Engineering and Technology Quarterly Reviews

Mociran, H. A., & Lapuste, A. V. (2025), Structural Health Monitoring of Buildings Using Computer Vision: A State-of-the-Art Review. In: *Engineering and Technology Quarterly Reviews*, Vol.8, No.2, 35-45.

ISSN 2622-9374

The online version of this article can be found at:
https://www.asianinstituteofresearch.org/

Published by:
The Asian Institute of Research

The Asian Institute of Research
Engineering and Technology Quarterly Reviews
Vol.8, No.2, 2025: 35-45
ISSN 2622-9374

Structural Health Monitoring of Buildings Using Computer Vision: A State-of-the-Art Review

Horatiu-Alin Mociran[1], Adina-Victorița Lăpuște[2]

[1,2] Faculty of Civil Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Correspondence: Horatiu-Alin Mociran, Faculty of Civil Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, St. Memorandumului 28, Romania. E-mail: horatiu.mociran@mecon.utcluj.ro

Abstract
Structural Health Monitoring (SHM) is essential for building safety, durability, and functionality. Buildings, as key components of the built environment, suffer from cracking, spalling, corrosion, and moisture damage. Traditional SHM approaches like vibration-based measurements, non-destructive testing (NDT), and manual inspections are reliable. However, they are expensive, slow, and difficult to use at scale. Recent developments in computer vision (CV), powered by advances in machine learning (ML) and deep learning (DL), have enabled modern, automated, and contactless inspection systems capable of detecting structural defects with high precision. This paper reviews the state of the art in computer vision applications for SHM of buildings. It focuses on the evolution of image processing, ML and DL architectures, and new 3D and multimodal systems. The paper categorizes common building defects, lists datasets for algorithm training and validation, and gives examples from recent studies. Finally, the review identifies current obstacles and suggests future research directions. It focuses on integration with drones, Building Information Modelling (BIM), the Internet of Things (IoT), and Digital Twin technologies.

Keywords: Structural Health Monitoring, Computer Vision, Deep Learning, Defect Detection, Buildings, Civil Engineering

1. Introduction

Ensuring the safety and long-term performance of civil engineering structures is one of the fundamental goals of structural engineering practice. Among these structures, buildings represent the backbone of the built environment, accommodating residential, commercial, educational, and industrial activities. Preserving their integrity and reliability throughout their service life is therefore a matter of public safety, economic value, and social well-being.

Structural Health Monitoring (SHM) has become a multidisciplinary approach dedicated to assessing the condition of structures, identifying early signs of deterioration, and supporting decisions regarding maintenance and rehabilitation. Its ultimate objective is to extend the service life of structures, reduce life-cycle costs, and improve their resilience against natural or anthropogenic hazards (Farrar & Worden, 2012).

Conventional SHM practices rely heavily on non-destructive testing (NDT) techniques such as ultrasonic pulse velocity, infrared thermography, rebound hammer testing, and strain gauge measurements, often complemented by manual visual inspections. While these methods are well established, they also face inherent limitations: manual inspections are time-consuming and subjective, whereas sensor-based systems require extensive setup, calibration, and maintenance, which increase operational costs.

In recent years, the convergence of computer vision (CV), machine learning (ML), and deep learning (DL) has triggered a paradigm shift in how structural health can be monitored. With the availability of high-resolution cameras, affordable unmanned aerial vehicles (UAVs), and increasingly powerful computational tools, vision-based systems have made it possible to perform scalable, automated, and non-contact inspections. These systems are particularly suitable for buildings, where most degradation processes—such as cracks or spalling—are visually perceptible and can be effectively analysed using digital imagery (Spencer et al., 2025; Zhuang et al., 2025).

The present paper aims to provide a comprehensive synthesis of computer vision applications in building SHM. The main contributions are as follows:

- Methodological overview: A detailed summary of computer vision approaches ranging from traditional image processing to deep learning and multimodal frameworks used for defect detection.
- Application-oriented review: An examination of practical case studies focusing on typical building defects such as cracks, spalling, corrosion, and moisture damage.
- Research perspectives: A discussion of current challenges and future directions, including integration with BIM, IoT, and Digital Twin environments.

2. Fundamentals of Structural Health Monitoring

2.1 Definition and Objectives

It is both conventional and expedient to divide the Method section into labeled subsections. These usually include a section with descriptions of the participants or subjects and a section describing the procedures used in the study. The latter section often includes description of (a) any experimental manipulations or interventions used and how they were delivered-for example, any mechanical apparatus used to deliver them; (b) sampling procedures and sample size and precision; (c) measurement approaches (including the psychometric properties of the instruments used); and (d) the research design. If the design of the study is complex or the stimuli require detailed description, additional subsections or subheadings to divide the subsections may be warranted to help readers find specific information.

Include in these subsections the information essential to comprehend and replicate the study. Insufficient detail leaves the reader with questions; too much detail burdens the reader with irrelevant information. Consider using appendices and/or a supplemental website for more detailed information.

Structural Health Monitoring (SHM) refers to the continuous or periodic observation of a structure through the acquisition and interpretation of data obtained from sensors, measurements, or visual inputs. Its purpose is to assess the current condition of a structure, detect potential damage, and predict its future performance. The concept of SHM emerged in the late 20th century as an evolution of traditional Non-Destructive Evaluation (NDE), shifting the focus from localized inspections toward global and continuous assessments (Sohn et al., 2001; Farrar & Worden, 2007).

In the case of buildings, SHM serves several essential objectives:

- Safety: Detecting early signs of damage before structural integrity is compromised.
- Serviceability: Ensuring that the building continues to perform its intended function effectively.
- Lifecycle management: Supporting preventive maintenance and timely repairs to extend service life.
- Post-event assessment: Providing rapid evaluations following earthquakes, fires, or extreme weather

conditions.

An effective SHM system thus provides engineers and facility managers with actionable information that enables data-driven decision-making and enhances the reliability of the built environment throughout its life cycle.

2.2 Traditional SHM Methods

Before the emergence of vision-based approaches, SHM systems relied primarily on three categories of methods: vibration-based measurements, non-destructive testing (NDT), and manual inspections.

Vibration-Based Methods - These techniques are founded on the principle that structural damage modifies the dynamic properties of a system—such as its natural frequencies, damping ratios, or mode shapes (Doebling et al., 1996). Accelerometers and strain gauges are typically employed to record such variations, especially in tall buildings or after seismic events. Advantages: sensitive to global damage and useful for dynamic behaviour assessment. Limitations: require dense sensor networks, complex modal analysis, and are less effective for localized surface-level defects.

Non-Destructive Testing (NDT) - Techniques including ultrasonic pulse velocity, infrared thermography, and rebound hammer testing are widely applied for localized inspections and subsurface assessment. Advantages: high precision in identifying internal or material-level damage. Limitations: labour-intensive, dependent on specialized equipment, and difficult to scale for large areas.

Manual Visual Inspection - Still the most common method in practice, manual inspection relies on the experience of engineers or inspectors who visually examine accessible parts of the structure. Advantages: straightforward, inexpensive, and does not require complex instrumentation. Limitations: subjective, time-consuming, and limited by accessibility and safety conditions.

While these traditional techniques remain valuable, their use on a large scale is constrained by high operational costs, potential human bias, and limited scalability (Balageas et al., 2006). The increasing complexity and ageing of urban infrastructure have made the need for more automated and efficient monitoring approaches evident.

2.3 Computer Vision vs. Conventional Methods

Over the last decade, computer vision (CV) has emerged as a promising alternative and complement to traditional SHM methods. Unlike sensor-based systems that depend on physical contact, CV relies on visual data—images or videos—acquired using cameras placed on tripods, handheld devices, or unmanned aerial vehicles (UAVs).

Main advantages of CV-based SHM include: non-contact monitoring; cost efficiency as cameras and UAVs become affordable; scalability for inspecting large façades or multiple structures; automation through ML/DL algorithms; and integration potential with BIM and Digital Twin platforms. Limitations include sensitivity to lighting and environmental factors, dependence on image quality, the need for large annotated datasets, and the difficulty of detecting subsurface defects without complementary NDT methods.

Despite these drawbacks, numerous studies indicate that CV-based techniques are maturing rapidly. In many cases, they now outperform conventional methods in terms of efficiency, automation, and scalability (Hoskere et al., 2018; Dong & Catbas, 2018). Hybrid frameworks that combine sensor data and vision analytics are also emerging as a balanced approach for comprehensive SHM (Mardanshahi et al., 2025).

2.4 Summary

The evolution of SHM represents a shift from localized, sensor-dependent evaluations toward image-based and data-driven methodologies. Traditional methods continue to play an important role, especially for subsurface defect analysis. However, computer vision provides major advantages in scalability, automation, and cost reduction. Given that most building defects—such as cracks, corrosion, or spalling—manifest on visible surfaces, vision-based systems are naturally suited to this domain.

3. Computer Vision Techniques for SHM

Computer Vision (CV) has become an increasingly important tool in Structural Health Monitoring (SHM) because it can extract relevant information from visual data and automate structural defects identification. Over the past two decades, CV techniques have evolved from classical image processing to advanced machine learning (ML) and deep learning (DL) methods, bringing substantial improvements in accuracy, robustness, and scalability.

3.1 Image Processing Techniques

Early vision-based SHM systems relied mainly on traditional image processing algorithms designed to enhance visible defects—such as cracks, spalling, or discoloration—by analysing variations in pixel intensity and texture. Common approaches include edge detection (e.g., Canny, Sobel), thresholding (e.g., Otsu), morphological operations (erosion, dilation, skeletonization), and texture analysis (GLCM). These methods are efficient and easy to implement, but are sensitive to lighting variations, surface textures, and noise. For instance, stains or paint patterns may be misclassified as cracks. Parrany et al. (2022) showed that with careful tuning, classical methods can still perform well under variable lighting. Today, image processing is often used as a preprocessing stage rather than a standalone detector.

3.2 Machine Learning Approaches

As labelled datasets became more available, machine learning (ML) methods emerged as a data-driven alternative to handcrafted rules. Algorithms such as Support Vector Machines (SVM), k-Nearest Neighbours (k-NN), and ensemble methods like Decision Trees and Random Forests classify patterns associated with defects using engineered features (e.g., HOG, LBP, wavelets). ML models are more flexible than classical image processing but depend strongly on feature quality and data diversity. Jahanshahi & Masri (2012) demonstrated an SVM-based system for crack detection in reinforced concrete with robust results in practical settings.

3.3 Deep Learning Approaches

Deep Learning (DL) has transformed CV applications in SHM by learning hierarchical features directly from raw images. Major architectures include classification networks (AlexNet, VGG, ResNet), object detection frameworks (Faster R-CNN, YOLO, SSD), segmentation models (U-Net, SegNet, Mask R-CNN), and transformer-based models (Vision Transformers and hybrid CNN–Transformer designs). DL offers high accuracy and robustness but requires large annotated datasets and considerable computation, and may struggle to generalize to unseen conditions (Zhang et al., 2018). Cha et al. (2017) achieved >98% crack detection accuracy using CNNs; Yasmin et al. (2023) demonstrated effective spalling detection via deep segmentation. Reviews by Cha et al. (2024) and Gao et al. (2023) underline DL's dominant role in current SHM research.

3.4 3D Vision and Multi-Modal Approaches

Beyond 2D imagery, 3D vision and multimodal systems combine visual, geometric, and thermal information for richer assessment. Photogrammetry and Structure-from-Motion reconstruct 3D façade models onto which defects can be mapped. LiDAR–CV integration enables accurate geometry and spatial localization (Tang et al., 2010; Olsen et al., 2010). Depth sensors (e.g., RGB-D) assist in detecting spalling and surface deformation, while RGB–thermal fusion supports moisture and subsurface defect identification. These benefits come with higher equipment costs and more complex data processing pipelines.

3.5 Summary

Computer vision for SHM has progressed from pixel-based processing to deep learning and multimodal 3D frameworks. While classical processing remains valuable for enhancement and preprocessing, DL currently leads in automation and precision. Integrating LiDAR, 3D reconstruction, and thermal imaging paves the way for hybrid monitoring systems that merge visual, geometric, and physical data.

4. Applications in Buildings

The adoption of CV for SHM in buildings has accelerated due to the prevalence of surface-visible defects and the practicality of image-based methods. Yet, challenges such as variable textures, illumination, and limited access—particularly for high-rise façades—persist. This section reviews typical building defects and representative case studies.

4.1 Typical Defects in Buildings

4.1.1 Cracks in Concrete and Masonry

Cracks arise from shrinkage, thermal effects, settlement, overloading, or seismic actions. Although fine cracks may be superficial, their propagation can affect serviceability and safety. Detection approaches include classical edge-based and thresholding methods (Yamaguchi & Hashimoto, 2010), DL models (e.g., CNNs) with accuracies above 95% (Cha et al., 2017), and emerging Vision Transformers that improve robustness to lighting and noise.

4.1.2 Spalling of Concrete

Spalling involves the flaking or detachment of concrete cover, often exposing reinforcement and accelerating corrosion. DL-based segmentation (U-Net, Mask R-CNN) on RGB images is widely used (Yasmin et al., 2019). Combining RGB with depth or 3D photogrammetry improves quantification of depth and extent. UAV imaging is practical for upper façades and difficult-to-reach areas.

4.1.3 Corrosion of Reinforcement and Metallic Components

Corrosion reduces the cross-sectional area of steel elements, leading to capacity loss. Colour-based segmentation using hue/saturation and DL models trained on corrosion datasets have shown strong performance (Gao & Mosalam, 2018). Infrared–RGB fusion can reveal corrosion beneath coatings in some scenarios.

4.1.4 Moisture and Water-Induced Damage

Moisture ingress leads to staining, efflorescence, biological growth, and coating delamination. Thermal infrared combined with CV detects temperature anomalies linked to moisture; RGB methods capture discoloration, and CNNs improve robustness. Linking moisture maps with BIM facilitates preventive maintenance planning.

4.2 Case Studies from Literature

4.2.1 UAV-Based High-Rise Façade Inspection

Dorafshan et al. (2018) compared edge-based crack detection with CNNs on UAV imagery, finding higher accuracy and faster inspection using deep learning—improving safety and efficiency in façade assessments.

4.2.2 Deep Learning for Crack Detection in Walls

Cha et al. (2017) developed a CNN-based framework for detecting cracks on concrete surfaces, achieving accuracy above 98% and demonstrating feasibility for real inspections.

4.2.3 UAV Inspection with Infrared and Visual Imaging

UAVs equipped with RGB and infrared cameras effectively assess building envelopes. Thermal imagery identifies insulation defects and moisture ingress, while RGB supports detailed crack analysis (Fox et al., 2016; Hoskere et al., 2018).

4.2.4 Crack Mapping Using Fully Convolutional Networks (FCNs)

Dung & Anh (2019) proposed an FCN for pixel-level crack segmentation, enabling quantitative evaluation and monitoring of crack evolution.

4.2.5 Spalling Detection in Reinforced Concrete Buildings

Yasmin et al. (2023) applied deep segmentation for spalling detection and, by incorporating depth information, estimated severity as well as extent.

4.2.6 Transfer Learning for Corrosion Recognition

Gao & Mosalam (2018) used deep transfer learning to recognize corrosion in reinforced concrete, showing that pre-trained CNNs can perform well even with limited training data.

4.2.7 BIM-Integrated Computer Vision for Maintenance

Brilakis et al. (2010) integrated CV-based defect detection with BIM, mapping defects onto 3D models to support predictive maintenance. Recent work extends this toward Digital Twin frameworks (Xu et al., 2023).

4.2.8 Hybrid Thermal and RGB Imaging for Moisture Detection

Fox et al. (2016) combined thermal and RGB imaging to detect moisture-induced deterioration, providing early warning of insulation failures and potential mould growth.

4.2.9 Post-Earthquake Building Assessment

Zhuang et al. (2025) developed a deep learning framework for post-earthquake damage classification with 96.1% accuracy, using Grad-CAM for interpretability and prioritization in emergency response.

4.3 Summary

CV applications in building SHM span cracks, spalling, corrosion, and moisture-related defects. Field studies, including UAV-based inspections, confirm gains in efficiency and safety. Persistent challenges include environmental variability, dataset standardization, and multimodal integration. Addressing these is key to robust, scalable, automated monitoring.

5. Datasets and Benchmarks

The performance and reliability of CV techniques for SHM depend on dataset quality and diversity. Large, well-annotated datasets are essential for robust ML/DL models capable of handling varied materials, defect types, and conditions.

5.1 Importance of Datasets in CV – based SHM

DL algorithms rely on large labelled datasets to achieve generalizable performance. Ideal SHM datasets include multiple defect categories (cracks, spalling, corrosion, moisture), various materials (concrete, masonry, steel), and diverse environmental conditions. Without such diversity, models overfit and fail in new scenarios (Zhang et al., 2018; Özgenel, 2018). Dataset creation is hindered by manual annotation effort, class imbalance, and limited diversity.

5.2 Publicly Available Datasets

5.2.1 CrackForest Dataset (CFD)

Originally for road surfaces, CrackForest (Shi et al., 2016) includes 118 annotated concrete images with pixel-level crack labels and is widely reused for building studies.

5.2.2 SDNET2018

SDNET2018 (Dorafshan et al., 2018) provides over 56,000 images of decks, walls, and pavements under diverse conditions. Though bridge-oriented, its vertical imagery suits façade inspection.

5.2.3 Concrete Crack Dataset (Özgenel, 2018)

Comprising 40,000 cropped patches labelled as crack/non-crack, this dataset is common for binary CNN training, though it lacks pixel-level annotations for segmentation.

5.2.4 Masonry Crack Dataset (MCD)

MCD (Dung & Anh, 2019) targets masonry walls and heritage buildings with labelled cracked and intact regions—useful beyond concrete, albeit smaller in scale.

5.2.5 Additional Specialized Datasets

Thermal datasets (moisture, subsurface defects) are often proprietary. Small corrosion datasets (e.g., Gao & Mosalam, 2018) support transfer learning. Spalling datasets exist but lack a standard public benchmark.

5.3 Benchmarking Practices

Common metrics: Accuracy and F1-score (classification), IoU (segmentation/localization), Precision–Recall (class imbalance), and FPS (real-time/UAV contexts). Comparisons across studies are difficult due to differing datasets, preprocessing, and protocols; the field lacks ImageNet-like standardized benchmarks (Hoskere et al., 2018).

5.4 Limitations of Current Datasets

Key limitations: small scale and narrow scope (focus on cracks), limited diversity (materials, lighting, weather), inconsistent annotation formats, and restricted access to industry data. These hinder generalization and slow progress toward deployable SHM systems.

5.5 Summary

Public datasets such as CrackForest, SDNET2018, and Özgenel's collection have shaped SHM research, but broader, multi-defect, multimodal datasets are needed. Standardized benchmarks and open challenges would accelerate progress and improve reproducibility.

6. Challenges and Future Directions

CV enables efficient, automated, non-contact SHM for buildings, yet adoption in practice is limited by environmental sensitivity, data scarcity, generalization issues, and computational demands. Concurrently, advances in AI, sensing, and digital construction create new opportunities.

6.1 Current Challenges

6.1.1 Environmental Sensitivity

Algorithms are sensitive to lighting, shadows, textures, and weather. Crack detection suffers under uneven illumination or stain-like backgrounds, limiting robustness (Chen et al., 2017).

6.1.2 Data Limitations

Large, diverse labelled data are scarce for buildings; popular datasets emphasize cracks. Models trained on one dataset often generalize poorly due to bias (Dong & Catbas, 2018).

6.1.3 Generalization and Transferability

Performance drops when moving from lab to field because of changing textures and lighting. Transfer learning and domain adaptation help but remain underexplored in SHM (Gao & Mosalam, 2018).

6.1.4 Integration with Structural Engineering Knowledge

Many CV systems lack explicit links between visual detections and structural significance (e.g., load paths, safety margins). Tighter coupling with mechanics would improve interpretability and usefulness (Hoskere et al., 2018).

6.1.5 Computational Demands and Real-Time Operation

Training and inference are resource-intensive, challenging real-time UAV deployments. Lightweight, energy-efficient models for edge devices are an active research area.

6.2 Integration with Emerging Technologies

6.2.1 Unmanned Aerial Vehicles (UAVs)

UAVs provide rapid, safe access to façades, roofs, and hard-to-reach zones. With RGB/IR/depth sensors, they capture high-resolution data at scale; paired with CV, they enable automated, repeatable monitoring (Dorafshan et al., 2018).

6.2.2 Internet of Things (IoT) and Wireless Sensor Networks

Hybrid SHM merges visual data with strain, vibration, and humidity sensing via IoT platforms, enabling richer diagnostics and multi-sensor fusion (Seo et al., 2015).

6.2.3 Building Information Modeling (BIM) and Digital Twins

CV detections can be mapped onto BIM for spatial records and maintenance planning. Digital Twins link physical assets to virtual models for real-time visualization, simulation, and prediction (Brilakis et al., 2010; Torzoni et al., 2024; Sun et al., 2025).

6.2.4 Augmented and Virtual Reality (AR/VR)

AR/VR enhances on-site decision-making by overlaying defect information on real scenes, enabling access to historical records and facilitating repair simulations.

6.3 Research Gaps and Future Trends

Promising directions include: large, multi-defect open datasets; compact models for edge/UAV deployment; self-supervised and semi-supervised learning to reduce labelling needs; explainable AI to link visual cues with structural behaviour; integration with lifecycle asset management; and advanced transformer-based architectures (Hu et al., 2025) for context-aware analysis.

6.4 Summary

Despite clear benefits, CV-based SHM faces barriers related to data, environment, and computation. Progress will rely on tighter integration with UAVs, IoT, BIM, and Digital Twins, along with explainable, lightweight AI models. These advances will shift practice from reactive inspection to proactive, predictive maintenance.

7. Conclusions

Structural Health Monitoring (SHM) plays a vital role in ensuring the safety, functionality, and durability of buildings throughout their service life. Traditional inspection methods—such as manual surveys, non-destructive testing (NDT), and sensor-based systems—have served the profession effectively for decades, yet their limitations in cost, scalability, and subjectivity motivate more automated solutions. Computer vision (CV) enables non-contact, efficient, and scalable monitoring.

This review summarized computer vision–based SHM approaches for buildings, covering foundational concepts, methods from classical image processing to deep learning and multimodal 3D systems, applications to typical defects, datasets and benchmarking practices, and key challenges with future directions. DL models—especially CNNs and segmentation networks—currently offer the most accurate detection and quantification of surface defects. Field studies confirm the maturity and feasibility of CV-based inspections.

While public datasets such as CrackForest, SDNET2018, and Özgenel's collection have accelerated progress, broader, standardized multi-defect datasets are needed for fair comparison and reproducibility. Future development will likely be shaped by integration with UAVs, IoT, BIM, and Digital Twins, together with advances in explainable, lightweight AI.

By leveraging modern AI, multimodal sensing, and digital construction workflows, CV-based SHM can deliver accurate, real-time, and cost-efficient assessments—supporting a shift from reactive inspection to proactive, predictive maintenance. This review aims to serve researchers and practitioners as a concise reference and to encourage scalable frameworks for smart, resilient, and sustainable cities.

Author Contributions: All authors contributed to this research.

Funding: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: The authors declare that ChatGPT (OpenAI, GPT-5 model) was used to assist in language refinement and editing of the manuscript. The use of this tool was limited to improving clarity, coherence, and consistency of expression. The authors reviewed and approved all content prior to submission and take full responsibility for the scientific integrity and accuracy of the manuscript.

References

Balageas, D., Fritzen, C. P., & Güemes, A. (2006). *Structural health monitoring*. John Wiley & Sons.

Brilakis, I., Lourakis, M., Sacks, R., Savarese, S., Christodoulou, S., Teizer, J., & Makhmalbaf, A. (2010). Toward automated generation of parametric BIMs based on hybrid video and laser scanning data. Advanced Engineering Informatics, 24(4), 456–465. https://doi.org/10.1016/j.aei.2010.06.006.

Cha, Y. J., Ali, R., Lewis, J., Büyüköztürk, O. (2024). Deep learning-based structural health monitoring. Automation in Construction, 161, 105328. https://doi.org/10.1016/j.autcon.2024.105328.

Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack detection using convolutional neural networks. Computer-Aided Civil and Infrastructure Engineering, 32(5), 361–378. https://doi.org/10.1111/mice.12263.

Chen, F. C., Jahanshahi, M. R., Wu, R. T., & Joffre, C. (2017). A texture-based video processing methodology using Bayesian data fusion for autonomous crack detection on metallic surfaces. Computer-Aided Civil and Infrastructure Engineering, 32(4), 271–287. https://doi.org/10.1111/mice.12256.

Doebling, S. W., Farrar, C. R., Prime, M. B., & Shevitz, D. W. (1996). *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review*. Los Alamos National Laboratory Report LA-13070-MS. https://doi.org/10.2172/249299.

Dong, C. Z., & Catbas, F. N. (2020). A review of computer vision–based structural health monitoring at local and global levels. Structural Health Monitoring, 20(2). https://doi.org/10.1177/1475921720935585.

Dorafshan, S., Thomas, R. J., & Maguire, M. (2018). SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks. Data in Brief, 21, 1664–1668. https://doi.org/10.1016/j.dib.2018.11.015.

Dorafshan, S., Thomas, R. J., & Maguire, M. (2018). Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. Construction and Building Materials, 186, 1031–1045. https://doi.org/10.1016/j.conbuildmat.2018.08.011.

Dung, C. V., & Anh, L. D. (2019). Autonomous concrete crack detection using deep fully convolutional neural networks. Automation in Construction, 99, 52–58. https://doi.org/10.1016/j.autcon.2018.11.028.

Farrar, C. R., & Worden, K. (2007). An introduction to structural health monitoring. Philosophical Transactions of the Royal Society A, 365(1851), 303–315. https://doi.org/10.1098/rsta.2006.1928.

Farrar, C. R., & Worden, K. (2012). *Structural health monitoring: A machine learning perspective*. John Wiley & Sons.

Fox, M., Goodhew, S., & De Wilde, P. (2016). Building defect detection: External versus internal thermography. Building and Environment, 105, 317–331. https://doi.org/10.1016/j.buildenv.2016.06.011.

Gao, Y., & Mosalam, K. M. (2018). Deep transfer learning for image-based structural damage recognition. Computer-Aided Civil and Infrastructure Engineering, 33(9), 748–768. https://doi.org/10.1111/mice.12363.

Gao, Y., Xu, W., Yang, J., Qian, H., Mosalam, K. M. (2023). Multiattribute multitask transformer framework for vision-based structural health monitoring. Computer-Aided Civil and Infrastructure Engineering, 38(12), 1578–1596. https://doi.org/10.1111/mice.13067.

Hoskere, V., Narazaki, Y., Hoang, T. A., & Spencer, B. F. (2018). *Vision-based structural inspection using Multiscale Deep Convolutional Neural Networks*. 3rd Huixian International Forum on Earthquake Engineering for Young Researchers, University of Illinois, Urbana-Champaign. https://doi.org/10.48550/arXiv.1805.01055.

Hu, D., Lin, Y., Li, S., Wu, J., & Ma, H. (2025). Hierarchical attention transformer-based sensor anomaly detection in structural health monitoring. Sensors, 25(16), 4959. https://doi.org/10.3390/s25164959.

Jahanshahi, M. R., & Masri, S. F. (2012). Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. Automation in Construction, 22, 567–576. https://doi.org/10.1016/j.autcon.2011.11.018.

Mardanshahi, A., Sreekumar, A., Yang, X., Barman, S. K., & Chronopoulos, D. (2025). Sensing Techniques for Structural Health Monitoring: A State-of-the-Art Review on Performance Criteria and New-Generation Technologies. Sensors, 25(5), 1424. https://doi.org/10.3390/s25051424.

Olsen, M. J., Kuester, F., Chang, B. J., & Hutchinson, T. C. (2010). Terrestrial laser scanning-based structural damage assessment. Journal of Computing in Civil Engineering, 24(3), 264–272. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000028.

Özgenel, Ç. F. (2018). *Concrete crack images for classification. Mendeley Data*, V1. https://doi.org/10.17632/5y9wdsg2zt.1.

Park, J. S., Joohyun, A., & Park, H. S. (2023). Computer Vision-based Structural Health Monitoring: A Review. International Journal of High-Rise Buildings, 12(4), 321–333. https://doi.org/10.21022/IJHRB.2023.12.4.321.

Parrany, R., Yazdani, N., & Dey, S. (2022). A new image processing strategy for surface crack identification in building structures under non-uniform illumination. IET Image Processing, 16(2), 407–415. https://doi.org/10.1049/ipr2.12357.

Pan, X., Yang, T. T. Y., Li, J., Ventura, C., Málaga-Chuquitaype, C., Li, C., Su, R. K. L., & Brzev, S. (2025). A review of recent advances in data-driven computer vision methods for structural damage evaluation: algorithms, applications, challenges, and future opportunities. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-025-10279-8.

Seo, J., Han, S., Lee, S., & Kim, H. (2015). Computer vision techniques for construction safety and health monitoring. Advanced Engineering Informatics, 29(2), 239–251. https://doi.org/10.1016/j.aei.2015.02.001.

Shi, Y., Cui, L., Qi, Z., Meng, F., & Chen, Z. (2016). Automatic road crack detection using random structured forests. IEEE Transactions on Intelligent Transportation Systems, 17(12), 3434–3445. https://doi.org/10.1109/TITS.2016.2552248.

Sohn, H., Farrar, C. R., Hemez, F. M., & Czarnecki, J. J. (2001). *A review of structural health monitoring literature: 1996–2001*. Los Alamos National Laboratory Report LA-UR-02-2095.

Spencer Jr., B. F., Sim, S. H., Kim, R. E., & Yoon, H. (2025). Advances in artificial intelligence for structural health monitoring: A comprehensive review. KSCE Journal of Civil Engineering, 29(3), 100203. https://doi.org/10.1016/j.kscej.2025.100203.

Sun, Z., Jayasinghe, S., Sidiq, A., Shahrivar, F., Mahmoodian, M., & Setunge, S. (2025). Approach towards the development of digital twin for structural health monitoring of civil infrastructure: A comprehensive review. Sensors, 25(1), 59. https://doi.org/10.3390/s25010059.

Tang, P., Huber, D., Akinci, B., Lipman, R., & Lytle, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. Automation in Construction, 19(7), 829–843. https://doi.org/10.1016/j.autcon.2010.06.007.

Torzoni, M., Tezzele, M., Mariani, S., Manzoni, A., & Willcox, K. E. (2024). A digital twin framework for civil engineering structures. Computer Methods in Applied Mechanics and Engineering, 418(Part B), 116584. https://doi.org/10.1016/j.cma.2023.116584.

Xu, J., Shu, X., Qiao, P., Li, S., & Xu, J. (2023). Developing a digital twin model for monitoring building structural health by combining a building information model and a real-scene 3D model. Measurement, 217, 112955. https://doi.org/10.1016/j.measurement.2023.112955.

Yamaguchi, T., & Hashimoto, S. (2010). Fast crack detection method for large-size concrete surface images using percolation-based image processing. Machine Vision and Applications, 21, 797–809. https://doi.org/10.1007/s00138-009-0189-8.

Yasmin, T., La, D., La, K., Nguyen, M. T., & La, H. M. (2023). Concrete spalling detection system based on semantic segmentation using deep architectures. Computers & Structures, 300. https://doi.org/10.1016/j.compstruc.2024.107398.

Zhang, A., Wang, K. C. P., Fei, Y., Liu, Y., Tao, S., Chen, C., Li, J. Q., & Li, B. (2018). Deep Learning–Based Fully Automated Pavement Crack Detection on 3D Asphalt Surfaces with an Improved CrackNet. Journal of Computing in Civil Engineering, 32(5). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000775.