AI-Powered Digital Twinning and Monitoring of Concrete Dams with Drones and Photogrammetry

VAHIDREZA GHAREHBAGHI, JIAN LI, HANG ZHAO,
CAROLINE R. BENNETT and REMY D. LEQUESNE

ABSTRACT

Concrete dams are critical infrastructure assets that deteriorate over time, making effective monitoring and inspection essential. Traditional manual inspections are often time-consuming, costly, and risky. In contrast, drones provide a faster, safer, and more reliable solution for structural health monitoring. This research develops a digital twin of the dam using aerial imaging and photogrammetry techniques. Creating a high-quality 3D model presents challenges due to variations in lighting conditions, camera motion, and shooting angles, which can result in stitching errors, distortions, and blurriness. To address these issues, a preprocessing pipeline is introduced, including image filtering, outlier removal, mesh smoothing, illumination normalization, and local feature enhancement. For damage detection, a human-AI collaboration approach is employed, using an iterative procedure that reduces annotation efforts and improves detection accuracy through human supervision of small data batches. To map detected damage onto the 3D model, a texture updating method is proposed, which avoids the need for model reconstruction by using damage-segmented images or a virtual camera to recapture imagery. Furthermore, a comprehensive dam inspection platform is developed, featuring a user-friendly application that visualizes the 3D model, analyzes and measures various types of damage, integrates GPS coordinates for spatial context, and generates detailed inspection reports. This robust system provides a practical and effective solution for automated dam condition assessment, documentation, and management.

INTRODUCTION

Dams are critical infrastructure for hydropower, flood control, and water supply. Due to exposure to harsh environmental and operational conditions, surface defects like cracks and spalling often signal underlying degradation, such as alkali-silica reaction (ASR), freeze-thaw damage, and reinforcement corrosion. Traditional inspection methods, relying on manual visual assessments, are hazardous, time-consuming, and prone to variability and missed detections [1]. Recent advances in computer vision, particularly deep learning, offer a scalable and automated alternative for structural health monitoring.

Vahidreza Gharehbaghi, PhD Student. Jian Li, Professor, jianli@ku.edu. Hang Zhao, Postdoctoral Researcher. Caroline R. Bennett, Professor. Rémy D. Lequesne, Professor. Department of Civil, Environmental & Architectural Engineering, The University of Kansas, USA

Deep learning techniques are now increasingly applied for crack detection in roads[2], ASR identification in bridges[3], and surface damage segmentation in buildings and dams[4, 5]. However, the effectiveness of these models heavily depends on the quality and diversity of the training data, especially when dealing with irregular crack patterns on textured surfaces[6].

A major challenge is that most deep learning models provide 2D results, lacking spatial context. To overcome this limitation, integrating deep learning with photogrammetry, which creates high-quality 3D models from 2D images through structure-from-motion and multi-view stereo techniques[7], enables inspectors to localize damage within a 3D spatial framework. This study proposes an AI-driven health assessment pipeline for concrete dams, combining UAV-based image acquisition, deep learning-based damage segmentation, and 3D modeling, while addressing key challenges in training data preparation, 3D model generation, and 2D-to-3D damage visualization.

RESEARCH METHODOLOGY

The overall workflow for crack detection and mapping is illustrated in Figure 1. The methodology integrates aerial imaging, damage segmentation, and 3D reconstruction into a cohesive pipeline aimed at generating damage maps on the surface of concrete dam.

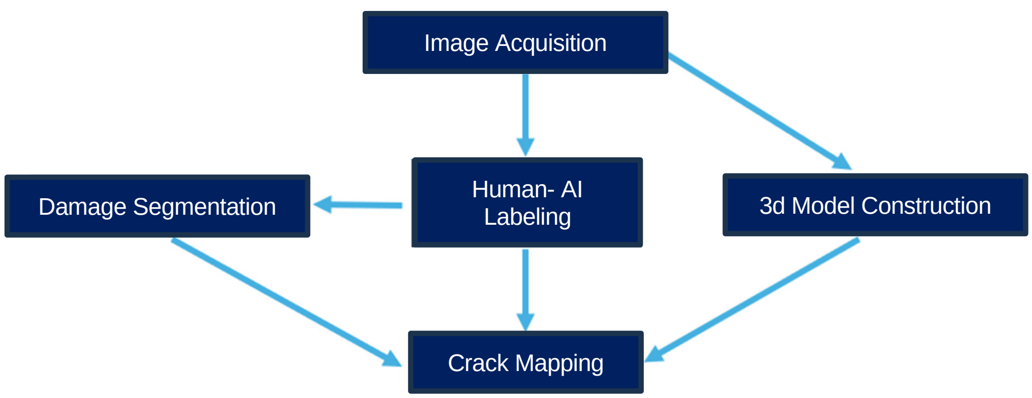

Figure 1. Overall research methodology.

The Upriver Dam, located on the Spokane River in Spokane, Washington (Figure 2), was originally constructed in 1894 and rehabilitated in 1936. This 38-foot-tall, 725-foot-long concrete gravity dam has extensive surface cracking and deterioration; hence, it was selected as a case study for crack detection and mapping.

A total of 2,300 images were collected using smartphone cameras for close-range imaging and a DJI Mavic 3 Enterprise (M3E) drone for aerial photogrammetry. The dataset includes close-up shots of deteriorated areas and aerial images captured from various angles and altitudes to ensure full surface coverage. To provide the data for deep learning models, the full-resolution images were divided into 512 × 512-pixel patches, creating approximately 84,000 patches. These patches were used in an iterative Human-AI annotation and training process to progressively improve model accuracy.

Figure 2. Upriver Dam.

For crack detection, a supervised learning method was employed to train a segmentation model. While supervised methods can achieve high accuracy, they require pixel-wise annotated masks for each damage instance, a process that is labor-intensive and prone to subjectivity, especially for fine, irregular crack patterns. To address these challenges, a collaborative Human-AI labeling framework was implemented, where initial pseudo-labels generated by a pre-trained model are refined by human supervisors through an iterative feedback loop, as illustrated in Figure 3.

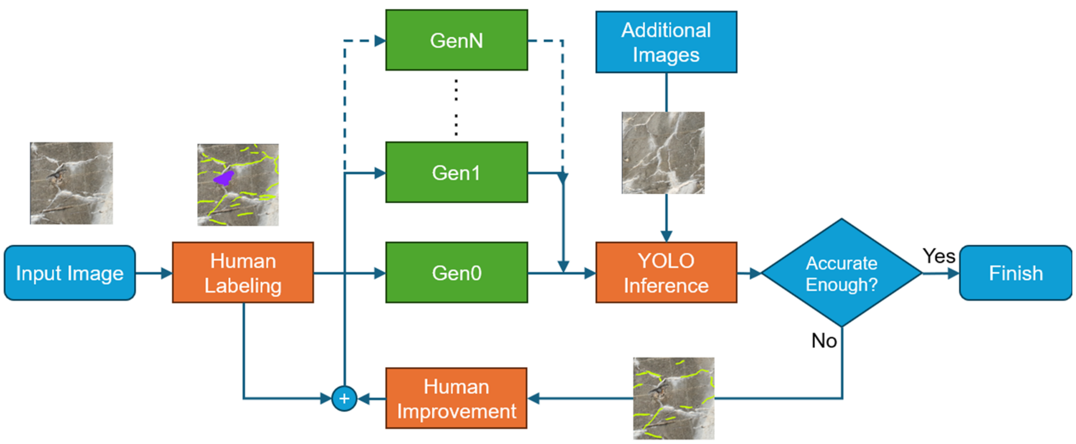

Figure 3. Human-AI collaboration for training data preparation.

To initiate the training process, 2,000 image patches were manually annotated for training the initial model generation, referred to as Gen0. Annotators, trained to identify cracks and spalling, created polygonal masks for each defect. All annotations were subsequently reviewed by a domain expert to ensure accuracy and consistency. Roboflow, a web-based platform, was used to segment regions of interest and assign class labels. Figure 4 shows an annotated image patch containing 31 crack instances (Class 0) and one spalling region (Class 1). Annotating fine-scale features, such as tiny cracks, proved to be time-consuming and labor-intensive, highlighting the need for an AI-assisted labeling approach to improve scalability and reduce manual effort.

The AI model used in this study is YOLOv9, a recent advancement in the YOLO family of real-time object detection models originally introduced in 2016[8]. YOLO architectures are built on convolutional neural networks and consist of three primary components: the backbone (for extracting hierarchical features), the neck (which aggregates multi-scale features using structures like feature pyramid networks), and the head (responsible for object classification and localization at multiple scales). YOLOv9, developed by Wang et al. and released in 2024[9], introduces innovative modules such as Programmable Gradient Information (PGI) and Generalized Efficient Layer Aggregation Network (GELAN), which enhance detection accuracy and learning efficiency by mitigating information degradation across deep network layers.

Figure 4. Roboflow user interface for damage annotation.

Within the proposed Human-AI collaborative framework, an iterative training and labeling process was used to improve both the dataset quality and model performance. Initially, the YOLO model was trained on a manually annotated batch of 2,000 images to create the initial model generation (Gen0). This model then generated segmentation masks for 500 unlabeled images, which were reviewed and refined by human annotators to correct false positives and false negatives. The corrected images were then added to the Gen0, expanding the data set to 2,500 images, which were then used for the development of the next model generation (Gen1). This process was repeated iteratively, with each round incorporating 500 new AI-labeled and human-refined images, until the third-generation model (Gen3) was trained on a total of 3,500 accurately labeled image patches.

Photogrammetry uses a series of photographs to measure distances between objects. Structure-from-Motion (SfM), a key component of photogrammetry, reconstructs the 3D model of an object from 2D images. The SfM process involves three main stages: (i) detecting and matching visual features (keypoints, lines) across images, (ii) estimating the relative motion of cameras based on these correspondences, and (iii) reconstructing the 3D scene by triangulating feature points and minimizing reprojection errors[10].

For the 3D reconstruction of the dam, SfM was implemented using Agisoft Metashape, a widely used photogrammetric software platform. The software reconstructs 3D geometry by matching features such as edges, corners, and textures across overlapping images from different viewpoints. Ensuring sufficient image overlap is essential for reliable feature matching and accurate camera pose estimation.

RESULTS AND DISCUSSION

Crack segmentation was performed using the latest-generation model (Gen3), which was trained on a curated dataset of 3,500 labeled images through the proposed Human-AI collaboration process. Leveraging AI-assisted inference during successive training iterations (from Gen1 to Gen3) substantially reduced the manual annotation time. As shown in Figure 5, the annotation time, calculated by averaging the annotation times for 20 sample images, decreased from approximately 234 seconds during the initial human-only annotation phase (Gen0) to 40 seconds in Gen3. This improvement is largely attributed to the incorporation of model-generated pseudo-labels, which were refined by human annotators, forming an efficient human-AI collaboration loop.

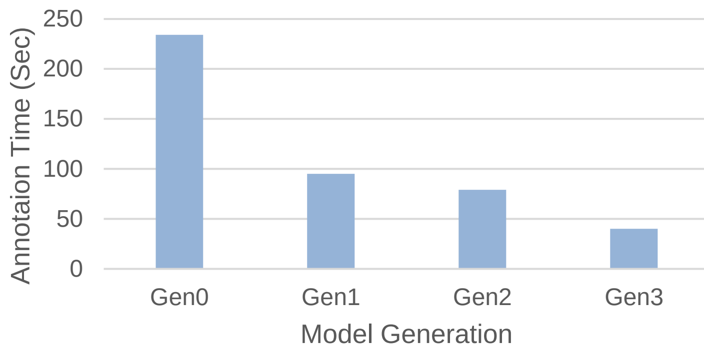

Figure 5. Annotation time comparison.

The model was evaluated using mean Average Precision at an Intersection-over-Union (IoU) threshold of 0.50 (mAP@50), which measures the accuracy of object detection or segmentation by evaluating the overlap between predicted and ground truth regions. mAP@50 reflects the percentage of predictions that achieve at least 50% IoU with the corresponding ground truth annotations.

Figure 6 shows the model's performance across different damage types. For crack segmentation, the mAP@50 improved from 48% in the baseline model to 71% in the Gen3 model, indicating a substantial accuracy boost. Similarly, for spalling detection, the mAP@50 increased from 37% to 74%, representing a 100% relative improvement.

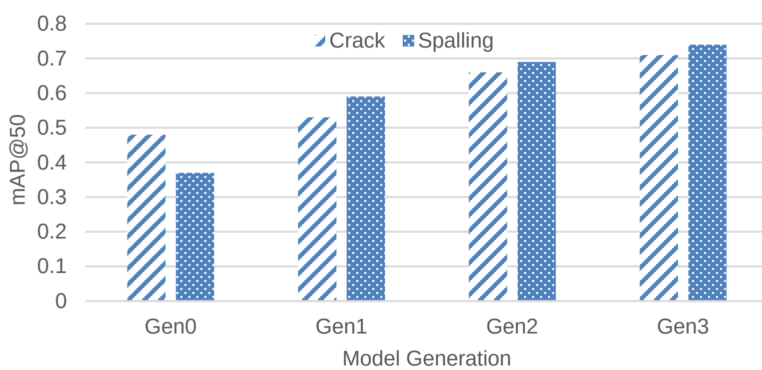

Figure 6. Model performance across generations.

To qualitatively assess the model’s performance, a representative image was segmented using four different model generations. As shown in Figure 7, the initial model exhibits substantial false negative errors, missing a huge portion of both cracks and spalling regions. With each successive generation, damage localization progressively improves, with an increasing number of correctly segmented crack patterns and spalling regions. This improvement indicates enhanced generalization and learning capacity. The Gen3 model demonstrates the highest segmentation accuracy, successfully detecting most visible damage with minimal errors.

Gen0 Gen1 Gen2 Gen3

Figure 7. Model inferences through model generations.

Figure 8a depicts a small segment of the dam. As depicted, the initial reconstruction exhibits significant surface noise and point cloud artifacts, which reduce visual quality and distort vital surface details such as cracks and material textures. These artifacts are primarily due to factors such as poor lighting conditions, low-resolution imagery, and point cloud outliers.

To address these deficiencies, a refinement pipeline was proposed, incorporating both image-level and model-level refinements. During preprocessing, low-quality images with poor lighting, motion blur, or underexposure were excluded. Additionally, point cloud denoising and mesh smoothing approaches were applied to decrease surface noise and geometric inaccuracies. Figures 8b and 8c present the results of the 3D model improvement. The refined model exhibits enhanced texture projection and clearer delineation of surface defects, such as cracks and spalling.

a. No Refinement b. Refinement Phase 1 c. Refinement Phase 2

Figure 8. Improvement of 3D reconstruction.

The projection of 2D segmented damage regions onto the 3D concrete surface was carried out using a standalone application developed in the .NET environment. This custom software supports a range of functionalities, including 3D navigation, damage visualization, filtering by damage type, and automated documentation, providing an interactive platform for dam management. Figure 9 shows the application's user interface alongside the textured 3D mesh of the dam.

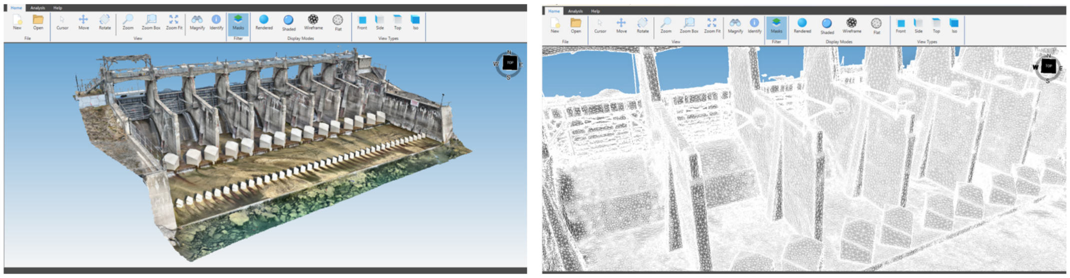

Figure 9. Software user interface.

Crack mapping was performed by integrating the high-resolution texture of the structure with segmentation inferences from the Gen3 model. Damage spots, such as cracks and spalling, were detected using Yolov9 model and projected onto the 3D concrete surface through UV unwrapping, aligning the 2D texture with the 3D mesh surface. Figure 10 illustrates the projected damage on the dam surface, with cracks and spalling highlighted in distinct colors, enabling clear localization and supporting more informed decision-making for monitoring and maintenance planning.

Figure 10. Crack mapping on the dam surface.

CONCLUSIONS

This study presents an AI-powered methodology for monitoring concrete dams, integrating UAV-based imaging, deep learning models, and photogrammetric 3D reconstruction. As a case study, we tested the proposed method on a real concrete dam and demonstrated how automated image acquisition and SfM techniques can reconstruct high-fidelity 3D models for digital twinning. The iterative human-AI collaborative labeling strategy demonstrated a reduction in annotation time and improved labeling quality for complex damage patterns like concrete cracks. Successive model generations (Gen0–Gen3) exhibited significant performance enhancements, validating the effectiveness of iterative refinement for large-scale damage detection. In addition, a custom software tool was developed to project 2D segmentation inferences onto the 3D dam surface, providing inspectors with an interactive, spatially aware environment for visualizing and assessing concrete dam surface damage.

ACKNOWLEDGEMENT

We sincerely acknowledge the funding support and contributions from the U.S. Army Engineer Research and Development Center (ERDC), U.S. Army Corps of Engineers, and the Department of Homeland Security.

REFERENCES

[1] V. R. Gharehbaghi *et al.*, "A critical review on structural health monitoring: Definitions, methods, and perspectives," *Archives of computational methods in engineering,* vol. 29, no. 4, pp. 2209-2235, 2022.

[2] T. Ahmad, V. Gharehbaghi, J. Li, C. Bennett, and R. Lequesne, "Crack segmentation in the wild using convolutional neural networks and bootstrapping," *Earthquake Engineering and Resilience,* vol. 2, no. 3, pp. 348-363, 2023.

[3] A. Nguyen, V. Gharehbaghi, N. T. Le, L. Sterling, U. I. Chaudhry, and S. Crawford, "ASR crack identification in bridges using deep learning and texture analysis," in *Structures*, 2023, vol. 50: Elsevier, pp. 494-507.

[4] N. Wang, X. Zhao, P. Zhao, Y. Zhang, Z. Zou, and J. Ou, "Automatic damage detection of historic masonry buildings based on mobile deep learning," *Automation in Construction,* vol. 103, pp. 53-66, 2019.

[5] B. Xu, Z. Chen, H. Su, and H. Zhang, "A deep learning method for predicting the displacement of concrete arch dams considering the effect of cracks," *Advanced Engineering Informatics,* vol. 62, p. 102574, 2024.

[6] P. Skondras, N. Zotos, D. Lagios, P. Zervas, K. C. Giotopoulos, and G. Tzimas, "Deep Learning Approaches for Big Data-Driven Metadata Extraction in Online Job Postings," *Information,* vol. 14, no. 11, p. 585, 2023.

[7] E. Berra and M. Peppa, "Advances and challenges of UAV SFM MVS photogrammetry and remote sensing: Short review," in *2020 ieee latin american grss & isprs remote sensing conference (lagirs)*, 2020: IEEE, pp. 533-538.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779-788.

[9] C.-Y. Wang, I.-H. Yeh, and H.-Y. Mark Liao, "Yolov9: Learning what you want to learn using programmable gradient information," in *European conference on computer vision*, 2024: Springer, pp. 1-21.

[10] O. Özyeşil, V. Voroninski, R. Basri, and A. Singer, "A survey of structure from motion*," *Acta Numerica,* vol. 26, pp. 305-364, 2017.