Pixels to Perception; Advancing Reality Capture for AI-Driven Decision Making

Mohsen MIRI, Denmark

Ludvig EMGÅRD, Sweden

Key words: imagery, ultra-high resolution, photogrammetry, 3D digital models, reality capture, AI-driven analysis, crack detection, infrastructure inspection

ABSTRACT

Highly accurate, image-based products have become essential tools in decision-making processes across various aspects of daily life. From the inspection of critical infrastructure to environmental monitoring applications, reality capture technologies combined with AI-based analysis is providing the fundamental information required for rapid and insightful decision making. How much detail from reality do we need to capture and how much is good enough? This paper embarks on a journey from 'Pixel to Perception,' exploring sub-millimeter image-based 3D meshes to country-wide mapping applications based on Phase One's cutting-edge technologies, discussing the premium quality and high accuracy of the results.

1. INTRODUCTION

For more than three decades, Phase One, a global company based in Copenhagen, has developed core imaging technologies and a range of digital cameras and imaging modules, setting new standards for image quality in terms of resolution, dynamic range, color fidelity, and geometric accuracy.

Focusing on the geospatial industry, Phase One provides imaging solutions for UAV, aerial, and space platforms. Its product range includes cameras with global shutter technology, such as the P3 payload, capable of achieving sub-millimeter resolutions for inspection and damage detection of critical structures. For larger areas, the PAS 280 and PAS 150 camera systems support forestry applications and corridor mapping. For city modeling and country-wide mapping, PAS 880 and PAS Pana are the proper solutions respectively. Furthermore, Phase One cameras are integrated as the imaging component in hybrid-systems together with Lidar and Thermal sensors. The acquired data from these systems can be integrated with AI engines to address various dimensions and area sizes, enabling advanced feature extraction and data analysis. Whether mapping expansive urban landscapes or detecting fine-scale structural defects, these cameras exemplify the versatility and precision required for modern decision-making processes. Figure 1 shows the automatically extracted features, e.g. single trees and building roofs, using Geo-AI tools in Esri, based on the 3D dataset of city of Ljubljana, captured by PAS 880 camera system.

Figure 1: Building and single tree extraction with Esri AI-tools based on the Digital Twin of Ljubljana generated by ArcGIS Reality Studio, captured by Phase One PAS 880 camera system. Data courtesy of Phase One, Esri, FlyCom and Surveying and Mapping Authority of Slovenia.

1.1 Accuracy and Data Quality in Reality Capture

Determining the required level of detail and accuracy in data acquisition involves both relative and absolute accuracy. Relative accuracy defines how closely the geometry, and the texture of a digital twin resemble the physical object. In simpler terms, it assesses the fidelity of the digital 3D representation compared to the real-world counterpart. Absolute accuracy, on the other hand, pertains to the georeferencing of the digital twin, indicating how precisely the positioning and orientation of the final model align with a predefined (usually global) coordinate system.

Photogrammetric techniques are inherently influenced by lighting conditions and surface texture, both of which significantly impact the quality of the final results. Low-texture surfaces and curved geometries present challenges for dense matching algorithms, as they lack sufficient visual features for robust reconstruction. To mitigate these limitations, high-resolution imaging and global shutter cameras play a crucial role in enhancing the signal-to-noise (S/N) ratio in the final outputs, such as point clouds, 3D meshes, and orthophotos. This study focuses on structural monitoring of concrete surfaces, which are considered as challenging surfaces, due to their low-texture characteristics.

The Phase One GS120 camera, integrated with the P3 payload, has been used to account for varying lighting conditions. The study prioritizes relative accuracy by leveraging ultra-high-resolution imaging to achieve sub-millimeter precision, enabling automated crack detection on such surfaces.

The effectiveness of AI-based feature extraction relies on three key factors: (1) the quality of input data, (2) the accuracy and reliability of output products, and (3) the strength of the AI engine, which utilizes deep learning models. In this study, advanced AI algorithms developed by Spotscale (www.spotscale.com) are integrated for defect recognition. These algorithms are applied in a range of structural monitoring scenarios, including bridge and dam inspections, as well as urban infrastructure and power plant assessments.

1.2 Limitations of Manual Inspection Methods

According to Knyazkov et al., (2019), traditional *non-destructive testing* (NDT) methods for assessing reinforced concrete structures are mainly in manual ways, which rely predominantly on visual assessments based on direct measurement on the physical object. These methods present challenges in balancing key factors, including time, efficiency, accuracy, reliability, safety, risk, and scalability

Time-consuming and labor-intensive – Manual inspections are time-consuming and inefficient for large-scale infrastructure assessments.
Prone to human error – The accuracy of visual assessments is dependent on the inspector's expertise, leading to potential inconsistencies in defect identification.
Safety risks – Inspectors often operate in hazardous environments, increasing the likelihood of safety risks.
Limited access to critical areas – Structural constraints may prevent inspectors from thoroughly examining certain areas, resulting in undetected defects.

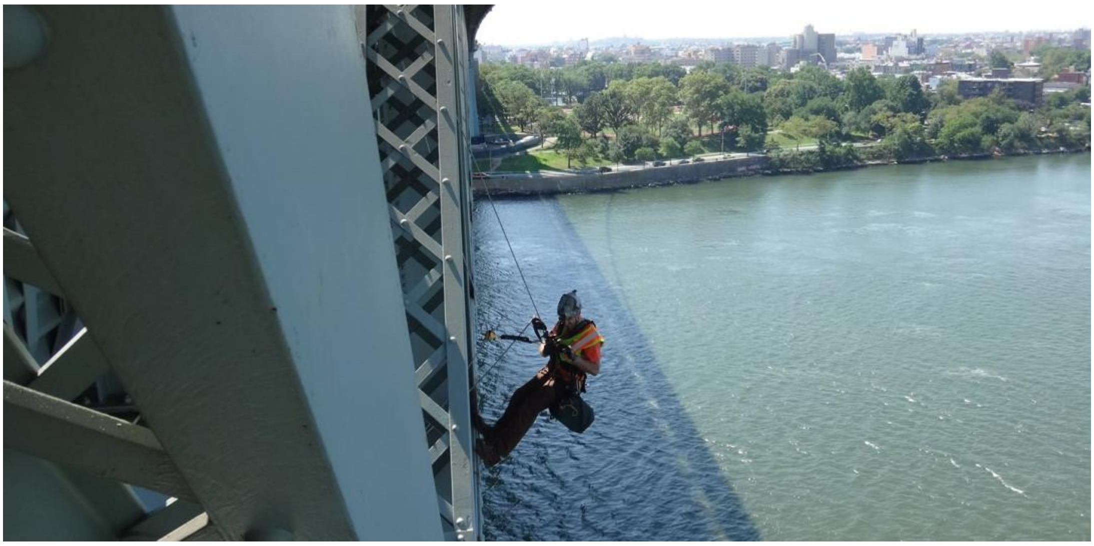

Figure 2: Elevated safety risks and reduced measurement reliability due to accessibility challenges.

1.3 UAV-based Reality Capturing and AI

Recent advancements in crack detection using Unmanned Aerial Vehicles (UAV) have demonstrated significant improvements in both efficiency and accuracy, compared to conventional inspection techniques. Over the past decades, studies have consistently shown that photogrammetry enhances structural monitoring precision and offers additional advantages over traditional methods (Valença et al., 2012).

The limitations in current methods using imagery include environmental factors, such as moisture or debris, that can obscure cracks. Recommendations for overcoming these challenges include developing more robust imaging systems and integrating AI algorithms capable of compensating for environmental variability.

Furthermore, lower-resolution imaging methods often fail to detect early-stage defects, delaying critical maintenance. In addition to the importance of accuracy in UAV-based imagery, one should consider increasing efficiency and productivity, for data acquisition of large areas or large objects, by using more automation in the workflow.

The table below compares *manual* and *automatic* (AI-based) concrete crack detection, focusing on key factors like accuracy, efficiency, cost, and safety.

Aspect	Manual Detection	Automatic (AI-Based) Detection
Accuracy	Subjective; prone to errors depends on expertise;	High accuracy; consistent results using AI and ML models.
Speed/Efficiency	Time-consuming; especially for large structures.	Fast processing; large areas analyzed in less time.
Cost	High labor costs; requires frequent site visits.	Higher initial setup cost; lower long-term operational cost.
Safety	High risk, especially for tall or remote structures.	Safer; drones and remote sensing eliminate on-site risks.
Coverage	Limited by accessibility; some areas hard to inspect.	Comprehensive; imagery drones cover hard-to-reach areas.
Data Consistency	Inconsistent due to human judgment.	Highly consistent; reduces human bias in detection.
Detail Level	May miss micro-cracks; dependent on human eyesight.	Detects micro-cracks with high-resolution imagery.
Documentation	Manual reports; time-consuming documentation.	Automated reports; integrated data analysis.
Scalability	Difficult to scale for large infrastructures.	Easily scalable with drone fleets and cloud processing.
Environmental Impact	Minimal; but frequent travel adds to emissions.	Efficient; fewer trips needed, reducing carbon footprint.
Real-Time Monitoring	Not feasible; periodic checks only.	Enables real-time monitoring and predictive maintenance.
Technology Dependency	Low; basic tools required.	High; requires drones, AI software, and computing power.

Table 1: Comparison Table: Manual vs. Automatic Concrete Crack Detection

1.4 Standards and the Need for Sub-millimeter Accuracy

The classification and assessment of concrete cracks are based on key parameters such as *width*, *depth*, *length*, and *location*, which determine their severity, impact on structural integrity, and the necessary repair actions. Various international standards and guidelines establish requirements for the design and production of concrete, such as *Eurocode 2*, while the inspection criteria for as-built concrete structures are often governed by national regulations.

National standards define procedures for concrete damage investigation and typically specify imaging resolution, crack width thresholds, and data validation. According to established guidelines, including those set by the *American Concrete Institute (2001)*, the *European Committee for Standardization (2004)*, and the *British Standards Institution (1997)*, crack width classification ranges from *0.1 mm* to *1.0 mm*. Achieving sub-millimeter accuracy in crack detection is therefore critical for ensuring compliance with these standards and enabling precise condition assessments of concrete structures. Table 2 presents the classification of crack severity levels based on crack width measurements.

Crack Width	Classification	Significance	Recommended Action
< 0.1 mm	Hairline / Micro-cracks	Surface-level, aesthetic; may result from shrinkage.	Monitor; no repair typically required.
0.1 mm – 0.3 mm	Fine cracks	Minor exposure risks; may allow moisture ingress.	Seal if exposed to aggressive environments.
0.3 mm – 0.5 mm	Moderate cracks	May compromise durability; risk of reinforcement corrosion.	Repair via sealing or epoxy injection.
0.5 mm – 1 mm	Major cracks	Potential structural impact; durability concerns.	Detailed assessment and structural repair.
> 1 mm	Critical cracks	Structural integrity at risk; significant damage.	Immediate repair or structural reinforcement.

Table 2: Classification of crack width in different risk levels

Sub-millimeter detection enables the identification of micro-defects that, if left unaddressed, can develop into critical structural issues. This underscores the necessity of adopting advanced imaging technologies for proactive maintenance, allowing for the precise localization and assessment of damaged areas requiring further destructive testing and eventual repairs.

To achieve accurate crack width detection, according to Nyquist-Shannon Sampling Theorem, the ground sample distance (GSD) must be at least half the target crack width. For instance, detecting a crack width of 0.2 mm requires a pixel resolution of 0.1 mm to ensure reliable measurement and analysis.

2. PHASE ONE UAV CAMERAS

Phase One's P3 payload offers unparalleled imagery capabilities for infrastructure monitoring based on GS120 and iXM100 cameras. Equipped with the GS120 camera with Global Shutter (GS) technology and more than 120 mega pixel resolution, this camera captures images with highest quality even in challenging illumination situations. This capability is a key factor in success and completeness of automatic detection of sub-millimeter structural cracks using AI-based algorithms. In addition to P3 payloads, Phase One offers the P5 camera, with 120 mega pixel resolution, which is suitable for sub-centimeter mapping and inspection applications, on fixed-wing drones. This article delves into the highest resolution with inspection cameras, e.g. the P3-GS120 camera.

2.1 Phase One GS120 camera on the P3 payload

The Phase One GS120 camera can be installed in any aircraft type (fixed-wing or copter) for high resolution imagery. This camera, which in this study is installed on the P3 payload, features a high dynamic range sensor based on Bayer Pattern technology, customizable lenses, and integration with drone systems for flexible deployment.

Due to high resolution (120 mega pixel), high dynamic range and the Global Shutter (GS) in GS120 camera, this advanced imaging technology, can capture sub-millimeter details from approximately 10 meters distance in challenging conditions, such as low light, higher speed flight or extreme weather.

Sensor type	CMOS Global Shutter
High Dynamic range (dB)	80
Pixel size (microns)	3.45
Sensitivity (ISO)	200
Global Shutter speed (sec)	1/16000
Color options	Color or Monochrome
Resolution	12768 x 9564
Max. field of view (°)	63
Continuous frame rate (fps)	Up to 5
RAW file compression (IIQ)	approx. 100MB

Table 03: Phase One GS120 camera technical specifications (up) and the features (down)

Figure 3: Phase One P3 payload for GS120 and iXM100 cameras on DJI-M350

2.2 Data Capture

High-resolution imaging using the P3-GS120 setup involves pre-planned flight paths or starting coordinates for drones or stationary setups for static structures. This ensures complete coverage and minimizes data gaps. The image collection for the needed accuracy is performed with a minimum of 80% overlap, with 7-10 meters distance to the object, ensuring that every defect is visible in at least 5 different images from different angles.

3. PROCESSING WORKFLOW

Measuring crack properties requires image representations with scale. The scale is achieved through projection of the crack pixels on a high accuracy geometry based on a photogrammetric 3D mesh. To achieve this, after the lab calibration of the cameras at Phase One facilities, the photogrammetry workflow in Spotscale software is utilized. This step includes image alignment, based on structure for motion (SFM) algorithm, followed by point cloud generation, and image rectification, 3D mesh generation, and finally, texturing the geometry with the original GSD. Parallel to this progress, a machine learning segmentation will generate the 2D image-based cracks. At the end the AI-based crack segments will be projected on the 3D mesh model. Advanced software tools optimize these steps for large datasets, ensuring high-quality outputs.

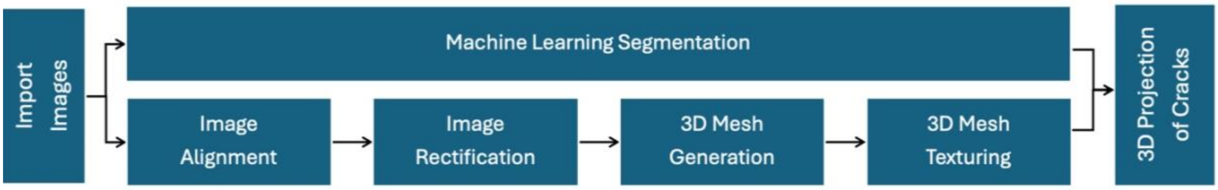

Figure 5: Photogrammetry and AI workflow for high-quality AI-based crack detection

3.1 High Dynamic Range and Intelligent Image Quality

Phase One cameras keep the image content and quality within a significant **high dynamic range (HDR)** in its smartly compressed raw data format, called Intelligent Image Quality (IIQ) format. This format ensures exceptional image quality across various lighting conditions. Figure 6 shows the HDR capability of Phase One's cameras to capture detailed information in both the brightest highlights and the deepest shadows, which is crucial for applications like aerial mapping and industrial inspections. This strong feature of Phase One images are across all its component cameras such as iXM-RS150, iXM-GS120, iXM-100MP and P5.

Figure 6: Raw IIQ images of a roof and a façade of a building in a challenging illumination condition with high contrast between extremely bright and dark areas (left, raw image), and radiometrically adjustment via HDR capabilities (right, adjusted image)

Figure 4: To extract features in the multi surface objects, some parts of the object could be hidden in the shadow areas (left) and where high dynamic range helps an advanced radiometric adjustment (right) to balance these areas

3.2 AI-based Analysis

AI algorithms detect and classify cracks based on predefined criteria, such as width and orientation. Each single crack is interpreted as a separate object and represented either in projected raster on the geometry or as a 3D polyline along the center line (medial axis) of the crack. To achieve this geometry, the crack pixels are analyzed from several different viewpoints and compared before projected on the mesh. This analysis significantly enhances the robustness of the Spotscale approach. The ability to transform the pixels to 3D polylines (vectors) enables the possibility to automatically determine medium and max width over the crack length and establish the length in three dimensions.

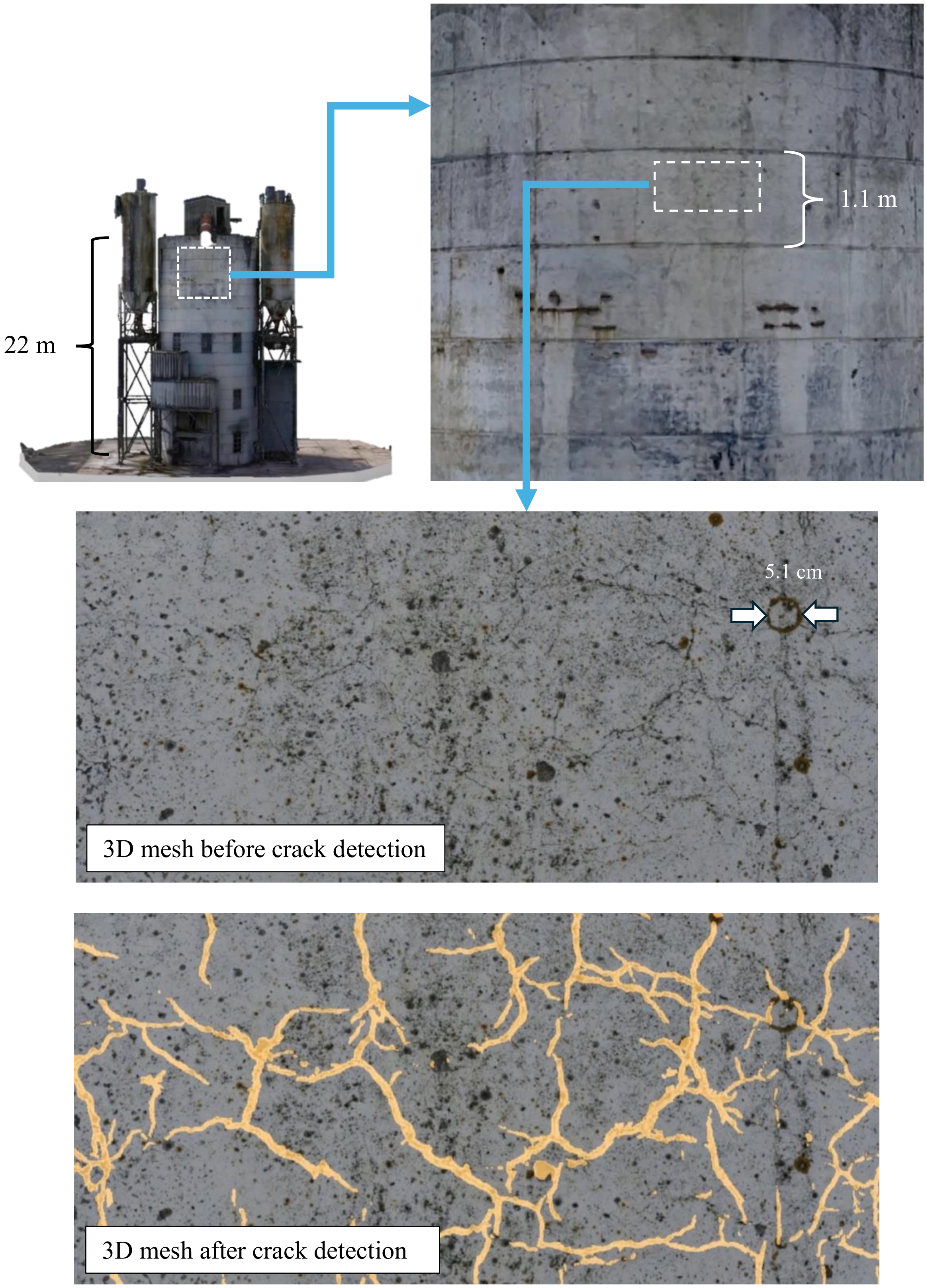

Figure 5: 3D detection single cracks in the shadowed side of a silo construction in risk of damages

Figure 6: 3D mesh (up) and the results from crack detection (down), identifying classes of cracks (yellow), spallings (pink) and rust (purple), highlighting an emerging spalling

3.3 Evaluation and Quality Control

Quality control measurements include cross-referencing AI results with manual annotations and evaluating metrics such as precision, recall, and F1 scores to enhance reliability. For this evaluation, some single tracks were measured manually within an arm length, to be used as benchmarks. This assessment is important to illustrate if the high-quality and ultra-resolution images of Phase One cameras could be successfully modeled in the geometry of 3D textured mesh, followed by AI results in *Spotscale*.

3.4 Accuracy and Results

Phase One's P3-GS120 camera and Spotscale algorithms achieve exceptional accuracy in crack detection in comparison with smaller camera sensors. The results show that with a flight distance of 4-7 meters, the target GSD of 0.1 mm which is half size of the commonly expected standards is achieved.

Metrics such as detection rate, false positive rate, and processing speed have been evaluated in a project under realistic circumstances. For example, the results based on the P3-GS120 camera achieved a 98.6% detection of the total crack length that was labelled by human labelling of the 3D reconstruction.

In contrast to manual inspection, there are virtually no limits in how many defects can be detected and presented by the software. In some concrete structures, thousands of individual cracks have been detected, measuring up to a kilometer in length.

In another investigation, the P5 camera from Phase One is used for pavement inspection of roads. In this research, the importance of geometry for automatic crack detection can be seen, to differentiate the open cracks from the repaired (filled) cracks.

3.5 Visual Illustrations

Visual examples of detected defects are provided in Figure 5 and 6, showing annotated images and 3D models. Figure 7 shows that the crack layers can be projected on the 3D mesh geometry, both in raster or vector formats, showing the length and width of the damage. By applying the AI interpretation, concrete experts can predict future decay such as concrete loss from the surface (spallings). The Spotscale software further enables the expert to measure depth on spallings that occurred in a sub-mm accuracy depth representation.

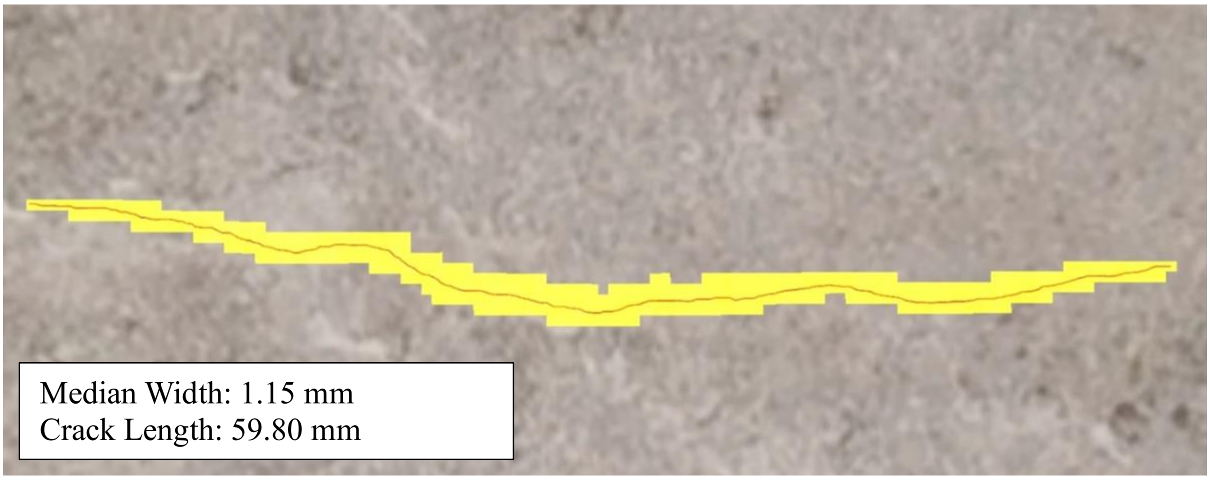

Figure 7: Results from crack detection, identifying classes of cracks in raster format(yellow), and representation of the crack axis in vector format (red line)

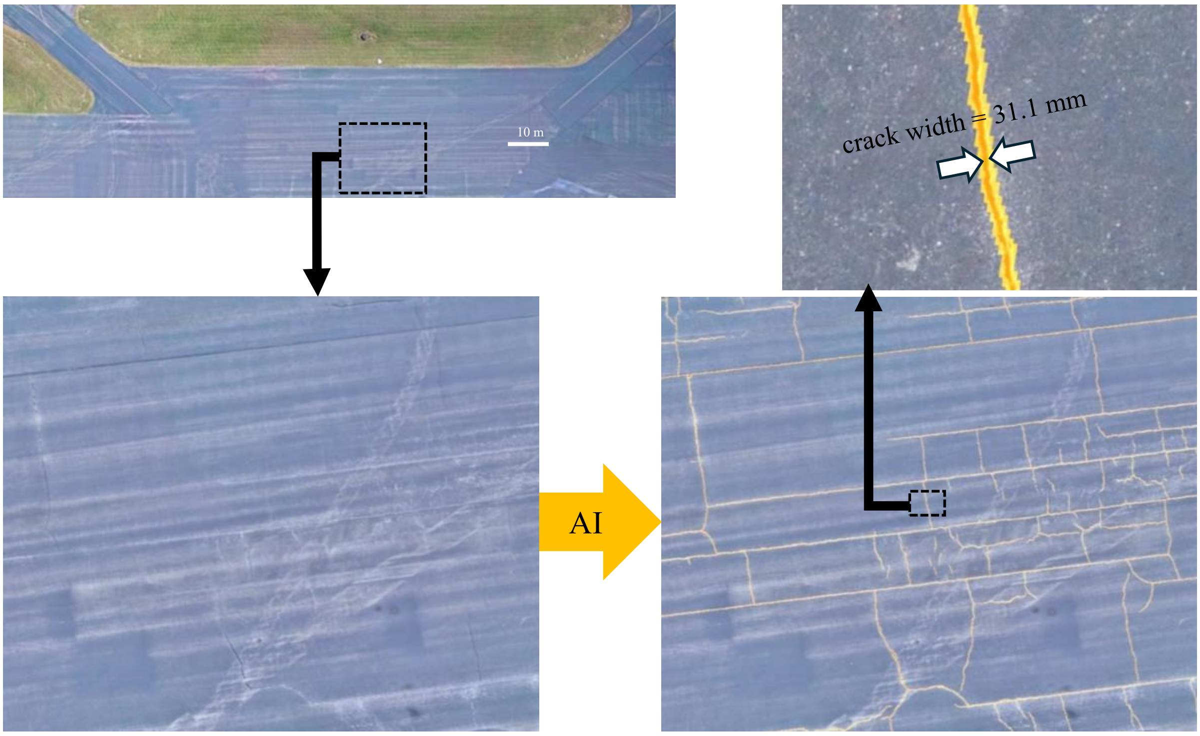

Figure 8: Results from crack detection for pavement inspection of roads and airport areas, based on P5 camera, before (left) and after (right) crack detection

4. CONCLUSION AND OUTLOOK

The integration of sub-millimeter imaging and AI-based analysis represents a significant advancement in non-destructive visual structural monitoring. By delivering high accuracy and meeting industry standards, this technology reduces risks, enhances maintenance efforts, and facilitates informed decision-making, including further destructive testing. Future directions include expanding the application scope to new domains, enhancing AI algorithms for broader defect recognition, and developing more cost-effective solutions for large-scale implementation.

Future research in AI-based crack detection should focus on developing advanced AI models capable of identifying a broader range of structural defects. Integrating data from multiple sensors, such as thermal imaging, radar, or ultrasonic sensors, can enhance deep learning models by providing additional information, thereby increasing the probability of accurate damage detection and improving overall reliability.

To facilitate large-scale deployment, a market analysis across various asset types and applications is essential to optimize the implementation of these technologies, ensuring a more targeted and efficient approach in different structural monitoring scenarios.

ACKNOWLEDGEMENT

The authors would like to thank the whole teams at Phase One and Spotscale, as well as company Kiwa (www.kiwa.com), for their support in this study.

BIOGRAPHY OF AUTHORS

Dr. Mohsen Miri is the Director of Strategic Partnerships at Phase One (www.phaseone.com). With his long-term experience in software and hardware product management, he leverages his extensive knowledge in geospatial business and strategic collaboration. He holds a PhD in Machine Learning and Feature Extraction, MSc in Photogrammetry Engineering and a BSc in Geodesy and Geomatics. As sustaining member of ISPRS, ASPRS, DGPF, EAASI and MAPPS, his career spans multiple commercial companies and research institutes, connecting solutions in AI, image processing, photogrammetry, surveying, and Geospatial Information Systems

Ludvig Emgård is Founder and CPO at Spotscale (www.spotscale.com), a tech entrepreneur and product leader specialized in 3D mapping and imaging technology. As the Founder of Spotscale, he built and led the company, now focusing on product and technology strategy in special applications like crack detection of critical infrastructure. With long-term product management experience in the reality capturing industry, and his strong background in geospatial imaging and AI-driven insights, he is shaping the future of this industry with state-of-the-art intelligent end-to-end solutions.

REFERENCES

Knyazkov, A. V., Kolpashchikov, V. L., & Kuchumov, R. A. (2019). *About the possible limitations in the usage of the non-destructive testing methods for assessment of reinforced concrete structures*. Buildings, 9(9), 202. https://doi.org/10.3390/buildings9090202

Valença, J., Júlio, E., & Dias-da-Costa, D. (2012). Application of photogrammetry to structural assessment. *Strain*, 48(5), 433–439. https://doi.org/10.1111/j.1475-1305.2011.00831.x

American Concrete Institute. (2001). Control of cracking in concrete structures (ACI 224R-01). https://www.concrete.org/store/productdetail.aspx?ItemID=22401

European Committee for Standardization. (2004). *Eurocode 2: Design of concrete structures – Part 1-1: General rules and rules for buildings* (EN 1992-1-1). Brussels, Belgium: CEN

British Standards Institution. (1997). *Structural use of concrete – Part 1: Code of practice for design and construction* (BS 8110-1:1997). London, UK: British Standards Institution.

Nyquist-Shannon Theorem (wikipedia update 2025) https://en.wikipedia.org/wiki/Nyquist%E2%80%93Shannon_sampling_theorem

REFERENCES OF FIGURES

Figure 1: ArcGIS Reality Experience, Phase One PAS 880 dataset, Data courtesy of Phase One, Esri, FlyCom and Surveying and Mapping Authority of Slovenia.

Figure 2: Bridge Inspections Combine Modern Technology with Traditional Techniques, WSP website (June 2024, update February 2025), https://www.wsp.com/en-gb/insights/2024-bridge-inspections-for-safety-and-integrity

Figure 3: Phase One website (update, February, 2025), P3 Payload, https://www.phaseone.com/solutions/geospatial-solutions/uav-payloads/p3-payload/

Figure 4,5,6,7 and 8: Phase One images and the results of the data sets (images, 3D meshes, AI cracks) in Spotscale software.

CONTACTS

Dr. **Mohsen Miri**

Director of Strategic Partnerships

Phase One A/S,

Copenhagen, Denmark

mir@phaseone.com

www.phaseone.com

Ludvig Emgard

Founder and CPO

Spotscale

Linkoeping, Sweden

ludvig.emgard@spotscale.com

www.spotscale.com