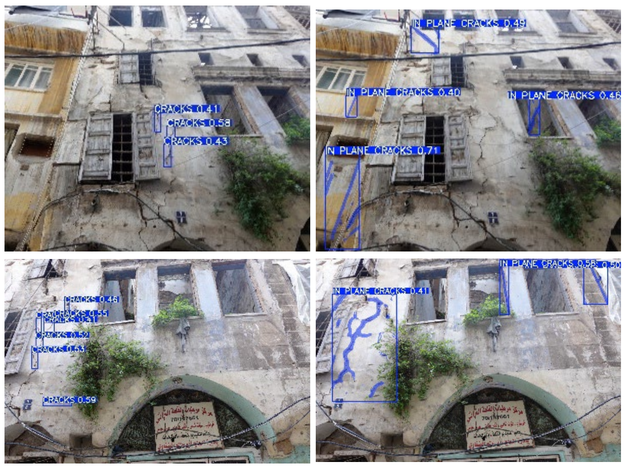

Figure 6. YOLO11m results (left) and YOLO11m-seg results (right) on 2D imageries in the case study.

The box detection results from YOLO11m were then used to indicate SAM2 model for the segmentation of the crack. Then the outcome and the YOLO11m-seg results are used to recreate the texture of the 3D mesh model (Figure 7). In both results, regions with value generally correspond to the spatial distribution of cracks. The YOLO11m-Seg outcome shows higher value accumulation along crack locations; however, it also highlights several incorrect regions and covers a broader extent of irrelevant areas, as indicated by the widespread blue regions.

Comparing to the 3D manual annotated ground truth, YOLO-SAM and YOLO11m-seg results reaches overall accuracy 0.97 and 0.85 with setting confidence of 40% telling apart the crack and the background surface. If consider the grayscale value [0-255] in 3D outcome as confidence, and map F1 score at threshold of each degree of confidence, this curve shows at what degree of confidence the methods arrive to highest recall on segmenting the 3D. Practically, at the peak point, the outputted crack distribution is most reliable. From this F1-confidence curve, the F1 of the mapping reaches 0.15 and 0.025 when confidence reaches 2.5 and 5.0 in respect to YOLO-SAM and YOLO11m-seg (Figure 8).

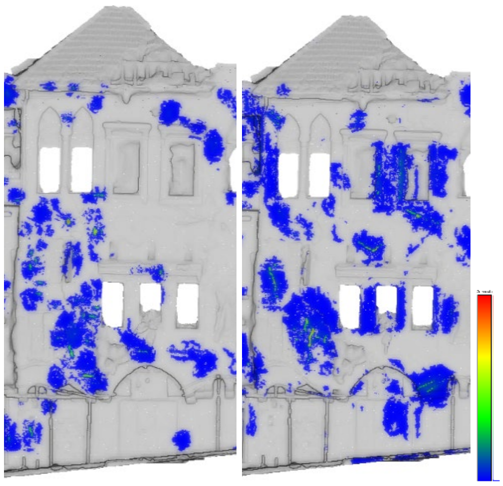

Figure 7. YOLO-SAM result (left) and YOLO11m-seg results (right) projected in 3D in the case study

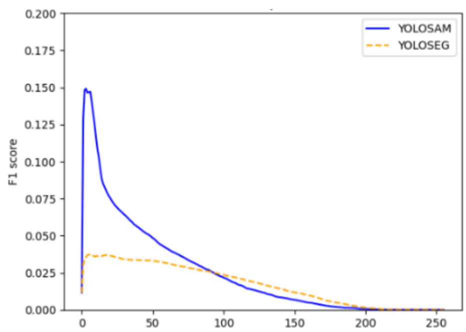

Figure 8. F1-confidence curve of the two methods

4.2 The second case study

The second case study is also a three-storey residential structure, situated within a highly compact urban fabric, where adjacent residential structures are directly connected or separated by minimal gaps (Figure 9). It was built primarily with masonry load-bearing walls, Baghdadi timber ceilings, and cement-based plaster finishes resulting from later interventions. While maintaining the traditional structural typology, the use of modern materials has introduced material incompatibilities that affect its structural behaviour and state of conservation.

Figure 9. Image of the case study

The building is primarily affected by distinct in-plane cracking along the main façade, particularly within the mortar plaster covering the stone masonry, together with localized partial collapse of the ceilings. These damage patterns indicate a progressive loss of material cohesion and a reduction in load-transfer capacity within the structural system. As a result, both the architectural coherence and structural continuity of the building have been significantly compromised.

The survey work of this residential building was conducted in 2025, which involved photogrammetric survey and laser scanning. The overall dataset with 316 images at 5772*3648 resolution for 3D reconstruction purpose. All the images are directly collected from the site.

The two methods are evaluated firstly on the collected 2D images and then projected to 3D, making comparison with annotated 3D ground truth. While YOLO11m detection arrives at box precision 56.2%, 0.312 mAP at IoU 0.5, the recall reaches 9.1% and 24.7% at 0.4 and 0.2 confidence threshold correspondingly. The YOLO11m-seg detection arrives at 16.2% mAP at IoU 50%, while box F1 reaches 0.15, the mask F1 reaches only 0.14 at confidence 0.171. Inspecting the plotted results, the YOLO11m model accurately mapped the cracks on the decaying surfaces, while it lacks on mapping the spatial continuity of the crack geometry. The YOLO11m-Seg model achieves continuous segmentation of cracks and still prone to misclassifying and distraction.

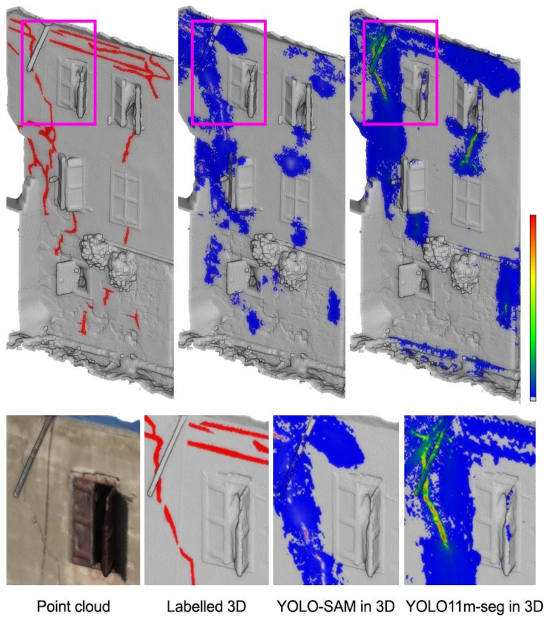

Figure 10. YOLO-SAM result (in the centre) and YOLO11m-seg (on the right) results projected in 3D in the case study comparing with the ground truth (on the left).

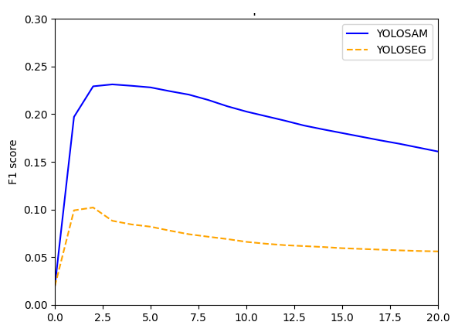

Figure 11. F1-confidence curve of the two methodsss

The detection results from YOLO11m were used to indicate SAM2 model for the segmentation of the crack. The outcome and the YOLO11m-seg results are used to recreate the texture of the 3D mesh model (Figure 10). Comparing to the annotated 3D ground truth, YOLO-SAM and YOLO11m-seg results reaches overall accuracy 0.92 and 0.89 with setting confidence of 40%.

From the precision-confidence curve, both curves achieve higher performance compared to the first case study. The smoothness of the curves indicates that the data feature exhibit limited variation across viewing angles (Figure 11).
For the applications, both YOLO11m model and YOLO11m-seg model were adequately trained and achieved optimal performance on the evaluation set (see 3.2). However, performance degradation exhibit when applied to real-world data. One contributing factor is the limited representation of negative samples in the training set. Specifically, examples that illustrate what does *not* constitute the target category, i.e. in-plane crack in this research. In both case studies, YOLO11m-Seg frequently misclassified wires and architraves as cracks; these instances represent valuable negative examples that could be used to refine the model. Generally, it is difficult to include all visually similar but irrelevant samples in the process of dataset curation, but such samples are critical for improving robustness.
Another contributing factor is the inherent ambiguity of the concept “crack”, as a physical phenomenon. It leads a wide range of geometric and visual characteristics. The diversity of materials, dimensions, orientations, and surface conditions leads to overlapping features that challenge model generalization.

5. Conclusion

This study evaluated AI methods within modern surveying workflows for post-disaster scenarios, enabling the extraction of semantic information and spatial localization in 3D scenes. While the challenges of achieving strong generalization in detection models under real-world conditions are acknowledged, the study nonetheless provides valuable insights and lessons for future development and application.
The case study applications demonstrate the potential of these AI methods and their integration with 3D reconstruction. Based on the 3D test results, the YOLO–SAM approach outperforms the YOLO11m-Seg model in crack localization, although both methods exhibit low recall. However, the comparison is insufficient to establish the general performance of YOLO–SAM. Despite this, its higher training efficiency and reduced data annotation time make it a more practical approach. Considering that the 2D segmentation on 2K resolution imagery arrives to within 100ms using Cuda, the general process allows the visual inspection of the detection within hours, supporting timely damage assessment. Moreover, the probability information embedded in the results quantifies the confidence of crack presence across the 3D scene, facilitating more informed interpretation and prioritization.
There are multiple lessons to be noted. Among all, data availability remains a primary concern, as operative activities access may be restricted due to reality constraints. For the monitoring of the individual case study, model shall be trained on the data from target domain. To be generalized across different application scenarios, more data are required and should be consolidated into a unified dataset to train models. Particularly, typical negative samples corresponding to the target domain must be included, enhancing direct post-disaster deployment and analysis. Future research should also focus on defining crack subcategories based on attributes such as material type, scale, and orientation, combined with targeted data collection and standardized processing protocols, to improve overall model performance and generalizability.

References

Bai, Y., Sezen, H., Yilmaz, A., 2021. Detecting Cracks and Spalling Automatically in Extreme Events by End-to-end Deep Learning Frameworks. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. V-2–2021, 161–168. https://doi.org/10.5194/isprs-annals-V-2-2021-161-2021

Chen, J., Tang, H., Ge, J., Pan, Y., 2022. Rapid Assessment of Building Damage Using Multi-Source Data: A Case Study of April 2015 Nepal Earthquake. Remote Sens. 14, 1358. https://doi.org/10.3390/rs14061358

Cheng, C.-S., Khajwal, A.B., Behzadan, A.H., Noshadravan, A., 2023. A Probabilistic Crowd–AI Framework for Reducing Uncertainty in Postdisaster Building Damage Assessment. J. Eng. Mech. 149, 04023059. https://doi.org/10.1061/JENMDT.EMENG-6992

Dogan, G., Hakan Arslan, M., Ilki, A., 2023. Detection of damages caused by earthquake and reinforcement corrosion in RC buildings with Deep Transfer Learning. Eng. Struct. 279, 115629. https://doi.org/10.1016/j.engstruct.2023.115629

Dondi, P., Gullotti, A., Inchingolo, M., Senaldi, I., Casarotti, C., Lombardi, L., Piastra, M., 2025. Improving Post-Earthquake Crack Detection using Semi-Synthetic Generated Images. pp. 19–35. https://doi.org/10.1007/978-3-031-91907-7_2

Dong, X., Liu, Y., Dai, J., Dong, X., Liu, Y., Dai, J., 2024. Concrete Surface Crack Detection Algorithm Based on Improved YOLOv8. Sensors 24. https://doi.org/10.3390/s24165252

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD’96. AAAI Press, Portland, Oregon, pp. 226–231.

Feng, D.-C., Yi, X., Deger, Z.T., Liu, H., Chen, S.-Z., Wu, G., 2024. Rapid post-earthquake damage assessment of building portfolios through deep learning-based component-level image recognition. J. Build. Eng. 98, 111380. https://doi.org/10.1016/j.jobe.2024.111380

Gros, A., Guillem, A., De Luca, L., Baillieul, É., Duvocelle, B., Malavergne, O., Leroux, L., Zimmer, T., 2023. Faceting the post-disaster built heritage reconstruction process within the digital twin framework for Notre-Dame de Paris. Sci. Rep. 13, 5981. https://doi.org/10.1038/s41598-023-32504-9

Gu, J., Xie, Z., Zhang, J., He, X., Gu, J., Xie, Z., Zhang, J., He, X., 2024. Advances in Rapid Damage Identification Methods for Post-Disaster Regional Buildings Based on Remote Sensing Images: A Survey. Buildings 14. https://doi.org/10.3390/buildings14040898

Jia, J., Ye, W., 2023. Deep Learning for Earthquake Disaster Assessment: Objects, Data, Models, Stages, Challenges, and Opportunities. Remote Sens. 15, undefined-undefined. https://doi.org/10.3390/rs15164098

Jiya, E.A., Anwar, N.S.N., Abdullah, M.Z., 2016. Detection of Cracks in Concrete Structure Using Microwave Imaging Technique. Int. J. Microw. Sci. Technol. 2016, 3195716. https://doi.org/10.1155/2016/3195716

Jocher, G., Qiu, J., 2024. Ultralytics YOLO11.

Kaveh, H., Alhajj, R., 2024. Recent advances in crack detection technologies for structures: a survey of 2022-2023 literature. Front. Built Environ. 10. https://doi.org/10.3389/fbuil.2024.1321634

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. https://doi.org/10.48550/arXiv.2304.02643

Ling, L., Ma, G., Hwang, H.-J., Tan, X., 2025. Post-earthquake detection of surface spalling and cracks in masonry buildings based on computer vision. Structures 78, 109226. https://doi.org/10.1016/j.istruc.2025.109226

Luiz Carvalho Ottoni, A., Toledo Cordeiro Ottoni, L., 2025. ImageOP-Crack. Mendeley Data 1. https://doi.org/10.17632/m8f87ztdyw.1

Macchiarulo, V., Giardina, G., Milillo, P., Aktas, Y.D., Whitworth, M.R.Z., 2025. Integrating post-event very high resolution SAR imagery and machine learning for building-level earthquake damage assessment. Bull. Earthq. Eng. 23, 5021–5047. https://doi.org/10.1007/s10518-024-01877-1

Mayya, A.M., Alkayem, N.F., Saii, M., Ahmad, M.H., Bayat, M., Asteris, P.G., Cao, M., 2025. Post-earthquake concrete crack dataset (PECCD). Mendeley Data 1. https://doi.org/10.17632/w7549ryvx2.1

Mouaddib, E.M., Pamart, A., Pierrot-Deseilligny, M., Girardeau-Montaut, D., 2024. 2D/3D data fusion for the comparative analysis of the vaults of Notre-Dame de Paris before and after the fire. J. Cult. Herit., Notre-Dame de Paris : a multidisciplinary scientific site 65, 221–231. https://doi.org/10.1016/j.culher.2023.06.012

Nyathi, M.A., Bai, J., Wilson, I.D., Nyathi, M.A., Bai, J., Wilson, I.D., 2024. Deep Learning for Concrete Crack Detection and Measurement. Metrology 4, 66–81. https://doi.org/10.3390/metrology4010005

Pantoja Rosero, B., Oner, D., Kozinski, M., Achanta, R., Fua, P., Perez-Cruz, F., Beyer, K., 2022. TOPO-Loss for continuity-preserving crack detection using deep learning. Constr. Build. Mater. 344, 128264. https://doi.org/10.1016/j.conbuildmat.2022.128264

Ranieri, A., Palmieri, G., Biasotti, S., 2025. Automated Monitoring of Cultural Heritage Artifacts Using Semantic Segmentation.

Rao, A., Jung, J., Silva, V., Molinario, G., Yun, S.-H., 2023. Earthquake building damage detection based on synthetic-aperture-radar imagery and machine learning. Nat. Hazards Earth Syst. Sci. 23, 789–807. https://doi.org/10.5194/nhess-23-789-2023

Raushan, R., Singhal, V., Kumar Jha, R., 2024. Damage Detection Dataset for Concrete Structures with Multi-Feature Backgrounds. Mendeley Data 1. https://doi.org/10.17632/9t3y6hhddk.1

Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K.V., Carion, N., Wu, C.-Y., Girshick, R., Dollár, P., Feichtenhofer, C., 2024. SAM 2: Segment Anything in Images and Videos. https://doi.org/10.48550/arXiv.2408.00714

Réby, K., Guilhelm, A., De Luca, L., 2023. Semantic Segmentation using Foundation Models for Cultural Heritage: an Experimental Study on Notre-Dame de Paris, in: 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Presented at the 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), pp. 1681–1689. https://doi.org/10.1109/ICCVW60793.2023.00184

Reis, H.C., Turk, V., Ustuner, M., Kaya Yildiz, C.M., Tatli, R., 2025. Post-seismic structural assessment: advanced crack detection through complex feature extraction using pre-trained deep learning and machine learning integration. Earth Sci. Inform. 18, 133. https://doi.org/10.1007/s12145-024-01574-2

Saquella, S., Scarpiniti, M., Pedone, L., Angelucci, G., Francioli, M., Matteoni, M., Pampanin, S., 2025. Building damage level classification using deep learning: a CNN-based approach for post-earthquake structural assessment. Institute of Structural Analysis and Antiseismic Research School of Civil Engineering National Technical University of Athens (NTUA).

Shafian, S.A., Hu, D., Shafian, S.A., Hu, D., 2024. Integrating Machine Learning and Remote Sensing in Disaster Management: A Decadal Review of Post-Disaster Building Damage Assessment. Buildings 14. https://doi.org/10.3390/buildings14082344

Singh, D.K., Hoskere, V., Singh, D.K., Hoskere, V., 2023. Post Disaster Damage Assessment Using Ultra-High-Resolution Aerial Imagery with Semi-Supervised Transformers. Sensors 23. https://doi.org/10.3390/s23198235

UNESCO World Heritage Centre, 2023. UNESCO World Heritage Centre - List of World Heritage in Danger [WWW Document]. UNESCO World Herit. Cent. URL https://whc.unesco.org/en/danger-list/ (accessed 1.7.26).

University, 2022. crack Dataset. Roboflow Universe.

Wang, N., Zhao, X., Zou, Z., Zhao, P., Qi, F., 2020. Autonomous damage segmentation and measurement of glazed tiles in historic buildings via deep learning. Comput.-Aided Civ. Infrastruct. Eng. 35, 277–291. https://doi.org/10.1111/mice.12488

Xue, Z., Zhang, X., Prevatt, D.O., Bridge, J., Xu, S., Zhao, X., 2024. Post-hurricane building damage assessment using street-view imagery and structured data: A multi-modal deep learning approach. https://doi.org/10.48550/arXiv.2404.07399

Zhang, K., Mea, C., Fiorillo, F., Fassi, F., 2024. Classification And Object Detection for Architectural Pathology: Practical Tests with Training Set. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLVIII-2-W4-2024, 477–484. https://doi.org/10.5194/isprs-archives-XLVIII-2-W4-2024-477-2024

Zhuang, X., Tran, T.V., Nguyen-Xuan, H., Rabczuk, T., 2025. Deep learning-based post-earthquake structural damage level recognition. Comput. Struct. 315, 107761. https://doi.org/10.1016/j.compstruc.2025.107761