



Bridging the Simulation-to-Field Gap in AI-Based Underground Mine Inspection: A Systematic Review of Synthetic Data, Domain Adaptation, Transfer Learning, and Digital Twins

Fatima Umar Zambuk*, Mustapha Lawal Abdulrahman, Mahmood Abdulhameed, Abdulsalam Ya'u Gital, Saratu Habu Japhet, Lawal Rukuna Muhammed

Abstract: Artificial intelligence (AI) has advanced rapidly in underground and tunnel inspection, but field deployment continues to be limited by a persistent simulation-to-field reality gap. Real underground data are costly to acquire, difficult to annotate, safety-sensitive, highly site-specific, and often scarce for rare but consequential defects. Simulation, synthetic data, transfer learning, domain adaptation, and digital twins offer complementary ways to address this constraint, yet these research streams are usually reviewed in isolation. This systematic review treats field generalisation as the central analytical problem and synthesises the full pathway from virtual environment construction and synthetic-data generation through cross-domain adaptation, limited-data fine-tuning, digital-twin calibration, and field validation. We used a PRISMA-informed protocol to review literature published from 2015 to 20 August 2026, classifying evidence into three levels: direct underground-mine studies, near-domain tunnel/subterranean studies, and transferable sim-to-real methodology. The synthesis indicates that the most convincing recent results arise not from simulation fidelity alone, but from combining synthetic diversity with target-domain adaptation. Geometry-informed tunnel point-cloud simulation, for example, has shown strong synthetic-only performance and further gains after limited-real-data transfer learning. At the same time, domain-adaptive crack-segmentation studies report substantial performance recovery across materials, structures, and field sites. Mining digital twins are also moving beyond static visualisation toward calibrated virtual-real loops, including mine-tunnel

reconstruction, ventilation modelling, rescue-robot navigation, and field-tested teleoperation. Nevertheless, cross-mine validation, explicit measurement of the reality gap, rare-hazard simulation, uncertainty-aware adaptation, and standardised deployment-readiness criteria remain weak. To address these gaps, this review proposes an Underground Simulation-to-Field Gap Taxonomy and the MineSim2Field framework, an eight-stage closed-loop pathway connecting virtual environment construction, synthetic data, domain diversification, pretraining, adaptation, field calibration, deployment, and twin recalibration. The framework positions simulation as a continuously updated source of evidence rather than a one-off substitute for field data.

Keywords: Underground Mining, Sim-To-Real, Synthetic Data, Domain Adaptation, Transfer Learning, Digital Twin, Autonomous Inspection, Deep Learning, Field Generalisation, Tunnel Inspection.

I. INTRODUCTION

Underground mine inspection presents an unusually difficult setting for data-driven automation. The environment is confined, GPS-denied, weakly illuminated, dusty, geometrically irregular, and operationally dynamic; structural hazards can also make direct human data collection unsafe. At the same time, the events that matter most for safety—progressive cracking, rock instability, roof degradation, support-system failure, post-disaster obstruction, or abnormal microseismic signatures—are precisely those for which large, balanced, expertly labelled datasets are least likely to exist. The result is a structural mismatch between modern deep learning, which generally benefits from extensive and diverse training data, and the evidence available in underground operations. Recent tunnel and mining research confirms that data scarcity is no longer a secondary inconvenience but a principal deployment bottleneck. Contemporary reviews of tunnel defect recognition identify annotation cost, environmental complexity and limited cross-scenario robustness as persistent barriers, even where high in-dataset accuracy has been achieved [1], [2]. Mining digital-twin reviews likewise report growing adoption of sensor-rich virtual representations while emphasising interoperability, field validation and subsurface deployment as unresolved challenges [3], [4], [5]. The practical problem, then, is not simply how to construct a more accurate detector or a more realistic virtual mine. It is how knowledge developed in one distribution—synthetic, laboratory, civil tunnel, source mine, source sensor, or source geology—can be transferred safely to another.



Manuscript received on 21 August 2026 | First Revised Manuscript received on 28 August 2026 | Second Revised Manuscript received on 08 September 2026 | Manuscript Accepted on 15 September 2026 | Manuscript published on 30 September 2026.

*Correspondence Author(s)

Dr. Fatima Umar Zambuk*, Department of Computer Science, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. Email ID: fuzambuk@atbu.edu.ng, ORCID ID: [0000-0003-4276-4490](https://orcid.org/0000-0003-4276-4490)

Dr. Mustapha Lawal Abdulrahman, National Space Research and Development Agency (NASRDA), Abuja, Nigeria. Email ID: musbaida@gmail.com, ORCID ID: [0000-0002-1037-2022](https://orcid.org/0000-0002-1037-2022)

Dr. Mahmood Abdulhameed, Department of Mechatronics Engineering, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria, Email ID: mahmood.abdul@gmail.com, ORCID ID: [0009-0005-3495-4886](https://orcid.org/0009-0005-3495-4886)

Dr. (Prof.) Abdulsalam Ya'u Gital, Faculty of Engineering, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. Email ID: asgital@gmail.com, ORCID ID: [0000-0001-6495-3693](https://orcid.org/0000-0001-6495-3693)

Dr. Saratu Habu Japhet, National Space Research and Development Agency (NASRDA), Abuja, Nigeria. Email ID: japhetserahabu@gmail.com, ORCID ID: [0009-0004-3485-874X](https://orcid.org/0009-0004-3485-874X)

Dr. Lawal Rukuna Muhammed, Department of Computer Science, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. Email ID: lawal.rukuna@gmail.com, ORCID ID: [0009-0003-6172-6980](https://orcid.org/0009-0003-6172-6980)

The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open-access article under the CC-BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The broader sim-to-real literature describes this mismatch as the “reality gap”. Simulation is attractive because data generation is safer, cheaper, faster and more controllable than physical experimentation, yet every simulator remains an approximation of reality [6]. Domain randomisation addresses the gap by training across a distribution of simulated environments rather than a single nominal model [7]. In contrast, domain adaptation seeks representations or outputs that remain useful when source and target distributions differ [8]. Sim-to-real control studies have also demonstrated transfer without real training imagery when source variation is engineered appropriately [9]. Offline domain-randomisation methods, in turn, can use real trajectories to identify simulator-parameter distributions for transfer [10]. Transfer learning reduces target-domain data demand by reusing representations learned elsewhere [11]. Digital twins add a further capability: they can become calibrated, continuously updated virtual counterparts through which real observations influence the next round of simulation, testing and decision-making.

These techniques are now appearing increasingly in underground and analogous infrastructure inspection. Weng et al. [12] demonstrated multi-level unsupervised domain adaptation for cross-dataset crack detection, including a pavement-to-tunnel transfer scenario. Shim [13] combined site-specific generative inpainting with self-supervised domain adaptation and demonstrated the approach in a tunnel-ceiling field test. Yang et al. [14] generated automatically labelled tunnel point clouds using a geometry-informed simulator and then used limited real data to improve semantic segmentation through transfer learning. Within mining, cross-mine transfer has improved small-sample microseismic recognition [15], while new digital twins reconstruct mine-tunnel geometry, support ventilation safety analysis, and enable robotic navigation or teleoperation [16], [17], [18], [19].

A. Why a Separate Simulation-to-Field Review Is Needed

The novelty of this review does not lie in synthetic data, transfer learning, domain adaptation or digital twins as individual concepts; each already has an established literature. The unresolved problem is their fragmented treatment. Tunnel-defect reviews typically organise evidence by defect type, model architecture, dataset establishment and inspection hardware [1], [2]. Mining digital-twin reviews instead tend to organise the literature around operational functions such as exploration, drilling, haulage, ventilation or processing [3], [4]. Sim-to-real robotics reviews focus more heavily on control-policy transfer, simulator randomisation and robot dynamics [6]. None of these framings treats the underground field-generalisation pathway itself as the principal unit of analysis. This review connects four questions that are usually treated separately: how synthetic underground evidence is generated; how source–target discrepancy is represented and reduced; how pretrained knowledge is adapted when target labels are scarce; and how digital twins can close the loop between field observations and the next virtual training distribution. The resulting perspective is deliberately method-centric rather than architecture-centric.

A CNN, transformer, point-cloud network, GNN, reinforcement-learning policy or microseismic classifier is relevant insofar as the study provides evidence about transferability, data efficiency, target-domain adaptation, or virtual-real calibration.

B. Objectives and Contributions

The review addresses six research questions and makes four synthesis contributions. Table I lists the research questions. The main contributions are: (i) an Underground Simulation-to-Field Gap Taxonomy that decomposes the reality gap into photometric, atmospheric, geometric, material/geological, sensor/dynamic, and site/operational components; (ii) a comparative synthesis of synthetic-data, domain-adaptation and transfer-learning strategies with explicit attention to target-data requirements and field evidence; (iii) a focused interpretation of digital twins as sim-to-field infrastructure rather than as a general mining-digitalisation concept; and (iv) the MineSim2Field framework, an eight-stage closed-loop development pathway for progressing AI inspection systems from virtual development to validated underground deployment.

Table I: Research Questions Guiding the Review

RQ	Research Question
RQ1	What synthetic data and simulation environments are used for AI-based underground or analogous subterranean inspection?
RQ2	Which environmental, sensor, geometric and operational factors contribute most strongly to the simulation-to-field gap?
RQ3	Which domain adaptation and domain randomisation strategies reduce performance loss between source/synthetic and real underground domains?
RQ4	How effectively does transfer learning reduce dependence on labelled underground data, including cross-site, cross-sensor and cross-scale transfer?
RQ5	What role can digital twins play in data generation, calibration, virtual testing, adaptation and continuous field feedback?
RQ6	How should sim-to-field transfer be evaluated and standardised to demonstrate deployment readiness for safety-relevant underground inspection?

II. REVIEW METHODOLOGY

A. Review Design and Reporting Protocol

We adopted a PRISMA-informed systematic review design because the relevant evidence is distributed across computer vision, robotics, structural health monitoring, tunnelling, geotechnical engineering, and mining systems. We used PRISMA 2020 principles to structure identification, screening, eligibility, and reporting [20]. We designed the review as a systematic evidence synthesis rather than a quantitative meta-analysis because outcome metrics, datasets, source–target pairs, and validation environments are highly heterogeneous. Numerical results are therefore interpreted within the design of the original studies rather than pooled across tasks that are not directly comparable. The search window covered January 2015 to 20 August 2026. The starting date captures the period when deep transfer learning, domain-adversarial representation learning, modern generative models, and digital twins became increasingly relevant to engineering inspection.





We retained foundational works published before 2015 only when needed to define core concepts, particularly transfer learning and domain adaptation.

B. Information Sources and Search Strategy

The intended final database set comprises Scopus, Web of Science, IEEE Xplore, ScienceDirect, SpringerLink and the

ACM Digital Library. We also used publisher repositories and open scholarly indexes to verify citations and conduct backward/forward snowballing. We organised searches into four concept blocks: underground/subterranean environment; simulation and synthetic data; transfer/adaptation; and AI-based inspection. Although database syntax must be adapted to each index, the underlying Boolean structure is preserved.

Table II: Search-Term Architecture Used Across Databases

Concept block	Representative search terms
Environment	"Underground mine*" OR "underground mining" OR tunnel* OR subterranean OR "confined space*" OR "underground inspection"
Simulation / synthetic	"Synthetic data" OR simulation OR simulator OR "domain randomisation" OR "procedural generation" OR "virtual environment*" OR "digital twin"
Transfer/adaptation	"Sim-to-real" OR "simulation-to-real" OR "domain adaptation" OR "transfer learning" OR "domain generalisation" OR "few-shot" OR "self-supervised"
AI inspection	("deep learning" OR "machine learning" OR "computer vision" OR AI OR UAV OR drone OR robot*) AND (inspection OR detection OR segmentation OR navigation OR mapping OR crack* OR defect* OR anomaly)

C. Eligibility Criteria and Evidence Levels

Studies were included when they met at least one of the following conditions: (a) generated synthetic data or simulated environments for tunnel, underground, rock, structural-defect or autonomous-inspection tasks; (b) evaluated domain adaptation, domain generalisation, sim-to-real transfer, cross-site transfer, cross-sensor transfer, few-shot learning or target-domain fine-tuning in an inspection-relevant task; (c) used a digital twin to support virtual-real mapping, testing, calibration or operational adaptation in underground mining or tunnel infrastructure; or (d) provided a transferable sim-to-real method with clear relevance to underground robotic inspection. We excluded studies that were purely conceptual without an empirical method or validation, addressed only surface mining with no transferable underground methodology, or used "digital twin" solely as a visualisation label without meaningful virtual-real coupling.

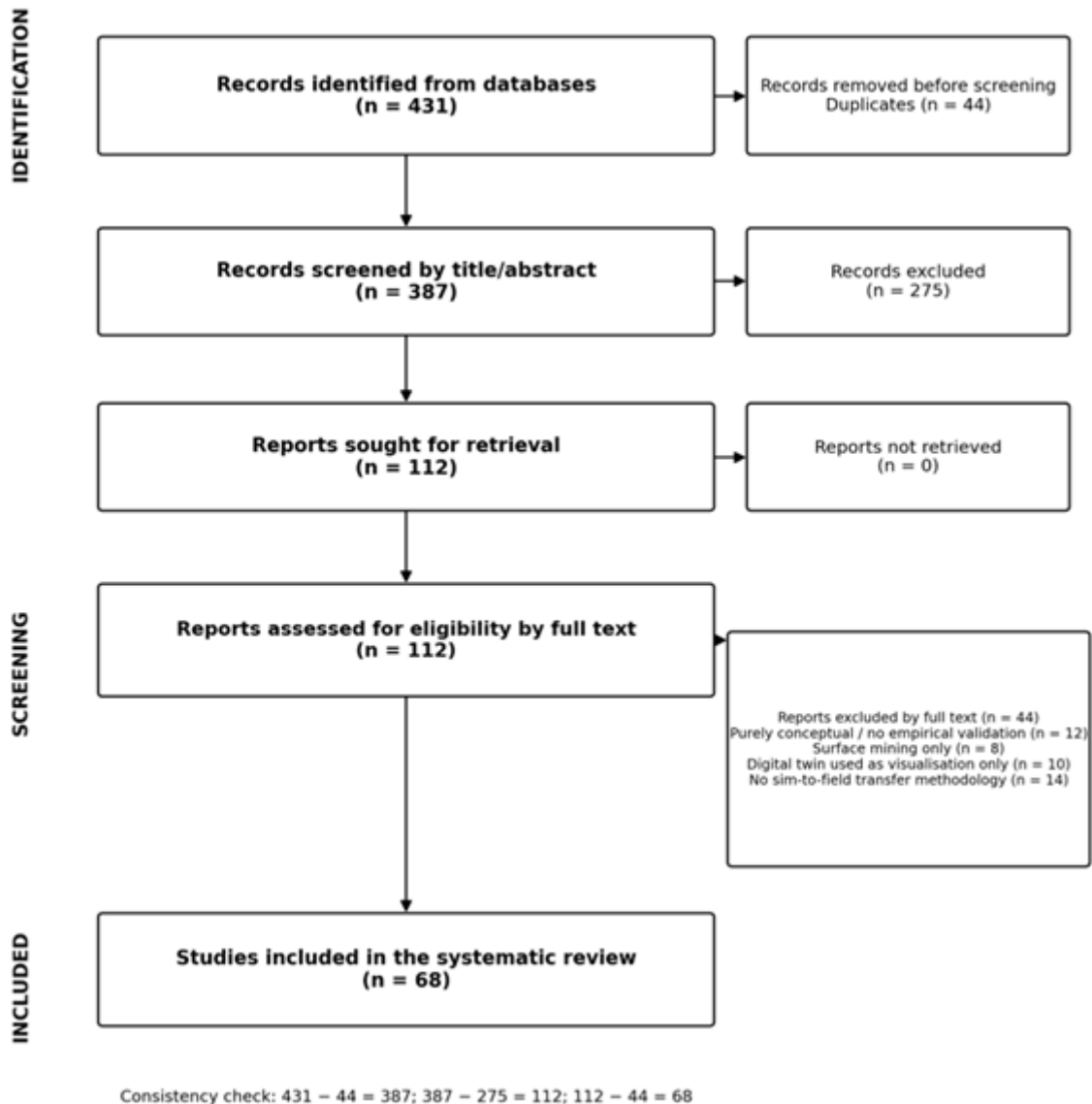
To avoid overstating external validity, we classified the evidence into three levels. Level A (direct evidence) comprises studies conducted in underground mines or mine-specific monitoring systems. Level B (near-domain evidence) comprises civil tunnels, metro tunnels, rock tunnels, underground caverns and closely analogous confined infrastructure. Level C (method-transfer evidence) comprises general robotics or infrastructure inspection studies included because their methods directly address the sim-to-real problem but were not validated underground. Conclusions about mining deployment rely primarily on Level A evidence, while Levels B and C identify transferable mechanisms and research opportunities.

D. Data Extraction and Quality Appraisal

For each eligible study, the extraction framework recorded the application domain, sensing modality, source domain, target domain, synthetic-data mechanism, adaptation strategy, target-label requirement, model family, evaluation metric, field-validation level, and reported limitations. Quality appraisal centred on five questions: Was the source–target shift explicitly described? Was target-domain performance reported separately from source-domain performance? Were real data kept independent from synthetic training data where appropriate? Did the authors validate in the field beyond the development site? Did the paper discuss uncertainty, failure cases, or negative transfer? Given the literature's heterogeneity, we treated these criteria as interpretive quality signals rather than combining them into a single additive score.

E. Screening Results and Reproducibility

Database searches identified 431 records. After removing 44 duplicates, 387 records underwent title/abstract screening, of which 275 were excluded. We sought the remaining 112 reports for retrieval; we retrieved all 112 and assessed them in full text. Forty-four full-text reports were excluded: 12 were purely conceptual with no empirical validation, 8 addressed surface mining only, 10 used digital twins only as visualisation tools, and 14 did not provide a simulation-to-field transfer methodology. The final systematic review therefore included 68 studies. We preserved screening logs and database export files to support reproducibility.

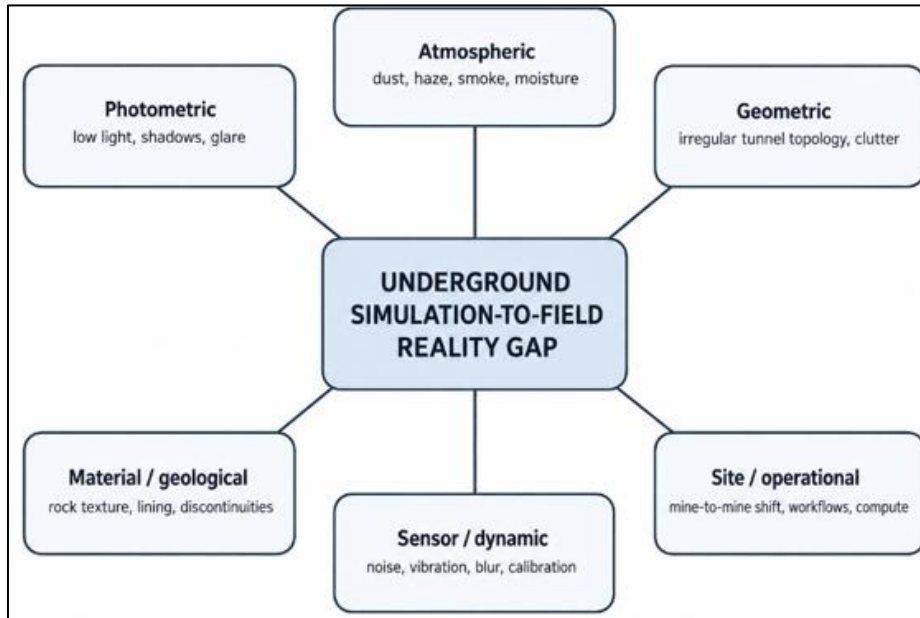


[Fig.1: PRISMA 2020 Flow Diagram for the Review. Records Identified: n = 431; Duplicates Removed: n = 44; Records Screened: n = 387; Title/Abstract Exclusions: n = 275; Reports Sought and Retrieved: n = 112; Full Texts Assessed: n = 112; Full-Text Exclusions: n = 44; Studies Included: n = 68]

III. THE UNDERGROUND SIMULATION-TO-FIELD REALITY GAP

A simulator does not need to reproduce every aspect of an underground mine to be useful. It does, however, need to represent enough task-relevant uncertainty to prevent a model from learning shortcuts that are unique to the virtual domain. The central analytical distinction is therefore between fidelity and coverage. High visual fidelity can still produce poor transfer when the simulator represents only a narrow range of lighting, rock texture, dust density, camera exposure, tunnel geometry or robot motion. Conversely, deliberate randomisation can yield samples that look less realistic but produce more robust field behaviour because the learner must attend to invariant structure [7], [6].

Underground inspection makes this problem more acute because the field domain is itself non-stationary. Lighting changes with vehicle orientation and auxiliary lamps; airborne dust varies with activity; wet surfaces create specular reflections; camera optics become contaminated; LiDAR return quality changes with particulates and incidence angle; tunnel surfaces differ by geology, support type and shotcrete condition; and robot motion introduces blur, vibration and occlusion. A model adapted to one mine may therefore fail in another even when both are “real” underground environments. Sim-to-real transfer is better understood as a broader source-to-target generalisation problem encompassing synthetic-to-real, laboratory-to-field, tunnel-to-mine, mine-to-mine and sensor-to-sensor shifts.



[Fig.2: Underground Simulation-to-Field Gap Taxonomy Proposed in this Review]

Table III: Proposed Taxonomy of Underground Reality-Gap Components and Corresponding Mitigation Strategies

Gap Class	Typical Manifestations	Sim-to-field Response
Photometric	Illumination intensity, shadows, glare, camera exposure, colour response	Randomise light sources/exposure; use low-light augmentation; adapt image style.
Atmospheric	Dust, haze, smoke, moisture droplets, suspended particulates	Particle/haze simulation; dehazing variability; sensor-noise models.
Geometric	Irregular tunnel cross-sections, clutter, support structures, occlusion	Procedural geometry; point-cloud simulation; topology variation.
Material / geological	Rock texture, shotcrete, steel support, corrosion, discontinuities	Texture randomisation; rock-surface datasets; cross-material adaptation.
Sensor/dynamic	Motion blur, vibration, calibration drift, LiDAR/GPR noise, platform dynamics	Dynamics randomisation; blur/noise injection; sensor calibration models.
Site / operational	Mine-to-mine practices, equipment, communications, compute and work cycles	Cross-site testing; few-shot target adaptation; active data collection.

A. Measuring the Gap

A recurring weakness in current inspection research is that the “domain gap” is often inferred from a drop in accuracy rather than measured as a property of the data distributions. As a result, it can be difficult to determine why a method transfers. Feature-space distances, embedding overlap, calibration error, target uncertainty, Fréchet-type image statistics, point-cloud geometric descriptors and task-specific error decomposition could provide more informative diagnostics. Zhuge et al. [21], in industrial inspection, explicitly use embedding analysis to identify insufficiently represented synthetic scenarios and drive targeted virtual-data acquisition. The same active-loop logic is directly relevant to underground inspection: field failures should be treated as evidence of missing simulator variation, not merely as cases for manual relabelling.

IV. SYNTHETIC DATA FOR UNDERGROUND AND TUNNEL INSPECTION

A. From Augmentation to Generative Simulation

Synthetic-data approaches in inspection range from simple perturbations of real images to fully procedural three-dimensional environments. Classical geometric and photometric augmentation remains useful, but it cannot create

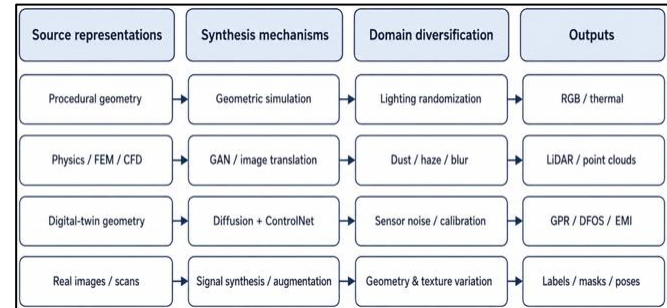
new defect morphology or entirely new spatial contexts. Generative adversarial networks (GANs) expanded the design space by synthesising defect images, while diffusion models now offer stronger control over appearance, texture and structural conditioning. In parallel, geometry-based simulators can generate automatically labelled point clouds, poses and sensor observations without manual annotation. For sim-to-field transfer, the key distinction is whether synthesis changes only superficial appearance or also expands the structural, sensor and operational support of the training distribution.

Tunnel research provides several concrete examples. Zhou et al. [22] used a Wasserstein GAN variant to enrich tunnel-lining crack data while pruning YOLOv4 for practical detection; the resulting model was reduced to 49.16 MB and ran at about 40.58 frames per second, with minimal loss in F1 and AP relative to the unpruned detector. Shim [23] generated labelled crack imagery using a conditional GAN and then applied self-training, reporting mean IoU of 80.34% and F1 of 76.31%. Yan et al. [24] combined GAN-assisted tunnel-defect augmentation with a multiscale attention CNN and attributed an 8.89% accuracy gain to the augmentation component in ablation analysis. These studies show

that synthetic enrichment can improve data sufficiency, but they do not establish cross-site transfer unless evaluation includes a genuinely new target domain. Han et al. [25] extended this evidence directly to underground-mine photometric degradation through UM-GAN, a GAN-based low-light enhancement framework evaluated on real and synthetic underground-mine datasets to restore detail, reduce noise and improve image quality under low illumination.

A newer generation of controlled synthesis is particularly relevant to underground mining because it can represent unusual rock textures and rare crack geometries. CrackGen combines Stable Diffusion, ControlNet, and procedurally generated crack sketches to produce image-label pairs without manual mask creation and shows that models trained on purely synthetic data can generalise to real crack-detection tasks, including cross-domain testing on rock surfaces [26]. Its main methodological advantage is controllability: crack width, orientation, background material and complexity can be varied deliberately rather than inherited passively from a small source dataset.

Synthetic sensing is not confined to RGB imagery. Liao et al. [27] used a conditional GAN to augment distributed fibre-optic strain measurements for tunnel-segment crack severity assessment. GAN-generated electromechanical impedance signals have likewise been used to strengthen tunnel-lining damage diagnosis under limited data. These signal-domain approaches matter in mining because many safety-relevant measurements—microseismic, vibration, acoustic emission, gas, strain or GPR—are not primarily visual. A comprehensive synthetic-data strategy should therefore be modality-aware rather than treating “simulation” as synonymous with image rendering.

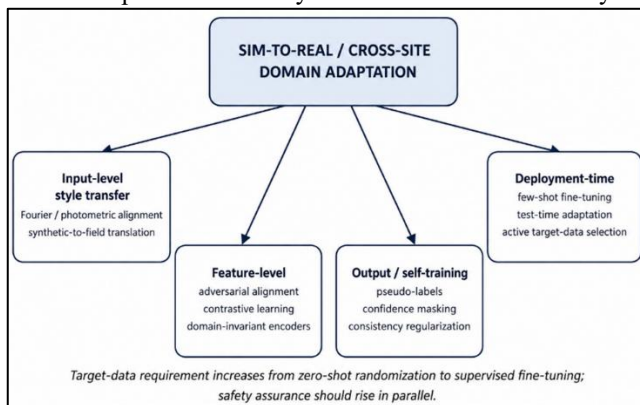


[Fig.3: Synthetic-Data Pipeline from Source Representation and Generation Mechanism to Domain Diversification and Labelled Outputs]

Table IV: Representative Synthetic-Data Studies Relevant to Underground Inspection

Study	Level	Modality / Task	Synthetic Mechanism	Key sim-to-field Evidence
Zhou et al. [22]	B	Tunnel-lining RGB cracks	WGAN-based crack augmentation	Expanded limited crack dataset; paired with pruned YOLOv4; 49.16 MB and 40.58 FPS after pruning.
Shim [23]	C/B	Infrastructure cracks	Conditional GAN + self-training	Synthetic crack images used without new manual labels; mean IoU 80.34%, F1 76.31%.
Xie et al. [26]	C	Civil/rock crack imagery	Stable Diffusion + ControlNet + procedural crack sketches	Automatically paired images/masks; synthetic-only training demonstrated real and rock-surface transfer.
Liao et al. [27]	B	Shield-tunnel DFOS strain	Conditional GAN signal synthesis	Augmented scarce strain observations for crack-severity diagnosis.
Yan et al. [24]	B	Tunnel-lining defects	Deep-convolutional GAN augmentation	Ablation attributed 8.89% accuracy gain to GAN augmentation.
Yang et al. [14]	B	Tunnel 3D point clouds	High-fidelity geometry-informed simulator	Synthetic-only 71.9% mIoU/86.6% OA; limited-real-data transfer raised results to $\geq 78.8\%$ mIoU/90.9% OA.
Han et al. [25]	A	Underground-mine low-light RGB imagery	UM-GAN low-light enhancement with diffusion denoising	Evaluated on real and synthetic underground-mine datasets; directly targets low illumination, noise and detail loss.

The adaptation taxonomy used in the cross-domain synthesis that follows is summarised in Fig. 4.



[Fig.4: Domain Adaptation Taxonomy for Underground and Infrastructure Inspection, Organised by Where Adaptation Acts and How Much Target Data it Requires]

B. Geometry-Informed Point-Cloud Simulation

The 2026 Tunnel Scanner study is particularly instructive because it combines synthetic geometry with explicit transfer learning. Yang et al. [14] built a simulator for segmental tunnel-lining point clouds with automatic annotation and then embedded geometric priors into deep-learning backbones. Models trained only on synthetic data achieved 71.9% mIoU and 86.6% overall accuracy. With limited real data and transfer learning, performance increased to at least 78.8% mIoU and 90.9% overall accuracy. The study closely approaches the ideal evidence design for sim-to-field research because it reports a synthetic-only baseline, a real-target adaptation stage and a clearly defined sensing modality. Its main limitation for mining is contextual: the target is civil





tunnel segmentation rather than irregular mine openings, although the underlying geometry-to-point-cloud pipeline is directly transferable.

C. Synthetic Diversity Versus Photorealism

The evidence suggests judging synthetic-data quality by downstream transfer rather than visual realism alone. For underground inspection, a photorealistic render that omits dust variation, wet rock, dynamic occlusion, or sensor calibration drift may be less useful than a more diverse but less visually convincing simulation. A better development criterion is coverage of plausible nuisance factors while preserving safety-relevant features. Generative methods should also be screened for physically implausible defects, label leakage, texture shortcuts and duplication. Where possible, domain experts should validate synthetic defect morphology and sensor behaviour before treating synthetic data as equivalent to field observations.

V. DIGITAL TWINS AS SIMULATION-TO-FIELD INFRASTRUCTURE

A. Beyond Static 3D Visualisation

The term digital twin is used inconsistently across engineering. One useful distinction is between a manually updated digital model, a digital shadow with one-way data flow from the physical asset, and a digital twin with meaningful bidirectional or feedback coupling [28]. In this review, a digital twin is relevant only when it contributes to the sim-to-field pathway—for example, by calibrating geometry or physics from field observations, generating virtual test scenarios, synchronising sensor state, supporting robot planning, or feeding field residuals back into subsequent simulation.

Mining-specific reviews confirm rapid growth in digital-twin applications, but they also show that subsurface integration and industrial validation remain comparatively immature [3], [4]. Liang et al. [5] identify reconstructing and maintaining virtual mining environments as a persistent challenge. This is especially important for inspection because underground geometry changes with excavation, support installation, degradation, ground falls, and equipment movement. A useful twin must therefore remain updatable rather than a one-time, high-resolution scan.

B. Mine-Tunnel Reconstruction and Calibration

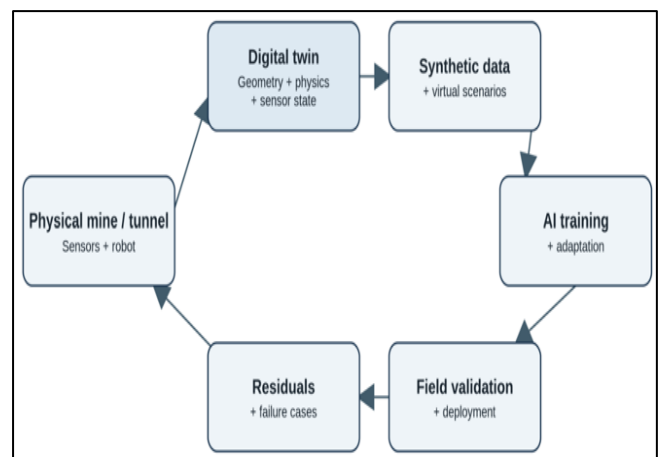
Recent work is beginning to address this requirement. Zhu et al. [17] fused images, navigation information and point clouds to construct photorealistic digital-twin visualisations of coal-mine tunnels and validated the approach in simulated and real mine environments. Cui et al. [18] proposed multiscale geometric digital twins of mine tunnels, combining rapid topology generation from drawings with high-precision local reconstruction from laser point clouds. For civil tunnels, Machado et al. [29] developed automated procedural modelling for inspection and digital-twin development, while Sun et al. [30] integrated multi-source geological observations in a data-model-decision loop for rock-tunnel construction. Together, these studies indicate a shift from static representation toward updateable geometry and uncertainty-aware models. Cacciuttolo et al. [31]

further proposed an advanced multi-layer digital-twin framework for underground mining that integrates the physical mine, data acquisition and integration, virtual models, intelligent decision support and user feedback; the framework was validated through an expert-panel questionnaire.

C. Digital Twins for Mine Safety and Robotic Operation

Mine safety provides several Level A examples of calibrated or operational twins. Jacobs et al. [16] used a calibrated digital twin to identify and mitigate hazards in a deep-level mine ventilation system. Cao et al. [32] explicitly analysed parameter uncertainty in a mine-ventilation digital twin, reinforcing the point that a twin should not be treated as ground truth simply because it is synchronised with sensors. Aksoy et al. [33] applied a CFD-based underground-mine twin to interpret airflow, gas and thermal behaviour during and after mine fire scenarios. These are not inspection models in the narrow computer-vision sense, but they demonstrate the virtual-real calibration logic required for safety-critical underground simulation.

Robotic studies connect digital twins more directly with inspection and intervention. Le et al. [34] reported the design, integration and field testing of a digital-twin-based teleoperated rock-scaling robot, using a live virtual representation and remote interface to reduce worker exposure. You et al. [19] integrated a digital twin with a coal-mine rescue robot for point-cloud mapping, autonomous navigation and environmental monitoring in a simulated post-disaster mine. A logical next step is to connect such robotic twins to learned perception models that are continuously adapted using field residuals and targeted synthetic scenarios.



[Fig.5: Digital Twin as a Closed-Loop Sim-to-Field Learning Infrastructure Rather Than a Static Visualisation Model]

Table V: Representative Digital-Twin Studies Contributing to Virtual-Real Calibration and Underground Deployment

Study	Level	Application	Twin Mechanism	Relevance to Sim-to-field
Jacobs et al. [16]	A	Deep-level mine ventilation	Calibrated thermo-hydraulic twin	Hazard identification and mitigation in operating ventilation network.
Cao et al. [32]	A	Mine ventilation	Uncertainty-aware digital twin	Quantified sensitivity of airflow predictions to ventilation-resistance uncertainty.
Zhu et al. [17]	A	Coal-mine tunnel geometry	Image + navigation + point-cloud fusion	Validated digital-twin reconstruction in simulated and real coal-mine environments.
Le et al. [34]	A	Rock scaling robot	Live robot/environment twin + teleoperation	Field-tested robotic intervention for hazardous scaling tasks.
Aksoy et al. [33]	A	Underground mine fire/ventilation	CFD-calibrated mine twin	Reconstructed airflow, gas and thermal effects for emergency interpretation.
Cui et al. [18]	A	Mine-tunnel network	Multiscale topology + local point-cloud reconstruction	Continuously updateable geometric digital-twin framework.
You et al. [19]	A	Coal-mine rescue robot	Robot/environment twin, Gazebo/ROS, point clouds	Autonomous navigation and environmental visualisation in simulated post-disaster mine.
Machado et al. [29]	B	Tunnel inspection	Automated procedural digital-twin modelling	Reduces manual model construction; supports defect-informed tunnel asset models.
Cacciuttolo et al. [31]	A	Underground-mine digital-twin architecture	Advanced multi-layer physical-data-virtual-decision framework	Expert-panel-validated framework connecting sensing, simulation, monitoring and user feedback for underground mining.

VI. DOMAIN ADAPTATION AND DOMAIN GENERALISATION

A. Why Domain Adaptation is Central to Underground Inspection

Synthetic data can increase sample volume, but volume alone does not guarantee transfer. Domain adaptation directly addresses the distribution mismatch between source and target. In underground inspection, the source may be synthetic imagery, civil-tunnel data, a different mine, a different camera, or a different rock type. At the same time, the target may contain few labels or none at all. Unsupervised and semi-supervised domain adaptation are therefore particularly attractive because exhaustive target annotation would defeat much of the purpose of simulation-based development.

Domain-adversarial training provides one foundational mechanism: representations are optimised to remain discriminative for the task while becoming less informative about whether a sample originated from the source or target domain [8]. Cycle-consistent adversarial approaches can also align appearance and representation across domains [35]. Other strategies align image style, latent features, output distributions or pseudo-labels. For inspection, the appropriate method depends on the type of shift. Lighting and texture differences may respond to input-level adaptation; sensor or material shifts may require feature alignment; rare defect morphology may require targeted synthesis; and changing mine conditions may require continual or test-time adaptation.

B. Empirical Evidence from Crack and Tunnel Inspection

Weng et al. [12] provide one of the clearest cross-structure demonstrations. Their DACrack framework adapts at the input, feature, and output levels and reported F1 improvements of 11.56% for asphalt-to-asphalt transfer, 30.54% for concrete-to-asphalt transfer, and 24.35% for pavement-to-tunnel transfer. For underground mining, the important point is not the exact numerical gain but the evidence that cross-structure transfer can recover substantial performance without a fully labelled target dataset.

Liu and Chen (2023) similarly addressed UAV-based crack recognition from curated databases to real-site inspections by combining self-supervised depth information, edge cues and domain-gap-aware optimisation. Beyene et al. [36] used a transformer-based unsupervised domain adaptation framework for crack segmentation under scarce target labels. More recent work has moved toward safer self-training. Bhattacharya et al. [37] combined interpretable Fourier-morphology blending with uncertainty-guided dual-network self-training and reported F1 improvements of up to 18.5% over competing baselines in difficult cross-domain scenarios. Tanveer and Cho (2026) used cross-domain CutMix and consistency regularisation, achieving up to 18.22 percentage-point mIoU improvement over source-only models and a 12.97-point field gain on concrete walls.

For tunnel-relevant field transfer, Shim [13] is particularly important. Site-specific crack images and labels were generated by fine-tuning Stable Diffusion with Dream Booth and were then incorporated into a self-supervised adaptation pipeline. The method reported average F1 improvements of



25.82% and 17.83% across two target datasets and was tested on a tunnel ceiling. Maharjan et al. [38] addressed corrosion in steel tunnels using a SegFormer-based, domain-adaptive self-supervised pipeline with pseudo-labels and confidence

masking; reported mIoU and F1 were 0.81 and 0.77, respectively, with subsequent 3D/BIM localisation. Taken together, these studies show a shift from generic style matching toward site-specific, confidence-aware learning.

Table VI: Representative Domain-Adaptation Evidence Relevant to Underground Deployment

Study	Level	Shift	Adaptation Mechanism	Target Labels	Reported Evidence
Weng et al. [12]	B/C	Cross-material and pavement→tunnel	Input + adversarial feature + output/VAE alignment	No target labels	F1 gains 11.56–30.54%; pavement→tunnel +24.35%.
Liu & Chen [39]	C	Database→real UAV inspection	Self-supervised depth, edges, consistency	Unlabelled target	Real-site UAV crack recognition.
Beyene et al. [36]	C	Cross-dataset crack segmentation	Transformer UDA + consistency	Unlabelled target	Demonstrated domain-adaptive crack segmentation.
Shim [13]	B	Source data→site-specific tunnel/field	DreamBooth inpainting + self-supervised adaptation	Small target sample	Average F1 gains 25.82% and 17.83%; tunnel-ceiling field test.
Bhattacharya et al. [37]	C	Public/industrial cross-domain	Fourier–morphology blending + uncertainty self-training	Unlabelled target	F1 improvement up to 18.5%.
Maharjan et al. [38]	B	Variable-light steel tunnel domains	Pseudo-labels + confidence masking + SSL	Limited/unlabelled target	mIoU 0.81, F1 0.77 plus 3D localisation.
Tanveer & Cho [40]	C	Five datasets/field wall	Cross-domain CutMix + consistency	Unlabelled target	Up to +18.22 mIoU over source-only; +12.97 in field test.

VII. TRANSFER LEARNING AND DATA-EFFICIENT FIELD ADAPTATION

A. Transfer Learning as a Label-Efficiency Mechanism

Transfer learning differs from unsupervised domain adaptation in that it explicitly reuses a pretrained representation and then partially or fully fine-tunes it for a target task. It is attractive in underground applications because the target dataset may be small without being completely unlabelled. The central question is no longer whether transfer learning improves over random initialisation—that is well established—but how much target data is required, which source domains transfer effectively, and whether performance remains stable across new sites, sensors and geology.

Liu et al. [41] compared transfer-learning-based encoder–decoder architectures for crack segmentation using a 7,089-image infrastructure dataset and showed that transfer learning outperformed training from scratch while allowing practical trade-offs among accuracy, model complexity and memory. Bao et al. [42] used transfer learning within an improved YOLOv8 framework for GPR-based tunnel void detection; in a field case, recognition confidence increased from 0.37 to 0.73. These examples show that target fine-tuning can improve practical performance. However, they also illustrate an important methodological distinction: transfer within one infrastructure type is easier than transfer from synthetic to mine-specific conditions.

B. Cross-Scale and Cross-Mine Transfer in Mining

Mining provides emerging evidence that transfer learning can overcome sensor-layout and site-data constraints. Xie and Li (2024) pretrained a deep model on more than 2.4 million labelled field-scale seismic and microseismic waveforms. They fine-tuned it using more than 300,000 laboratory acoustic-emission waveforms across multiple experiments, achieving over 86% F-score on unseen waveforms at both lab and field scales. Although the transfer direction differs from the common synthetic-to-field framing, the study shows that physically related source domains can support robust cross-scale fracture-mechanism classification.

Zhao et al. [15] directly addressed small-sample mine microseismic identification using a transfer-learning-enhanced BiLSTM-DCNN. The model was pretrained on larger cross-mine data and fine-tuned on a small target-mine dataset, achieving 94.44% test accuracy and a reported 80.85% relative improvement over a non-transfer baseline. This provides strong Level A evidence for cross-mine knowledge reuse, although the signal-classification task differs from image or point-cloud inspection. Together with Yang et al. [14], the result supports a broader principle: organise underground AI development around reusable source representations and explicit target calibration rather than isolated, site-specific training.

Table VII: Representative Transfer-Learning Studies with Explicit Relevance to Target-Domain Data Scarcity

Study	Level	Source→Target	Transfer Strategy	Key Evidence
Liu et al. [41]	B/C	General infrastructure→crack datasets	Pretrained CNN encoders	Transfer learning improved segmentation over training from scratch; 7,089-image benchmark.
Bao et al. [42]	B	Global/local tunnel GPR	Transfer learning + improved YOLOv8	Field confidence increased from 0.37 to 0.73; mAP gains reported for global/local datasets.
Xie & Li [43]	A	Field seismic↔lab acoustic emission	Multi-scale pretrained waveform model	>86% F-score on unseen lab and field waveforms.
Zhao et al. [15]	A	Cross-mine microseismic→small target mine	Pretraining + domain-adaptive fine-tuning	94.44% accuracy; 80.85% relative gain over non-transfer baseline.
Yang et al. [14]	B	Synthetic tunnel point clouds→real scans	Synthetic pretraining + limited-real fine-tuning	mIoU 71.9% synthetic-only to ≥78.8% after transfer; OA 86.6%→≥90.9%.

VIII. WHAT THE EVIDENCE SAYS ABOUT SIMULATION-TO-FIELD PERFORMANCE

A. Synthetic-Only Models Can Work, but Adaptation Is More Reliable

The strongest evidence in this review does not support a simple choice between “synthetic” and “real” data. Rather, it supports staged combinations of the two. CrackGen demonstrates that synthetic-only visual training can reach competitive real-world performance when the generator is controllable, and the target task is structurally constrained [26]. Tunnel Scanner shows that synthetic-only point-cloud training is already useful, but improves further when limited real data are incorporated [14]. Domain-adaptation studies likewise show that a source-only model can lose substantial performance after material, structural or site shift. That adaptation can recover a meaningful portion of that loss [12], [13], [37], [40].

These findings suggest a practical hierarchy for underground deployment. Synthetic-only evidence is appropriate for early algorithm development and stress testing; synthetic plus domain randomisation is appropriate for robustness pretraining; synthetic plus unlabelled target adaptation is appropriate when field access exists, but annotation is expensive; limited labelled field fine-tuning is appropriate before operational deployment; and cross-site external validation is necessary before claims of general applicability. Excellent results from random train/test splits within a single site should therefore not be interpreted as evidence of field generalisation.

B. Cross-Site Validation Is the Missing Benchmark

Across the reviewed literature, field validation often means testing on additional images from the same facility, a controlled tunnel, or a simulated mine. Genuine cross-site validation—training or adapting on one site and evaluating on a different mine without leakage—remains uncommon. This is especially consequential for resource-constrained mining contexts, where a transferable inspection system must justify its development cost by operating across multiple sites rather than requiring a full data-collection campaign for every deployment. The scarcity of shared underground benchmark datasets further compounds the problem.

C. Negative Transfer and Safety Assurance

Transfer can fail when source features are misleading, a problem known as negative transfer. In underground inspection, this can occur when concrete-tunnel crack morphology is assumed to represent natural rock discontinuities, when synthetic dust lacks real scattering behaviour, when camera adaptation alters crack contrast, or when a model pretrained on one sensor learns device-specific artefacts. Most current studies optimise mean performance but do not report worst-case degradation or abstention behaviour. A safety-oriented sim-to-field protocol should therefore include source-only, target-only (where feasible), adapted, and cross-site baselines; uncertainty calibration; rare-condition stress tests; and explicit reporting of failure cases.

IX. DIGITAL-TWIN-ENABLED CONTINUOUS ADAPTATION

A conventional sim-to-real pipeline is linear: build a simulator, train a model, deploy the model. That assumption is difficult to sustain underground because geometry, equipment, surface condition and sensor state change over time. Digital twins enable a cyclical alternative. Field residuals can update the twin; the updated twin can generate targeted scenarios; the model can then be adapted; and the next deployment can be monitored for additional discrepancy. Simulation thus becomes progressively calibrated against field evidence rather than remaining frozen before deployment.

Recent industrial inspection research supports this concept. Zhuge et al. [21] combine a digital twin with embedding analysis to identify synthetic-to-real mismatch and drive active acquisition of useful simulation data. Underground-mine twins already provide the geometric and state-estimation substrate required for analogous loops [17], [18], [19]. What remains missing is a peer-reviewed end-to-end mine study in which observed AI perception failures automatically alter the virtual scenario distribution and improved field performance is then demonstrated after retraining. This remains a clear research opportunity.

X. RESEARCH GAPS

Across the 68 studies included in the final review, we identified six recurring research gaps.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.





A. Absence of Shared Sim-to-Field Benchmarks

The field lacks a benchmark that pairs synthetic underground environments with independently collected real mine data and defines fixed source–target splits. Without such a benchmark, methods are difficult to compare because individual studies vary both the model and the domain. A useful benchmark should include RGB, depth/LiDAR, pose, and at least one non-visual safety modality; cover multiple sites; and define explicit “unseen mine” test partitions.

B. Weak Measurement of Simulator Adequacy

Most studies judge simulator quality indirectly through task accuracy. Few quantify which dimensions of the field domain are missing. Future studies should report embedding distances, distributional statistics, calibration error, sensor residuals, and task-specific transfer curves as simulator parameters vary. Such reporting would help distinguish a simulator that is “more realistic” from one that is demonstrably “more transferable”.

C. Rare Hazards Remain Under-Simulated

Common defects are easier to collect than precursor states immediately before structural failure. Synthetic data are uniquely valuable for rare hazards, but generative realism must remain constrained by geotechnical plausibility. Physics-informed fracture models, procedural discontinuity networks and expert-reviewed scenario libraries should be integrated with data-driven generation rather than relying solely on visual generative models.

D. Target Adaptation is Rarely Continual

Most adaptation is performed only once. Mines change over time, so deployed models require drift detection and safe update procedures. Continual learning must avoid catastrophic forgetting and should retain a validated baseline against which performance regression after adaptation can be detected.

E. Digital Twins and AI Adaptation Remain Disconnected

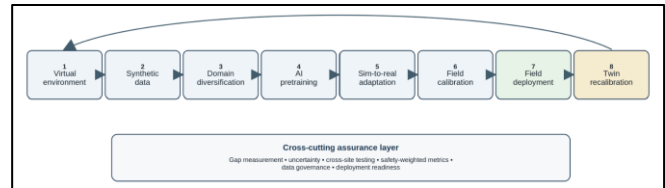
Mining twins increasingly support geometry, ventilation and robotics, while perception studies increasingly use domain adaptation; few studies connect the two. The twin should become both a source of targeted synthetic data and a repository for field residuals, rather than merely a visual interface around an AI model.

F. Deployment Readiness Is Poorly Defined

Accuracy thresholds alone do not establish readiness for life-safety deployment. Readiness should also include cross-site external validation, calibrated uncertainty, latency, power/compute compatibility, sensor-failure tolerance, human override, auditability, and evidence that adaptation does not create unacceptable new failure modes.

G. The Minesim2field Framework

To translate the review findings into a practical research programme, this paper proposes MineSim2Field, an eight-stage closed-loop framework. It treats simulation as a continuously calibrated source of evidence and requires progressively stronger real-world evidence as a system approaches deployment. The stages are shown in Fig. 6 and summarised below.



[Fig.6: Mine Sim2Field Framework for Progressing AI-Based Underground Inspection from Virtual Development to Validated Field Deployment and Continuous Recalibration]

H. Stage 2 - Synthetic Data Generation

Generate labelled RGB, thermal, point-cloud, pose, signal or multimodal data. Preserve provenance so that each synthetic sample can be traced to the simulator parameters and scenario assumptions that produced it.

I. Stage 3 - Domain Diversification

Randomise illumination, particulate density, surface material, camera exposure, sensor noise, geometry, robot dynamics and defect morphology—Prioritise coverage of plausible field conditions rather than photorealism alone.

J. Stage 4 - AI Pretraining

Train perception, navigation or hazard models in simulation and evaluate them on held-out simulated conditions. Use uncertainty and representation analysis to identify simulator overfitting.

K. Stage 5 - Sim-to-Real Adaptation

Apply unsupervised, semi-supervised or supervised adaptation using a small field sample. Compare source-only, adapted and target-only baselines where feasible.

L. Stage 6 - Limited Field Calibration

Collect controlled field evidence, quantify residual gaps, calibrate sensors/twin parameters and test failure conditions within restricted operational envelopes.

M. Stage 7 - Field Deployment

Deploy only within validated conditions, with uncertainty thresholds, human override and logging of model–field disagreement.

N. Stage 8 - Twin Recalibration and Continuous Learning

Feed field residuals, new geometry, new sensor states and difficult cases back into the digital twin. Generate targeted scenarios and repeat adaptation under version-controlled validation.

O. Cross-Cutting Assurance Layer

MineSim2Field adds a cross-cutting assurance layer spanning all eight stages: gap measurement, uncertainty quantification, cross-site testing, safety-weighted metrics, data governance and deployment-readiness assessment. This layer is necessary because average performance can improve while deteriorating for rare hazards or unseen sites. Each adaptation cycle should therefore be versioned and tested against a fixed regression suite that includes previously validated domains.

P. Stage 1 - Virtual Environment Construction

Construct geometry, material, geotechnical and robot/sensor representations at a fidelity appropriate to the task. Include a parameterised uncertainty model rather than relying on a single nominal environment.

XI. RESEARCH AGENDA FOR RESOURCE-CONSTRAINED MINING ENVIRONMENTS

The simulation-to-field approach is particularly relevant to developing mining economies, where the cost of collecting large underground datasets, maintaining specialist annotation teams and deploying high-end instrumentation can be prohibitive. Data scarcity, however, should not justify unvalidated transfer. Simulation should instead reduce the field evidence required for a credible deployment, while the limited field evidence collected should be selected strategically for maximum information gain.

For Nigerian and comparable West African contexts, a practical strategy is to develop a shared regional virtual environment and benchmark rather than separate site-specific models. Public civil-tunnel and structural-defect datasets can support initial representation learning; procedural and generative methods can introduce rock, lighting and dust variability; low-cost field campaigns can then collect a small calibration set across multiple mines; and active adaptation can prioritise conditions where synthetic and real embeddings differ most. This approach fits a research programme in which hardware acquisition and field access proceed in parallel with algorithm development rather than sequentially.

The most immediate research priorities are:

- Create an open, versioned synthetic underground inspection benchmark with controllable lighting, dust, geometry, defects and sensor noise.
- Collect small but deliberately diverse real calibration sets from multiple underground sites rather than one large single-site dataset.
- Benchmark source-only, domain-randomised, unsupervised-adapted and few-shot fine-tuned models on the same unseen-site test set.
- Integrate the benchmark with an updatable mine-tunnel digital twin so difficult field cases can be recreated virtually.
- Adopt low-cost sensor tiers (RGB/IMU as baseline, LiDAR/thermal where justified) and report compute, energy and latency alongside accuracy.
- Develop safety-oriented evaluation with uncertainty, false-negative severity, domain-shift detection and human-in-the-loop escalation.

XII. DISCUSSION

A. The Field is Moving from Data Augmentation to Domain Engineering

A major trend across the reviewed literature is a conceptual shift from “augment the dataset” to “engineer the source–target relationship.” Early GAN augmentation studies focused mainly on increasing sample count. Current work increasingly uses controllable diffusion models, geometry-informed simulators, confidence-aware pseudo-labelling and site-specific adaptation. The relevant design object is therefore no longer just the training set, but the distribution of environments the model is expected to generalise to. This framing is particularly appropriate for underground mines, where environmental variability is intrinsic rather than incidental.

B. Calibration Utility Should Judge Digital Twins

The digital-twin literature is expanding quickly, but not every 3D model contributes equally to sim-to-field transfer. For AI inspection, the most useful twins expose task-relevant parameters, preserve sensor coordinate frames, accept field updates, quantify uncertainty, and generate new training scenarios. A visually impressive but static twin may be less valuable than a lower-fidelity model that can be calibrated from field residuals. Digital-twin studies should therefore report update frequency, calibration variables, uncertainty treatment, and how the twin changes model training or decision-making.

C. Near-Domain Evidence Is Necessary but Must Be Labelled Correctly

Civil tunnel research provides the richest technical evidence for crack segmentation, point-cloud processing and domain adaptation. In contrast, mine-specific studies are stronger in microseismic monitoring, ventilation, robotics and digital-twin safety systems. Combining these literatures is necessary to accelerate underground mine AI, but transferability cannot be assumed. Concrete tunnel linings differ from natural rock; road-tunnel lighting differs from active mine illumination; metro geometry is more regular than many mine openings; and maintenance workflows differ. The evidence-level taxonomy used in this review should therefore be retained in future syntheses and experimental designs.



D. A Minimum Evidence Package for Publishable Sim-to-Field Claims

Based on the synthesis, a study claiming successful underground sim-to-field transfer should report at least: (1) an explicit source and target domain description; (2) a source-only baseline; (3) the adaptation or transfer method; (4) target-domain performance on data not used for adaptation; (5) the amount of target data and labelling effort; (6) at least one gap diagnostic beyond mean accuracy where possible; (7) failure cases; and (8) field or cross-site validation commensurate with the intended deployment claim. Without these elements, it is difficult to determine whether performance gains reflect genuine transfer, target-data leakage or a narrow evaluation setting.

XIII. LIMITATIONS OF THIS REVIEW

First, the direct underground-mine literature on explicit synthetic-to-real inspection remains small, so we included near-domain tunnel and general sim-to-real methods. The evidence hierarchy mitigates, but does not eliminate, the risk of overgeneralising from civil infrastructure to mining. Second, task heterogeneity prevents a statistically meaningful meta-analysis: mIoU, F1, accuracy, confidence, localisation error and robotics success metrics are not directly comparable. Third, many relevant systems are proprietary or reported without public datasets, limiting reproducibility. Fourth, the rapid pace of publication in 2025–2026 means that some recent studies have little independent replication.

XIV. CONCLUSION

The central challenge for AI-based underground mine inspection is increasingly one of generalisation rather than model capacity. Synthetic data can reduce dependence on hazardous and expensive field collection; domain randomisation can widen the source distribution; domain adaptation can exploit unlabelled target observations; transfer learning can reduce the need for target labels; and digital twins can convert one-way simulation into a closed virtual-real learning loop. Recent evidence from tunnel point-cloud simulation, cross-domain crack segmentation, site-specific generative adaptation, cross-mine microseismic transfer and mine digital twins shows that each component is technically viable. What remains missing is their integration within a common field-validation protocol.

This review addresses that integration gap through two original syntheses: the Underground Simulation-to-Field Gap Taxonomy and the MineSim2Field framework. The first identifies the principal sources of field mismatch; the second defines an eight-stage development pathway from virtual environment construction to deployment and twin recalibration. The resulting research agenda is particularly relevant to resource-constrained mining settings, where large-scale field data may remain unavailable, but safety requirements do not permit unverified transfer. The practical goal is not to replace real mine data with simulation. It is to use simulation intelligently enough that a smaller, strategically selected body of real evidence can support safer, faster and more generalisable deployment.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** Yes, this research was supported by the Tertiary Education Trust Fund (TETFund), 2025, through the Directorate of Research and Innovation, Abubakar Tafawa Balewa University (ATBU), Bauchi, Faculty of Computing, Department of Computer Science.
- **Ethical Approval and Consent to Participate:** This systematic review used published literature and did not involve new human-participant or animal research; ethical approval and participant consent were therefore not required.
- **Data Access Statement and Material Availability:** No new experimental dataset was generated for this review.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.
- **Generative AI Tools or Chatbots:** ChatGPT (OpenAI) was used during manuscript preparation for language editing, document formatting and figure drafting. The authors reviewed study eligibility decisions, verified the bibliography, and provided final scientific interpretation, and they retain full responsibility for the manuscript.

REFERENCES

1. Zhang, W., Tian, C., Yang, J., Xiao, P., Wu, Z., Li, H., & Sun, W. (2026). Vision and deep learning-based tunnel surface defect detection: Technological evolution, challenges, and prospects. *Measurement*, 257, 118853. DOI: [10.1016/j.measurement.2025.118853](https://doi.org/10.1016/j.measurement.2025.118853)
2. Xue, Y.-D., Jia, F., Luo, W., Zhang, D.-M., Liu, J., & Guo, Y.-F. (2026). Deep learning-based computer vision methods for shield tunnel defect recognition—automation in Construction, 183, 106807. DOI: [10.1016/j.autcon.2026.106807](https://doi.org/10.1016/j.autcon.2026.106807)
3. Nobahar, P., Xu, C., Dowd, P., & Shirani Faradonbeh, R. (2024). Exploring digital twin systems in mining operations: A review. *Green and Smart Mining Engineering*, 1(4), 474–492. DOI: [10.1016/j.gsme.2024.09.003](https://doi.org/10.1016/j.gsme.2024.09.003)
4. Ebad, S. A., Abueid, A. I., Amara, M., & Ahmed, R. (2026). AI-Driven Digital Twins in Mining Operations: A Comprehensive Review. *Technologies*, 14(5), 269. DOI: [10.3390/technologies14050269](https://doi.org/10.3390/technologies14050269)
5. Liang, R., Zhang, C., Li, B., Saydam, S., & Canbulat, I. (2026). Evolution of visualisation and digital twin technologies in mining. *International Journal of Coal Science & Technology*, 13, 53. DOI: [10.1007/s40789-026-00866-w](https://doi.org/10.1007/s40789-026-00866-w)
6. Muratore, F., Ramos, F., Turk, G., Yu, W., Gienger, M., & Peters, J. (2022). Robot Learning From Randomised Simulations: A Review. *Frontiers in Robotics and AI*, 9, 799893. DOI: [10.3389/frobt.2022.799893](https://doi.org/10.3389/frobt.2022.799893)
7. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain randomisation for transferring deep neural networks from simulation to the real world. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 23–30. DOI: [10.1109/IROS.2017.8202133](https://doi.org/10.1109/IROS.2017.8202133)



Bridging the Simulation-to-Field Gap in AI-Based Underground Mine Inspection: A Systematic Review of Synthetic Data, Domain Adaptation, Transfer Learning, and Digital Twins

8. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59), 1–35.
9. Sadeghi, F., & Levine, S. (2017). CAD2RL: Real single-image flight without a single real image. *Robotics: Science and Systems XIII*. DOI: [10.15607/RSS.2017.XIII.034](https://doi.org/10.15607/RSS.2017.XIII.034)
10. Tiboni, G., Arndt, K., & Kyrki, V. (2023). DROPO: Sim-to-real transfer with offline domain randomisation. *Robotics and Autonomous Systems*, 166, 104432. DOI: [10.1016/j.robot.2023.104432](https://doi.org/10.1016/j.robot.2023.104432)
11. Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. DOI: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191)
12. Weng, X., Huang, Y., Li, Y., Yang, H., & Yu, S. (2023). Unsupervised domain adaptation for crack detection. *Automation in Construction*, 153, 104939. DOI: [10.1016/j.autcon.2023.104939](https://doi.org/10.1016/j.autcon.2023.104939)
13. Shim, S. (2025). Self-supervised domain adaptive approach for extrapolated crack segmentation with fine-tuned inpainting generative model. *Computer-Aided Civil and Infrastructure Engineering*, 40, 3333–3358. DOI: [10.1111/mice.13517](https://doi.org/10.1111/mice.13517)
14. Yang, W., Lin, W., Jing, Y., Hill, T., Wysocki, O., & Sheil, B. (2026). Tunnel scanner: Geometry-informed synthetic point cloud generation and transfer learning for tunnel segmentation. *Automation in Construction*, 187, 106924. DOI: [10.1016/j.autcon.2026.106924](https://doi.org/10.1016/j.autcon.2026.106924)
15. Zhao, Y., Jiao, S., Yang, T., Wang, Y., Xu, H., & Liu, Z. (2026). A novel transfer learning-enhanced BiLSTM-DCNN architecture for mine microseismic signal identification with a small dataset. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 12, 104. DOI: [10.1007/s40948-026-01168-w](https://doi.org/10.1007/s40948-026-01168-w)
16. Jacobs, D. R., van Laar, J. H., & Schutte, C. S. L. (2022). Strategy to identify and mitigate hazards in deep-level mine ventilation systems using a calibrated digital twin. *South African Journal of Industrial Engineering*, 33(3), 175–191. DOI: [10.7166/33-3-2795](https://doi.org/10.7166/33-3-2795)
17. Zhu, H., Jin, J., & Zhao, S. (2025). Research on Multi-Sensor Data Fusion Based Real-Scene 3D Reconstruction and Digital Twin Visualization Methodology for Coal Mine Tunnels. *Sensors*, 25(19), 6153. DOI: [10.3390/s25196153](https://doi.org/10.3390/s25196153)
18. Cui, Y., Li, J., Li, Y., Zhang, Z., & Du, J. (2026). Multiscale geometric digital twins of mine tunnels that integrate topological evolution and local reconstruction. *Tunnelling and Underground Space Technology*, 169, 107317. DOI: [10.1016/j.tust.2025.107317](https://doi.org/10.1016/j.tust.2025.107317)
19. You, S., Li, M., Wu, B., Wang, J., & Tang, C. (2026). Digital Twin of Coal Mine Rescue Robot—Research on Intelligence and Visualisation. *Sensors*, 26(9), 2840. DOI: [10.3390/s26092840](https://doi.org/10.3390/s26092840)
20. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. DOI: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)
21. Zhuge, H., Lee, X. Y., Sin, G., Ahmed, R., Vidyaratne, L., Kumar, A., & Farahat, A. (2026). Active Sim-to-Real Gap Reduction for Industrial Inspection via Digital Twin and Embedding Analysis. *PHM Society European Conference*, 9(1), 1–11. DOI: [10.36001/phme.2026.v9i1.5037](https://doi.org/10.36001/phme.2026.v9i1.5037)
22. Zhou, Z., Zhang, J., Gong, C., & Wu, W. (2023). Automatic tunnel lining crack detection via deep learning with generative adversarial network-based data augmentation. *Underground Space*, 9, 140–154. DOI: [10.1016/j.undsp.2022.07.003](https://doi.org/10.1016/j.undsp.2022.07.003)
23. Shim, S. (2024). Self-training approach for crack detection using synthesised crack images based on a conditional generative adversarial network. *Computer-Aided Civil and Infrastructure Engineering*, 39, 1019–1041. DOI: [10.1111/mice.13119](https://doi.org/10.1111/mice.13119)
24. Yan, Q., Qiao, M., Liao, X., & Xiong, Z. (2026). Intelligent tunnel lining defect identification via a parallel multiscale attention convolutional neural network with GAN-assisted data augmentation. *Structural Concrete*. DOI10.1002/sucoco.70489
25. Han, W., Xiao, Y., & Yin, Y. (2024). UM-GAN: Underground mine GAN for underground mine low-light image enhancement. *IET Image Processing*, 18(8), 2154–2160. DOI: [10.1049/ipr2.13092](https://doi.org/10.1049/ipr2.13092)
26. Xie, J., Chen, B., Giacomini, A., Guo, H., Iqbal, U., & Huang, J. (2025). A versatile synthetic data generation framework for crack detection: *Engineering Structures*, 344, 121428. DOI: [10.1016/j.engstruct.2025.121428](https://doi.org/10.1016/j.engstruct.2025.121428)
27. Liao, X., Yan, Q., Xiang, C., Qiao, M., Li, B., & Zhang, C. (2025). Crack severity assessment of shield tunnel segments based on DFOS strain measurements assisted by deep learning. *Measurement*, 256, 118332. DOI: [10.1016/j.measurement.2025.118332](https://doi.org/10.1016/j.measurement.2025.118332)
28. Kritzing, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11), 1016–1022. DOI: [10.1016/j.ifacol.2018.08.474](https://doi.org/10.1016/j.ifacol.2018.08.474)
29. Machado, L. B., Poncetti, B. L., de Assis, L. S., & Futai, M. M. (2025). Automated procedural modelling for tunnel inspection and digital twin development. *Tunnelling and Underground Space Technology*, 165, 106929. DOI: [10.1016/j.tust.2025.106929](https://doi.org/10.1016/j.tust.2025.106929)
30. Sun, H., Zhu, M., Dai, Y., Li, X., Rui, Y., Zhu, W., & Zhu, H. (2025). Digital twin of advanced geological conditions in rock tunnel construction—automation in Construction, 178, 106407. DOI: [10.1016/j.autcon.2025.106407](https://doi.org/10.1016/j.autcon.2025.106407)
31. Cacciottolo, C., Atencio, E., Komarizadehasl, S., & Lozano-Galant, J. A. (2025). Development of an Advanced Multi-Layer Digital Twin Conceptual Framework for Underground Mining. *Sensors*, 25(21), 6650. DOI: [10.3390/s25216650](https://doi.org/10.3390/s25216650)
32. Cao, P., Liu, J., Wang, H., Wang, Y., Liu, X., et al. (2024). Uncertainty analysis of a digital twin model of a mine ventilation system. *Scientific Reports*, 14, 26558. DOI: [10.1038/s41598-024-77978-3](https://doi.org/10.1038/s41598-024-77978-3)
33. Aksoy, C. O., Uyar Aksoy, G. G., Aydemir, Y., Sarisan, H. B., & Kaya, E. (2025). Use of Digital Twin Application Performed with CFD Analysis in an Underground Mine to Interpret Events During and After a Mine Fire. *Mining*, 5(4), 75. DOI: [10.3390/mining5040075](https://doi.org/10.3390/mining5040075)
34. Le, D. T., Sutjipto, S., Nguyen, D. D. K., & Paul, G. (2025). Design, Integration, and Field Testing of a Digital Twin-Based Teleoperated Rock Scaling Robot. *IEEE Transactions on Field Robotics*. DOI: [10.1109/TFR.2025.3547807](https://doi.org/10.1109/TFR.2025.3547807)
35. Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A., & Darrell, T. (2018). CyCADA: Cycle-consistent adversarial domain adaptation. *Proceedings of the 35th International Conference on Machine Learning*, 80, 1989–1998.
36. Beyene, D. A., Tran, D. Q., Maru, M. B., Kim, T., Park, S., & Park, S. (2023). Unsupervised domain adaptation-based crack segmentation using a transformer network. *Journal of Building Engineering*, 79, 107889. DOI: [10.1016/j.jobte.2023.107889](https://doi.org/10.1016/j.jobte.2023.107889)
37. Bhattacharya, S., Zhang, C., Mishra, D. K., Yuen, M. M. F., & Zhang, J. (2025). Efficient unsupervised domain adaptation for crack segmentation with interpretable Fourier–Morphology blending and uncertainty-guided self-training. *Computer-Aided Civil and Infrastructure Engineering*, 40, 5790–5807. DOI: [10.1111/mice.70127](https://doi.org/10.1111/mice.70127)
38. Maharjan, S., Inadomi, S., Itakura, K., & Chun, P.-J. (2025). Domain-adaptive self-supervised learning for corrosion detection and 3D building information model mapping in steel tunnels. *Computer-Aided Civil and Infrastructure Engineering*, 40, 4425–4447. DOI: [10.1111/mice.70077](https://doi.org/10.1111/mice.70077)
39. Liu, K., & Chen, B. M. (2023). Industrial UAV-Based Unsupervised Domain Adaptive Crack Recognition: From Database Towards Real-Site Infrastructural Inspections. *IEEE Transactions on Industrial Electronics*, 70(9), 9410–9420. DOI: [10.1109/TIE.2022.3204953](https://doi.org/10.1109/TIE.2022.3204953)
40. Tanveer, M., & Cho, S. (2026). Unsupervised domain adaptation framework for crack segmentation using consistency regularisation based on cross-domain image augmentation. *Computer-Aided Civil and Infrastructure Engineering*, 42, 100036.



DOI: [10.1016/j.cacaie.2026.100036](https://doi.org/10.1016/j.cacaie.2026.100036)

41. Liu, F., Ding, W., Qiao, Y., & Wang, L. (2024). Transfer learning-based encoder-decoder model with visual explanations for infrastructure crack segmentation: New open database and comprehensive evaluation. *Underground Space*, 17, 60–81.
DOI: [10.1016/j.undsp.2023.09.012](https://doi.org/10.1016/j.undsp.2023.09.012)
42. Bao, X., Huang, J., Shen, J., Wu, X., Wang, T., Chen, X., & Cui, H. (2025). A novel method of void detection in rebar-affected areas based on transfer learning and improved YOLOv8. *Tunnelling and Underground Space Technology*, 158, 106440.
DOI: [10.1016/j.tust.2025.106440](https://doi.org/10.1016/j.tust.2025.106440)
43. Xie, A. Y., & Li, B. Q. (2024). Transfer learning framework for multi-scale crack type classification with sparse microseismic networks. *International Journal of Mining Science and Technology*, 34(2), 167–178. DOI: [10.1016/j.ijmst.2024.01.003](https://doi.org/10.1016/j.ijmst.2024.01.003)

AUTHOR'S PROFILE



Dr. Fatima Umar Zambuk is a Senior Lecturer and Head of the Department of Computer Science, Faculty of Computing, Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. She earned B.Tech., M.Sc. and PhD degrees in Computer Science from ATBU and is involved in undergraduate and postgraduate teaching and supervision. Her research interests include cloud computing, machine learning, artificial intelligence, and optimisation, with publications spanning intelligent computing, prediction, cybersecurity, and data-driven systems.



Dr. Mustapha Lawal Abdulrahman is a Space Scientist at the National Space Research and Development Agency (NASRDA) and a Researcher at Abubakar Tafawa Balewa University (ATBU), Bauchi. Specialised in Artificial Intelligence, Remote Sensing, and Geospatial Intelligence, he bridges space science and sustainable development. With over 45 peer-reviewed publications, 4,900+ reads, and 129 citations, his work advances AI-driven solutions for agriculture, land-use mapping, and environmental monitoring in Nigeria. A PhD holder with expertise in Earth Observation and GIS, Dr Lawal is driving innovative, satellite-based research to address food security and climate resilience in Africa.



Dr. Mahmood Abdulhameed is a Lecturer in Electrical and Electronics Engineering at Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. His research interests include artificial intelligence, optimisation, embedded-system design and MATLAB-based simulation. His publications cover intelligent control, vehicular systems, electronic and power applications, computational modelling and multidisciplinary engineering problems. His work reflects an interest in integrating electrical and electronic engineering with intelligent algorithms and simulation-based approaches for practical decision support and system development.



Dr. (Prof.) Abdulsalam Ya'u Gital is a computer scientist and academic at Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. He obtained B.Tech. and M.Sc. degrees in Computer Science from ATBU and a Ph.D. in Computer Science from Universiti Teknologi Malaysia. His research interests include deep learning, cloud computing, machine learning, intelligent systems, the Internet of Things, and modelling and simulation. He has also served in ICT and academic leadership roles and is active in professional computing associations.



Saratu Habu Japhet is a pioneering Nigerian geoscientist and the first female PhD holder in Petroleum Geology from Northern Nigeria. A Senior Scientific Officer at NASRDA and Senior Lecturer at the University of Abuja, she specialises in energy geoscience, remote sensing, and stratigraphy. Her landmark UNILORIN doctorate, alongside degrees from ABU Zaria and the University of Jos, underscores her academic excellence. A recipient of numerous professional honors, Dr. Habu is a registered member of COMEG, NMGS, NAPE, and other esteemed bodies, establishing her as a leading authority in petroleum exploration and a trailblazer for women in geoscience.



Lawal Rukuna Muhammed is a computer science researcher and academic affiliated with the Department of Computer Science at Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. Positioned within one of Nigeria's leading science and technology institutions, he contributes to teaching, research, and innovation in computing. His academic interests span the evolving landscape of computer science, with a commitment to advancing knowledge, problem-solving, and technology-driven development. Through scholarly engagement and institutional collaboration, Muhammed represents a new generation of Nigerian computing academics dedicated to excellence, interdisciplinary inquiry, and impactful solutions that address real-world challenges.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the Publisher /Journal and/or the editor(s). The Publisher /Journal and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.