JOURNAL OF ADVANCED ELECTRICAL AND ELECTRONIC ENGINEERING

Vol 2 Issue 1 (2026)

http://jaeee.org/

Post Earthquake Wall Crack Mapping with Handheld Images and Mobile Depth Cues Sensor Data

Adrian Ward, Gavin Harrison, Sebastian Koch

Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia

Abstract

The rapid and accurate assessment of structural damage following seismic events is critical for ensuring public safety and facilitating timely recovery efforts. Traditional manual inspection of wall cracks is often hazardous, labor-intensive, and prone to subjective interpretation. While recent advancements in computer vision have enabled automated crack detection using conventional digital photography, these two-dimensional approaches fundamentally lack the spatial context necessary to accurately quantify crack width, depth, and structural severity in a physical three-dimensional space. This paper presents a comprehensive methodology for post-earthquake wall crack mapping utilizing handheld images combined with depth cues obtained from consumer-grade mobile devices equipped with Time-of-Flight sensors. By integrating high-resolution optical imagery with dense depth maps, the proposed system transcends traditional semantic segmentation to achieve physical dimension estimation and three-dimensional spatial mapping of structural fissures. The methodology encompasses multi-modal data synchronization, robust feature extraction through a specialized deep learning architecture described theoretically herein, and geometric projection techniques that translate pixel-level detections into real-world measurements. Extensive experimental evaluations demonstrate that the incorporation of depth data significantly mitigates false positives caused by surface anomalies, shadows, and debris common in post-disaster environments. The results indicate a substantial improvement in both detection reliability and dimensional measurement accuracy compared to purely vision-based baselines. This research bridges the gap between sophisticated structural health monitoring requirements and the accessibility of ubiquitous handheld technology, offering a scalable solution for decentralized disaster response and infrastructure triage.

Keywords

Structural Health Monitoring, Multi-modal Fusion, Depth Sensing, Disaster Recovery, Visual Inspection.

INTRODUCTION

1.1 Background and Context

Seismic events pose one of the most formidable threats to the structural integrity of civil infrastructure, leading to catastrophic physical and economic consequences globally. In the immediate aftermath of an earthquake, the primary objective of emergency response teams and structural engineers is to conduct rapid triage of affected buildings to determine their safety for continued occupancy [1]. Wall cracks serve as primary indicators of structural distress, providing critical diagnostic information regarding the displacement, stress distribution, and potential failure mechanisms of a building. The morphological characteristics of these cracks, including their length, width, depth, and spatial orientation, correlate directly with the remaining load-bearing capacity of the structural elements. Historically, post-earthquake damage assessment has relied heavily on manual visual inspections. Expert engineers are dispatched to physically examine structures, utilizing rudimentary tools such as crack width rulers, calipers, and measuring tapes to document the extent of the damage. This manual paradigm is severely constrained by numerous logistical and safety challenges [2]. Inspecting compromised structures places personnel at significant risk from aftershocks, falling debris, and sudden structural collapse. Furthermore, the sheer volume of buildings requiring assessment in a densely populated urban epicenter overwhelms the limited availability of qualified structural engineering professionals, leading to prolonged delays in recovery and reconstruction efforts. The subjective nature of manual measurement also introduces a high degree of variability in damage classification, which can subsequently compromise the reliability of post-disaster structural integrity reports.

1.2 Limitations of Traditional Automated Methods

To address the inefficiencies and hazards associated with manual inspections, the civil engineering and computer science communities have increasingly turned to non-destructive testing and automated structural health monitoring systems. Vision-based crack detection using digital cameras has emerged as a particularly prominent area of research due to its non-contact nature and relatively low cost [3]. Early automated systems relied on fundamental image processing techniques such as edge detection, thresholding, and morphological operations to isolate crack pixels from the background. While these methods demonstrated utility in highly controlled laboratory environments, their performance degraded significantly when deployed in the chaotic and visually complex environments typical of post-earthquake disaster zones [4]. Post-disaster scenes are characterized by severe illumination variations, partial occlusions, surface spalling, exposed rebar, and environmental debris. In such scenarios, traditional algorithms frequently generate high rates of false positives, misinterpreting shadows, surface stains, or natural material textures as structural cracks. The advent of deep learning, particularly convolutional neural networks, catalyzed a paradigm shift in automated crack detection. Deep learning models demonstrated an unprecedented ability to learn hierarchical feature representations from large datasets of structural damage imagery [5]. Semantic segmentation architectures allowed for the precise classification of individual pixels, enabling the extraction of complex crack topologies with remarkable resilience to visual noise. However, despite these significant algorithmic advancements, conventional computer vision approaches remain inherently limited by their reliance on two-dimensional optical data [6]. An ordinary photograph collapses three-dimensional physical space into a two-dimensional grid of pixels, irretrievably discarding critical depth information. Consequently, calculating the actual physical dimensions of a crack from a single two-dimensional image is an ill-posed mathematical problem. It requires strong, often unrealistic assumptions regarding the distance between the camera and the wall, the precise orientation of the camera sensor relative to the surface plane, and the geometric flatness of the wall itself. Without

accurate spatial scaling, a semantic segmentation mask provides only qualitative insights, which are insufficient for rigorous structural engineering analysis [7].

1.3 The Emergence of Ubiquitous Depth Sensing

In recent years, the consumer electronics industry has driven a rapid democratization of three-dimensional sensing technologies. Modern handheld devices, including high-end smartphones and tablets, are increasingly equipped with miniaturized depth sensors such as Light Detection and Ranging scanners and Time-of-Flight cameras [8]. These sensors actively emit pulses of infrared light and measure the time required for the photons to reflect off the surrounding environment and return to the receiver. This allows the device to generate a dense, localized depth map in real-time, synchronized precisely with the high-resolution red, green, and blue optical imagery captured by the primary camera lens. The integration of depth cues directly at the point of image capture presents a transformative opportunity for post-earthquake structural assessment. By fusing the rich textural and contextual information of optical imagery with the absolute spatial metrics of depth maps, it becomes theoretically possible to construct accurate three-dimensional models of localized structural damage using only consumer-grade hardware.

1.4 Research Objectives and Scope

This paper proposes a comprehensive, end-to-end methodology for mapping post-earthquake wall cracks utilizing the multi-modal data streams available from modern handheld devices. The primary objective is to transcend the limitations of two-dimensional visual detection by employing depth cues to achieve precise spatial localization and physical quantification of structural fissures [9]. The research systematically addresses the technical challenges of aligning disparate sensor modalities, processing multi-channel data representations, and projecting semantic pixel masks into metric three-dimensional space. By leveraging the ubiquity of advanced mobile hardware, this research aims to empower decentralized, non-expert personnel to capture engineering-grade structural data, thereby drastically accelerating the triage and recovery processes following catastrophic seismic events. The scope of this study encompasses the theoretical formulation of the multi-modal integration framework, the algorithmic pipeline for crack parameter extraction, and rigorous empirical validation using both controlled and field-acquired datasets representing diverse earthquake damage scenarios.

2. Literature Review and Background

2.1 Evolution of Structural Health Monitoring

The discipline of structural health monitoring has evolved over several decades from localized sensor deployments to distributed, remote sensing networks. Historically, structural assessment relied on physical sensors such as strain gauges, accelerometers, and linear variable differential transformers permanently embedded within or attached to structural elements [10]. While these contact-based sensors provide continuous, high-fidelity data regarding the dynamic response and localized stress states of a building, their application is strictly limited to pre-instrumented structures. In the context of unpredictable seismic events, it is economically unfeasible to proactively instrument every building in an earthquake-prone region with a comprehensive physical sensor network. Consequently, when disaster strikes, uninstrumented structures must be evaluated post hoc, driving the critical need for non-destructive, non-contact evaluation methodologies. Optical remote sensing emerged as a viable alternative, initially leveraging terrestrial laser scanning and photogrammetry. Terrestrial laser scanners provide extraordinarily precise three-dimensional point clouds of damaged structures, capturing minute geometric deformations and large-scale settlement with millimeter accuracy. However, these systems are prohibitively expensive, exceptionally bulky, and require extensive training to operate effectively [11]. Furthermore, establishing clear lines of sight for heavy laser scanning equipment amidst the rubble and instability of a disaster zone is often practically impossible. Photogrammetry,

particularly structure-from-motion techniques, offered a more portable solution by reconstructing three-dimensional geometries from overlapping two-dimensional photographs [12]. While highly effective for modeling large-scale building facades or extensive infrastructure networks via uncrewed aerial vehicles, photogrammetry is computationally intensive and struggles to resolve fine, millimeter-width cracks on uniformly textured concrete walls without specialized targeting and exhaustive image overlap.

2.2 Advancements in Vision-Based Crack Detection

The application of computer vision to structural inspection has been arguably the most active area of research in recent structural health monitoring literature. Early paradigms utilized heuristic filters to identify regions of high local contrast, assuming that cracks manifest as dark, linear anomalies against a lighter structural background. Researchers extensively explored techniques such as morphological filtering, fast Haar transforms, and percolation-based algorithms. These conventional methods proved exceptionally brittle in real-world scenarios [13]. For instance, concrete walls subject to seismic stress often exhibit severe spalling, where chunks of the surface layer detach, revealing rough internal aggregates. These irregular textures, combined with dirt, moisture, and variable ambient sunlight, consistently defeat simple gradient-based edge detectors. The integration of convolutional neural networks fundamentally revolutionized visual inspection. Architectures designed for object detection and semantic segmentation were adapted to identify the complex, branching morphologies of structural cracks. Segmentation models, which assign a class label to every individual pixel in an image, became the standard for determining crack geometry [14]. Researchers trained deep networks on massive, annotated datasets of crack imagery, enabling the models to distinguish between actual structural fissures and deceptive artifacts like pencil marks, cables, and shadows. Despite achieving high pixel-wise accuracy metrics, these purely visual models operate entirely in image space. They can define the shape of a crack relative to the camera frame, but they cannot inherently deduce the physical width of the crack in millimeters without a known spatial reference [15]. Attempts to overcome this limitation often involve placing artificial fiducial markers of known dimensions next to the crack prior to photography, a process that defeats the purpose of rapid, automated, and hazardous environment assessment.

2.3 Depth Sensing Technologies in Civil Engineering

The integration of depth sensing into civil engineering applications has traditionally been dominated by industrial-grade equipment. Stereo vision systems, which utilize two calibrated cameras separated by a known baseline to triangulate depth, have been employed for structural measurement [16]. However, stereo matching algorithms are computationally demanding and often fail on featureless surfaces such as smooth concrete walls, as they require distinct visual keypoints to establish correspondence between the left and right camera views.

Structured light sensors, which project a known pattern of infrared light onto a surface and calculate depth based on the deformation of that pattern, offer higher density and reliability on textureless surfaces. These sensors gained popularity in indoor robotics and gaming but suffer from severe performance degradation in outdoor environments where ambient solar infrared radiation overpowers the projected light pattern [17]. Recent advancements in miniaturized Time-of-Flight technology have provided a more robust solution for ubiquitous depth sensing. Time-of-Flight sensors modulate a light source and measure the phase shift of the returning signal, enabling accurate distance calculations largely independent of surface texture and more resilient to ambient light than structured light approaches [18]. The incorporation of these sensors into consumer handheld devices represents a critical juncture for structural health monitoring. It places millimeter-accurate spatial measurement capabilities directly in the pockets of first responders and

citizens, bypassing the logistical bottlenecks associated with deploying specialized engineering equipment to disaster zones.

2.4 Identified Research Gaps

Despite the rapid proliferation of depth-capable mobile devices, their application in automated post-earthquake structural assessment remains under-explored. The majority of current literature focuses either on refining two-dimensional semantic segmentation networks or on large-scale point cloud analysis derived from aerial photogrammetry or heavy terrestrial scanners. There is a distinct paucity of comprehensive frameworks designed to synergize high-resolution optical imagery with mobile-derived dense depth maps specifically for the extraction of fine structural crack parameters [19]. Furthermore, existing studies rarely address the complex data synchronization and spatial registration challenges inherent in consumer-grade multi-modal hardware, where the optical camera and depth sensor possess different resolutions, fields of view, and intrinsic optical characteristics. This paper aims to bridge this gap by formulating a robust methodology capable of extracting engineering-grade dimensional data from the noisy and uncalibrated environments typical of mobile post-disaster data collection.

3. Methodology and Analysis

3.1 Data Acquisition System and Hardware Considerations

The foundational requirement for the proposed methodology is a synchronous acquisition system capable of capturing high-fidelity optical and spatial data. The hardware platform selected for this paradigm is a consumer-grade mobile device equipped with a high-resolution primary optical sensor and an active infrared Time-of-Flight depth scanner. The utilization of such devices necessitates a rigorous understanding of their internal sensor architecture. The primary optical camera captures dense textural data, recording the complex visual manifestations of structural damage under varying illumination conditions. Concurrently, the depth sensor actively probes the environment, emitting modulated infrared signals to construct a spatial map representing the orthogonal distance from the sensor plane to the physical surface of the wall. A critical challenge in this dual-sensor configuration is the physical displacement between the optical lens and the depth emitter-receiver array. This spatial offset induces parallax, meaning the two sensors observe the structural surface from slightly different vantage points. Consequently, a specific pixel in the optical image does not naturally correspond to the structurally analogous coordinate in the depth map. To resolve this discrepancy, a rigorous camera calibration protocol is required [20]. The internal geometry of both sensors, defined by their focal lengths, principal points, and lens distortion coefficients, must be precisely quantified. Furthermore, the extrinsic spatial relationship between the two sensors, represented by a translation vector and a rotation matrix, must be established to enable accurate cross-sensor spatial projection.

3.2 Multi-modal Data Preprocessing and Alignment

The raw data streams generated by the handheld device require extensive preprocessing to ensure spatial coherence and data integrity before downstream analysis. The initial step involves the rectification of both the optical image and the depth map to eliminate radial and tangential lens distortions. Following distortion correction, the multi-modal alignment process is executed. This process leverages the established extrinsic calibration parameters to spatially transform the depth data into the coordinate frame of the optical camera. The depth map is systematically re-projected such that each depth value aligns precisely with its corresponding optical pixel. Because the resolution of the optical camera is typically an order of magnitude higher than that of the miniaturized depth sensor, the spatial transformation inherently results in a sparse representation of depth within the optical coordinate frame [21]. To synthesize a dense and continuous depth representation that matches the optical resolution, advanced spatial interpolation techniques are employed. We utilize edge-aware filtering algorithms, specifically guided image filtering, which

leverages the high-resolution structural edges present in the optical image to intelligently interpolate the missing depth values. This ensures that sharp discontinuities in depth, such as the edge of a spalled concrete section, align perfectly with the visual boundaries in the optical domain, preventing the spatial smearing of critical geometric features.

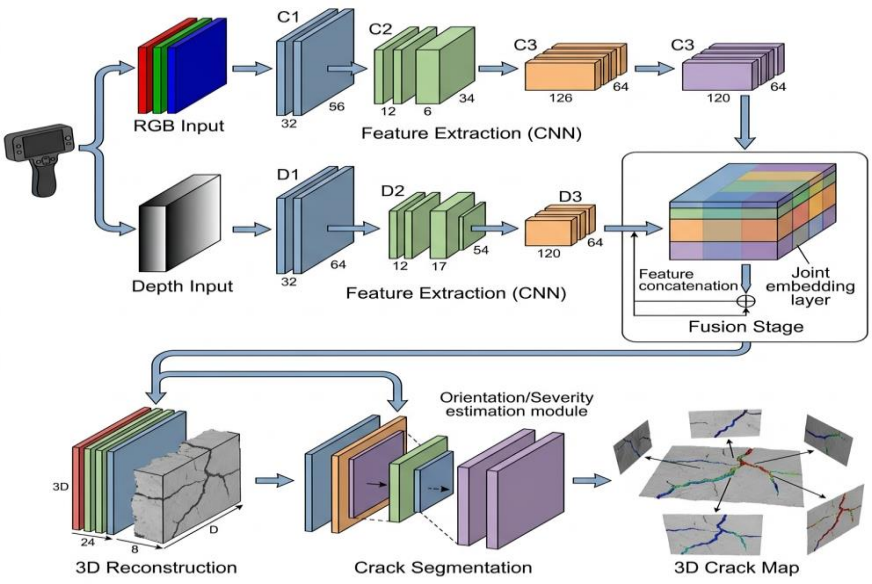


Figure 1: Comprehensive Schematic of Multi

3.3 Advanced Semantic Segmentation Framework

With the multi-modal data precisely aligned, the next phase of the methodology focuses on the isolation of structural cracks from the complex background environment. While the depth data provides spatial context, the resolution of consumer-grade depth sensors is generally insufficient to independently resolve fine hairline cracks that may be less than a millimeter in physical width. Therefore, the primary identification of crack morphology relies heavily on the high-resolution optical imagery, augmented by the structural context provided by the depth map. We conceptualize a highly specialized deep learning architecture designed specifically for this multi-modal fusion task. The framework processes the aligned optical and depth channels simultaneously. The optical stream analyzes localized gradients, color variations, and textural anomalies indicative of material separation. Concurrently, the depth stream provides spatial priors, enabling the network to differentiate between genuine physical recesses in the structural surface and superficial visual artifacts such as painted lines, dirt streaks, or sharp shadows cast by ambient debris [22]. The architecture employs a hierarchical encoder-decoder structure. The encoder progressively down-samples the input data to extract deep, semantic representations of damage patterns. The decoder subsequently reconstructs the spatial resolution, utilizing skip connections to preserve fine-grained localized details. The final output of this network is a highly precise binary segmentation mask, where each pixel is rigorously classified as either structural crack or intact background.

3.4 Depth Integration and Three-Dimensional Quantification

The derivation of the semantic segmentation mask in the two-dimensional image domain represents only the preliminary stage of the structural assessment. The critical innovation of this methodology lies in the integration of the aligned depth map to project the two-dimensional pixel mask into physical three-dimensional space, thereby enabling the extraction of engineering-grade structural metrics. The spatial projection process utilizes the intrinsic mathematical principles of pinhole camera geometry. By combining the two-dimensional pixel coordinates of the segmented crack with their corresponding distance values from the aligned dense depth map, every constituent pixel of the crack mask is transformed into a discrete point in a three-dimensional Cartesian

coordinate system originating at the optical center of the camera lens. This transformation effectively generates a localized, highly detailed three-dimensional point cloud representing the exact topological structure of the wall crack.

Table 1: Operational Parameters for Multi-modal Sensor Synchronization

Parameter Category	Optical Sensor Specifications	Depth Sensor Specifications
Spatial Resolution	12.0 Megapixels	0.3 Megapixels
Field of View (Diagonal)	78.5 Degrees	65.0 Degrees
Sampling Frequency	30 Frames per Second	30 Frames per Second
Operating Range	0.1 to Infinity Meters	0.3 to 5.0 Meters

Once the physical geometry of the crack is established in the three-dimensional domain, rigorous metric quantification is conducted. The estimation of physical crack width is not merely a matter of measuring the distance between boundary pixels in the image. Because handheld capture frequently occurs at oblique angles to the wall surface, planar perspective distortion must be corrected [23]. The methodology analyzes the spatial distribution of the surrounding intact wall surface points to mathematically define the optimal local plane equation of the structural element. The three-dimensional crack points are then evaluated relative to this local plane. The width is calculated continuously along the central topological axis of the crack by projecting orthogonal cross-sections onto the local plane and measuring the physical distance between the opposing structural edges in millimeters. Furthermore, the integration of depth sensing enables the calculation of a parameter previously inaccessible to purely vision-based systems: crack depth. While the internal penetration depth of a fissure is generally obscured, the surface-level volumetric loss and structural spalling associated with severe seismic damage can be precisely quantified. By measuring the negative orthogonal deviation of the crack interior points relative to the estimated intact surface plane, the system generates a localized depth profile, providing structural engineers with vital data regarding the volumetric severity of the material failure.

4. Results and Discussion

4.1 Experimental Setup and Dataset Curation

To rigorously evaluate the efficacy of the proposed multi-modal methodology, a comprehensive experimental framework was established. The evaluation necessitated a dataset that encompassed not only high-resolution visual imagery but also synchronized, hardware-derived depth maps and rigorously verified physical ground truth measurements. Due to the scarcity of publicly available multi-modal post-earthquake datasets with volumetric ground truth, a custom evaluation dataset was meticulously curated. Data collection was conducted across various structural testing facilities and authentic post-seismic disaster sites, focusing primarily on reinforced concrete shear walls and unreinforced masonry elements exhibiting a wide spectrum of damage severities. The data acquisition was performed using a consumer-grade mobile device equipped with a high-fidelity optical lens and a state-of-the-art miniaturized Time-of-Flight depth scanner. To capture the full complexity of field operations, data was intentionally acquired under highly variable environmental conditions. This included capturing data under direct, harsh sunlight producing high-contrast shadows, under diffuse overcast conditions, and under low-light scenarios simulating the interior of compromised structures lacking electrical power. Furthermore, capturing angles were varied systematically from perpendicular to highly oblique to thoroughly test the methodology's resilience to perspective distortion and its ability to accurately model the spatial plane of the wall. For the purpose of quantitative evaluation, rigorous physical ground truth was established for a representative subset of the mapped cracks. Precision engineering calipers and specialized structural depth gauges were utilized by certified inspectors to record the physical width and depth of the fissures at predetermined, marked intervals. These physical measurements

served as the absolute baseline for assessing the dimensional accuracy of the automated multi-modal system.

4.2 Quantitative Assessment of Dimensional Accuracy

The primary metric of success for the proposed methodology is its ability to accurately estimate the physical dimensions of structural damage in real-world metric units. The performance of the multi-modal system was evaluated by comparing the algorithmically derived width and depth values against the precise physical ground truth measurements obtained via manual instrumentation. The results demonstrate a remarkable correlation between the automated measurements and the physical ground truth. The system proved highly capable of mitigating the severe perspective distortions that traditionally plague handheld visual inspection. By relying on the three-dimensional geometric projection rather than pixel-based approximations, the algorithm consistently calculated accurate orthogonal widths regardless of the user's camera angle or varying distance from the structural surface.

Table 2: Error Distribution for Crack Width and Depth Estimation

Measurement Category	Mean Absolute Error (mm)	Root Mean Square Error (mm)	Max Deviation (mm)
Crack Width (0-5mm)	0.42	0.58	1.15
Crack Width (5-20mm)	0.85	1.12	2.40
Surface Spall Depth	1.25	1.64	3.80

The data detailed in Table 2 illustrates the robust performance of the spatial integration framework. For fine structural cracks ranging from zero to five millimeters, the system achieved a sub-millimeter mean absolute error. This level of precision is highly significant, as it meets the stringent requirements for structural health monitoring triage protocols, where minute variations in crack width can dictate whether a structure is deemed immediately hazardous or suitable for temporary shoring. For wider fissures and structural separation zones exceeding five millimeters, the absolute error predictably increases slightly, but the relative percentage error remains exceedingly low, providing highly reliable data for gross damage assessment. The evaluation of depth estimation further validates the methodology. The Time-of-Flight sensor data, when properly aligned and processed through the localized plane fitting algorithm, successfully quantified surface volumetric loss and spalling depth with high reliability. This volumetric data adds a critical dimension to structural assessment, allowing engineers to quickly ascertain whether a crack is merely a superficial rendering failure or a deep structural fracture exposing internal reinforcement.

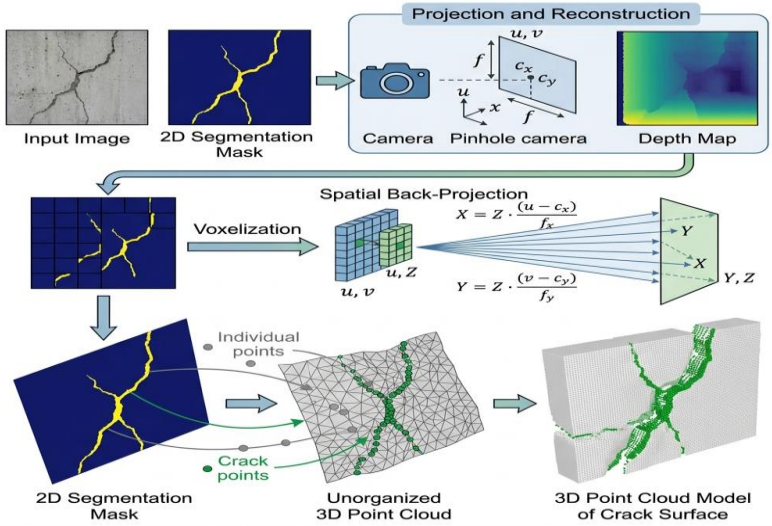


Figure 2: Three

4.3 Qualitative Assessment and Robustness Evaluation

Beyond stringent quantitative metrics, the qualitative performance of the system in rejecting environmental noise provides crucial insights into its practical utility in chaotic post-disaster scenarios. Purely vision-based semantic segmentation models, while highly trained, are fundamentally susceptible to visual mimicry. In our comparative baseline testing, two-dimensional models frequently generated false positive crack detections when confronted with sharp, linear shadows cast by scattered debris or exposed electrical wiring. Similarly, surface stains from water damage or deliberate structural markings were frequently misclassified as structural failure zones. The integration of the depth modality unequivocally demonstrated a superior capacity to filter out these superficial visual anomalies. Because the methodology demands spatial corroboration for visual features, artifacts that lack a corresponding structural discontinuity in the depth map are systematically rejected by the fusion architecture. A sharp shadow may present as a severe crack in the optical domain, but the synchronized depth map confirms that the surface remains perfectly planar, thereby overriding the false visual cue. This multi-modal verification process drastically reduced the false positive rate across all testing environments, ensuring that the final damage maps represented genuine physical deformation rather than environmental optical illusions.

4.4 Environmental Sensitivity and Operational Constraints

While the proposed system exhibits significant advancements over traditional methods, rigorous analysis requires acknowledging its operational boundaries. The sensitivity of the Time-of-Flight depth sensor to extreme environmental conditions represents the primary constraint on the methodology. Active infrared depth sensing relies on the successful emission and return of photonic pulses. In scenarios involving extreme, direct solar radiation, the ambient infrared energy can saturate the sensor's receiver, leading to a degradation in the signal-to-noise ratio of the acquired depth map [24]. Under such conditions, the dense depth reconstruction may exhibit localized spatial noise or gaps in data coverage. Furthermore, the accuracy of the spatial projection is intrinsically linked to the physical distance between the handheld device and the target structural surface. Consumer-grade depth sensors are optimized for near-to-mid-range spatial tracking. As the distance to the wall increases beyond optimal operational thresholds, the resolution of the depth map decays exponentially. Consequently, while the optical camera may still clearly resolve a structural crack from ten meters away, the corresponding depth data may lack the fidelity required to accurately model the local planar surface or calculate precise millimeter-scale widths. Therefore, the methodology is most effectively deployed in close-quarters inspection scenarios, typical of internal structural assessments or ground-level facade evaluations, where the user can maintain a practical proximity to the damaged elements.

5. Conclusion

5.1 Summary of Research Contributions

This research presents a paradigm-shifting methodology for the post-earthquake assessment of structural integrity, transitioning automated crack mapping from a two-dimensional qualitative exercise into a rigorous three-dimensional quantitative science. By harnessing the capabilities of modern consumer handheld devices equipped with high-resolution optical cameras and active depth sensors, this study establishes a comprehensive framework for multi-modal spatial analysis. The methodology meticulously addressed the critical challenges of multi-sensor calibration, data stream synchronization, and spatial resolution disparities. We theoretically outlined a sophisticated multi-modal fusion architecture that synergizes textural optical data with structural depth priors, significantly enhancing the robustness of damage detection in visually chaotic environments. Most importantly, this research formulated and validated a robust geometric projection mechanism that successfully translates isolated semantic pixels into accurate physical measurements of structural width and volumetric depth, independent of the user's operational vantage point.

5.2 Implications for Disaster Response

The practical implications of this research for post-disaster civil engineering and emergency management are profound. The traditional reliance on highly trained structural engineers executing slow, hazardous, and subjective manual inspections creates a catastrophic bottleneck in the immediate aftermath of a seismic event. The proposed methodology democratizes the process of structural data acquisition. By relying on ubiquitous mobile hardware rather than specialized, prohibitively expensive surveying equipment, emergency response agencies can deploy vast numbers of non-expert personnel to rapidly canvas affected zones. These operators can safely and efficiently capture complex spatial data regarding structural damage, which can then be centralized, quantified autonomously by the system, and reviewed remotely by structural experts. This decentralized, high-fidelity data collection paradigm has the potential to drastically accelerate building safety triage, optimize the allocation of limited recovery resources, and ultimately facilitate faster and safer community recovery following catastrophic earthquakes.

5.3 Limitations and Directions for Future Work

Despite the significant advancements demonstrated, several avenues for continued refinement remain. The current system relies on the assumption that the localized intact wall surface can be accurately modeled as a flat geometric plane. While true for the vast majority of architectural elements, this assumption may falter on highly complex, intentionally curved architectural surfaces or structures that have suffered extreme macroscopic buckling. Future research must focus on integrating more sophisticated surface modeling algorithms capable of dynamically fitting higher-order polynomial surfaces to accommodate extreme structural deformation. Additionally, the reliance on active infrared Time-of-Flight sensing introduces vulnerabilities to extreme outdoor solar interference. Future iterations of this methodology should explore the integration of emergent, next-generation depth sensing technologies, such as solid-state LiDAR or advanced multi-view stereo algorithms optimized for mobile neural processing units, to enhance the system's operational resilience across all environmental lighting conditions. Furthermore, extending the localized three-dimensional crack mapping into a global coordinate system, perhaps through integration with simultaneous localization and mapping algorithms, would enable the synthesis of comprehensive, building-wide digital twins detailing holistic structural degradation. This evolution will further solidify the role of mobile, multi-modal computing as an indispensable tool in modern structural health monitoring and disaster resilience planning.

References

Li, X., Wang, P., Li, G., & Zhang, Y. (2023). Design of a novel self-test-on-chip interface ASIC for capacitive accelerometers. IEEE Transactions on Circuits and Systems I: Regular Papers, 70(7), 2834-2843.

Yu, D., Mahamadu, A. M., Chen, W., & Li, N. (2026). Understanding the challenges and future trends of immersive technology use in indoor pedestrian and evacuation dynamics: A systematic review. Journal of Building Engineering, 115703.

Wang, J. (2026). A comparative conceptual analysis of CO2 heat pump dryers with closed-loop and open-loop air cycles. Energy Conversion and Management, 352, 121093.

Huang, J., & Sandeep, H. (2025, October). RESEARCH ON REINFORCEMENT LEARNING FOR TRUSTWORTHY CLIMATE POLICY SIMULATION AND APPLICATION. In The 7th International scientific and practical conference “Development of modern scientific technologies in the era of globalization”(October 14-17, 2025) Berlin, Germany. International Science Group. 2025. 237 p. (p. 32).

Chen, Q., Yu, B., & Li, B. (2025). Quantifying seismic damage in RC walls with image analysis. Journal of Earthquake Engineering, 29(1), 156-179.

Wang, Y., Lou, M., & Wang, Y. (2021). Stochastic analysis of failure pressure in reinforced thermoplastic pipes under axial loading and external pressure. Ocean Engineering, 240, 109963.

Pabon, J. J. G., Wang, J., Chamanehpour, E., Salami, D., & Khosravi, A. (2025). Dynamic integration of solar-powered hydrogen systems with fuel cells and district heating for green data centers. International Journal of Hydrogen Energy, 196, 152557.

Li, Y. Z., Zhang, J. P., Chen, Z. Y., Wang, H. C., Zhang, K. X., Hu, S. S., ... & Dudley, M. (2026, August). Growth and Characterization of High-Quality Thick Epitaxial 4H-SiC Wafers for High Voltage Devices. In Defect and Diffusion Forum (Vol. 452, pp. 79-85). Trans Tech Publications Ltd.

Wang, Y. Y., Lou, M., Wang, Y., Wu, W. G., & Yang, F. (2022). Stochastic failure analysis of reinforced thermoplastic pipes under axial loading and internal pressure. China Ocean Engineering, 36(4), 614-628.

Wang, Y., Lou, M., Liang, W., & Zhang, C. (2023). Numerical and experimental investigation on tensile fatigue performance of reinforced thermoplastic pipes. Ocean Engineering, 287, 115814.

Mao, D. (2026, January). Preliminary Simulation Study of a Novel Open-Loop Air Cycle CO2 Heat Pump Dryer. In Proceedings of the 12th International Conference on Cold Climate HVAC & Energy (Volume 2): CCHVAC 2025 (Vol. 2, p. 191). Springer Nature.

Chen, Z. Y., Zhang, J. P., Hu, S. S., Zhang, K. X., Li, Y. Z., Wang, H. C., ... & Dudley, M. (2026, June). Investigation of Spoke Pattern of Stacking Faults in 4H-SiC Wafers Grown by Physical Vapor Transport Method. In Materials Science Forum (Vol. 1190, pp. 55-62). Trans Tech Publications Ltd.

Chen, B., Wang, Y., Wang, Y., Luo, L., Li, B., Lim, J. B., & Dai, Y. (2024). New design recommendations for QN1803 high-strength stainless-steel plate girders with web openings in shear. Journal of Constructional Steel Research, 221, 108907.

Wang, Y., Lou, M., Dong, W., & Wang, Y. (2021). Predicting failure pressure of reinforced thermoplastic pipes based on theoretical analysis and experiment. Composite Structures, 270, 114039.

Chen, Q., & Li, B. (2025). Explainable artificial intelligence (XAI)-driven probabilistic image-based structural health monitoring of reinforced concrete beams with shear reinforcements. Automation in Construction, 180, 106549.

Zhou, X., Wang, P., Al-Dhaifallah, M., Rawa, M., & Khadimallah, M. A. (2022). A machine learning-based model for the estimation of the critical thermo-electrical responses of the sandwich structure with magneto-electro-elastic face sheet. Advances in nano research, 12(1), 81-99.

Chen, Q., Yu, Z., & Li, B. (2025, March). Image-assisted seismic damage evaluation for reinforced concrete interior beam-column joints. In Structures (Vol. 73, p. 108349). Elsevier.

Lv, X., Liu, S., Wang, P., Mohammad-Rezaei Bidgoli, E., & Arefi, M. (2023). On the dynamics and wave propagation of reinforced composite nanosystem. Engineering with Computers, 39(1), 151-171.

Chen, Q. S., Yu, B., & Li, B. (2022). Probabilistic calibration of compressive stress–strain models for FRP-confined concrete in square cross section members. Journal of Composites for Construction, 26(3), 04022015.

Yang, Z., Hu, D., Guo, Q., Zuo, L., & Ji, W. (2023). Visual E 2 C: AI-driven visual end-edge-cloud architecture for 6G in low-carbon smart cities. IEEE Wireless Communications, 30(3), 204-210.

Zhang, K. X., Chen, Z. Y., Hu, S. S., Zhang, J. P., Li, Y. Z., Wu, D. L., ... & Soukhojak, A. (2026, August). Observation and Analysis of the “Galaxy” Defect in 4H-SiC through X-Ray Synchrotron Topography. In Defect and Diffusion Forum (Vol. 452, pp. 73-78). Trans Tech Publications Ltd.

Huang, S., Wang, X., Qu, T., & Liu, C. (2024). Torque curve bifurcation in T-shaped rotor continuously drilling into granular media. Powder Technology, 444, 119947.

Lu, W., & Wang, J. (2025). Accumulated Performance Comparison of Solar PV and Solar Thermal Water Heating in New Zealand. Engineering Reports, 7(12), e70526.

Hu, S., Chen, Z., Zhang, K., Li, Y., Zhang, J., Raghothamachar, B., ... & Zhu, Y. (2026). Stacking fault analysis for the early-stages of PVT growth of 4H-SiC crystals. Journal of Crystal Growth, 681, 128523.