Automation in Construction 181 (2026) 106666

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Real-time vision-based defect detection for large-scale on-site earthen additive manufacturing: Annotated dataset and dual-model framework ☆

Hendrik Benz *, The Vinh Nguyen Trong, Katharina Klemt-Albert

Institute for Construction Management, Digital Engineering and Robotics in Construction, RWTH Aachen University, Germany

ARTICLE INFO

Dataset link: https://huggingface.co/datasets/ICoM-RWTH/3DPE-Crack22, https://github.com/ICoM-RWTH/3DPE-Crack22.git, https://doi.org/10.5281/zenodo.17554807

Keywords:
Additive manufacturing
Earthen 3D-printing (3DP-E)
Computer vision
Deep learning
Defect detection
Dataset

ABSTRACT

Additive manufacturing is transforming construction through automated, digitalized workflows, yet the layer-wise process remains prone to defects and geometric deviations during long-duration prints. Vision-based quality monitoring, when supported by domain-specific annotated datasets, enables deep learning models to detect such defects in real time. This paper introduces a publicly available annotated, material-specific dataset focused on large-scale earthen 3D-printing (3DP-E) under outdoor on-site conditions, comprising 2000 images labeled into three crack-related classes. A catalog of 14 visually distinguishable defect types is compiled, prioritizing those most relevant for computer vision-based monitoring. To demonstrate the dataset's applicability, a dual-model vision framework was implemented, stacking a nozzle-tracking model and a crack-segmentation model within a dynamic region of interest (ROI) that adapts to changing camera positions and printer motion. The framework confirms the feasibility of real-time defect detection and outlines a pathway for integrating vision-based supervision into feedback-driven process control workflows for on-site additive manufacturing.

1. Introduction

Advances in digital fabrication and automation are reshaping construction, with additive manufacturing (AM) emerging as a key enabler of material-efficient and design-flexible building processes. Despite rapid progress, the transition from laboratory-scale research to large-scale implementation on-site introduces new challenges in process control, quality assurance, and reliability.

AM is gaining traction in construction, but the layer-wise process faces challenges from complex automation, material variability, and long process durations, harming quality due to defects during deposition. This section discusses the current state of large-scale additive manufacturing, notes the limitations of existing monitoring methods and datasets, and emphasizes the need for a material-specific, vision-based framework for real-time defect detection.

1.1. Context

AM, commonly referred to as 3D-printing, is transforming the way physical structures are designed and produced [1]. Unlike subtractive manufacturing, AM fabricates components layer by layer directly from digital models, enabling greater geometric freedom and facilitating higher levels of automation through fully digital workflows [2]. The term *construction-scale additive manufacturing* has been used in early literature to describe extrusion-based fabrication systems capable of producing architectural components at building scale, notably in the context of concrete 3D-printing (3DCP) [3]. The adoption of AM in construction enables highly automated fabrication, expanded geometric freedom, and new opportunities for process optimization [4]. Within the construction industry, these developments have led to the emergence of additive construction, where large-scale structural components can be fabricated on-site without conventional formwork [5]. Combining automation through AM with the use of natural, low-carbon materials presents a promising pathway toward more responsible construction practices, where material efficiency and environmental impact are jointly optimized [6]. Under appropriate conditions, the adoption of earth-based AM — also referred to as earthen 3D-printing (3DP-E) — is not only considered a technical alternative but a strategic shift toward improving construction efficiency while contributing to a more sustainable practice [7–10]. Design-oriented investigations have explored the architectural potential of 3DP-E structures and material-driven form-finding strategies [11], while environmental assessments demonstrate the advantages of locally sourced, low-carbon materials, as illustrated by the life cycle analysis of the *TOVA* prototype. The paper quantified

☆ This article is part of a Special issue entitled: 'Next-Gen Construction' published in Automation in Construction.

* Corresponding author.
E-mail address: benz@icom.rwth-aachen.de (H. Benz).

https://doi.org/10.1016/j.autcon.2025.106666
Received 1 September 2025; Received in revised form 7 November 2025; Accepted 12 November 2025
Available online 19 November 2025

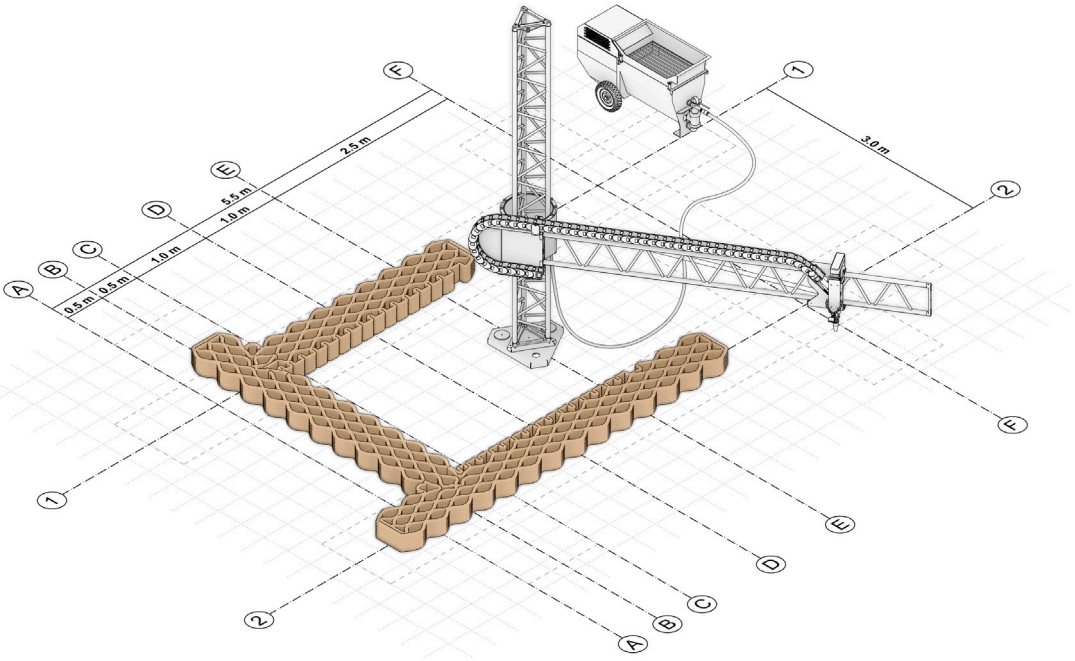

Fig. 1. Schematic overview of the *Crane WASP* on-site 3D-printing setup used for the fabrication of TOVA, a single-storey earthen unit.

a reduction of approximately 97% in CO_2 emissions compared to conventional fired-brick construction and up to 72% lower emissions than 3D-printed concrete. These improvements were primarily attributed to locally sourced materials (≈ 85%) and the elimination of formwork and transport-related impacts, underscoring the circular-economy potential of 3DP-E [12].

1.2. Motivation & challenges

The integration of 3DP-E into circular construction strategies highlights its potential for reducing CO_2 emissions and promoting material reuse [13]. However, compared to 3DCP, 3DP-E introduces unique challenges due to its distinct rheological behavior during continuous deposition [14]. Higher water content, slower hardening rates, and pronounced drying shrinkage make earthen materials more susceptible to defects such as cracks, insufficient interlayer adhesion, and geometric deformation [15]. If undetected, such defects may propagate across multiple layers, significantly compromising structural integrity and long-term durability.

Automated quality monitoring using computer vision (CV) and deep learning (DL) holds significant potential to address these challenges through early defect detection and enhanced process reliability. However, several technical barriers remain:

I *Dynamic process environment:* The on-site, large-scale printing setup leads to changing camera viewpoints and larger observed areas due to the unstructured environment compared to controlled laboratory-scale setups or industrial conveyor-based CV systems, complicating robust defect detection.

II *Outdoor environmental conditions:* Variable daylight, shadows, and occlusions affect image quality, requiring models to generalize under inconsistent visual conditions. This challenge was also evident during the *TOVA* project in Valldaura, Spain, printed outdoors over 24 days (≈ 6 h per day) under varying temperature, humidity, and sunlight exposure [12], where natural illumination caused shadows and reflections influencing defect visibility.

III *Limited data availability:* Publicly available datasets for construction-scale 3D-printing are scarce, particularly for earthen materials and defect-specific annotations, limiting the development and benchmarking of DL-based monitoring solutions.

IV *Process complexity:* Large-scale AM involves interacting factors, including material properties, nozzle positioning, and deposition direction, which must be considered jointly for accurate quality assessment.

This paper uses a dataset acquired during the on-site construction of the earthen 3D-printed demonstrator 'TOVA' in Spain, carried out within the 2021–2022 3D-Printing Architecture (3DPA) program at the Institute for Advanced Architecture of Catalonia (IAAC) using a crane WASP polar printer. (see Fig. 1).

Construction took place outdoors without any artificial lighting or enclosure, exposing the process to natural climate variability. The experiments were conducted in Barcelona, Spain, under typical Mediterranean conditions with clear skies and high sun exposure during most printing hours, as captured in Fig. 2. In addition to the direct influence of the sun, the polar kinematic configuration of the Crane WASP system introduced self-shadowing effects: as the articulated arm rotated around the central mast, parts of the machine intermittently blocked the sunlight, casting moving shadows across the freshly printed layers. These shifting occlusions produced transient dark regions independent of actual surface discontinuities.

Sun position changed continuously relative to the nozzle path, resulting in dynamic shadow casting both from environmental illumination and machine geometry. When the extruder arm moved with the sun behind it, surface tension cracks appeared with sharp contrast and were easily detectable. In contrast, when printing occurred against the light source or when the arm occluded the sun path, elongated shadow traces formed along filament edges, visually mimicking crack patterns even in the absence of material failure. These optical ambiguities, caused by both environmental lighting and machine-induced self-shadowing, directly motivated the definition of the *crack_shadow* label class to explicitly capture shadow-induced false positive crack detections.

While drying-induced cracking is a known phenomenon in earthen additive manufacturing and may occur due to climate-driven surface

Fig. 2. Dominant lighting scenarios during on-site printing: (a) direct sunlight and (b) diffuse lighting.

dehydration [9,10], the present paper focuses strictly on defects that occur during the active extrusion phase. At this stage, the material remains sufficiently plastic, and crack formation is primarily governed by extrusion pressure and deposition irregularities rather than post-deposition shrinkage. Drying-induced cracks may visually resemble some of the annotated defect classes; however, they are not explicitly included in the taxonomy, as the focus of this paper is limited to extrusion-phase irregularities. Only their indirect impact on visual appearance — primarily through light-driven contrast variations — was considered due to its relevance for vision-based detection.

In addition to optical variability, process-related material inconsistencies influenced crack formation during printing. Since earthen mixtures were prepared in small batching intervals, slight fluctuations in water content and mixing quality occurred throughout the build. Operators occasionally responded by manually increasing extrusion pressure to maintain filament continuity, which improved deposition flow but led to local over-extrusion zones prone to stress accumulation during drying. These interacting environmental and process conditions — sun exposure, self-shadowing due to gantry kinematics, and extrusion pressure adjustments — were intentionally preserved in the dataset to reflect the complexity of real on-site 3DP-E environments.

1.3. State of the art in defect detection in 3D-printing

Recent advances in computer vision have significantly improved process monitoring in AM, particularly in laboratory-scale settings where environmental conditions can be controlled. In construction-oriented AM, several sensing modalities have been investigated for defect-related monitoring tasks such as layer height estimation, surface deviation tracking, and geometric drift analysis. These approaches can be broadly categorized into vision-based and sensor-based strategies, each providing distinct benefits and trade-offs.

Deep learning for vision-based defect detection

Deep learning has become a key enabler in industrial automation due to its ability to extract meaningful representations from high-dimensional sensory data [16]. In predictive maintenance and health monitoring, deep learning methods have outperformed traditional signal processing pipelines by autonomously identifying early-stage failure patterns without handcrafted feature engineering [17]. In the context of visual inspection, convolutional neural networks (CNNs) have demonstrated strong performance in detecting surface anomalies in manufacturing environments where illumination and appearance variability challenge rule-based algorithms [18]. The integration of AI-driven perception systems into cyber–physical manufacturing architectures has been shown to improve process awareness and enable adaptive quality control in real time [19]. Specifically, CNN-based models have been adopted for real-time defect segmentation on production lines, where inference latency and robustness against material variability are critical [20]. These findings indicate that deep learning provides a suitable foundation for vision-based defect detection in AM workflows, especially under the dynamic and visually unstable conditions found in large-scale on-site AM.

RGB-based image monitoring

RGB image analysis has been widely explored for surface deformation detection, crack identification, and geometric consistency evaluation [10,21–23]. Its primary advantage lies in its simplicity and non-invasive deployment, allowing rapid dataset acquisition with minimal hardware complexity. However, RGB-only systems are highly sensitive to illumination changes and material reflectance—conditions that are exacerbated in outdoor 3DP-E, where surface color and texture vary with moisture content and drying progression.

RGB-D and depth-enhanced imaging

To address depth ambiguity inherent to RGB images, multi-view RGB-D sensing has been introduced for layer geometry reconstruction and height deviation detection [24]. Depth augmentation improves robustness against color-based segmentation errors, yet RGB-D cameras typically assume fixed calibration and stable backgrounds. This limits their applicability in mobile on-site setups where camera-to-object distance varies continuously and ambient sunlight interferes with active depth sensing.

Laser profilometry and ToF sensing

Laser-based systems such as 2D profilometers [25,26] and 1D time-of-flight (ToF) sensors [27] provide highly accurate geometric measurements of extruded filaments and layer height consistency. These systems excel at detecting deviations during controlled robotic extrusion. However, their narrow sensing field and dependency on precise alignment make them less effective for monitoring broader surface defects such as cracks, especially when printing materials exhibit irregular surface topology like earth-based mixtures.

Kinematic and motion-integrated monitoring

Motion-tracking and kinematic feedback systems have been used to correlate nozzle path deviations with resulting print irregularities [28]. These methods benefit from grounding defect detection in process-aware spatial context, yet they lack semantic interpretation of surface phenomena—making them insufficient on their own for identifying visually subtle defects such as fine cracks or shadow-induced anomalies.

Non-vision inline process sensing

Several studies have investigated non-vision feedback such as extrusion pressure monitoring to detect material inconsistencies and flow interruptions [29]. Such sensors are effective at detecting internal flow disturbances before defects visually manifest.

While these sensing strategies demonstrate strong performance within controlled environments, they do not directly translate to large-scale on-site 3DP-E, where camera viewpoints shift dynamically due to gantry motion and outdoor lighting conditions introduce shadow artifacts that mimic crack patterns [30].

1.4. Research gap

In large-scale on-site AM, particularly with gantry-based systems like the *Crane WASP*, the process environment is highly dynamic: camera perspectives shift due to machine kinematics, backgrounds change continuously, and outdoor lighting introduces shadow artifacts that challenge conventional defect detection strategies.

Unlike concrete-based AM research, where layer height deviation has been reported as a dominant geometric failure mode [26,27], the polar Crane WASP setup used in this project is mechanically stable in the Z-axis, resulting in consistently uniform vertical deposition. Consequently, geometric drift in build height was not considered a critical defect indicator in this context. Instead, in 3DP-E, material-related inconsistencies — primarily caused by fluctuating water content and small-batch mixing — are of higher relevance. These variations frequently result in manual compensation through increased extrusion

flow, which, while maintaining nominal layer geometry, introduces localized stress concentrations that manifest as surface cracks. For this reason, the present dataset and monitoring framework focus on defect types linked to material homogeneity and extrusion consistency, rather than layer height deviation. This aligns with the objective of developing a vision-based quality control strategy tailored to the process-specific failure modes of 3DP-E.

However, no publicly available dataset currently exists that captures these real-world visual challenges for 3DP-E under outdoor printing conditions.

This absence of benchmark data, combined with the limited real-time capability of existing defect detection frameworks, poses a significant barrier to developing robust, deployable quality monitoring systems.

There is therefore a clear need for:

I An annotated on-site dataset that reflects the visual complexity of 3DP-E processes, including shadow-induced crack artifacts.
II A lightweight, high-throughput detection framework capable of operating at real-time inference speeds during continuous nozzle movement.

1.5. Research objective

This paper aims to design and evaluate a computer vision-based monitoring system for detecting surface cracks during construction-scale 3DP-E. The approach integrates two YOLOv11 models: one for crack segmentation and one for nozzle tracking. By linking nozzle position to image-based predictions, a dynamic region of interest (ROI) is established, focusing detection on freshly printed layers.

The framework is evaluated on a custom annotated dataset created from on-site footage and assessed using standard performance metrics, including precision, recall, F1-score, mean average precision (mAP), and frames per second (FPS). The work is positioned as a proof-of-concept toward real-time quality monitoring, providing a foundation for extending the framework to additional defect types and multi-camera setups.

1.6. Research scope

This paper intentionally narrows its scope to enable a focused proof-of-concept investigation. The primary target is surface crack detection, selected due to its structural relevance and frequent occurrence in the dataset. The dataset originates from a single construction-scale printing project, limiting variability in material and geometry. Although the proposed monitoring framework is designed for real-time applications, evaluation is currently based on recorded video data.

Unlike controlled laboratory AM environments, large-scale on-site printing introduces constantly changing illumination conditions due to natural sunlight, dynamic nozzle orientation, and shifting surface reflectance as the material begins to dry. In particular, earthen materials exhibit strong visual variability because of moisture gradients and surface texture changes, which amplify shadow contrast and affect crack visibility. Instead of attempting to normalize or eliminate these environmental effects, the present work intentionally captures them as part of the dataset, establishing a realistic basis for vision-based defect detection under practical outdoor conditions.

2. Methodology

This paper proposes a *proof-of-concept framework* for camera-based quality monitoring in construction-scale AM using earthen materials. The framework demonstrates the feasibility of integrating DL-based defect detection into dynamic printing processes with continuously changing viewpoints, lighting conditions, and nozzle motion. As an initial use case, the current work focuses on *surface crack detection*, the most frequent and visually dominant defect in the dataset.

2.1. Defect catalog

A defect catalog was created to identify and describe 14 visually detectable defect types relevant for camera-based monitoring in construction-scale 3D-printing. Although the current dataset focuses on cracks, the catalog provides a foundation for future extensions and supports the development of automated defect detection strategies.

2.2. Physical setup

A dedicated vision dataset was captured during the fabrication of *TOVA*, a load-bearing earthen structure using the *Crane WASP* printing system under real on-site conditions in Spain. The printer operates in a polar kinematic configuration, consisting of one rotational degree of freedom around the central mast (θ), one radial extension axis (r), and a vertical axis (Z), with material extrusion following the direction of the extruders toolpath (E). and vertical positioning of the Extruder corresponds to Z-axis control. This kinematic layout results in non-linear camera-to-nozzle relationships during printing, making viewpoint diversity a relevant factor for dataset generalization.

To balance variability and calibration consistency, a single RGB camera (GoPro Hero5 Black, monocular 4K sensor) was used and sequentially repositioned rather than employing multiple unsynchronized cameras. Two mounting strategies were selected to systematically capture different observational contexts: (1) A wide-view mounting point, located on the Y-axis arm, capturing global deposition context, toolpath curvature, and background disturbances. (2) A near-nozzle mounting point, directly attached to the extrusion head, emphasizing material flow behavior, filament interface quality, and early crack initiation at close range.

Although the angular disparity between these two perspectives is limited, the change in spatial relation between nozzle, material deposition, and background results in distinct visual conditions. This controlled repositioning strategy increases viewpoint diversity without introducing temporal desynchronization, providing a balanced dataset for training vision models to operate under realistic robotic motion and changing occlusion patterns typical of outdoor AM scenarios.

To capture multiple viewpoints without introducing synchronization bias between camera feeds, a single RGB camera (GoPro Hero5 Black, monocular 4K sensor) was used and sequentially repositioned between different fixed mounting configurations. Mainly, two mounting points were used for this dataset:

A *wide-view* camera mounting position and a *near-nozzle* camera mounting position were employed to generate complementary visual perspectives, increasing dataset variability and supporting the robustness of the proposed machine vision–based defect detection workflow (Fig. 3).

2.3. Dataset acquisition

Raw videos, totaling one hour of on-site recording, were stored in MP4 format and processed into individual frames at a rate of one frame per second using a custom Python script. Each frame was divided into ten cropped regions of 640 × 640 pixels to match the YOLOv11 input format and to ensure coverage of nozzle path, deposited layers, and surrounding context [31]. A filtering step was then applied to retain only crops containing visible printed layers, while also including defect-free regions to improve model generalization. This process resulted in a balanced dataset focusing on structurally relevant areas for proof-of-concept evaluation.

2.4. Dataset preprocessing

All images were manually annotated using polygonal segmentation masks. Key statistics are summarized in Table 1. To improve consistency, image quality was enhanced through contrast and brightness normalization, providing more uniform input conditions for training.

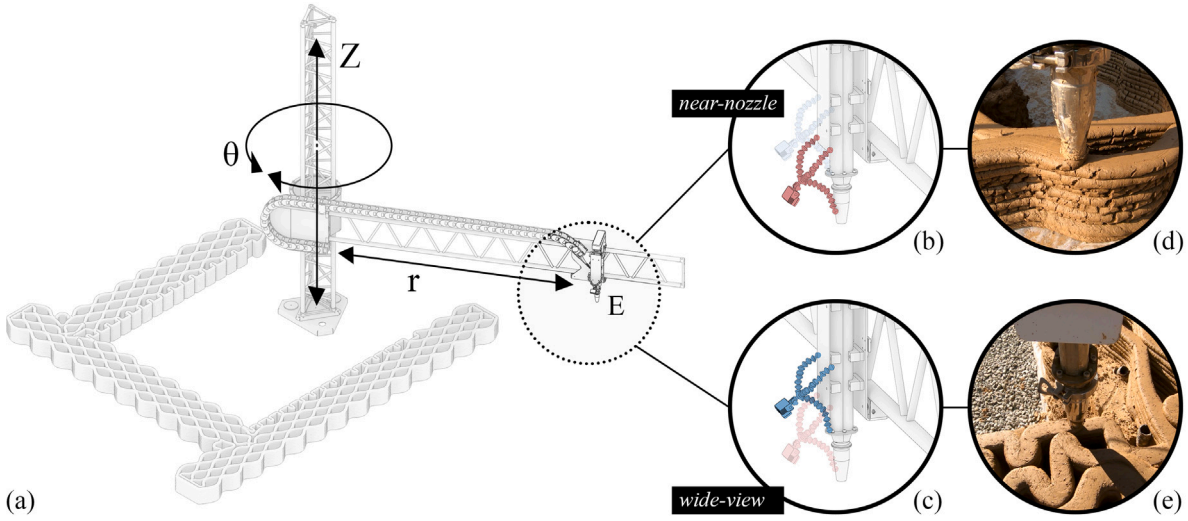

Fig. 3. Polar kinematic configuration of the *Crane WASP* system showing the rotational (θ), radial (r), vertical (Z), and extrusion (E) axes, along with the two camera mounting positions: *near-nozzle* (b) and *wide-view* (c), and their corresponding real on-site viewpoints (d–e).

Table 1
Parameters of experimental setup and annotated dataset statistics for defect detection in 3DP-E.

Aspect	Description
AM System	Crane WASP, modular polar 3D-printing system
AM method	Extrusion-based, continuous deposition (CD)
AM material	Earthen mix, on-site excavated soil
Environment	Outdoor; dynamic lighting and backgrounds
Sensor type	RGB-monocular 4K sensor
Model	GoPro Hero5 Black
Duration	60 min
Dataset size	2,000 extracted frames
Resolution	1680 ×720 pixels
Annotation Labels	*crack_b*, *crack_s*, *crack_shadow*
Lable Distribution	21,864 (9,197/12,033/634)
Ground Truth Format	Polygonal segmentation
Enhancement	Adaptive Contrast Enhancement

2.5. Annotation

The dataset was created from on-site recordings of large-scale earth-based AM processes under real construction conditions. Footage was captured using a single RGB-camera mounted on a adjustable base directly attached to the extruder or y-axis of a Crane WASP polar system, as shown in Fig. 4.

Annotation was performed manually to generate ground-truth labels for both nozzle detection and crack segmentation. Different annotation strategies were applied depending on the task: bounding boxes were used to localize the nozzle position in each frame, while instance segmentation masks were applied to capture the shape and extent of visible cracks.

Cracks were categorized into three classes: *crack_b* for long and wide cracks, *crack_s* for short and narrow cracks, and *crack_shadow* for cracks where more than 30% of their length was covered by shadows, as illustrated in Fig. 5. This classification accounts for visibility variations caused by lighting conditions and enables differentiation between defects with potential structural relevance and those with primarily aesthetic impact.

The initial dataset contained 2000 original images, which were divided into training, validation, and testing subsets. To enhance performance and robustness, data augmentation was applied, including variations in saturation (±15%) and exposure (±10%) to simulate lighting changes. This process expanded the dataset to 7593 images. A detailed distribution of the original and augmented datasets across the different splits is presented in Table 2. This approach increased dataset diversity while preserving realistic visual characteristics for accurate model training.

Table 2
Dataset composition before and after augmentation.

Split	Original	Augmented	Percentage
Training	1,400	6,994	70%/92%
Validation	400	401	20%/5%
Testing	200	198	10%/3%
Total	2,000	7,593	100%

To ensure sufficient training diversity, the dataset was designed with both statistical and practical requirements in mind. According to Yamane's formula (1977), the required sample size can be estimated as Eq. (1).

Symbol	Definition
n	Sample size
Z	Z-score
p	Assumed proportion (set to 0.5)
E	Acceptable margin of error

$$n = \frac{Z^2 \cdot p \cdot (1-p)}{E^2} \quad (1)$$

With a 99% confidence level ($Z = 2.58$) and a 5% margin of error ($E = 0.05$), the minimum required sample size per class was approximately 666 [32]. In practice, Ultralytics recommends at least 1500 labeled images per class and over 10,000 annotated instances in total to achieve stable YOLO performance [31]. The dataset used in this paper, consisting of 7593 augmented images, therefore represents a balanced compromise between statistical sufficiency, established best-practice guidelines, and the practical limitations of manual annotation.

2.6. Annotation tools

Two annotation platforms were used in this research: CVAT and Roboflow. The project initially began with CVAT for precise manual labeling, but the workflow was later transitioned to Roboflow to streamline dataset management and training. Roboflow provided an integrated environment that combined annotation, data augmentation, and model versioning, enabling faster iteration and simplified collaboration.

Fig. 4. Camera setup for data collection: (a) *Crane WASP* printer configuration during operation, (b) Camera mounted on the extruder with focus on nozzle to capture synchronized close-up views of freshly extruded layers.

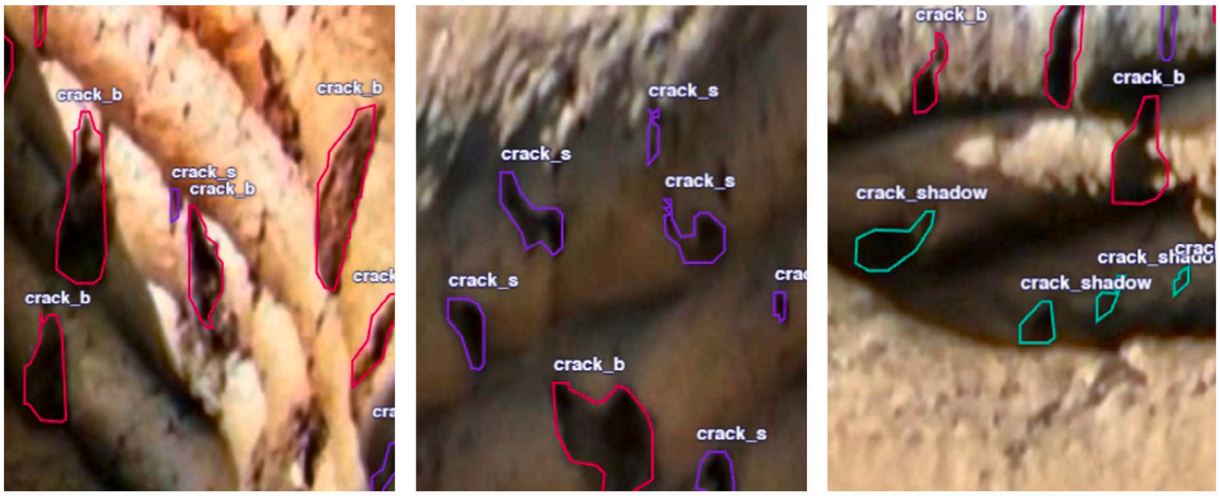

Fig. 5. Annotation examples for the three label classes in the dataset. Each panel highlights the dominant defect type: (a) *crack_b* showing larger size cracks (big), (b) *crack_s* illustrating fine surface cracks (small), and (c) *crack_shadow* representing shadow-induced crack-like artifacts. Polygonal masks distinguish true defects from lighting-induced false positives. Due to significant zooming to emphasize annotation regions, the visual resolution of the crops appears reduced.

The annotation process was conducted in Roboflow, applying bounding boxes for nozzle detection and segmentation masks for cracks. The AutoLabel function supported semi-automated labeling, improving speed and maintaining consistency across the three defect classes (*crack_b*, *crack_s*, *crack_shadow*). Additional augmentations, such as flipping, brightness variation, and contrast adjustment, were used to enhance model robustness under different daylight and surface conditions.

Roboflow was ultimately preferred over CVAT due to its integrated pipeline and rapid prototyping capabilities. While CVAT remains suitable for precise, privacy-sensitive projects, Roboflow's automation and training integration offered a more efficient workflow for this paper's iterative development process.

2.7. Monitoring framework

A dual-model YOLOv11 framework was implemented for crack detection. The first model detects the nozzle position, while the second segments surface cracks. A dynamic ROI is computed from the nozzle bounding box, restricting crack segmentation to a local area around the nozzle. This improves detection robustness in challenging conditions and reduces computational overhead.

3. Contributions

This paper advances automated quality monitoring for 3DP-E by providing four key contributions:

I **Annotated dataset:** The first construction-scale dataset for 3DP-E, comprising **2000 high-resolution images** collected during large-scale printing under real outdoor conditions, annotated for segmentation into three crack types.

II **Defect catalog:** A structured catalog of **14 visually identifiable defect types**, supporting the development of standardized computer vision-based monitoring frameworks and enabling future extensions beyond surface crack detection.

III **Dual-model detection framework:** A proof-of-concept deep learning pipeline integrating two YOLOv11 models for nozzle

Table 3
Defect catalog for AM in construction in alphabetical order.

Defect type	Description	Detection method & Suitability
Breakpoints	Discontinuities in extrusion due to flow interruption, clogging, or rheological mismatch [33,34].	Clearly visible as gaps or separations; suitable for RGB-based detection.
Bubbles	Surface damage from trapped air in the mix or poor degassing [33].	Detectable when close to the surface; small bubbles may require higher-resolution imaging.
Cracking	Linear fractures along layer surfaces, filament edges, or interfaces caused by shrinkage or thermal stress [35,36].	Easily detected in RGB imagery; critical for structural integrity monitoring.
Delamination	Poor adhesion or weak bonding between layers, visible gaps or delamination caused by lack of compaction or fast printing [35,37].	Visible through edge discontinuities and texture variations.
Elastic Buckling	Lateral deflection of slender elements due to insufficient stiffness [37].	Requires geometric monitoring or multi-view analysis to detect deformation trends.
Excessive Flowability	Sagging or spreading layers due to high water content or low viscosity [33,37].	Easily observed as uncontrolled layer deformation in RGB imagery.
Inconsistent Layer Height	Variations in geometry caused by over-extrusion, unsynchronized flow rates, or heterogeneous material properties [22,38].	Detected via height profile analysis and curvature-based segmentation.
Misalignment	Horizontal misalignment between adjacent layers, offsets or skewed lines caused by e.g. machine failure, material sagging, calibration errors, vibrations, or flow-out [22].	Detected using line tracking and contour comparison.
Over-Extrusion	Excessive material deposition in localized regions.	Produces visible surface irregularities and nozzle collisions; identifiable via contour tracking.
Plastic Buckling	Structural collapse from low early-age strength or short time delays between layers [37,39].	Strongly visible through bulging or deformation during printing.
Pores and Voids	Internal or interlayer cavities caused by air pockets or mismatched print speed [40,41].	Often undetectable on the surface; full detection may require CT or fracture surface analysis.
Process Downtime	Unplanned pauses in extrusion or toolpath execution.	Detected indirectly via time-lapse image sequences and weak interlayer bonding.
Temperature Fluctuations	Uneven thermal performance caused by variable printing parameters [42].	Detection requires thermal or infrared imaging; limited detectability via RGB cameras.
Under-Extrusion	Missing or incomplete extrusion in deposited layers; discontinuities in extrusion due to flow interruption, clogging, or rheological mismatch [33,34].	Visible gaps between deposited paths; well-suited for RGB-based detection.

tracking and crack segmentation within a dynamic region of interest (ROI), enabling localized, real-time detection.

IV **Evaluation under construction conditions:** Demonstrates robust performance in unstructured outdoor environments, achieving **52 FPS** and a precision of **0.82** for large-crack detection, validating the framework's feasibility for **feedback-driven robotic workflows**.

4. Defect catalog

This section provides a detailed overview of the 14 defect types included in the proposed defect catalog (Table 3). The defects are categorized based on their visual detectability, detection method and suitability within 3DP-E.

The defect catalog was compiled based on literature on extrusion-based AM in construction, focusing on structural impact, visual detectability, and relevance for real-time monitoring in Table 3 [22,35,37]. Although the catalog is derived from defects in 3D-printed concrete, these issues are also relevant in earth-based printing.

4.1. Breakpoints

Breakpoints refer to complete discontinuities in the extruded filament, caused by flow interruptions, nozzle clogs, or material inconsistencies [33,34]. They typically appear as visible gaps and are among the easiest defects to detect using RGB-based monitoring and real-time segmentation.

4.2. Bubbles

Bubbles form when trapped air escapes during extrusion, producing small surface "pop-outs" or irregularities [33]. They are moderately detectable via high-resolution RGB imagery but may require closer inspection when small.

4.3. Cracking

Cracks are linear fractures along layer surfaces, filament edges, or interfaces caused by shrinkage, thermal stress, or restrained deformation [35,36]. They are highly relevant due to their structural impact and are reliably detectable using edge tracking, morphological segmentation, or deep learning-based instance segmentation.

4.4. Delamination

Delamination occurs when interlayer adhesion is insufficient, leading to separation or visible gaps between layers [35,37]. It is easily detectable visually through edge discontinuities and textural inconsistencies.

4.5. Elastic buckling

Elastic buckling refers to lateral deflection of tall or slender printed elements under compressive loads, even without material yielding [37]. Detection requires geometric monitoring or multi-view camera setups, as subtle deformations are not always apparent in single-view RGB imagery.

4.6. Excessive flowability

Excessive flowability results in sagging, spreading, or layer collapse due to low yield stress, high water content, or improper mix design [33,37]. Deformations are typically visible during printing and can be effectively monitored via RGB cameras.

4.7. Inconsistent layer height

This defect describes uneven layer geometry caused by improper yield stress, over-extrusion, or mismatched nozzle speeds [22,38]. DL-based thickness segmentation or height-profile analysis enables reliable detection.

4.8. Misalignment

Misalignment occurs when printed layers are horizontally offset or skewed due to calibration errors, vibrations, or excessive nozzle speeds [22]. Line tracking and contour comparison methods provide robust detection.

4.9. Over-extrusion

Over-extrusion refers to the localized over-deposition of material, resulting in surface bulges or nozzle collisions. The defect is highly visible and can be identified through contour-based monitoring and RGB-based detection [43,44].

4.10. Plastic buckling

Plastic buckling occurs when fresh lower layers collapse under the weight of subsequent layers due to insufficient early-age strength [37, 39]. Bulging or folding deformations make this defect visually obvious during printing.

4.11. Pores and voids

Pores are small internal cavities formed by entrapped air or poor compaction, while voids are larger interlayer gaps caused by mismatched extrusion rates [40,41]. Voids are visually detectable on the surface, whereas pore detection typically requires advanced imaging such as micro-CT.

4.12. Process downtime

Process downtime refers to unplanned pauses during printing, typically caused by extrusion errors, mechanical malfunctions, or system resets. It can be inferred indirectly from time-lapse image sequences and is associated with weak interlayer bonding.

4.13. Temperature fluctuations

Temperature fluctuations occur due to nonuniform extrusion rates, nozzle speeds, or environmental conditions, leading to irregular curing and thermal stresses [42]. Detection generally requires thermal or infrared imaging; visibility in RGB imagery is limited.

4.14. Under-extrusion

Under-extrusion occurs when insufficient material is deposited, resulting in filament gaps or incomplete layers [33,34]. It is easily detectable via RGB-based inspection due to visible separation and measurable gap size.

5. Defect detection pipeline

The methodology begins with the identification and prioritization of relevant defect types, followed by dataset preparation, model training, and system integration. The overall workflow is designed to enable targeted monitoring of freshly printed layers and provide performance evaluation under realistic conditions.

Cracking was identified as the primary target for this paper due to its high structural impact, strong visual detectability, and suitability for vision-based methods [36].

5.1. Model selection

The YOLO (You Only Look Once) architecture was selected due to its balanced trade-off between detection accuracy, inference speed, and computational efficiency. As a single-stage detector, YOLO predicts bounding boxes and class probabilities in one forward pass, eliminating the need for region proposals typical of two-stage networks such as Mask R-CNN [45]. This design enables substantially higher processing speeds while maintaining competitive precision [46].

Two-stage models like Mask R-CNN reach high segmentation accuracy but are limited by slower inference and higher computational cost [46]. In contrast, YOLO enables real-time detection and has proven effective for tasks where speed and adaptability are essential. Comparative studies confirm this trade-off, showing that YOLOv8 performs better in detecting small and complex objects, while DeepLabv3+ achieves slightly higher pixel-level accuracy for larger homogeneous areas [47]. Similarly, combining DeepLabv3+ with YOLOv5 further improves overall detection performance, demonstrating that YOLO provides a robust and efficient foundation for hybrid detection–segmentation tasks [48].

Research in biomedical imaging also supports the advantage of integrating YOLO with segmentation models. The combination of YOLOv8 and the Segment Anything Model (SAM) achieved high precision, and close agreement with manual annotation [49].

Based on these findings, YOLOv11 was employed in this paper as the most recent and optimized version available at the time of experimentation. It integrates architectural refinements from previous generations, offering improved feature extraction and anchor-free prediction while maintaining real-time capability. These properties make YOLOv11 suitable for continuous outdoor monitoring in construction-scale 3D-printing.

5.2. YOLOv11 training setup

Two YOLOv11 models were trained for different tasks. The first model performed instance segmentation to detect and classify cracks into the three defined classes: crack_b (big), crack_s (small), and crack_shadow, as illustrated in Fig. 5. The second model performed object detection to localize the printing nozzle in each frame. Combining these models allowed for spatially aware defect detection through nozzle path tracking and dynamic ROI monitoring.

All input images were resized to 640 × 640 pixels to match the YOLOv11 input format [31]. For the crack detection model, segmentation masks were used, while the nozzle detection model used bounding boxes. Both models employed standard augmentation techniques such as horizontal flips, brightness adjustments, and contrast variations in addition to the dataset-level augmentations described in Section 2.5.

Model training was conducted on Roboflow using the YOLOv11 architecture for both segmentation and detection tasks. The crack segmentation model focused on accurate crack morphology learning, while the nozzle detection model emphasized robustness to viewpoint and lighting variations.

The configuration summarized in Table 4 reflects the distinct optimization goals of both models: minimal augmentation for precise crack texture learning and extensive augmentation for robust nozzle detection under varying site conditions. The best-performing checkpoints were

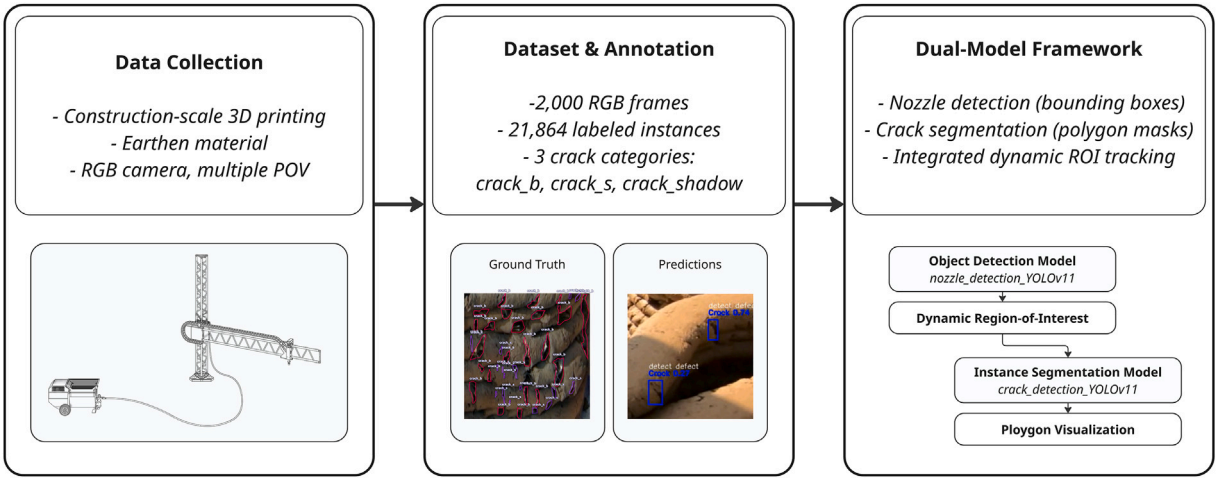

Fig. 6. Workflow of vision-based defect detection for construction-scale 3D-printing.

Table 4
Training parameters for crack segmentation and nozzle detection models.

Category	Parameter	Crack segmentation	Nozzle detection
Classes	model.names	crack_b, crack_s, crack_shadow	nozzle
Epochs	train.epochs	300	300
Batch size	train.batch	14	81
Image size	train.imgsz	640	640
Learning rate	lr.lr0	0.01	0.01
Weight decay	lr.weight_decay	0.0005	0.0005
Patience	train.patience	10	90
Optimizer	train.optimizer	auto	auto
Mosaic	aug.mosaic	0.0	1.0
Color augmentation (HSV h/s/v)	aug.hsv_h/s/v	0.0/0.0/0.0	0.015/0.7/0.4
Flip (left–right)	aug.fliplr	0.0	0.5
Translate	aug.translate	0.0	0.1
Scale	aug.scale	0.0	0.5
Erasing	aug.erasing	0.4	0.4

selected based on the highest validation mAP@0.5 and exported as PyTorch weights for inference.

The training process used a dynamic epoch configuration, where the number of training epochs, learning rate, optimizer, and other hyperparameters were automatically adjusted to optimize convergence. Transfer learning from YOLOv11 pre-trained weights was applied to improve accuracy and training efficiency, given the relatively limited dataset size.

All models were trained in a cloud environment with GPU acceleration, and the final model weights were exported in PyTorch (.pt) format for integration into the custom Python-based inference pipeline.

5.3. Dual-model framework and dynamic ROI logic

The detection system integrates two YOLOv11 models into a single framework, one for nozzle detection and one for crack segmentation, as described in Fig. 6. Each frame is first processed by the nozzle detection model, which outputs the bounding box coordinates of the nozzle. In parallel, the crack detection model performs instance segmentation and classifies defects into three categories: crack_b, crack_s, and crack_shadow.

The outputs of both models are merged by mapping crack detections relative to the nozzle position. This enables the definition of a dynamic red zone, positioned directly below the nozzle in the printing direction. The center of the red zone is offset from the nozzle center to match the location where cracks are most likely to appear on the freshly deposited layer. The radius of the red zone is scaled proportionally to the detected nozzle width, ensuring that its size adjusts to variations in perspective or zoom.

Spatial filtering is then applied so that only cracks whose centroids fall inside the red zone are kept. This ensures that the system focuses exclusively on defects appearing in the most recent layer, avoiding false detections from previous layers or background textures. Additionally, the red zone position is updated for every frame based on the nozzle movement, allowing it to follow curved or angled print paths.

Multi-stage computer vision pipelines are commonly used in industrial inspection and autonomous systems, where an initial detection stage is followed by localized refinement or segmentation [18,20]. However, such approaches typically assume static monitoring conditions, including fixed camera viewpoints, conveyor-aligned components, or globally stable regions of interest, which are not representative of on-site AM environments [50].

The proposed framework extends this paradigm by introducing a motion-aware dual-YOLO configuration designed for the polar kinematics and continuously shifting viewpoints of on-site AM. A YOLO-based nozzle detector dynamically conditions a crack segmentation model through a real-time updated region of interest (ROI) that follows the nozzle position rather than relying on a static field of view. This kinematically coupled ROI strategy enables spatially consistent crack monitoring under outdoor conditions with fluctuating illumination, material appearance shifts, and camera-to-surface perspective changes—advancing dual-stage vision concepts from static inspection settings toward process-aligned visual supervision in large-scale earthen 3D-printing.

For each confirmed detection, the system records the crack class, segmentation mask coordinates, confidence score, and frame timestamp. This information can be visualized in real-time by overlaying the nozzle position, red zone, and crack masks on the video feed, and can also be stored for later use in documentation for process optimization.

5.4. Evaluation metrics

The performance of the detection system is evaluated using standard object detection and segmentation metrics. For crack detection, precision, recall, and F1-score are calculated to assess the model's ability to correctly identify and classify defects. Precision (Eq. (2)) measures the proportion of correctly identified cracks among all predicted cracks, while recall (Eq. (3)) evaluates the proportion of detected cracks relative to all ground-truth cracks. The F1-score (Eq. (4)) provides a harmonic mean between precision and recall, ensuring a balanced evaluation of the model. The mean average precision at an Intersection over Union threshold of 0.5 ($mAP@0.5$) (Eq. (5)) is used as the main indicator of detection accuracy.

A confusion matrix is generated for the crack segmentation model to visualize classification performance across the three classes (crack_b, crack_s, and crack_shadow), highlighting cases of misclassification. For the nozzle detection model, $mAP@0.5$ is reported along with precision and recall.

In addition to accuracy metrics, the system's real-time capability is evaluated by measuring FPS during inference on the test hardware. The reported FPS includes the full pipeline, from frame capture and preprocessing to detection, ROI filtering, and visualization, to provide a realistic measure of operational performance. These combined metrics ensure that both detection quality and processing speed are considered in assessing the system's suitability for live monitoring in earth-based additive manufacturing.

Symbol	Definition
TP	True Positives
FP	False Positives
FN	False Negatives
AP_i	Average Precision for class i
n	Number of classes

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

$$mAP = \frac{1}{n} \sum_{i=1}^{n} AP_i \tag{5}$$

6. Evaluation results

This section evaluates the dual-model framework on the 3DP-E dataset, reporting quantitative results for nozzle tracking and crack segmentation, analyzing performance under different outdoor conditions, and illustrating success and failure cases.

6.1. Overview of experiments

The evaluation of the detection system was performed on the test dataset described in Section 2.5, which contains 198 annotated images representing different nozzle movements, lighting conditions, and defect types. This ensured that the system was tested under both controlled and challenging scenarios.

All experiments were carried out using the YOLOv11 models trained on a GPU-accelerated server and exported in PyTorch format. Inference was executed within a custom Python pipeline, which integrated nozzle detection, crack segmentation, red zone logic, and real-time visualization, as in Fig. 6. The test environment consisted of a standard workstation running Python 3.10, with OpenCV for image processing, Ultralytics' YOLOv11 framework for model inference. GPU acceleration was available during training in the cloud environment, while local inference tests were executed on CPU to validate compatibility and runtime efficiency.

Performance was measured using precision, recall, F1-score, $mAP@0.5$, and FPS, as defined in Section 5.4. In addition to these quantitative results, a qualitative evaluation was conducted by inspecting detection overlays in video frames to highlight correct detections, missed defects, and false positives.

Table 5
Crack detection performance on the test dataset.

Class	Precision	Recall	F1-score	$mAP@0.5$
crack_b	0.82	0.83	0.82	0.83
crack_s	0.71	0.65	0.68	0.65
crack_shadow	0.48	0.44	0.46	0.40
Overall	0.67	0.63	0.65	0.62

6.2. Crack detection performance

The performance of the crack detection model was evaluated on the test dataset using precision, recall, F1-score, and $mAP@0.5$. These metrics were obtained through post-processing of segmentation results from the validation pipeline, where predictions were compared against ground-truth masks for each crack class. The representative outcomes used for quantitative evaluation are shown in Fig. 8, illustrating typical segmentation accuracy and error pattern across the full range of confident scores. Table 5 presents the results, showing both class-wise performance and the overall system accuracy.

The results show that large cracks (crack_b) achieved the highest detection performance with balanced precision and recall. Smaller cracks (crack_s) were detected with moderate accuracy, while crack_shadow achieved the lowest scores, confirming the difficulty of detecting partially obscured cracks under low-contrast conditions.

A confusion matrix in Fig. 7 illustrates class-wise detection accuracy and misclassifications. Most errors occurred between crack_s and crack_shadow, reflecting the challenge of separating fine cracks from shadow artifacts in earth-based printing. In addition, crack_shadow was frequently misinterpreted as background, which reduced recall for this class. The matrix also shows that crack_b instances were generally classified reliably, with only minor confusion toward other classes. Overall, the confusion matrix highlights the main strengths and weaknesses of the model: robust detection of larger cracks, moderate success with smaller ones, and limitations in distinguishing shadow-influenced defects from surface textures.

Together, both representations provide a comprehensive view of the model performance. Table 5 quantifies detection quality by aggregating raw counts into statistical indicators such as precision, recall, and F1-score, while Fig. 7 visualizes where misclassifications occur at the default confidence threshold. Both originate from the same evaluation output but emphasize different analytical aspects: the table quantifies their statistical impact on the overall model performance, whereas the confusion matrix visualizes categorical confusion. This distinction clarifies the model's operational behavior under real-world conditions.

In Fig. 7 shows qualitative results that complement the quantitative evaluation. Correct detections were consistently observed for well-defined cracks, while false positives were typically associated with surface roughness and strong shadows. Missed detections occurred primarily for very fine or low-contrast cracks, underlining the limitations of visual detection under variable lighting conditions.

6.3. Nozzle detection performance

The nozzle detection model was evaluated on the test dataset using precision, recall, F1-score, and $mAP@0.5$. Table 6 presents the results, showing both class-wise performance and overall system accuracy.

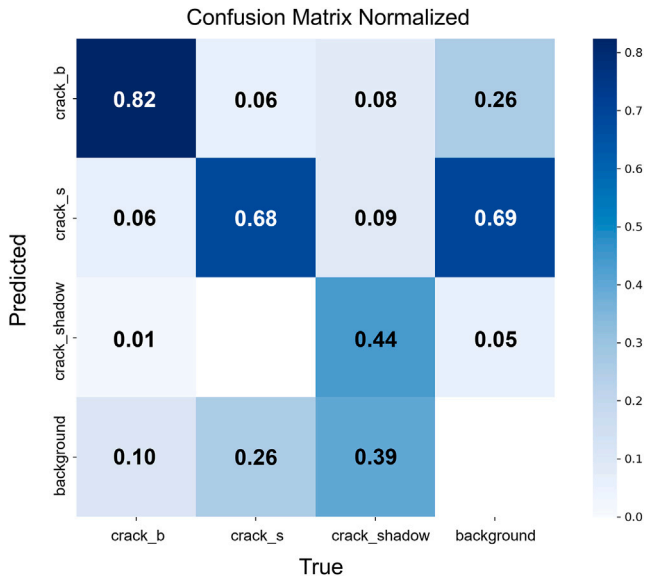

Fig. 7. Normalized confusion matrix from crack detection.

Table 6
Nozzle detection performance on the test dataset.

Class	Precision	Recall	F1-score	$mAP@0.5$
nozzle	0.97	0.97	0.97	0.98

The results show that the nozzle was detected with very high precision and recall, confirming the robustness of the YOLOv11 detection model for this task. The stable performance is explained by the nozzle's distinct visual features and consistent appearance across the dataset, making it easier to detect compared to cracks.

Visual inspection of detection overlays in Fig. 9 supports the quantitative evaluation. The nozzle was consistently localized with accurate bounding boxes, even under variable lighting conditions and during fast movements. Occasional deviations occurred when the nozzle partially overlapped with the printed layer or when motion blur was present, but these cases were rare and did not significantly affect overall performance.

6.4. Real-time performance

In practical applications, a real-time detection system must not only be accurate but also stable and responsive during continuous operation. The proposed framework was therefore evaluated with a focus on runtime performance and spatial consistency during live monitoring of the printing process.

The system achieved a sustained inference speed of *52 FPS*, confirming its suitability for real-time defect detection in earth-based additive manufacturing. The complete framework, including nozzle detection, crack segmentation, ROI logic, and visualization, maintained smooth frame-by-frame processing without delays, stuttering, or frame drops across the full validation sequence.

Qualitative inspection further underlined these results. Live inference frames (Fig. 10) demonstrate the real-time integration of nozzle tracking, defect segmentation, and visualization. Cracks were correctly detected within the monitoring region, with confidence scores remaining stable across frames. Even under challenging conditions such as motion blur or overlapping material layers, detections remained consistent with minimal misalignment.

Overall, the evaluation shows that the system is not only fast and accurate but also operationally robust, making it suitable for the integration into real-time automated defect detection and feedback-driven workflows.

6.5. Error analysis

Although the system achieved good overall performance, several recurring error types were observed during evaluation. These were grouped into false negatives, false positives, and overlapping predictions. Their main causes and impacts are summarized in Table 7, and examples are shown in Fig. 12.

These error patterns show that while the system is robust in most cases, it still struggles under challenging conditions such as poor lighting, complex textures, and irregular geometries. Further improvements in dataset diversity, annotation quality, and segmentation refinement are expected to reduce these errors.

6.6. Effect of dynamic ROI filtering

The impact of the dynamic ROI filtering was evaluated by comparing system performance with and without this component. As shown in Fig. 13, without ROI, the detections were scattered across the frame and focused mainly on a few large cracks. Smaller cracks and shadow-affected defects were often missed, and irrelevant background areas were mistakenly identified as cracks.

When ROI filtering was applied, as shown in Fig. 11, the system concentrated directly beneath the nozzle. The monitoring zone followed the nozzle's movement, ensuring that detections were limited to the freshly deposited layer. This improved coverage of both large and small cracks, stabilized detections across frames, and reduced false positives caused by surface textures and background noise.

The comparison highlights that dynamic ROI filtering increases detection precision, ensures spatial consistency, and keeps the monitoring process focused on the critical extrusion region.

7. Discussion

This paper demonstrates the feasibility of using camera-based monitoring for construction-scale AM under highly dynamic outdoor conditions. By separating nozzle localization from crack segmentation, the proposed dual-model YOLOv11 framework achieves robust performance despite changing backgrounds, multiple camera viewpoints, and variable lighting. The introduced dataset, which captures diverse viewpoints and challenging environmental conditions, provides an initial step toward establishing standardized benchmarks for visual quality monitoring in 3DP-E.

However, several limitations must be acknowledged. The current dataset annotation focuses exclusively on surface crack detection, while other visually detectable defects from the proposed catalog remain unannotated. Furthermore, the framework has not yet been tested in real-time deployment, and its generalization to other printers, materials, and sites requires further evaluation. Despite these constraints, the presented approach addresses key challenges not covered in related works, which often rely on controlled laboratory environments or synchronized multi-camera setups [24,30]. Future developments could integrate multimodal sensors such as RGB-D or thermal cameras to complement visual monitoring, enabling more robust detection under variable lighting and enhancing the discrimination of surface and subsurface defects [22,24,42].

The results suggest that computer vision has significant potential to enhance quality control in construction-scale AM. By integrating defect detection directly into the printing process, monitoring can become more efficient and less dependent on manual inspection. This creates opportunities for adaptive process control and lays the groundwork for scalable, automated quality assurance in large-scale additive manufacturing.

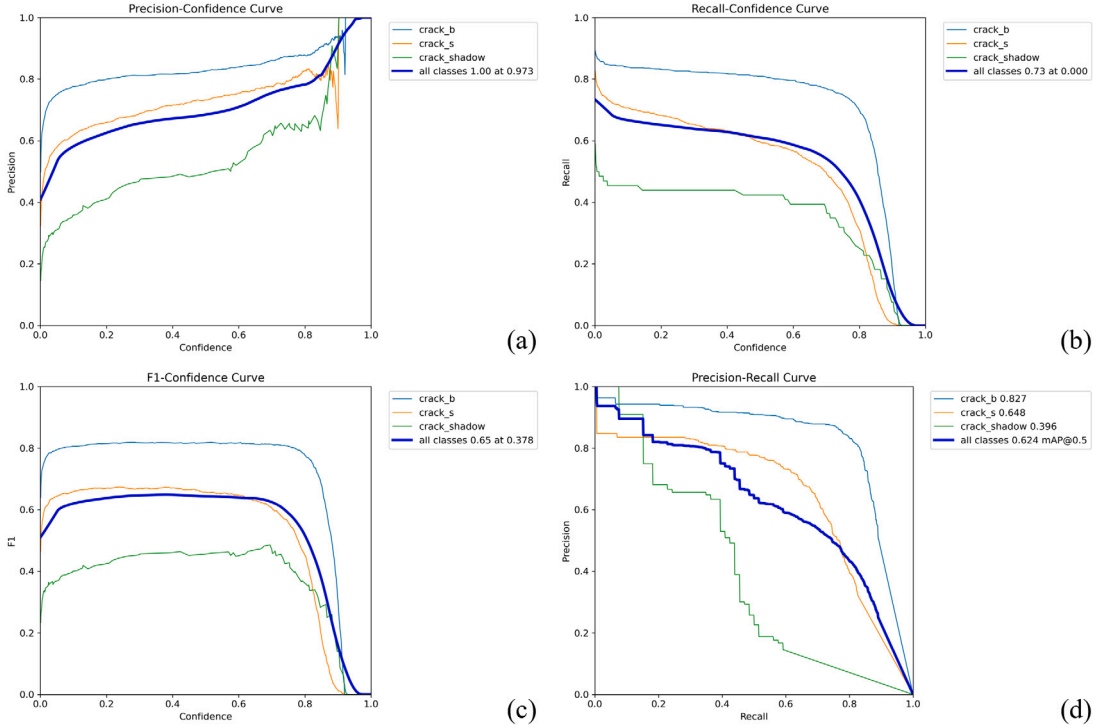

Fig. 8. Visualized crack detection performance across all confidence scores: (a) precision–confidence curve, (b) recall–confidence curve, (c) F1–confidence curve, and (d) precision–recall curve.

Table 7
Overview of error types, causes, and impact.

Error type	Main causes	Impact on system performance
False Negatives	Low-contrast cracks, shadows, partial obstruction	Missed critical defects, reduced recall
False Positives	Surface textures, reflections, lighting distortions	Reduced precision, unnecessary alerts
Overlapping Predictions	Irregular crack geometry, interrupted contours	Redundant detections, cluttered output

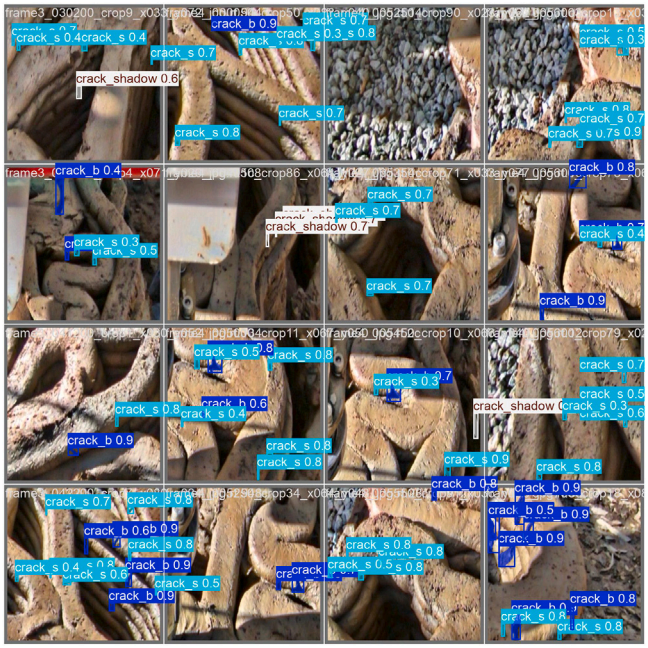

Fig. 9. Visual prediction results.

8. Conclusions

This section highlights the key results of the dataset and dual-model vision framework for real-time defect detection in earthen AM, along with future directions for practical deployment in digital construction workflows.

8.1. Conclusions

This paper presented a proof-of-concept computer vision framework for real-time defect detection in construction-scale 3DP-E. A dual-model YOLOv11 framework was implemented, combining instance segmentation for surface crack detection with nozzle tracking and dynamic ROI filtering. The models were trained on a custom annotated dataset of 7593 images and evaluated using precision, recall, F1-score, $mAP@0.5$, and FPS.

The results demonstrate reliable detection of surface cracks, achieving high accuracy for larger, visually dominant cracks and moderate performance for smaller or shadow-affected defects. Nozzle localization achieved consistently high precision and recall, ensuring stable tracking during printing. The complete framework operates at *52 FPS*, confirming its feasibility for real-time monitoring, while dynamic ROI filtering further improves detection stability and reduces false positives.

Overall, this paper provides a practical proof-of-concept for AI-driven quality control in construction-scale 3DP-E. By enabling continuous process monitoring and early defect detection, the proposed approach contributes to improving reliability and reducing failure risks in large-scale additive manufacturing. Future work will extend the framework to additional defect types, multi-sensor monitoring, and full integration into digital construction workflows.

(a) (b)

Fig. 10. Visual prediction of the nozzle from different camera angles: (a) *wide-view* camera, (b) *near-nozzle* camera.

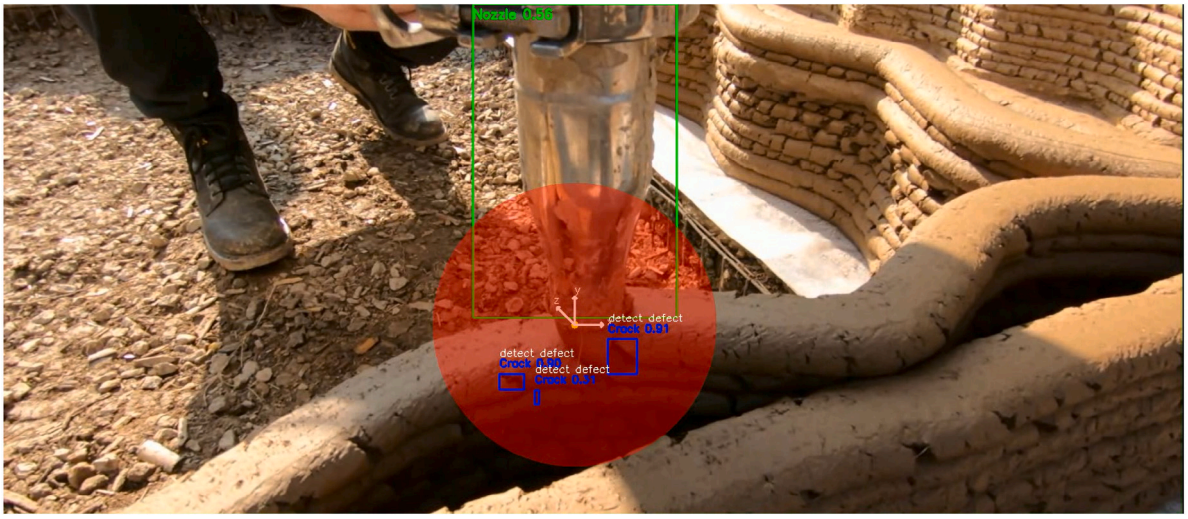

Fig. 11. Live inference frames showing integration of nozzle tracking, crack segmentation, and dynamic ROI filtering in real-time during printing.

(a) (b) (c)

Fig. 12. Examples of error: (a) false negative, (b) false positive, (c) overlapping predictions.

8.2. Future outlook

The development of more comprehensive, annotated datasets that capture a wider range of defect types, material conditions, and printing scenarios will be key to improving model generalization across different machine configurations and environmental contexts. In particular, integrating *multimodal sensing* — combining RGB vision with depth, thermal, or near-infrared imaging — offers strong potential to reduce ambiguity between material defects, background structures, and shadow-induced artifacts while improving robustness against illumination variability. Fusing sensor modalities with complementary characteristics could enable more reliable detection of subsurface defects, moisture-related anomalies, and early-stage crack initiation that are not always distinguishable in RGB imagery alone.

Beyond dataset expansion, future methodological advances are expected to leverage *multi-camera configurations* and synchronized sensor fusion pipelines that link perception directly to the kinematic state and deposition control of the printer. Such multimodal, process-aware integration would support more accurate detection of geometric drift, layer displacement, and deposition inconsistencies during extended prints.

Fig. 13. Live inference frames without ROI filtering.

A key research direction lies in embedding defect detection outputs into *closed-loop digital construction workflows*. Real-time feedback mechanisms capable of pausing printing, adjusting extrusion rates, compensating material feed, or dynamically adapting toolpaths represent an emerging pathway toward autonomous process control in large-scale AM. In addition, automated defect logging and structured quality reporting could support digital twin environments and progress monitoring frameworks, contributing to traceable and data-driven quality assurance in on-site additive construction.

CRediT authorship contribution statement

Hendrik Benz: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **The Vinh Nguyen Trong:** Visualization, Validation, Software, Formal analysis, Data curation, Writing – review & editing. **Katharina Klemt-Albert:** Writing – review & editing, Funding acquisition, Supervision.

Funding

This research was partially supported by the Federal Ministry for Economic Affairs and Climate Action (Bundesministerium für Wirtschaft und Klimaschutz, BMWK; later renamed Bundesministerium für Wirtschaft und Energie, BMWE) under the Central Innovation Programme for Small and Medium-Sized Enterprises (Zentrales Innovationsprogramm Mittelstand, ZIM), within the framework of the research project ClayX at RWTH Aachen University. The dataset was acquired during the construction of the demonstrator TOVA within the 2021–2022 3D-printing Architecture (3DPA) program at the Institute for Advanced Architecture of Catalonia (IAAC) in Barcelona. IAAC's 3DPA program participated in the Living Prototypes collaborative project, supported by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt- und Raumforschung, BBSR) on behalf of the Federal Ministry for Housing, Urban Development and Building (Bundesministerium für Wohnen, Stadtentwicklung und Bauwesen, BMWSB), through the Zukunft Bau research funding program.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hendrik Benz reports financial support was provided by Zentrale Innovationsprogramm Mittelstand (ZIM). The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The footage used in this dataset was recorded during the construction of the research prototype TOVA, conducted within the 2021/22 3DPA program at the Institute for Advanced Architecture of Catalonia (IAAC), under the direction of Edouard Cabay and Alexandre Dubor.

Researchers involved in the design and construction of TOVA include Marwa Abdelrahim, Mariam Arwa Al-Hachami, Adel Atassi, Michelle Bezik, Leonardo Bin, Hendrik Benz, Seni Dara, Deena El-Mahdy, Mehdi Harrak, Michelle Isoldi, Mouad Laalou, Eugene Marais, Charles Musyoki, Orestis Pavlidis, Juliana Rodriguez Torres, Nareh Sarnaghi, Nawaal Saksouk, and Aslinur Taskin.

IAAC Faculty involved in the program include Chamorro, Vincent Huyghe, Ashkan Foroughi, Elisabetta Carnavale, Francesco Polvi, David Škaroupka, Bruno Ganem Coutinho, Marielena Papandreou, Guillem Baraut, Gloria Font, Lukas Fertig, Nikol Kirova, Lili Tayefi, and Josep Alcover.

Data availability

The dataset 3DPE-Crack22 is publicly available at https://huggingface.co/datasets/ICoM-RWTH/3DPE-Crack22and is mirrored on GitHub at https://github.com/ICoM-RWTH/3DPE-Crack22.git. A citable DOI is provided via Zenodo: https://doi.org/10.5281/zenodo.17554807.

References

[1] B. Khoshnevis, D. Hwang, K.-T. Yao, Z. Yeh, Mega-scale fabrication by contour crafting, Int. J. Ind. Syst. Eng. (2006) http://dx.doi.org/10.1504/IJISE.2006.009791.

[2] ISO/ASTM, Additive Manufacturing – General Principles: Fundamentals and Vocabulary (ISO/ASTM 52900:2021), two thousand and twenty first ed., (52900) Beuth Verlag GmbH, 2021, URL https://www.dinmedia.de/en/standard/iso-astm-52900/348364425.

[3] S. Lim, R.A. Buswell, T.T. Le, S.A. Austin, A.G.F. Gibb, T. Thorpe, Developments in construction-scale additive manufacturing processes, Autom. Constr. 21 (2012) 262–268, http://dx.doi.org/10.1016/j.autcon.2011.06.010.

[4] A. Paolini, S. Kollmannsberger, E. Rank, Additive manufacturing in construction: A review on processes, applications, and digital planning methods, Addit. Manuf. 30 (2019) 100894, http://dx.doi.org/10.1016/j.addma.2019.100894.

[5] N. Labonnote, A. Rønnquist, B. Manum, P. Rüther, Additive construction: State-of-the-art, challenges and opportunities, Autom. Constr. 72 (2016) 347–366, http://dx.doi.org/10.1016/j.autcon.2016.08.026.

[6] O.B. Carcassi, L. Ben-Alon, Additive manufacturing of natural materials, Autom. Constr. 167 (2024) 105703, http://dx.doi.org/10.1016/j.autcon.2024.105703.

[7] A. Dubor, E. Cabay, A. Chronis, Energy efficient design for 3D printed earth architecture, in: Humanizing Digital Reality: Design Modelling Symposium Paris 2017, Springer Singapore, Singapore, 2018, pp. 383–393, http://dx.doi.org/10.1007/978-981-10-6611-5_33.

[8] H. Alhumayani, M. Gomaa, V. Soebarto, W. Jabi, Environmental assessment of large-scale 3D printing in construction: A comparative study between cob and concrete, J. Clean. Prod. 270 (2020) 122463, http://dx.doi.org/10.1016/j.jclepro.2020.122463.

[9] M. Gomaa, W. Jabi, A. Veliz Reyes, V. Soebarto, 3D printing system for earth-based construction: Case study of cob, Autom. Constr. 124 (2021) 103577, http://dx.doi.org/10.1016/j.autcon.2021.103577.

[10] O. Asaf, A. Bentur, P. Larianovsky, A. Sprecher, From soil to printed structures: A systematic approach to designing clay-based materials for 3D printing in construction and architecture, Constr. Build. Mater. 408 (2023) 133783, http://dx.doi.org/10.1016/j.conbuildmat.2023.133783.

[11] A. Dubor, E. Cabay, Y. Tayoun, O. Carrasco, E. Carnevale, E. Chamorro, V. Huygues, 3D-printed earth architecture: Design approach for a performative habitat, in: P. Ayres, M.R. Thomsen, B. Sheil, M. Skavara (Eds.), Fabricate 2024, UCL Press, 2024, http://dx.doi.org/10.2307/jj.11374766.27.

[12] D. El-Mahdy, E. Marais, M. AbdelRahim, A. Dubor, E. Cabay, Life cycle assessment of earth-based residential unit TOVA: A 3D printed on-site load-bearing structure, Case Stud. Constr. Mater. 23 (2025) e04925, http://dx.doi.org/10.1016/j.cscm.2025.e04925.

[13] K. Chadha, A. Dubor, E. Cabay, Y. Tayoun, L. Naldoni, M. Moretti, Additive manufacturing for the circular built environment: Towards circular construction with earth-based materials, in: C. De Wolf, S. Çetin, N.M.P. Bocken (Eds.), A Circular Built Environment in the Digital Age, in: Circular Economy and Sustainability, Springer, 2024, pp. 111–128, http://dx.doi.org/10.1007/978-3-031-39675-5_7.

[14] A. Perrot, D. Rangeard, E. Courteille, 3D printing of earth-based materials: Processing aspects, Constr. Build. Mater. 172 (2018) 670–676, http://dx.doi.org/10.1016/j.conbuildmat.2018.04.017.

[15] Y. Ji, P. Poullain, N. Leklou, The selection and design of earthen materials for 3D printing, Constr. Build. Mater. 404 (2023) 133114, http://dx.doi.org/10.1016/j.conbuildmat.2023.133114.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444, http://dx.doi.org/10.1038/nature14539.

[17] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R.X. Gao, Deep learning and its applications to machine health monitoring, Mech. Syst. Signal Process. 115 (2020) 213–237, http://dx.doi.org/10.1016/j.ymssp.2019.03.012.

[18] X. Zhang, Y. Xu, X. Jiang, Y. Liu, Deep learning-based visual inspection in manufacturing: A survey, J. Manuf. Syst. 54 (2019) 190–210, http://dx.doi.org/10.1016/j.jmsy.2019.01.002.

[19] J. Lee, B. Bagheri, H.-A. Kao, A cyber-physical systems architecture for industry 4.0-based manufacturing systems, Engineering 3 (2018) 5–12, http://dx.doi.org/10.1016/J.ENG.2017.05.015.

[20] L. Ren, W. Wang, Y. Yuan, Y. Sun, CNN-based real-time quality assessment in robotic manufacturing, Robot. Comput.-Integr. Manuf. 68 (2021) 102056, http://dx.doi.org/10.1016/j.rcim.2020.102056.

[21] S. Zimmermann, D. Griego, R.J. Flatt, Visualizing defects of concrete 3D printed structures with augmented reality based on machine learning-driven image analysis, in: D. Lowke, N. Freund, D. Böhler, F. Herding (Eds.), Fourth RILEM International Conference on Concrete and Digital Fabrication, DC 2024, in: RILEM Bookseries, vol. 53, Springer, 2024, http://dx.doi.org/10.1007/978-3-031-70031-6_18.

[22] R. Rill-García, E. Dokladalova, P. Dokládal, J.-F. Caron, R. Mesnil, P. Margerit, M. Charrier, Inline monitoring of 3D concrete printing using computer vision, Addit. Manuf. 60 (2022) 103175, http://dx.doi.org/10.1016/j.addma.2022.103175.

[23] J. Jhun, D.-H. Lee, A. Ur Rehman, S. Kang, J.-H. Kim, Development of a real-time geometric quality monitoring system for extruded filaments of 3D concrete printing construction, IEEE Access 12 (2024) 69981–69999, http://dx.doi.org/10.1109/ACCESS.2024.3401472.

[24] A. Zaki, L. Fischer, D. Müller, M. Schäfer, Monitoring of 3D concrete printing quality through multi-view RGB-D images, in: Proceedings of the 41st International Symposium on Automation and Robotics in Construction, ISARC, ISARC, 2024, http://dx.doi.org/10.22260/ISARC2024/0127.

[25] O.A. Jeyifous, E. Schönsee, C. Strangfeld, G. Hüsken, Correlation of continuously measured in-line process parameters and extruded geometry in 3D concrete printing experiments, in: D. Lowke, N. Freund, D. Böhler, F. Herding (Eds.), Fourth RILEM International Conference on Concrete and Digital Fabrication, DC 2024, in: RILEM Bookseries, vol. 53, Springer, 2024, http://dx.doi.org/10.1007/978-3-031-70031-6_15.

[26] L. Lachmayer, J. Quantz, H. Heeren, T. Recker, R. Dörrie, H. Kloft, A. Raatz, A spatial multi-layer control concept for strand geometry control in robot-based additive manufacturing processes, in: D. Lowke, N. Freund, D. Böhler, F. Herding (Eds.), Fourth RILEM International Conference on Concrete and Digital Fabrication, DC 2024, in: RILEM Bookseries, vol. 53, Springer, 2024, http://dx.doi.org/10.1007/978-3-031-70031-6_14.

[27] R.J.M. Wolfs, F.P. Bos, E.C.F. van Strien, T.A.M. Salet, A real-time height measurement and feedback system for 3D concrete printing, in: D. Hordijk, M. Luković (Eds.), High Tech Concrete: Where Technology and Engineering Meet, Springer, 2018, http://dx.doi.org/10.1007/978-3-319-59471-2_282.

[28] C. Gill, L. Justham, N. Lohse, A. Haynes, J. Dobrzanski, R. Buswell, P. Kinnell, Measurement and control of serial manipulator robots for 3D concrete printing, in: D. Lowke, N. Freund, D. Böhler, F. Herding (Eds.), Fourth RILEM International Conference on Concrete and Digital Fabrication, DC 2024, in: RILEM Bookseries, vol. 53, Springer, 2024, http://dx.doi.org/10.1007/978-3-031-70031-6_11.

[29] A. Kazemian, B. Khoshnevis, Real-time extrusion quality monitoring techniques for construction 3D printing, Constr. Build. Mater. 303 (2021) 124520, http://dx.doi.org/10.1016/j.conbuildmat.2021.124520.

[30] Y. Chen, K. Li, H. Zhou, X. Yang, Image processing and computer vision techniques for defect detection in additive manufacturing: A review, Int. J. Adv. Manuf. Technol. 126 (2025) 1123–1145, http://dx.doi.org/10.1007/s00170-025-16382-1.

[31] Ultralytics, Tips for best training results, 2023, URL https://docs.ultralytics.com/yolov5/tutorials/tips_for_best_training_results/.

[32] J. Gill, P. Johnson, Research Methods for Managers, Sage, 2010, URL http://digital.casalini.it/9780857023483.

[33] K. Wang, K. Wi, S. Laflamme, S. Sritharan, P. Taylor, H. Qin, Feasibility Study of 3D Printing of Concrete for Transportation Infrastructure, Institute for Transportation, Iowa State University, 2020, URL https://rosap.ntl.bts.gov/view/dot/56729, Final Report.

[34] H. Zhao, J. Sun, X. Wang, Y. Wang, Y. Su, J. Wang, L. Wang, Real-time and high-accuracy defect monitoring for 3D concrete printing using transformer networks, Autom. Constr. 170 (2025) 105925, http://dx.doi.org/10.1016/j.autcon.2024.105925.

[35] Y. Zhao, W. Meng, P. Wang, D. Qian, W. Cheng, Z. Jia, Research progress of concrete 3D printing technology and its equipment system, material, and molding defect control, J. Eng. 2022 (2022) 1–22, http://dx.doi.org/10.1155/2022/6882386.

[36] G.M. Moelich, J. Kruger, R. Combrinck, Plastic shrinkage cracking in 3D printed concrete, Compos. Part B: Eng. 200 (2020) 108313, http://dx.doi.org/10.1016/j.compositesb.2020.108313.

[37] F.Z. Oulkhir, I. Akhrif, M. El Jai, 3D concrete printing success: An exhaustive diagnosis and failure modes analysis, Prog. Addit. Manuf. 10 (1) (2025) 517–559, http://dx.doi.org/10.1007/s40964-024-00638-5.

[38] A.V. Rahul, M. Santhanam, H. Meena, Z. Ghani, 3D printable concrete: Mixture design and test methods, Cem. Concr. Compos. 97 (2019) 13–23, http://dx.doi.org/10.1016/j.cemconcomp.2018.12.014.

[39] V. Sergis, C. Ouellet-Plamondon, An optimum mix design method for 3D concrete printing applications, in: A. Jędrzejewska, F. Kanavaris, M. Azenha, F. Benboudjema, D. Schlicke (Eds.), International RILEM Conference on Synergising Expertise Towards Sustainability and Robustness of Cement-Based Materials and Concrete Structures, in: RILEM Bookseries, vol. 44, Springer Nature Switzerland, Cham, 2023, pp. 665–672, http://dx.doi.org/10.1007/978-3-031-33187-9_61.

[40] N. Alanazi, J.T. Kolawole, R. Buswell, L. Susmel, The theory of critical distances to assess the effect of cracks/manufacturing defects on the static strength of 3D-printed concrete, Eng. Fract. Mech. 269 (2022) 108563, http://dx.doi.org/10.1016/j.engfracmech.2022.108563.

[41] J. Kruger, A. Du Plessis, G. Van Zijl, An investigation into the porosity of extrusion-based 3D printed concrete, Addit. Manuf. 37 (2021) 101740, http://dx.doi.org/10.1016/j.addma.2020.101740.

[42] M. Sovetova, J.K. Calautit, Influence of printing parameters on the thermal properties of 3D-printed construction structures, Energy 305 (2024) 132265, http://dx.doi.org/10.1016/j.energy.2024.132265.

[43] M.M.R. Kiyani, S.M. Hussain, R. Emaan, M.B. Kamal, S. Ur Rehman, R.D. Riaz, M. Usman, Influence of process parameters on 3D concrete printing: A step towards standardized approaches, Case Stud. Constr. Mater. 23 (2025) e05212, http://dx.doi.org/10.1016/j.cscm.2025.e05212.

[44] O. Disu, S.O. Ismail, L. Wood, A. Chrysanthou, A. Kanellopoulos, Experimental study on buildability of 3D-printed cement-based structures using aluminium sulphate, Case Stud. Constr. Mater. 23 (2025) e05192, http://dx.doi.org/10.1016/j.cscm.2025.e05192.

[45] T. Diwan, G. Anirudh, J.V. Tembhurne, Object detection using YOLO: Challenges, architectural successors, datasets and applications, Multimedia Tools Appl. 82 (6) (2023) 9243–9275, http://dx.doi.org/10.1007/s11042-022-13644-y.

[46] P. Bharati, A. Pramanik, Deep Learning Techniques—R-CNN to Mask R-CNN: A Survey, in: A.K. Das, J. Nayak, B. Naik, S.K. Pati, D. Pelusi (Eds.), Computational Intelligence in Pattern Recognition, in: Advances in Intelligent Systems and Computing, vol. 999, Springer Singapore, Singapore, 2020, pp. 657–668, http://dx.doi.org/10.1007/978-981-13-9042-5_56.

[47] Z. Zhang, K. Guo, M. Bai, Y. Lyu, H. Xu, W. Guo, D. Li, Comparative Analysis of YOLOv8 and DeepLabv3+ on WildScenes Dataset: Evaluating mIoU Performance, in: Y. Tan, Y. Shi (Eds.), Data Mining and Big Data, in: Communications in Computer and Information Science, vol. 2356, Springer Nature Singapore, Singapore, 2025, pp. 289–301, http://dx.doi.org/10.1007/978-981-96-7175-5_24.

[48] Y. Wang, M. Chu, X. Kang, G. Liu, A deep learning approach combining DeepLabV3+ and improved YOLOv5 to detect dairy cow mastitis, Comput. Electron. Agric. 216 (2024) 108507, http://dx.doi.org/10.1016/j.compag.2023.108507.

[49] S.A. Da Silva, L.E.F. Berton, M.F.T. Carvalho, C.C.O. Bernardo, J.V.C.M. Perles, J.N. Zanoni, L. Nanni, A.L.R.G. Sevilha, G.Z. Felipe, F.C. Flores, Y.M.G. Costa, Nuclei Segmentation in Hepatocytes Using YOLO and SAM, in: 2024 31st International Conference on Systems, Signals and Image Processing, IWSSIP, IEEE, 2024, pp. 1–6, http://dx.doi.org/10.1109/IWSSIP62407.2024.10634020.

[50] Q. Zou, Y. Cao, Q. Li, Q. Wang, Deep learning-based detection and segmentation of pavement cracks under fixed-mounted camera conditions, Autom. Constr. 113 (2020) 103147, http://dx.doi.org/10.1016/j.autcon.2020.103147.