

Review

Not peer-reviewed version

Vehicle Autonomy to Ecosystem Intelligence: A Systematic Review of Dynamic Vision Architectures in Surface Mining Operations

[Nana Yaa Damtewaa Anti](#)^{*}, [Samuel Frimpong](#), [Muhammad Azeem Raza](#)

Posted Date: 1 June 2026

doi: 10.20944/preprints202606.0102.v1

Keywords: autonomous haulage systems; dynamic vision; surface mining; cooperative perception; digital twin; sensor fusion; ecosystem intelligence; open-pit; perception



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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 MDPI and/or the editor(s). MDPI 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.

Review

Vehicle Autonomy to Ecosystem Intelligence: A Systematic Review of Dynamic Vision Architectures in Surface Mining Operations

Nana Yaa Damtewaa Anti *, Samuel Frimpong and Muhammad Azeem Raza

Department of Mining and Explosives, Missouri University of Science and Technology, Rolla, USA

* Correspondence: nanfc@mst.edu

Abstract

Autonomous Haulage Systems (AHS) have significantly transformed surface mining operations by improving safety, productivity, and operational consistency. Currently, implementations predominantly rely on vehicle-centric perception architectures. Onboard LiDAR, radar, cameras, and Global Navigation Satellite Systems (GNSS) perform sensing, interpretation, and decision-making in individual systems. Decision-making is done using onboard LiDAR, radar, and cameras, and Global Navigation Satellite Systems (GNSS) in individual systems. These approaches enable collision avoidance and path tracking. They remain limited in their ability to account for the broader, dynamic mining environment characterized by dust, terrain degradation, geotechnical instability, heterogeneous traffic, and rapidly evolving operational conditions. This paper presents a systematic review of dynamic vision systems deployed in surface mining. It critically analyses the transition from solitary vehicle autonomy to interconnected, ecosystem-aware intelligence. The review synthesizes literature from mining automation, robotics, intelligent transportation systems, and multi-agent perception. This is to assess sensing technologies, perception algorithms, sensor fusion strategies, and environmental robustness techniques. Attention is given to the limitations of ego-centric perception models in complex open-pit ecosystems. Building on identified gaps, the paper proposes a conceptual framework for Ecosystem-Centric Dynamic Vision (ECDV). This perception is augmented through integration with fleet communication networks, dispatch systems, digital twins, geotechnical monitoring platforms, and environmental sensing infrastructure. The framework outlines a multi-layer architecture enabling cooperative perception, predictive hazard modeling, and risk-aware decision support at the mine-wide level. The review concludes by defining a research agenda for transitioning from vehicle autonomy to ecosystem intelligence in surface mining. It highlights opportunities in cooperative perception, adaptive sensor fusion under degraded visibility, and digital twin integrated predictive safety systems.

Keywords: autonomous haulage systems; dynamic vision; surface mining; cooperative perception; digital twin; sensor fusion; ecosystem intelligence; open-pit; perception

1. Introduction

Surface mining occupies a uniquely hazardous position among industrial sectors. It is characterized by the interaction of large-scale mobile equipment, complex operating conditions, and the persistent exposure of personnel to high-risk environments. The combination of heavy haul trucks, monotonous work cycles, and challenging terrain introduces fatigue, human error, and accident potential. This makes operational failures both severe and often irreversible [1]. The emergence of Autonomous Haulage Systems (AHS) represents one of the most significant technological transitions in modern mining. This is driven by the need to eliminate the "human factor," improve safety, and enhance productivity. AHS has evolved from early experimental systems to widely implemented industrial solutions. Since their first operational deployment in the late 2000s, autonomous haul trucks have demonstrated increased utilization, extended operating hours, and performance improvements of up to 20% compared to conventional systems [1].

Today, AHS is recognized as one of the most mature automation technologies in mining. Integrating advanced control systems, communication networks, and intelligent decision-making enables safer and more efficient haulage operations [2].

Despite these achievements, fundamental architectural limitations persist. Perception in current AHS is predominantly egocentric. Contemporary systems maintain situational awareness through onboard sensors such as LiDAR, radar, and cameras supported by GNSS positioning and fleet management software for routing and scheduling [3]. This vehicle-centric standard is proven effective within carefully delineated operational design domains. It carries structural constraints that become increasingly pronounced as mines grow larger, deeper, and operationally more complex.

The challenge is fundamentally both perceptual and contextual. The open-pit mining environment presents some of the most difficult operating conditions for autonomous systems from a sensing standpoint. According to [4], dust clouds produced by haul trucks travelling over unconsolidated material seriously impair LiDAR performance. They introduce significant noise into point cloud data and drastically reduce detection range, frequently by more than half in extreme situations. Vision-based sensors are similarly constrained, as they are highly sensitive to illumination variability. This includes glare from direct sunlight, low-light conditions at night, and motion-induced image distortion caused by continuous vehicle vibration [5]. Collectively, these factors create persistent visibility degradation scenarios where current egocentric sensor fusion approaches struggle to mitigate effectively.

At a broader contextual level, conventional vehicle-centric perception frameworks conceptualize each autonomous unit as an independent agent. It relies exclusively on its onboard sensing systems to support navigation and safety-critical decision-making. This inherently localized perception paradigm creates significant blind spots across the operational environment. Sensor range, line-of-sight, and environmental conditions constrain system awareness. Hazards such as slope instability, residual blast dust, localized ground weakening, or the presence of manned equipment may exist beyond the perceptual horizon of a single vehicle. These limitations are well recognized in autonomous systems research. Perception is shown to be restricted to the immediate sensing field and susceptible to environmental degradation, resulting in incomplete situational awareness [6]. As a result, autonomous haulage operations often require supplementary input such as externally defined operational constraints or supervisory control to compensate for gaps in perception. This underscores a fundamental limitation of current vehicle-centric autonomy paradigms.

Similar limitations have been found in the field of intelligent transportation systems, where complex, real-world environments have shown that vehicle-based autonomy is insufficient. In response, research has advanced toward cooperative perception frameworks such as Vehicle-to-Everything (V2X). This allows vehicles and infrastructure to exchange sensory and contextual data. It thereby extends situational awareness beyond individual sensing limits [7,8]. Simultaneously, the advent of digital twins and real-time geotechnical monitoring in mining contexts illustrates the capacity for incorporating extensive environmental intelligence into operational decision-making [9]. However, a cohesive integration of these ecosystem-level technologies into the perception architectures of Autonomous Haulage Systems (AHS) remains underdeveloped.

This paper is structured as follows. Section 2 describes the systematic review methodology. Section 3 reviews dynamic vision architecture in detail. Sections 4 and 5 examine vehicle-level autonomy and fleet-scale ecosystem integration, respectively. Section 6 presents the ECDV framework. Section 7 addresses open challenges and emerging research directions. Sections 8 and 9 present discussions and conclusions.

2. Methodology

2.1. Review Design

This systematic review follows the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [10]. The review protocol was designed to synthesize evidence on dynamic vision architectures deployed and evaluated in contexts of surface mining autonomy.

Also, to identify the architectural gap between current ego-centric implementations and ecosystem-level intelligence requirements.

2.2. Search Protocol

Searches were conducted across five electronic databases: IEEE Xplore, Scopus, Web of Science, ACM Digital Library, and Google Scholar, covering the period from January 2010 to date. The 2010 lower bound was selected to capture the first generation of deep learning-enabled perception systems while excluding pre-deep-learning literature that is methodologically distinct. The search was structured around three concept clusters combined with Boolean operators:

Cluster A (domain): "surface mining" OR "open pit" OR "open-cut" OR "AHS" OR "autonomous haulage"

Cluster B (technology): "perception" OR "vision" OR "LiDAR" OR "sensor fusion" OR "deep learning" OR "object detection" OR "semantic segmentation"

Cluster C (systems scope): "autonomous" OR "autonomy" OR "cooperative" OR "fleet" OR "ecosystem" OR "digital twin" OR "V2X"

2.3. Inclusion and Exclusion Criteria

Studies were included if they;

- addressed sensing, perception, or decision-making for autonomous or semi-autonomous vehicles or systems in surface mining environments;
- reported original empirical results, system designs, or rigorous simulation experiments;
- were published in peer-reviewed journals, major conference proceedings (IEEE ICRA, IROS, ITSC, Mine Automation), or substantive industry technical reports from named OEMs and research institutes; and
- were available in English. Studies were excluded if they addressed exclusively under-ground mining, considered only teleoperation without autonomous perception, re-ported purely theoretical modeling without validation, or had appeared only as abstracts or extended abstracts. Supplementary searches on geotechnical monitoring integration, digital twin applications in mining, and V2X cooperative perception were conducted.

2.4. Screening and Data Extraction

Title and abstract screening were conducted. Full-text screening applied the inclusion and exclusion criteria to the remaining papers. Data extraction captured the architecture family, sensor modality, validation environment (field, test track, simulation), performance metrics reported, operational scale (number of vehicles or site type), safety outcomes, and explicitly stated limitations. The PRISMA flow diagram is provided in Figure 1.

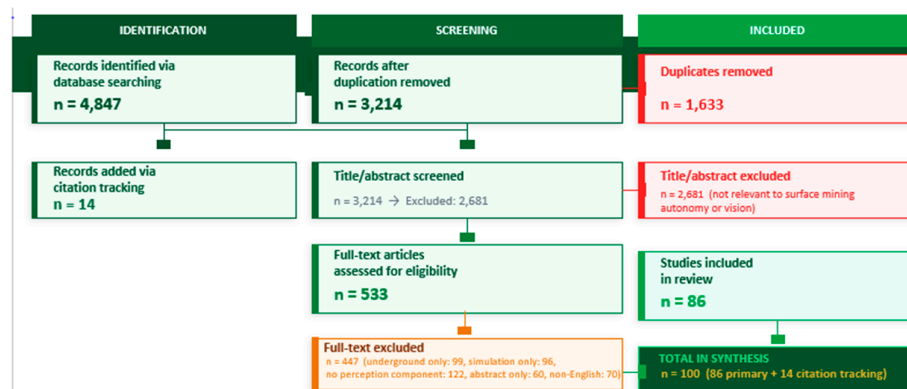


Figure 1. PRISMA 2020 flow diagram illustrating the study selection process for systematic review.

Figure 1 presents the PRISMA 2020 flow diagram used to illustrate the systematic screening and selection process adopted in this review. A total of 4,847 records were initially identified through database searching, while an additional 14 studies were obtained through citation tracking and expert recommendations. After duplicate removal, 3,214 unique records remained for title and abstract screening. The screening stage, 2,681 records were excluded because they were not directly relevant to surface mining autonomy or vision-based systems. This resulted in 533 full-text articles being assessed for eligibility. Following detailed evaluation, 447 studies were excluded for several reasons. These included studies focused exclusively on underground mining ($n = 99$), simulation-only studies without hardware validation ($n = 96$), studies lacking a perception or vision component ($n = 122$), conference abstracts without sufficient technical detail ($n = 60$), and non-English publications ($n = 70$). Ultimately, 86 primary studies satisfied the inclusion criteria and were retained for the review. Together with 14 supplementary studies identified through citation tracking, a total of 100 studies were included in the final synthesis. The selection process ensured that the review focused specifically on validated research related to autonomous haulage, perception systems, and ecosystem-level intelligence within surface mining environments.

2.5. Quality Assessment

Each included study was assessed on four criteria:

- (i) clarity of experimental or validation methodology;
- (ii) generalizability of results beyond the specific test site or simulation parameters;
- (iii) transparency of dataset or benchmark used; and
- (iv) whether reported metrics were reproducible from the described methods. Studies rated low on all four criteria were retained for discussion but flagged in the synthesis tables. This assessment informed the differentiation between deployment-validated evidence and research-prototype evidence throughout the review.

3. Results

This section presents the results of the systematic review. The technical architectures, performance metrics, and identified limitations are evaluated for of dynamic vision systems within surface mining contexts. Results are organized across four thematic areas. This includes single-frame perception, sensor fusion, temporal and sequential perception, and edge deployment. The findings draw on 100 included studies identified through the PRISMA-compliant review process.

3.1. Single-Frame Perception Systems

3.1.1. Two-Dimensional Object Detection

Two-dimensional object detection from monocular and stereo cameras remains the dominant visual perception primitive in current autonomous haulage systems (AHS). Most deployments use region-based detectors, such as Faster R-CNN or single-stage YOLO-family networks. This is because they provide a balance of accuracy and real-time performance on embedded platforms [11,12]. In automotive and road-driving benchmarks such as KITTI, Waymo Open Dataset, and BDD100K. Modern one-stage detectors such as YOLO v3 – v8 typically reach mAP values in the 70 – 90% range at IoU 0.5 while operating near real time, making them attractive for onboard mining hardware where compute is constrained [11,13,14].

However, evidence of open-pit mining analogs highlights substantial performance degradation in adverse visibility. A YOLO v5-based detector required extensive architectural modification and specialized Mine ExDark data to reach 71.9% mAP 0.5. This outperformed baseline YOLOv5 by 4.4 percentage points but still reflecting the difficulty of the domain [15]. More broadly, comparative studies consistently show that two-stage detectors such as Faster R-CNN retain an advantage in accuracy and minority-class robustness. This is especially true for small or partially occluded objects.

In contrast, YOLO variants provide superior speed and are therefore preferred when strict real-time constraints dominate, including industrial and vehicular applications [11,12,16].

Across this literature, a recurrent limitation is domain shift. This is because most training still relies on generic automotive or urban datasets. Meanwhile, the visual statistics of operational mines, such as dust plumes, headlamps, sparse background structure, and different object classes and scales, are poorly represented. This leads to reduced generalization without additional adaptation [15,17].

Recent domain-adaptive YOLO frameworks combine synthetic target-style data, semi-supervised labeling, and feature-level adaptation. These approaches show that much of this gap can be closed with limited labeled target data. This underscores that the main bottleneck for robust 2D perception in mining AHS is the lack of large, publicly available, labeled datasets from real open-pit operations rather than fundamental deficiencies in current detector architectures [17,18].

3.1.2. Three-Dimensional LiDAR Object Detection

Three-dimensional LiDAR-based object detection serves as the primary safety-critical perception modality in deployed AHS. Owing to LiDAR's robustness to ambient lighting conditions and its capacity to directly provide three-dimensional metric structure. Mining-focused adaptations of automotive 3D detectors such as PointPillars applied to hybrid solid-state LiDAR, achieve close to 90% vehicle recognition accuracy in open-pit mine conditions [19]. Semantic-geometric fusion pipelines developed specifically for mining datasets achieve processing rates of approximately 51 ms per frame within operational real-time latency requirements [17]. Nevertheless, point-cloud sparsity with range and the small physical size of pedestrians remains key challenges for reliable detection at long stopping distances [20]. Dust further degrades performance in a way that differs from simple occlusion. Controlled experiments show that airborne particulates produce systematic foreground returns. Frequency grows with optical depth, affecting ranging once transmittance drops below roughly 71–74% [21]. In mining and off-road settings, this has motivated dedicated dust-filtering algorithms that exploit LiDAR intensity and local spatial structure to remove sparse dust points while preserving obstacles [22]. Learning-based and RGB–LiDAR fusion approaches have also been developed to improve dust classification and F1 scores over conventional filters [23]. Together, these studies indicate that current voxel-based or pillar-based architecture can meet real-time latency budgets on embedded hardware. However, maintaining safe detection performance in dusty, long-range mining environments requires explicit modeling and filtering of dust-induced returns rather than relying on clean-weather automotive training.

3.2. Sensor Fusion Architectures

Multi-sensor fusion in mining leverages the complementary strengths of cameras, LiDAR, radar, GNSS, and IMU to achieve robust perception and localization under dust, vibration, and long-range visibility requirements. Classical surveys for road vehicles define the standard fusion taxonomy as high-level fusion, low-level fusion, and mid-level fusion [24]. Recent deep-learning reviews further refine this into BEV-centric fusion and cross-modal attention paradigms. This highlights unified BEV grids as an effective common space for integrating heterogeneous modalities [25]. BEV fusion exemplifies this trend by projecting camera and LiDAR into a shared BEV representation. This preserves both semantic richness and geometric accuracy while remaining computationally efficient for multi-task perception [25]. In mining-specific settings, PV Fusion shows that perspective-view fusion with depth densification and attentional feature fusion can better support >200 m perception in surface mines than conventional BEV models tuned for urban ranges [26]. Robustness to dust and adverse weather increasingly motivates radar integration. 4D mmWave radar is identified as a key upgrade for “Mining 5.0,” offering richer doppler and elevation information than 3D radars and improving autonomy in open-pit operations [27]. Practical open-pit deployments demonstrate LiDAR–radar fusion with adaptive confidence re-weighting to filter dust and stably detect 30–40 cm obstacles at 60 m on unpaved roads. Broader mmWave radar–vision fusion reviews emphasize data, feature, and decision-level schemes, along with demanding calibration requirements [28]. Across all

architectures, accurate, preferably online, multi-sensor calibration and synchronization are repeatedly highlighted as foundational for any reliable fusion system in autonomous vehicles and mining trucks [29].

3.3. Temporal and Sequential Perception

Temporal and sequential perception focuses on understanding how a scene evolves. These systems can track moving agents, estimate changing surface conditions, and anticipate rare but dangerous events. In mining-like settings, this includes predicting future positions of personnel and vehicles, monitoring haul-road degradation from tyre tracks, and detecting anomalous motions such as spoil pile collapses or bench failures. Methods from autonomous driving and robotics offer directly transferable tools for these tasks. Optical flow captures pixel-level or point-level motion between frames and underpins motion understanding and tracking. Recurrent and temporal fusion networks like Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), and temporal transformers integrate sequences of accelerometers, camera, or LiDAR data. This is to improve state estimation and classification over single-frame approaches. This is shown for traffic flow prediction, video action recognition, accident anticipation, and multi-object tracking [30,31]. The models exploit temporal context to better handle occlusions, noise, and complex interactions between multiple moving agents. 4D spatiotemporal occupancy and sequence-based 3D detection extend 3D perception with an explicit time dimension. Architectures such as OccFormer and RenderOcc represent scenes as voxelized 3D occupancy with semantics. Spatiotemporal networks predict future occupancy grids several seconds ahead without explicit object tracking [32]. Object-centric temporal detectors and trackers propagate object queries or proposals through time to model motion and interactions efficiently, improving accuracy and robustness for dynamic scenes [15]. These 4D-style approaches match the described “four-dimensional spatiotemporal detection” concept and are well suited to forecasting bench failure propagation or haul-road degradation once trained on mining-specific data.

3.4. Edge Deployment and Real-Time Inference

Real-time perception on autonomous haul trucks must satisfy strict end-to-end latency while running on sealed, vibration-tolerant embedded platforms. Across vision, LiDAR, and vision-language pipelines, work on Jetson/AGX Orin-class devices shows that meeting ~20–100 ms per-frame budgets is feasible only with aggressive, hardware-aware compression and optimization of perception models [13,33]. To fit within onboard, compute budgets while preserving safety-relevant accuracy, deployments combine structured pruning, post training quantization, and knowledge distillation. Surveys highlight these as the core compression tools for edge deployment, often used jointly [34,35]. INT8 PTQ tailored for LiDAR achieves up to ~3× speedup with almost no accuracy loss on CenterPoint, directly targeting edge devices. Mixed-precision PointPillars with TensorRT achieves up to a 2.5× latency reduction compared to FP32, while fully integer PTQ for PointPillars maintains FP32-level accuracy and enables low-latency hardware acceleration. For PointPillars-like 3D detectors, FPGA implementations with 8/2-bit hybrid quantization reach ~15.6 FPS while keeping detection quality acceptable.

Knowledge distillation further narrows the gap between compact students and heavier teachers: structured KD for 3D detection can compress PointPillars 4× while improving mAP over the teacher, and KD frameworks for 3D detectors reduce FLOPs by more than half while preserving or surpassing teacher accuracy and achieving >2× runtime speedup on high-end GPUs [36]. Combined schemes such as POK co-optimize pruning, quantization, and KD to produce lightweight models explicitly aimed at constrained devices [34]. Collectively, these results support deploying compressed 2D, 3D, and open-vocabulary perception models on Orin-class platforms within tight latency envelopes, with only modest accuracy trade-offs when compression is carefully designed and calibrated [37].

Table 1. Taxonomy of dynamic vision architecture families reviewed by sensor modality, representative models, validation environment, key performance metrics, and mining-specific limitations.

Architecture Family	Sensor Modality	Representative Models	Validation Environment	Key Performance Metrics	Mining-Specific Limitations
2D Object Detection	Camera	YOLOv8, Faster R-CNN, SSD	Controlled test tracks; synthetic datasets	mAP 78–92% (clean)	Severe degradation under dust; night sensitivity
3D Object Detection	LiDAR	Point Pillars, CenterPoint, VoxelNet	KITTI, nuScenes (road-domain)	mAP 55–82% (3D IoU)	Point cloud sparsity >50m; dust returns; vibration noise
Semantic Segmentation	Camera + LiDAR	Deep Lab v3+, RandLA-Net, SqueezeSegV3	Mining terrain datasets (limited)	mIoU 68–85%	Limited labelled mining data; terrain class imbalance
Multi-modal Fusion (early/mid/late)	LiDAR + Camera + Radar	BEV Fusion, Transfusion, Point Painting	Autonomous driving benchmarks	NDS 0.65–0.71	Cross-modal calibration drift; rain/dust degrades early fusion
Temporal / Sequential Modeling	LiDAR sequence, camera video	4D-Occ, BEV-Flow, ConvLSTM	Simulated mine environments	Velocity error <0.3 m/s	Latency accumulation; no mining-specific benchmarks
Transformer-based (ViT/BEV)	Camera (multi-view)	BEV Former, DETR3D, PETR	nuScenes; road data	NDS ~0.56–0.62	Compute-intensive; unproven in dust/vibration
Edge-deployed / Compressed Models	Camera, LiDAR	Pruned YOLOv8, TensorRT-quantised PointPillars	Onboard GPU (Orin, TX2)	Latency <50ms; 10–30% accuracy trade-off	Memory constraints limit model depth; calibration complexity

4. Vehicle Autonomy and Perception Challenges in Surface Mining

Autonomous Haulage Systems (AHS) represent the most mature and widely deployed form of vehicle autonomy in surface mining. Major mining companies, particularly Komatsu and Caterpillar, have successfully implemented autonomous fleets across large-scale iron ore, copper, coal, and oil sands operations. These systems operate within tightly controlled operational design domains (ODDs) that include geofenced mine boundaries, pre-surveyed haul roads, and centralized fleet management systems. Continuous monitoring from remote operation centers enables reliable production with minimal human intervention [38].

Current AHS architecture remains fundamentally egocentric. Each vehicle relies primarily on onboard LiDAR, radar, cameras, and GNSS to perceive and interpret its surroundings. While this approach performs effectively within structured environments, situational awareness remains limited to the vehicle's local sensing horizon. As mine operations become larger, deeper, and more complex, this localized perception model increasingly struggles to account for dynamic

environmental conditions, mixed-traffic interactions, and hazards beyond direct line-of-sight. The limitations of vehicle-centric autonomy become more apparent when considering the broader range of autonomous mining equipment now emerging in industry. Autonomous drills require highly accurate positioning under severe GNSS multipath interference near pit walls, while autonomous dozers depend on continuous terrain reconstruction and re-al-time blade load estimation. Auxiliary vehicles such as water trucks, fuel bowsers, and grade-control vehicles introduce additional complexity because they operate with unpredictable trajectories and frequently interact with both autonomous and manned equipment in mixed-traffic environments. These operational realities highlight an important distinction between mining autonomy and conventional road vehicle automation. Although current AHS are commonly categorized as SAE Level 4 systems, mining autonomy depends less on increasing vehicle-level automation and more on improving the information architecture supporting situational awareness [2,39]. The transition toward ecosystem intelligence therefore represents a shift from isolated onboard perception toward distributed, mine-wide awareness supported by fleet management systems, cooperative perception, infrastructure sensing, and shared environmental intelligence.

4.2. Environmental Perception Challenges

The reviewed evidence identifies six primary environmental factors that systematically degrade AHS onboard perception performance. These are documented in Table 2 with current mitigation strategies and residual gaps.

Table 2. Environmental factors affecting AHS onboard perception performance, current mitigation strategies, and residual capability gaps.

Environmental Factor	Affected Sensor(s)	Observed Performance Impact	Current Mitigation Strategy	Residual Gap
Airborne Dust (PM10/PM2.5)	LiDAR, Camera, Radar	LiDAR range reduced to 75%; camera contrast degraded >60%	Adaptive thresholding; multi-return LiDAR	No real-time dynamic compensation; no mine-specific benchmarks
Mud & Water Occlusion	Camera, LiDAR window	False positive rate elevated 3x; sensor window contamination	Compressed air cleaning; redundant cameras	No predictive contamination modelling
Direct Solar / Night Glare	Camera	Detection mAP drops ~40% in direct sun and is near zero at night without IR	IR cameras; HDR imaging; LiDAR primary	Mixed lighting transitions unhandled
High Vibration (haul roads)	Camera, LiDAR	Image blur; LiDAR point drift; calibration drift over hours	Shock-mounted housing; periodic recalibration	Real-time in-motion recalibration unsolved
GNSS Outage / Multipath (pit walls)	GNSS	Position error > 5 m; path planning failure	IMU dead reckoning; HD map prior	Prolonged outages degrade SLAM convergence
Geotechnical Instability	None (ego sensors blind)	No onboard detection; sudden bench failure	Periodic human inspection; slope radar (fixed)	No real-time integration with AHS perception pipeline

5. Fleet Intelligence and Ecosystem Integration

5.1. Fleet Management Systems and Perception Coupling

Fleet management systems constitute the existing layer of mine-wide intelligence in automated surface mining operations. Platforms such as Wenco International Mining Systems Modular Mining's dispatch, and Sandvik's OptiMine aggregate GPS-derived vehicle positions, payload measurements, fuel consumption data, and maintenance alerts. This helps to optimize cycle times, re-duce queue congestion, and maximize shovel utilization. However, these systems operate at a logistics and scheduling level. They do not currently consume or broadcast real-time perceptual data. AHS trucks generate LiDAR point clouds, camera feeds, and radar detections that remain siloed within each vehicle's onboard computer platform [40,41].

This architectural separation means that safety-critical environmental information detected by one vehicle, like an obstacle at a dump and a damaged berm section is not shared with other vehicles approaching the same location. Each vehicle independently detects, or fails to detect, the same hazard. The gap between FMS-level situational awareness and vehicle-level perceptual awareness constitutes the primary architectural bottleneck in the transition toward ecosystem intelligence.

5.2. Cooperative Perception and V2X in Mining

Cooperative perception refers to the sharing of sensory data or derived perception outputs between vehicles and infrastructure to construct a shared environmental model. This has advanced substantially on the road to the autonomous driving domain. V2X architectures encompassing Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), and Vehicle-to-Network (V2N) communication have been standardized through IEEE 802.11p (DSRC) and C-V2X (cellular) protocols and validated in road platooning and intersection management applications [40,42].

The translation of cooperative perception to surface mining is at an early stage. Key distinctions from the road context include the physical scale of mine operations (pit diameters of 1–10 km with significant elevation change); the severity of the communication environment (pit wall reflections, dust attenuation of wireless signals, limited 5G coverage in deep pits), and the heterogeneity of the vehicle fleet (autonomous trucks, manned light vehicles, manual dozers, shovel operators, drones). [15] demonstrated a cooperative perception prototype at a research mine site in China, using LTE V2X to share compressed LiDAR feature maps between two autonomous trucks and an RSU-mounted camera, achieving a 40% reduction in blind-zone area in the loading zone without exceeding available bandwidth constraints. This proof-of-concept represents the current frontier of cooperative perception in mining; no production deployments were identified in the reviewed literature.

5.3. Infrastructure Sensing Integration

Fixed sensing infrastructure which are the cameras and radar systems mounted at dump points, shovel pedestals, pit entry ramps, and berm edges offers a complementary perception angle. This is inherently free of the dust and vibration constraints affecting vehicle-mounted sensors. Several tier-1 mining companies have deployed fixed camera arrays for traffic management and personnel safety monitoring at high-risk locations. However, these systems operate as independent safety overlays rather than as data sources integrated into the AHS perception pipeline. The technical pathway for infrastructure integration involves two components: a communication interface from fixed sensor nodes to a central fusion server (achievable with existing LTE/5G private network infrastructure), and a data fusion architecture that can combine fixed-infrastructure detections with vehicle-ego detections into a coherent shared representation. The latter requires geometric calibration of fixed sensor frames to the mine coordinate system and temporal synchronization across heterogeneous hardware.

5.4. Digital Twin Integration

Several tier-1 operators have embraced mine-scale digital twins, which are constantly updating virtual copies of the actual mine environment, as tools for operational management and planning. Platforms including Hexagon's HxGN MineOperate OP Pro, Trimble's MineStar, and Bentley's iTwin provide real-time pit models incorporating survey data, equipment positions, blast designs, and resource block models. These twins are updated through periodic drone surveys, LiDAR scan-to-BIM workflows, and equipment telemetry.

The integration of digital twin states into the AHS perception pipeline represents a significant opportunity that is currently unrealized. A digital twin that incorporates up-to-date assessments of haul road conditions, geotechnical risk zones, blast exclusion boundaries, and anticipated traffic patterns can act as a high-level prior that constrains and enhances vehicle-level perception. This is like the role of HD maps in autonomous road driving, but with the additional capability of being dynamically updated at mine-operational timescales. The critical technical challenge is latency. Current digital twin update cycles range from minutes to hours, whereas the AHS perception pipeline operates on 50–100 ms cycles. Bridging this gap requires selective real-time data injection into the twin for safety-critical parameters.

5.5. Geotechnical Monitoring as a Perception Input

Geotechnical instability such as slope failure, bench subsidence, and tension cracking represents the highest-consequence hazard class in open-pit mining. Current monitoring approaches include slope stability radar systems (GroundProbe, Reutech), satellite InSAR (Interferometric Synthetic Aperture Radar), MEMS tiltmeters, and distributed fiber optic sensing [43,44]. These systems create constant streams of data. The data shows the rates and trends of displacement across the monitored slope sections. Alarm thresholds can trigger human inspection or personnel exclusion zones. Critically, geotechnical monitoring data is not currently integrated with AHS operational systems.

A slope entering an accelerating displacement phase does not automatically modify the safety envelope or exclusion zone of approaching autonomous trucks. Human operators usually manually implement speed restrictions or access prohibitions. This represents a significant latency gap. That is, the time between a geotechnical threshold being exceeded and a consequent change in AHS behavior may span minutes to hours, depending on shift oversight arrangements. Integration of geotechnical monitoring outputs as real-time inputs to the AHS safety planning layer is technically feasible. This includes flagging at-risk bench sections as dynamic no-go zones. With existing communication infrastructure, this represents one of the highest-value near-term integration opportunities identified in this review.

Table 3. Ecosystem integration technologies: current deployment maturity, contribution to eco-system intelligence, and key research gaps identified in the reviewed literature.

Integration Layer	Technology / Platform	Current Deployment Maturity	Contribution to Ecosystem Intelligence	Key Research Gaps
Fleet Management Systems (FMS)	Wenco, Modular Mining, Dispatch	Industry-standard; widely deployed	Route optimization, traffic conflict resolution, shift scheduling	No real-time perceptual feedback loop from trucks to FMS
V2X / Wireless Communication	4G LTE, 5G private networks, DSRC	4G deployed; 5G emerging	Low-latency data sharing; remote oversight	Bandwidth limits for HD point cloud sharing; coverage in pit walls

Infrastructure Sensing	Fixed cameras, radar at berms/dumps	Pilot deployments	Extended situational awareness beyond onboard sensors	No standardized integration protocol with AHS perception
Digital Twin Platforms	Hexagon, Trimble, Bentley iTwin	Operational in some tier-1 mines	Real-time pit model; planning support; simulation	Latency of twin update; feedback to vehicle perception pipeline
Geotechnical Monitoring	Slope radar (GroundProbe), MEMS, InSAR	Widely adopted for hazard warning	Bench stability data; subsidence mapping	Not integrated with AHS safety envelope; no real-time trigger
Cooperative / Shared Perception	V2V raw or feature sharing (research)	Research prototypes only	Occlusion resolution; extended detection range	No production deployment; bandwidth; latency; trust

6. Ecosystem-Centric Dynamic Vision (ECDV): A Conceptual Framework

6.1. Motivation and Design Principles

This paper proposes the Ecosystem-Centric Dynamic Vision (ECDV) framework, which is motivated by a specific architectural claim. That is the primary barrier to the next generation of AHS capability is not just vehicle-level perception accuracy. This has reached a sufficient baseline for geofenced operation. But the absence of a structured information architecture connecting vehicle perception to the broader mine system. ECDV is not a replacement for onboard perception. It is an augmentation layer that systematically addresses the blind spots of ego-centric systems by integrating mine-wide information into the vehicle's situational awareness.

The framework is grounded in four design principles. First is graceful degradation. Ecosystem-level information augments but does not re-place onboard perception as the primary safety-critical input. The system must operate safely when external data sources are unavailable. Second is latency stratification. Different information sources operate on different update timescales. The architecture must route time-critical data, such as V2V cooperative detections, through low-latency paths. Slower-updating data, such as digital twin state information, must be routed through higher-latency but richer pathways. Third is uncertainty propagation. All externally sourced data carries uncertainty estimates. These estimates must be propagated through the fusion pipeline and reflected in the vehicle's risk assessment. Fourth is explainability. To support incident investigation and regulatory audits, decisions influenced by external data must be traceable back to their original source.

6.2. ECDV Layer Architecture

The ECDV framework comprises five functional layers arranged in a hierarchical information architecture, summarized in Table 4. Each layer consumes outputs from the layer below and provides enriched representations to the layer above, while maintaining independence that allows for graceful degradation when any layer is unavailable. Layer 1 (L1: Onboard Perception) encompasses the vehicle's existing egocentric sensing and detection pipeline that is LiDAR, camera, radar, and GNSS/IMU, producing a local occupancy map and object list at operational frame rates. This layer is unchanged from current AHS architecture. Layer 2 (L2: Cooperative Perception) introduces external perceptual data from V2X communication, compressed point cloud features, detection lists, or occupancy map segments shared by neighboring vehicles and RSU infrastructure sensors. A cooperative BEV fusion module combines L1 and L2 data in a shared spatial reference frame. This extension increases the effective detection range and improves the resolution of occluded regions. Communication scheduling algorithms prioritize data from vehicles in geometric positions that

maximally complement the ego vehicle's field of view. Layer 3 (L3: Ecosystem Context) integrates non-perceptual mine-system data that includes the digital twin state, FMS traffic intent, geotechnical risk indices, blast schedule, and weather/dust monitoring outputs into a risk-annotated environment model. This layer operates on timescales of seconds to minutes for most parameters, with sub-second updates for geotechnical alarm states and traffic intent signals.

Layer 4 (L4: Predictive Safety and Planning) consumes the fused L1–L3 representation to perform predictive hazard modeling and risk-aware path planning. A probabilistic hazard model estimates the likelihood and consequence of candidate hazard scenarios. This includes personnel in path, slope failure, and road washout, giving current sensor evidence and ecosystem context. Risk-aware planning algorithms generate velocity profiles and path alternatives that maintain acceptable risk levels under uncertainty, including conditions where sensor data is degraded. Layer 5 (L5: Human-Machine Interface) presents synthesized situational awareness to control room operators, supports intervention request management, and maintains audit logs attributing decisions to their data sources. This layer also manages the interface with regulatory reporting systems.

Table 4. ECDV framework; five-layer architecture with components, data sources, functions, and inter-layer outputs.

ECDV Layer	Primary Components	Data Sources	Key Functions	Output to Next Layer
L1: Onboard Perception	LiDAR, camera, radar, GNSS/IMU	Vehicle sensors	Real-time object detection, segmentation, ego-motion estimation Shared BEV map	Local occupancy map; object list; ego-state vector
L2: Cooperative Perception	V2X links; edge servers; RSU cameras	Multi-vehicle infrastructure	+ construction; occlusion filling; conflict-zone awareness	Extended fused occupancy map; shared hazard layer
L3: Ecosystem Context Layer	Digital twin; FMS; geotechnical feeds	Mine-wide databases; monitoring platforms	Terrain state; slope risk index; traffic intent; blast schedule	Risk-annotated environment model; intent-aware route
L4: Predictive Safety & Planning	ML hazard models; probabilistic planners	L1–L3 fused data	Predictive hazard modeling, risk-aware path planning, and proactive speed management	Safe velocity profile; hazard alerts; maintenance triggers
L5: Human-Machine Interface	Control room dashboards; operator alerts	L4 outputs	Situation display; intervention requests; audit logging	Operator decisions; system overrides; regulatory records

7. Challenges and Emerging Directions

7.1. Open Technical Challenges

Several open challenges currently limit the transition from vehicle autonomy to eco-system intelligence in surface mining. The absence of mining-specific perception of benchmarks is the most consistently cited barrier across the reviewed literature. The field risks optimizing automotive benchmarks that do not reflect mining reality. Rigorous evaluation and comparison of perception architectures are impossible without large-scale, publicly available datasets that capture the full range of operational conditions. These conditions include dust, night operations, rain, vibration, and

geotechnical events. Without such datasets, the field risks optimizing automotive benchmarks that do not accurately reflect mining reality.

Safety certification for machine learning-based perception functions presents a regulatory and methodological challenge without established precedent. Functional safety standards, including IEC 61508 and the mining-specific ISO 17757 (Earth-moving Machinery Autonomous and Semi-Autonomous Machine System Safety) were developed primarily for deterministic control systems. The application of these standards to neural network-based perception, whose failure modes are difficult to characterize exhaustively. This requires the development of new verification and validation methodologies combining formal methods, statistical testing, runtime monitoring, and operational data analysis [45].

Communication architecture for cooperative perception in deep open-pit mines presents specific challenges not encountered in road V2X contexts. Pit wall reflections cause multipath interference that degrades DSRC and C-V2X reliability. 5G private network deployments in operational mines are beginning to address coverage. The bandwidth required for sharing compressed point cloud features is estimated at 5–20 Mbps per vehicle pair at 10Hz. Update rates approach the capacity limits of early 5G deployments at large multi-truck sites.

7.2. Emerging Technology Directions

Foundation models and vision-language models (VLMs) offer a potential pathway to reducing the labeled data dependency that limits perception model performance in mining environments. Models pre-trained on large and diverse image datasets have demonstrated substantial zero-shot and few-shot generalization capability across visual domain [10,11]. Preliminary evaluations of the Segment Anything Model (SAM) for mining terrain segmentation by [12] demonstrated competitive performance with specialized models using only 5% of the labeled data. This suggests that foundation model adaptation could substantially reduce the labeling burden for new mine sites.

Neuromorphic and event-based sensing offers a promising response to the high-vibration, high-dynamic-range challenge of the mining environment. Event cameras generate asynchronous pixel-level responses to luminance changes rather than fixed-frame images. They provide microsecond temporal resolution and a dynamic range exceeding 120 dB. This performance is substantially superior to frame-based cameras in direct solar and artificial lighting transitions encountered on mine sites [13]. Integration of event cameras with LiDAR for high-speed personnel detection in high-vibration zones has been proposed but not yet validated in mining.

Federated learning provides a potential route to training perception models on the aggregate operational data of multiple mine sites without centralizing proprietary production data. A federated training protocol in which each site trains locally and shares only model weight gradients with raw data could enable the development of mining-specific perception models with dataset diversity equivalent to hundreds of site years of operation. In this approach, each site trains locally and shares only model weight gradients rather than raw data. This demonstrates privacy-preserving federated learning with differential privacy guarantees in comparable industrial IoT contexts.

7.3. Regulatory and Standardization Pathways

The regulatory landscape for AHS is evolving. ISO 17757 (2019) establishes a framework for autonomous and semi-autonomous machine system safety in earth-moving applications, but implementation guidance for specific perception system requirements remains limited. Australia's Work Health and Safety (WHS) regulations, which govern most current large-scale AHS deployments, require operators to demonstrate that risks are reduced to as low as reasonably practicable (ALARP), a principle that accommodates technology evolution but requires robust evidence frameworks. The extension of AHS operational design domains now encompasses mixed-traffic zones, degraded visibility conditions, and geotechnically unstable areas. All scenarios that ecosystem-centric perception is designed to address will require updated regulatory guidance that acknowledges the contribution of multi-source, distributed perception to overall system safety.

8. Discussion

8.2. Comparison with Adjacent Domains

The trajectory of surface mining autonomy parallels, with a lag of approximately five years. The trajectory of road autonomous driving: the transition from individual vehicle autonomy to infrastructure-connected cooperative intelligence. The road AV industry has learned, at significant cost, that Level 4 capability in geofenced domains does not naturally scale to more complex environments without qualitative architectural changes [14]. Surface mining should internalize this lesson proactively rather than encountering it reactively at higher consequences.

The parallel with precision agriculture is also instructive. Agricultural autonomy has moved rapidly toward ecosystem-centric architectures: UAV-based field sensing feeds ground vehicle path planning, soil moisture networks inform irrigation robot decisions, and satellite imagery informs harvest timing [15]. The integration patterns are directly analogous to what the ECDV framework proposes for mining, and the agricultural implementation experience offers practical design guidance.

8.1. Synthesis of Findings

This systematic review examined the dynamic vision architectures underpinning surface mining autonomy across a fifteen-year literature spanning 2010–2024. Three overarching findings emerge from the synthesis.

Vehicle-level perception has advanced substantially in its technical maturity, but this advancement has been concentrated within a narrow operational envelope. The reviewed deployment literature demonstrates that AHS operating within geofenced, pre-mapped, FMS-managed environments under moderate environmental conditions achieves safety and productivity outcomes that justify large-scale commercial adoption. This is a genuine and important achievement. However, the same literature consistently reveals that performance degrades sharply outside this envelope that is under heavy dust, in mixed-traffic zones, near geotechnically active slopes, or when GNSS is degraded. Ego-centric architecture provides no mechanism for compensating for these failures from external information sources.

Current commercial architecture largely fails to address the significant disparity in ecosystem-level data available at a mine site. Digital twins, geotechnical monitoring systems, FMS traffic data, and fixed infrastructure cameras are all present at tier-1 mines. They simply do not communicate with the AHS perception and planning layers in real time. This is not primarily a technology gap but an integration and standardization gap. The ECDV framework proposed in Section 6 is intended as a structured articulation of how these existing data sources could be connected to vehicle perception.

The research literature on cooperative perception, while rapidly advancing in automotive contexts, has produced only sparse and limited mining-specific contributions. The physical and operational characteristics of surface mines include the scale, depth, communication environment, and fleet heterogeneity.

9. Conclusion

This paper has presented a systematic review of dynamic vision architectures in surface mining. It critically examines the evolution of vehicle-level egocentric perception in relation to the requirements of ecosystem-level intelligence. The review has identified a fundamental architectural gap. Current AHS perception is effective within its operational design domain but structurally incapable of accounting for the broader, dynamic mining environment. This limitation constrains the safety, operational envelope, and adaptability of autonomous haulage systems. These challenges cannot be resolved through incremental sensor improvement or deeper onboard learning alone. This paper proposes the Ecosystem-Centric Dynamic Vision (ECDV) framework as a structured response to this gap. This is achieved by articulating a five-layer architecture that connects onboard vehicle perception to cooperative V2X sensing, digital twin state, geotechnical monitoring, and risk-aware

planning. ECDV provides a conceptual foundation for the next generation of intelligent haulage systems. These systems can operate safely and efficiently within complex, data-rich, and dynamically evolving mining environments.

The research agenda identified in this review prioritizes seven areas. These include the creation of mining-specific perception benchmarks and the development and validation of cooperative perception architectures for mining V2X. Additional priorities include adaptive sensor fusion under dynamic contamination, digital twin-AHS perception integration, safety certification pathways for ML-based perception, federated learning for fleet-wide model training, and the evaluation of foundation models for mining perception tasks. Progress in these areas, pursued collaboratively between mining industry operators, technology developers, and the academic research community, will determine the pace at which ecosystem intelligence in surface mining becomes a reality.

The transition from vehicle autonomy to ecosystem intelligence is not merely a technical evolution. It represents a fundamental reconceptualization of what safety means in an automated mine. Safety is not the property of individual vehicles. It is a property of the system. Realizing this insight within the architecture of mining automation is the defining challenge of the next decade

Author Contributions: N.A.; original draft preparation and editing, S.F; supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Center for Disease Control and Prevention (CDC) and the National Institute for Occupational Safety and Health (NIOSH).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this study.

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

Abbreviations

The following abbreviations are used in this manuscript:

Acronym	Definition
AHS	Autonomous Haulage System
BEV	Bird's Eye View
DSRC	Dedicated Short-Range Communications
ECDV	Ecosystem-Centric Dynamic Vision
FMS	Fleet Management System
GNSS	Global Navigation Satellite System
HD Map	High-Definition Map
IMU	Inertial Measurement Unit
InSAR	Interferometric Synthetic Aperture Radar
IoU	Intersection over Union
LiDAR	Light Detection and Ranging
mAP	Mean Average Precision
MEMS	Micro-Electro-Mechanical Systems
PTQ	Post-Training Quantization
RSU	Road-Side Unit
SAE	Society of Automotive Engineers
SLAM	Simultaneous Localization and Mapping
VLM	Vision-Language Model
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
ViT	Vision Transformer

References

1. Voronov, Y.; Voronov, A.; Makhambayev, D. Current State and Development Prospects of Autonomous Haulage at Surface Mines. In Proceedings of the E3S Web of Conferences; EDP Sciences, June 18 2020; Vol. 174.
2. Long, M.; Schafrik, S.; Kolapo, P.; Agioutantis, Z.; Sottile, J. Equipment and Operations Automation in Mining: A Review. *Machines* 2024, 12.
3. Zhang, X.; Zhang, A.; Sun, J.; Zhu, X.; Guo, Y.E.; Qian, F.; Mao, Z.M. EMP: Edge-Assisted Multi-Vehicle Perception. In Proceedings of the Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM; Association for Computing Machinery, October 25 2021; pp. 545–558.
4. Dreissig, M.; Scheuble, D.; Piewak, F.; Boedecker, J. Survey on LiDAR Perception in Adverse Weather Conditions. In Proceedings of the IEEE Intelligent Vehicles Symposium, Proceedings; Institute of Electrical and Electronics Engineers Inc., 2023; Vol. 2023-June.
5. Sakaridis, C.; Dai, D.; Van Gool, L. Semantic Foggy Scene Understanding with Synthetic Data. *Int. J. Comput. Vis.* 2018, 126, 973–992, doi:10.1007/s11263-018-1072-8.
6. Zhang, Y.; Carballo, A.; Yang, H.; Takeda, K. Perception and Sensing for Autonomous Vehicles under Adverse Weather Conditions: A Survey. *ISPRS Journal of Photogrammetry and Remote Sensing* 2023, 196, 146–177.
7. Wang, J.; Shao, Y.; Ge, Y.; Yu, R. A Survey of Vehicle to Everything (V2X) Testing. *Sensors (Switzerland)* 2019, 19.
8. Feng, D.; Haase-Schutz, C.; Rosenbaum, L.; Hertlein, H.; Glaser, C.; Timm, F.; Wiesbeck, W.; Dietmayer, K. Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges. *IEEE Transactions on Intelligent Transportation Systems* 2021, 22, 1341–1360, doi:10.1109/TITS.2020.2972974.
9. Lu, Y.; Liu, C.; Wang, K.I.K.; Huang, H.; Xu, X. Digital Twin-Driven Smart Manufacturing: Connotation, Reference Model, Applications and Research Issues. *Robot. Comput. Integr. Manuf.* 2020, 61.
10. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *PLoS Med.* 2021, 18.
11. Carranza-García, M.; Torres-Mateo, J.; Lara-Benítez, P.; García-Gutiérrez, J. On the Performance of One-Stage and Two-Stage Object Detectors in Autonomous Vehicles Using Camera Data. *Remote Sens. (Basel)* 2021, 13, 1–23, doi:10.3390/rs13010089.
12. Diwan, T.; Anirudh, G.; Tembhurne, J. V. Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Applications. *Multimed. Tools Appl.* 2023, 82, 9243–9275, doi:10.1007/s11042-022-13644-y.
13. Liang, T.; Glossner, J.; Wang, L.; Shi, S.; Zhang, X. Pruning and Quantization for Deep Neural Network Acceleration: A Survey. *Neurocomputing* 2021, 461, 370–403, doi:10.1016/j.neucom.2021.07.045.
14. Cao, Y.; Li, C.; Peng, Y.; Ru, H. MCS-YOLO: A Multiscale Object Detection Method for Autonomous Driving Road Environment Recognition. *IEEE Access* 2023, 11, 22342–22354, doi:10.1109/ACCESS.2023.3252021.
15. Wang, S.; Liu, Y.; Wang, T.; Li, Y.; Zhang, X. Exploring Object-Centric Temporal Modeling for Efficient Multi-View 3D Object Detection. 2023.
16. Yadav, S.P.; Jindal, M.; Rani, P.; de Albuquerque, V.H.C.; dos Santos Nascimento, C.; Kumar, M. An Improved Deep Learning-Based Optimal Object Detection System from Images. *Multimed. Tools Appl.* 2024, 83, 30045–30072, doi:10.1007/s11042-023-16736-5.
17. Li, H.; Wang, Z.; Yu, G.; Gong, Z.; Zhou, B.; Chen, P.; Zhao, F. 3DSG: A 3D LiDAR-Based Object Detection Method for Autonomous Mining Trucks Fusing Semantic and Geometric Features. *Applied Sciences (Switzerland)* 2022, 12, doi:10.3390/app122312444.
18. Goma, A.; Abdalrazik, A. Novel Deep Learning Domain Adaptation Approach for Object Detection Using Semi-Self Building Dataset and Modified YOLOv4. *World Electric Vehicle Journal* 2024, 15, doi:10.3390/wevj15060255.

19. Li, C.; Yao, G.; Long, T.; Yuan, X.; Li, P. A Novel Method for 3D Object Detection in Open-Pit Mine Based on Hybrid Solid-State LiDAR Point Cloud. *J. Sens.* 2024, 2024, doi:10.1155/2024/5854745.
20. Fernandes, D.; Silva, A.; Névoa, R.; Simões, C.; Gonzalez, D.; Guevara, M.; Novais, P.; Mon-teiro, J.; Melo-Pinto, P. Point-Cloud Based 3D Object Detection and Classification Methods for Self-Driving Applications: A Survey and Taxonomy. *Information Fusion* 2021, 68, 161–191, doi:10.1016/j.inffus.2020.11.002.
21. Phillips, T.G.; Guenther, N.; McAree, P.R. When the Dust Settles: The Four Behaviors of LiDAR in the Presence of Fine Airborne Particulates. *J. Field Robot.* 2017, 34, 985–1009, doi:10.1002/rob.21701.
22. Afzalaghaeinaeini, A.; Seo, J.; Lee, D.; Lee, H. Design of Dust-Filtering Algorithms for LiDAR Sensors Using Intensity and Range Information in Off-Road Vehicles†. *Sensors* 2022, 22, doi:10.3390/s22114051.
23. Parsons, T.; Seo, J.; Kim, B.; Lee, H.; Kim, J.C.; Cha, M. Dust De-Filtering in LiDAR Applications with Conventional and CNN Filtering Methods. *IEEE Access* 2024, 12, 22032–22042, doi:10.1109/ACCESS.2024.3362804.
24. Yeong, D.J.; Velasco-hernandez, G.; Barry, J.; Walsh, J. Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review. *Sensors* 2021, 21, 1–37.
25. Liu, Z.; Tang, H.; Amini, A.; Yang, X.; Mao, H.; Rus, D.; Han, S. BEVFusion: Multi-Task Multi-Sensor Fusion with Unified Bird’s-Eye View Representation. 2024.
26. Ai, Y.; Yang, X.; Song, R.; Cui, C.; Li, X.; Cheng, Q.; Tian, B.; Chen, L. LiDAR-Camera Fusion in Perspective View for 3D Object Detection in Surface Mine. *IEEE Transactions on Intelligent Vehicles* 2024, 9, 3721–3730, doi:10.1109/TIV.2023.3343377.
27. Yang, J.; Gui, T.; Zhang, Y.; Ge, S.; Huang, Q.; Zhao, G. Enhancement Technology for Perception in Smart Mining Vehicles: 4D Millimeter-Wave Radar and Multi-Sensor Fusion. *IEEE Transactions on Intelligent Vehicles* 2024, 9, 5009–5013, doi:10.1109/TIV.2024.3427718.
28. Wei, Z.; Zhang, F.; Chang, S.; Liu, Y.; Wu, H.; Feng, Z. MmWave Radar and Vision Fusion for Object Detection in Autonomous Driving: A Review. *Sensors* 2022, 22.
29. Yeong, D.J.; Velasco-hernandez, G.; Barry, J.; Walsh, J. Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review. *Sensors* 2021, 21, 1–37.
30. He, R.; Zhang, C.; Xiao, Y.; Lu, X.; Zhang, S.; Liu, Y. Deep Spatio-Temporal 3D Dilated Dense Neural Network for Traffic Flow Prediction. *Expert Syst. Appl.* 2024, 237, doi:10.1016/j.eswa.2023.121394.
31. Pang, Z.; Li, J.; Tokmakov, P.; Chen, D.; Zagoruyko, S.; Wang, Y.-X. Standing Between Past and Future: Spatio-Temporal Modeling for Multi-Camera 3D Multi-Object Tracking. 2023.
32. Pan, M.; Liu, J.; Zhang, R.; Huang, P.; Li, X.; Wang, B.; Xie, H.; Liu, L.; Zhang, S. RenderOcc: Vision-Centric 3D Occupancy Prediction with 2D Rendering Supervision. 2024.
33. Peng, J.; Wang, T.; Pang, J.; Shen, Y. Towards Latency-Aware 3D Streaming Perception for Autonomous Driving. 2025.
34. Kim, J.; Chang, S.; Kwak, N. PQQ: Model Compression via Pruning, Quantization, and Knowledge Distillation. 2021.
35. Han, S.; Mao, H.; Dally, W.J. Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding. 2016.
36. Yang, J.; Shi, S.; Ding, R.; Wang, Z.; Qi, X. Towards Efficient 3D Object Detection with Knowledge Distillation. 2022.
37. Zhou, S.; Li, L.; Zhang, X.; Zhang, B.; Bai, S.; Sun, M.; Zhao, Z.; Lu, X.; Chu, X. LiDAR-PTQ: Post-Training Quantization for Point Cloud 3D Object Detection. 2024.
38. Bellamy, D.; Pravica, L. Assessing the Impact of Driverless Haul Trucks in Australian Surface Mining. *Resources Policy* 2011, 36, 149–158, doi:10.1016/j.resourpol.2010.09.002.
39. Vagia, M.; Transeth, A.A.; Fjerdingen, S.A. A Literature Review on the Levels of Automation during the Years. What Are the Different Taxonomies That Have Been Proposed? *Appl. Ergon.* 2016, 53, 190–202.
40. Chen, Q.; Tang, S.; Yang, Q.; Fu, S. Cooper: Cooperative Perception for Connected Autonomous Vehicles Based on 3D Point Clouds. 2019.
41. Liu, H.; Wu, C.; Wang, H. Real Time Object Detection Using LiDAR and Camera Fusion for Autonomous Driving. *Sci. Rep.* 2023, 13, doi:10.1038/s41598-023-35170-z.

42. Ren, S.; Lei, Z.; Wang, Z.; Dianati, M.; Wang, Y.; Chen, S.; Zhang, W. Interruption-Aware Cooperative Perception for V2X Communication-Aided Autonomous Driving. *IEEE Transactions on Intelligent Vehicles* 2024, 9, 4698–4714, doi:10.1109/TIV.2024.3371974.
43. Francioni, M.; Salvini, R.; Stead, D.; Coggan, J. Improvements in the Integration of Remote Sensing and Rock Slope Modelling. *Natural Hazards* 2018, 90, 975–1004.
44. Dick, G.J.; Eberhardt, E.; Cabrejo-Liévano, A.G.; Stead, D.; Rose, N.D. Development of an Early-Warning Time-of-Failure Analysis Methodology for Open-Pit Mine Slopes Utilizing Ground-Based Slope Stability Radar Monitoring Data. *Canadian Geotechnical Journal* 2015, 52, 515–529, doi:10.1139/cgj-2014-0028.
45. Fremont, D.J.; Chiu, J.; Margineantu, D.D.; Osipychev, D.; Seshia, S.A. Formal Analysis and Redesign of a Neural Network-Based Aircraft Taxiing System with VerifAI. In *Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer, 2020; Vol. 12224 LNCS, pp. 122–134.
46. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. *Segment Anything*. 2023.
47. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. *Learning Transferable Visual Models From Natural Language Supervision*. 2021.
48. Gallego, G.; Delbruck, T.; Orchard, G.; Bartolozzi, C.; Taba, B.; Censi, A.; Leutenegger, S.; Davison, A.J.; Conradt, J.; Daniilidis, K.; et al. Event-Based Vision: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 2022, 44, 154–180, doi:10.1109/TPAMI.2020.3008413.
49. Koopman, P.; Wagner, M. Autonomous Vehicle Safety: An Interdisciplinary Challenge. *IEEE Intelligent Transportation Systems Magazine* 2017, 9, 90–96, doi:10.1109/MITS.2016.2583491.
50. Liu, G.; Lei, J.; Guo, Z.; Chai, S.; Ren, C. Lightweight Obstacle Detection for Unmanned Mining Trucks in Open-Pit Mines. *Sci. Rep.* 2025, 15, doi:10.1038/s41598-025-93639-5.

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 MDPI and/or the editor(s). MDPI 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.