

AI-Enabled Fire Detection and Predictive Safety Monitoring for Smart Buildings in Saudi Arabia under Vision 2030: A State-of-the-Art Review and Framework

Abstract

Fire safety in smart buildings is moving from isolated threshold alarms towards data-rich systems that can recognise weak precursors, estimate fire development and support coordinated response. This state-of-the-art review synthesises recent peer-reviewed evidence on artificial intelligence, computer vision, multimodal sensing, building information modelling, digital twins, Internet of Things architectures and predictive fault monitoring, with specific interpretation for Saudi smart buildings. The evidence indicates that deep-learning models can improve visual smoke and flame recognition, while multimodal sensing and data-driven fire-state estimation can add information on location, intensity and likely evolution. Their safety value, however, depends on resilient sensing, edge processing, verified building context, calibrated uncertainty, cybersecurity and human oversight. The paper proposes a layered architecture in which AI augments rather than replaces approved deterministic fire-alarm functions, and it embeds a lifecycle assurance loop for commissioning, monitoring and revalidation. Saudi deployment must also align with the applicable Saudi Building Code, Civil Defense approval and witnessing processes, relevant HCIS requirements for in-scope assets, and recognised fire-alarm documentation and inspection practices such as NFPA 72. The resulting framework clarifies how predictive intelligence can be introduced without creating a hidden single point of life-safety failure, while identifying the validation evidence still required before consequential automated action is justified.

Keywords: artificial intelligence; fire detection; predictive safety; digital twin; smart buildings; Saudi Arabia; Vision 2030

1.Introduction

Smart buildings increasingly combine fire alarm systems, cameras, environmental sensors, building management systems, access control, occupancy analytics, and cloud or edge computing. This convergence creates an opportunity to detect fire-related anomalies earlier and to interpret them in context, but it also changes the nature of fire-safety assurance. Conventional detectors are intentionally simple and deterministic: a device senses a physical phenomenon and

triggers a predefined sequence once it meets a threshold or logic condition. Artificial intelligence can add pattern recognition and forecasting, yet a probabilistic model may also fail silently, drift when conditions change, or produce persuasive but incorrect classifications. The central question is therefore not whether AI can classify fire images, but how to engineer it as a dependable safety layer around established protection functions.

Recent building-fire studies show a clear progression from image classification toward integrated situational intelligence. Computer-vision research has demonstrated early smoke and flame recognition from surveillance streams, including structured deep networks and unified flame-smoke models (Pincott et al., 2022; Hosseini et al., 2022; Li et al., 2022). Sensor-based learning has moved beyond binary alarms by estimating fire source location and intensity or forecasting compartment conditions (Qu et al., 2022; Zhang et al., 2022a). At the system level, BIM, IoT, and digital-twin approaches attach dynamic sensor data to spatial and asset information so alarms can be interpreted as events occurring in a particular room, ventilation zone, escape route, or equipment state (Zhou et al., 2022; Jiang et al., 2023). These developments point toward predictive safety monitoring, where the objective is not merely to identify an existing fire, but to estimate risk, confidence, likely evolution, and appropriate response.

The Saudi context makes this transition particularly relevant. Research on BIM adoption and sustainable digital construction in Saudi Arabia identifies growing interest in integrated information models while also documenting organisational, capability, and interoperability barriers (Alaboud & Alshahrani, 2023; Hasanain & Nawari, 2022; Iqbal et al., 2025). Large smart-city and mixed-use developments create operational environments with extensive mechanical systems, dense sensing, sophisticated control networks, and high expectations for data-enabled management. At the same time, environmental exposure, dust, solar glare, temperature extremes, and heavy dependence on mechanical cooling can alter sensor behaviour and camera imagery. These conditions mean that models validated in laboratories or temperate buildings cannot simply be transferred without local testing.

This review treats AI-enabled fire safety as a cyber-physical assurance problem. It connects three often-separated literatures: fire detection and forecasting, digital building representation, and operational anomaly monitoring. The resulting synthesis distinguishes detection accuracy from system reliability, emphasises the interaction between AI and deterministic protection controls,

and develops a deployment logic suitable for Saudi smart buildings. The paper does not claim that AI is a regulatory substitute for approved fire alarm systems. Instead, it examines where AI can add earlier warning, richer diagnosis, and predictive decision support while preserving fail-safe principles and clear accountability.

2.Aim of the Study

This study aims to critically synthesise current evidence on AI-enabled fire detection and predictive safety monitoring and translate it into a technically defensible framework for smart buildings in Saudi Arabia. The study focuses on how computer vision, multimodal sensing, AIoT, BIM, digital twins, and building-system analytics can complement conventional life-safety functions, with particular attention to reliability, interoperability, environmental robustness, cybersecurity, human oversight, and lifecycle validation.

3.Research Objectives

Five objectives guide the review. First, it maps the principal AI methods used for fire, smoke, heat, and abnormal-condition detection in buildings. Second, it compares single-modality, multimodal, edge, cloud, BIM, and digital-twin approaches in terms of information value and failure exposure. Third, it identifies the technical and organisational conditions that affect the transfer of evidence to Saudi smart buildings. Fourth, it develops a layered architecture that integrates prediction with code-compliant deterministic safety systems. Fifth, it defines a research and implementation agenda to validate predictive safety monitoring across the building lifecycle.

4.Research Questions

The review addresses five questions. What forms of AI currently provide the strongest evidence for early fire recognition and fire-state estimation? How does combining visual, environmental, building-system, and spatial data change the quality of safety decisions? Which failure modes become important when AI is connected to real-time building controls? What Saudi environmental, project-delivery, and operational conditions should shape deployment and validation? Finally, what governance, assurance, and research practices are required before predictive outputs can support consequential life-safety decisions?

5.Review Methodology

This paper uses a state-of-the-art review and framework-synthesis method because the topic spans fire engineering, computer vision, building automation, digital twins and construction informatics. Searches targeted Scopus, Web of Science Core Collection, ScienceDirect, IEEE Xplore, SpringerLink and publisher databases, with Crossref and publisher records used to verify bibliographic details and DOIs. The principal publication window was January 2020 through December 2025. Search strings combined fire phenomena, AI methods, building integration and safety outcomes, with Saudi Arabia, smart city, Vision 2030 and BIM-adoption terms used for contextual searches. The purpose was thematic coverage and critical comparison rather than a systematic-review claim; therefore, no reconstructed retrieval or screening counts are presented.

Eligible evidence was peer-reviewed, clearly connected to buildings or built-environment safety, methodologically interpretable, and supported by traceable bibliographic metadata. Wildfire studies were excluded unless they supplied a directly transferable edge-AI or sensing principle, and purely promotional concepts without technical analysis were omitted. Saudi-context studies were retained where they illuminated digital adoption, information integration or smart-city implementation. Evidence was reviewed iteratively by title, abstract and full text where available, supported by citation chaining and DOI verification, until the principal technical themes were adequately represented.

Quality appraisal emphasized methodological transparency, realism of the validation environment, clarity of ground truth, diversity of operating conditions, treatment of false alarms and uncertainty, and relevance to integrated building operation. The synthesis did not treat experimental and simulation studies as equivalent: simulations are valuable for rare or hazardous scenarios, but they cannot establish field robustness on their own. The synthesis used a comparative evidence matrix and thematic coding. Studies were grouped into visual detection, multimodal sensing, fire-state prediction, digital-twin integration, evacuation decision support, and building-system anomaly monitoring. Thematic findings were then interpreted through a safety-engineering lens that separates information generation from safety actuation.

Formal codes, authority requirements and certification documents are treated as a separate regulatory layer rather than as peer-reviewed research evidence. For Saudi deployment, project teams should identify the applicable Saudi Building Code provisions, including SBC 201 and

fire-protection requirements under SBC 801, Civil Defense approval and witnessing obligations, and any HCIS directives that apply to the asset. NFPA 72 is also relevant where adopted or contractually invoked for fire-alarm documentation, record of completion, inspection, testing and maintenance. The review does not infer effectiveness from classification accuracy alone: curated datasets and controlled tests do not reproduce occlusion, dust, steam, reflections, maintenance faults, network interruption or changing occupancy. Greater interpretive weight is therefore given to studies that couple algorithms with real sensors, physical experiments, spatial models or system-level demonstrations.

6. Thematic Literature Review and Critical Analysis

Visual AI has become one of the most active strands of fire detection research because existing CCTV networks can provide wide-area coverage without installing a new detector at every observation point. Reviews and experiments show that convolutional networks can distinguish smoke or flame features that are difficult to encode through fixed rules (Pincott et al., 2022). UFS-Net integrates flame and smoke recognition in a unified architecture, while other end-to-end networks seek efficient inference suitable for surveillance deployment (Hosseini et al., 2022; Li et al., 2022). The advantage is rich spatial evidence: a model can indicate where a phenomenon appears and how it changes across frames. The limitation is domain sensitivity. Illumination, reflections, video compression, steam, dust, architectural screens, and camera position can move the input distribution away from the training data. A high test-set score therefore does not demonstrate a safety integrity level.

Temporal information partly addresses this weakness. Smoke is not only an object but a dynamic process so that spatio-temporal approaches can distinguish persistent growth and motion from static look-alikes. Hashemzadeh et al. (2022) show the value of combining convolutional representation with temporal cues, while Norkobil Saydirasulovich et al. (2023) demonstrate a modern YOLO-based detection route designed for real-time smart-city use. For Saudi buildings, temporal modelling is particularly important where haze, dust, sunlight, decorative lighting, or reflective finishes may trigger false visual similarities. Edge deployment can reduce latency and preserve privacy, but computational limits force trade-offs among model size, frame rate, thermal load, and redundancy. Visual AI is therefore strongest as corroborating evidence rather than as a single point of life-safety failure.

Environmental sensing offers a complementary evidence stream. Traditional smoke, heat, and gas sensors provide direct physical measurements but can struggle with nuisance conditions or delayed threshold crossing. Machine learning can exploit the joint trajectory of several variables. Qu et al. (2022) combine feature extraction with ensemble learning for multi-parameter fire detection, illustrating how correlated signals can improve discrimination. Xavier and Nanayakkara (2022) show that passive infrared data coupled with deep neural networks can support early fire recognition. The broader implication is that modality diversity matters more than merely adding sensors. Temperature, smoke obscuration, carbon monoxide, particle signatures, infrared change, and visual evidence have different failure modes; intelligently fused evidence can create diagnostic resilience if the channels are genuinely independent.

A second transition is from detection to state estimation. Qu et al. (2022) illustrate how multi-parameter sensing can support richer inference than a binary alarm, while Zhang et al. (2022a) use real-time deep learning to forecast compartment conditions and flashover-related behaviour. The AID-Fire demonstration further shows how data-driven fire models can be embedded in a real-scale system rather than remaining offline analytical tools (Zhang et al., 2022b). Predictive information changes the operational question from ‘is there a fire?’ to ‘what is happening, where is it, how fast is it developing, and what may happen next?’ This is valuable for smoke-control decisions and emergency response, but forecasting errors have asymmetric consequences. Predictive systems therefore need calibrated confidence, conservative escalation rules and fallback behaviour when inputs are missing.

Digital twins provide the context needed to turn detections into building-level decisions. A twin can associate sensor identity with geometry, compartment boundaries, doors, shafts, HVAC zones, fire dampers, stairs, and occupant routes. CloudFAS demonstrates a BIM-linked fire alarm concept in which alarm information is located and visualized in the building model (Zhou et al., 2022). Jiang et al. (2023) extend this logic by combining digital twins and semantic web technologies to control building fire protection intelligently. Chen et al. (2021) show how BIM, IoT, and immersive technologies can support fire-safety training and operational understanding. These studies indicate that the key benefit of a digital twin is not three-dimensional visualization itself; it is the structured relationship among assets, sensor states, control logic, and spatial consequences.

Fire-focused twins are becoming increasingly predictive. Ding et al. (2023) integrate real-time monitoring with evacuation intelligence, and Chen et al. (2024) combine computer vision with safety evaluation for dynamic tunnel conditions. Zhang et al. (2024) demonstrate an AIoT-enabled tunnel twin that predicts fire location and heat-release behaviour from live inputs and presents the result in a three-dimensional environment. Xie et al. (2025) advance the concept toward smart firefighting and super-real-time prediction. Although tunnel studies are not directly equivalent to buildings, they expose an important architecture: sensing must be time-synchronized, inference must be faster than the evolving hazard, and the digital representation must support operational action rather than merely replay data.

The predictive-safety concept should also include non-fire anomalies. A fire protection system depends on healthy networks, power supplies, dampers, fans, pumps, actuators, and environmental sensors. Building analytics literature shows that autoencoders and meta-learning can identify departures from learned normal operation, while attention-aware models can detect and correct sensor faults (Moon et al., 2023; Tariq et al., 2024). These methods matter because a sophisticated fire model is only as trustworthy as its inputs and actuators. Predictive monitoring should therefore watch sensor drift, unavailable devices, contradictory readings, network latency, and abnormal mechanical response. This expands fire safety from event recognition to continuous readiness assurance.

BIM and digital-twin research also warns against assuming a digital model is accurate just because it exists. Pan et al. (2022) demonstrate the effort required to enrich geometric twins with small objects by combining scanning and AI recognition. Fire safety depends on precisely such “small” assets: detectors, dampers, control modules, valves, speakers, doors, signs, and interface points. If asset location or state is stale, the twin can present false confidence. A safety twin therefore requires controlled information requirements, commissioning of asset identities and relationships, and disciplined update processes after fit-out or maintenance.

Table 1 compares the main sensing and intelligence approaches identified in the literature. The comparison shows that no modality dominates across sensitivity, context, latency, and robustness. A practical architecture should use deterministic approved detectors as the primary life-safety trigger and position AI as a parallel evidence and prediction layer. Where AI is

permitted to influence control, its authority should be bounded by validated scenarios and independent safeguards.

Saudi implementation adds a second layer of requirements. Studies of BIM adoption in the Kingdom report barriers related to skills, process standardisation, stakeholder coordination, and implementation maturity, while also identifying strong drivers for digital project delivery (Alaboud & Alshahrani, 2023; Iqbal et al., 2025). Hasanain and Nawari (2022) show how BIM can connect to Saudi sustainability objectives, and Yusuf and Abdulmohsen (2023) discuss NEOM as a test bed for technology-intensive urban development. These studies do not establish fire-safety performance, but they indicate that technical architecture must be matched by governance, data ownership, and capability development.

The regulatory and authority layer must be explicit. In Saudi projects, AI-enabled detection should be engineered around the applicable Saudi Building Code requirements, Civil Defense approval and witnessing processes, and project-specific HCIS obligations where relevant. Where NFPA 72 is adopted by code, authority, client standard or contract, its record-of-completion and inspection, testing and maintenance provisions provide an auditable baseline for the deterministic fire-alarm system. AI outputs should therefore be documented as advisory or bounded supplementary functions unless and until an authority-approved safety case establishes a different role.

Environmental localisation is equally important. Outdoor-adjacent cameras and atria may experience intense daylight contrast; dusty conditions can affect optics and particulate sensing; mechanical cooling produces large airflow fields that alter smoke transport; and large mixed-use buildings may contain kitchens, entertainment effects, high ceilings, transportation interfaces, and continuously changing tenant layouts. Local training data should therefore represent nuisance aerosols, cleaning activities, fog effects, seasonal dust, different lighting states, and HVAC modes. Validation must include degraded conditions, not only clean demonstrations.

The proposed architecture in Figure 1 uses five layers. The sensing layer combines approved fire detectors with cameras, environmental sensors, occupancy sources, and equipment-state data. The edge-assurance layer time-stamps data, checks device health, filters corrupt streams, and maintains local inference when cloud connectivity is unavailable. The intelligence layer performs visual recognition, multimodal fusion, anomaly detection, and uncertainty estimation. A

predictive digital-twin layer maps inferences to compartments, systems, and evacuation conditions, enabling scenario forecasting. The final action layer delivers operator decision support and, only where explicitly validated, bounded commands to building systems. A governance plane crosses every layer through cybersecurity, model versioning, access control, audit logs, and human authority.

This separation is critical because machine-learning confidence is not the same as safety confidence. A model can be statistically confident on an unfamiliar input. The system should therefore combine epistemic safeguards, such as out-of-distribution detection, with engineering safeguards, such as independent alarm paths and conservative defaults. Kim et al. (2024) show the value of combining digital twins with potential fire-risk assessment for emergency management. Their approach supports the broader principle that predictive safety should communicate risk and evidence, not just an opaque class label. The human operator needs to understand which sensors agree, what the model predicts, how uncertain it is, and which deterministic protections have activated.

Figure 2 extends this into a lifecycle assurance loop. Design defines required scenarios and data; commissioning verifies sensor identity, time synchronization, cause-and-effect relationships, and model baselines; operation accumulates normal and abnormal data; detection and prediction are continuously monitored; interventions are recorded; and incidents, nuisance alarms, maintenance changes, or model updates trigger revalidation. This lifecycle view is essential because AI performance can decay even when software remains unchanged. Cameras are moved, tenants alter spaces, sensors age, firmware changes, and operating schedules shift. Predictive fire safety must therefore be commissioned repeatedly as a living capability, not accepted once at handover.

The transition from isolated algorithms to safety services is also visible in work that links external smoke observations, user behaviour, and dynamic evacuation. Wang et al. (2022) used external smoke imagery to predict transient building-fire behaviour, illustrating that visual evidence can support estimation rather than only recognition. Ma and Wu (2020) connected BIM-based emergency management with occupants' behavioural decisions, showing why a fire model that ignores people may still produce operationally weak guidance. Ding et al. (2024) developed a BIM-based evacuation simulation system for large infrastructures, reinforcing the value of synchronized spatial, hazard, and movement information. These studies suggest that a

predictive platform should maintain two related forecasts: the physical hazard trajectory and the changing feasibility of protective actions. A route may be geometrically short yet unsafe because smoke conditions, door availability, crowding, or system status have changed.

This paper should be read alongside the companion commissioning-framework manuscript. The present review addresses detection, prediction and decision intelligence; the companion manuscript addresses the assurance regime by which interfaces, tests, acceptance evidence and lifecycle revalidation are governed. Stating this division of labour avoids conflating the two contributions and makes the shared Vision 2030 and digital-twin context explicit rather than duplicative.

Broader digital-twin scholarship helps clarify the maturity problem. Ketzler et al. (2020) describe city digital twins as data-intensive representations whose value depends on synchronization, modelling purpose, and data governance, while Shahat et al. (2021) emphasize the continuing ambiguity surrounding digital-twin concepts and the need for clearer research agendas. Yigitcanlar et al. (2020) similarly identify both the benefits and risks of AI in smarter cities. Transferred to fire safety, these observations caution against using the term “digital twin” for a static BIM viewer or a dashboard that lacks verified bidirectional links. A safety-relevant twin should have an explicit purpose, defined update frequency, trusted asset relationships, and measurable service levels for data freshness and availability.

A further implication concerns evidence hierarchy. Laboratory flames, synthetic smoke, CFD-generated sensor traces, and controlled full-scale tests are all useful, but they answer different questions. Synthetic data can expose a model to rare scenarios; physical tests reveal sensor and transport effects; operational data reveal nuisance sources, maintenance defects, and human workarounds. Robust Saudi deployment should combine these evidence types rather than privilege one. Model acceptance should therefore require traceability from dataset composition to scenario coverage, from scenario coverage to operational hazards, and from operational hazards to the decisions the model may influence. This chain creates a defensible safety case and makes residual uncertainty visible to owners and authorities. It enables reviewers to determine whether relevant operating conditions, failure tests, and repeatable evidence support safety claims.

Table 1. Study-oriented evidence and validation characteristics for AI-enabled fire detection and predictive monitoring

Representative evidence	AI / data basis	Validation environment	Dataset / scale reported	Performance indicators reported	Operational evidence gap	Assurance implication
Pincott et al. (2022); Hosseini et al. (2022); Li et al. (2022)	CCTV; CNN / unified smoke-flame detection	Curated datasets and focused experiments	Varies by study; dataset composition should be reported at source level	Accuracy / precision / recall commonly reported; operational false-alarm rate and lead time often absent	Domain shift from glare, steam, dust, occlusion and compression	Treat as corroborating evidence; validate on local scenarios
Qu et al. (2022); Xavier & Nanayakkara (2022)	Multi-parameter environmental sensing; ensemble / DNN	Laboratory or controlled sensor experiments	Sensor arrays and controlled fire / nuisance conditions	Classification performance reported; cross-study comparison limited by different protocols	Sensor drift, calibration dependence and correlated errors	Commission sensor health and channel independence
Zhang et al. (2022a, 2022b)	Time-series fire-state forecasting / real-scale data-driven modelling	Simulation plus real-scale demonstration	Scenario-dependent time-series datasets	Forecast error and real-time capability; safety-relevant	Simulation-to-reality gap; asymmetric underprediction on risk	Use for decision support with confidence bounds and fallback

				confidence calibration remains limited		
Zhou et al. (2022); Jiang et al. (2023)	BIM / semantic digital twin + live alarms	Prototype / system demonstration	Building-model and live point integration	System response demonstrated ; standardised latency and availability metrics not consistently reported	Stale geometry, identifier mismatch, network loss	Configuration-controlled situational intelligence
Ding et al. (2023); Zhang et al. (2024); Xie et al. (2025)	Digital twin + live sensing + predictive fire model	Tunnel / infrastructure demonstrations	Real-time or near-real-time scenario data	Location / heat-release / forecast outputs; direct transfer to occupied buildings not established	Transferability and operational-building validation	Require building-specific scenario validation before consequential use
Moon et al. (2023); Tariq et al. (2024)	Autoencoder / attention-based sensor and BMS anomaly analytics	Operational or experimental time series	Building-system sensor histories	Fault-detection metrics reported by study; rare-event coverage varies	Concept drift and rare fault classes	Continuous readiness monitoring and revalidation

Layered AI-Enabled Fire Safety Architecture with Assurance Feedback

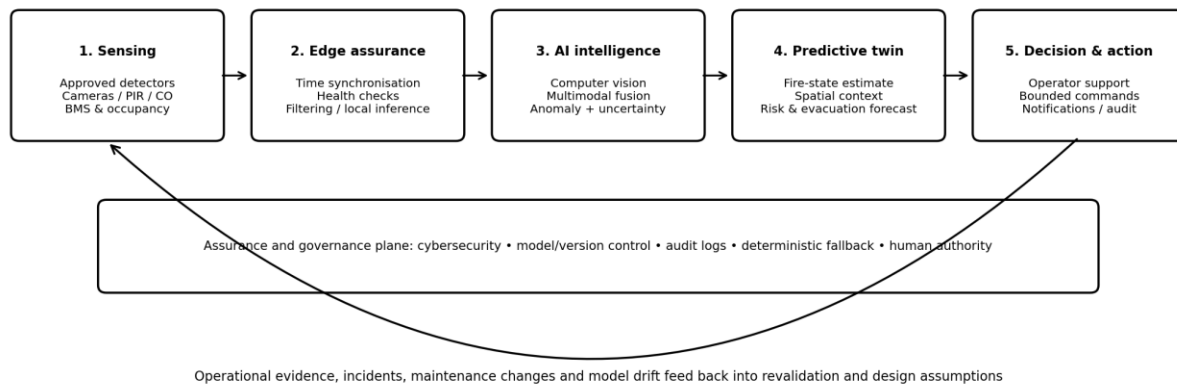


Figure 1. Layered AI-enabled fire safety architecture with governance and operational feedback for revalidation.

Lifecycle Assurance Loop for Predictive Fire Safety

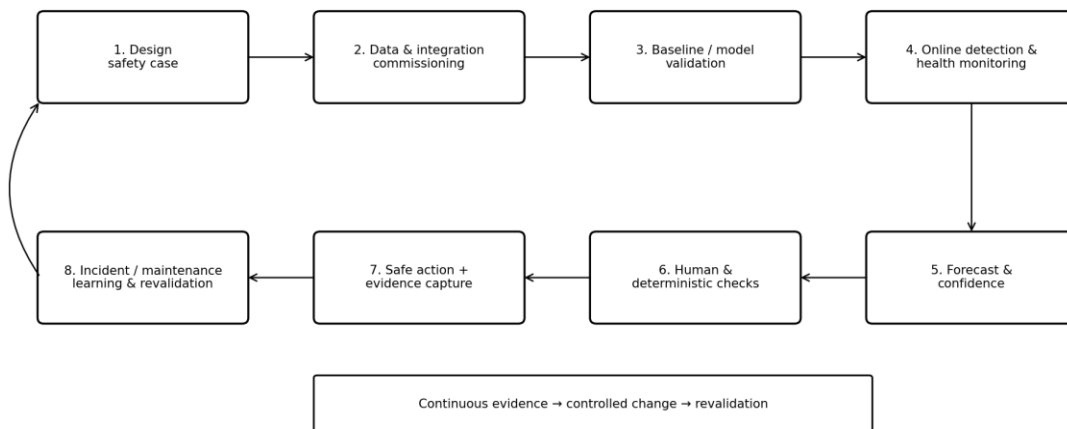


Figure 2. Lifecycle assurance loop for predictive fire-safety monitoring.

7. Discussion

The synthesis suggests that the strongest case for AI in building fire safety is not replacing conventional detection but expanding the information available before and after a confirmed

alarm. Visual models can provide spatial cues, multimodal learning can improve discrimination, inverse models can estimate source characteristics, and digital twins can place those outputs within the building's geometry and systems. Combined, these functions let operators move from isolated device alarms to a coherent picture of hazard development. The value is greatest in complex buildings where a single detector address provides too little information for rapid interpretation.

Nevertheless, evidence maturity is uneven. Many studies validate algorithms on curated datasets, simulations, or focused experiments, whereas safety systems must function through years of environmental variability and maintenance. The literature often emphasises sensitivity or classification metrics but pays less attention to availability, degraded modes, cybersecurity, calibration, model-update governance, or the consequences of false-negative and false-positive forecasts. This creates a gap between AI research performance and engineering assurance. In safety-critical deployment, a modest but well-characterized model with bounded authority may be more useful than a higher-accuracy model whose behaviour outside the training distribution is poorly understood.

Digital twins can narrow that gap if they are treated as verified information systems. The fire-safety twin should know not only geometry but also device identities, control dependencies, damper and fan states, power availability, escape routes, and current system impairments. That requirement links AI to commissioning and asset information management. It also explains why predictive safety is fundamentally socio-technical: data governance, naming conventions, maintenance discipline, cybersecurity roles, and operator training directly affect model credibility.

For Saudi Arabia, the strategic opportunity is to embed this assurance logic while smart-building platforms are being designed, rather than retrofitting it after fragmented systems are commissioned. Saudi BIM research indicates that adoption is advancing but remains sensitive to capability, standardisation, and coordination conditions. An implementation pathway should therefore prioritize common data semantics, edge resilience, controlled interfaces, and local validation before autonomous functions. Table 2 translates the evidence into deployment requirements. The framework is intentionally conservative: AI produces richer evidence and predictions, while deterministic fire-alarm logic remains available independently. This

338 architecture supports innovation without converting experimental models into hidden single
 339 points of failure.

340

341 **Table 2. Saudi deployment requirements and assurance evidence for AI-enabled fire safety.**

Saudi deployment issue	Required control	Validation evidence	Failure if neglected
Environmental localization	Train/test against dust, glare, heat and nuisance aerosols	Scenario library; sensitivity and false-alarm results	Excess nuisance alarms or missed events
Interoperability & semantics	Common tags, point mapping, room/asset identifiers and interface control	Reconciliation reports; end-to-end point tests	Incorrect context or wrong downstream action
Latency & resilience	Edge inference, independent fire-alarm operation, loss-of-cloud modes	Latency tests; network/power failure scenarios	Delayed or unavailable intelligence
Cybersecurity & privacy	Segmentation, least privilege, logging, secure update, privacy controls	Configuration audit; recovery/penetration tests	Manipulated, unavailable or untrusted safety data
Lifecycle data quality	Model/version baselines, maintenance update and drift monitoring	Revision history; periodic revalidation records	Stale twin or degraded model performance
Human governance	Bounded AI authority, explainable confidence, override and escalation rules	Operator drills; HMI acceptance; audit trails	Automation bias and unclear accountability

342

8. Research Gaps and Future Research

Six research gaps deserve priority. First, Saudi-specific multimodal datasets are scarce. Future datasets should include realistic nuisance sources, dust, glare, high-temperature boundary conditions, kitchens, atria, car parks, plant rooms, and diverse occupancy patterns, with synchronized ground truth across sensors and video. Second, benchmark protocols should measure not only accuracy but also detection lead time, false alarms per operating hour, uncertainty calibration, robustness to missing channels, and computational latency on deployable edge hardware.

Third, forecasting models need stronger physical grounding. Hybrid models that combine neural inference with zone- or CFD-informed constraints may reduce implausible predictions and improve transfer between buildings. Fourth, cyber-resilience should become part of fire-AI evaluation. Research should test spoofed sensor values, camera replay, network partitions, time-synchronization faults, compromised model updates, and denial-of-service conditions. Fifth, human-factors experiments are needed to determine how firefighters, control-room operators, and facility teams interpret probabilistic forecasts without over-trusting them. Evaluate explainability by decision quality, not visual attractiveness.

Finally, longitudinal research should examine model drift and maintenance economics after deployment. A credible safety case requires evidence that benefits persist through tenant changes, sensor replacement, software updates, and seasonal operating modes. Future Saudi pilot projects should publish anonymized performance data and lessons from failures. Such evidence would allow regulators, owners, and designers to distinguish demonstrable safety improvement from technology novelty.

9. Practical Managerial and Policy Implications

Owners should specify predictive fire intelligence as a layered service with explicit boundaries, not as a vague “AI fire system.” Requirements should define approved detectors and deterministic cause-and-effect logic, supplementary AI functions, data ownership, minimum edge capability, time synchronization, device-health monitoring, cybersecurity controls, model acceptance tests, revalidation triggers, and operator displays. Procurement should require

suppliers to disclose training assumptions, supported operating envelopes, model versions, and fallback behaviour.

Project managers should integrate the AI safety workstream with BIM, MEP, fire alarm, ICT, and facilities-management teams from design onward. Interface registers should identify every data source and control dependency. Commissioning should include adverse scenarios such as missing sensors, network loss, contradictory inputs, camera occlusion, nuisance aerosols, and changed HVAC modes. Facilities teams need procedures for model updates and for recording modifications that alter spatial or operational context.

At policy level, performance-based guidance could encourage innovation while requiring evidence proportional to consequence. Regulatory development should distinguish advisory AI from AI that initiates safety-relevant actions, define minimum evidence for local environmental validation, require auditable fallback behaviour, and preserve the independent authority of approved fire-alarm functions. Project-specific compliance mapping should name the applicable Saudi Building Code provisions, Civil Defense approval and witnessing requirements, HCIS directives where relevant, and any adopted NFPA 72 documentation and ITM obligations.

10.Limitations

This review is limited by the rapid evolution of AI methods and by the uneven availability of full operational evidence. Most peer-reviewed studies focus on prototypes, simulations, selected buildings, or bounded experiments, so conclusions about long-term reliability should remain cautious. English-language journal literature may underrepresent local engineering practice, proprietary fire-system data, and lessons held by authorities or contractors. The review also excludes standards and codes from the formal thirty-reference count, even though real deployment must comply with them. Finally, Saudi-specific peer-reviewed evidence on AI fire detection remains limited; contextual conclusions therefore combine direct fire-safety research with Saudi studies of BIM and digital construction. This makes the proposed framework a research-informed architecture requiring project-level validation, not a prescriptive compliance standard.

The proposed architecture has not yet been empirically validated as a complete system. A defensible next step is a structured Saudi expert-panel or Delphi study involving approximately

12–20 fire consultants, commissioning managers, systems integrators and engineers experienced with Civil-Defense-facing approvals. Participants should rate each layer for safety importance, feasibility, authority compatibility and current practice, followed by controlled scenario testing or an anonymised project retrospective. Until such validation is completed, the architecture should be interpreted as an evidence-informed design framework rather than a validated safety standard.

11. Conclusion

AI-enabled fire detection is evolving into predictive safety monitoring through the convergence of computer vision, multimodal sensors, fire-state estimation, BIM, digital twins, and building analytics. The literature demonstrates credible technical building blocks, but it does not justify replacing deterministic fire-protection functions with unconstrained machine learning. The most defensible approach is layered: approved life-safety systems retain independent operation, while AI adds corroboration, anomaly recognition, fire-state estimation, and predictive context. Digital twins can make those outputs operationally meaningful by linking them to compartments, assets, controls, and evacuation conditions.

For Saudi smart buildings, successful adoption depends on localisation and lifecycle assurance. Models must be tested against dust, glare, diverse nuisance sources, intensive HVAC operation, connectivity loss, and changing building configurations. Data quality, cybersecurity, asset-information accuracy, operator training, and revalidation are as important as model architecture. A national research agenda centred on realistic Saudi datasets, uncertainty-aware multimodal forecasting, edge resilience, and longitudinal field evidence would provide a stronger basis for future regulation and investment. AI can materially improve fire-safety intelligence when it is engineered as an auditable, bounded, continuously verified component of a wider life-safety system.

References

1. Wang, Z., Zhang, T., Wu, X., & Huang, X. (2022). Predicting transient building fire based on external smoke images and deep learning. *Journal of Building Engineering*, 47, 103823. <https://doi.org/10.1016/j.jobe.2021.103823>
2. Zhang, T., Wang, Z., Zeng, Y., Wu, X., Huang, X., & Xiao, F. (2022b). Building Artificial-Intelligence Digital Fire (AID-Fire) system: A real-scale demonstration. *Journal of Building Engineering*, 62, 105363. <https://doi.org/10.1016/j.jobe.2022.105363>
3. Pincott, J., Tien, P. W., Wei, S., & Calautit, J. K. (2022). Indoor fire detection utilizing computer vision-based strategies. *Journal of Building Engineering*, 61, 105154. <https://doi.org/10.1016/j.jobe.2022.105154>
4. Ding, Y., Zhang, Y., & Huang, X. (2023). Intelligent emergency digital twin system for monitoring building fire evacuation. *Journal of Building Engineering*, 77, 107416. <https://doi.org/10.1016/j.jobe.2023.107416>
5. Zhou, X., Li, H., Wang, J., Zhao, J., Xie, Q., Li, L., Liu, J., & Yu, J. (2022). CloudFAS: Cloud-based building fire alarm system using Building Information Modelling. *Journal of Building Engineering*, 53, 104571. <https://doi.org/10.1016/j.jobe.2022.104571>
6. Jiang, L., Shi, J., Wang, C., & Pan, Z. (2023). Intelligent control of building fire protection system using digital twins and semantic web technologies. *Automation in Construction*, 147, 104728. <https://doi.org/10.1016/j.autcon.2022.104728>
7. Chen, H., Hou, L., Zhang, G., & Moon, S. (2021). Development of BIM, IoT and AR/VR technologies for fire safety and upskilling. *Automation in Construction*, 125, 103631. <https://doi.org/10.1016/j.autcon.2021.103631>
8. <https://doi.org/10.1016/j.autcon.2019.102975>Ma, G., & Wu, Z. (2020). BIM-based building fire emergency management: Combining building users' behaviour decisions. *Automation in Construction*, 109, 102975. <https://doi.org/10.1016/j.autcon.2019.102975>
9. Qu, N., Li, Z., Li, X., Zhang, S., & Zheng, T. (2022). Multi-parameter fire detection method based on feature depth extraction and stacking ensemble learning model. *Fire Safety Journal*, 128, 103541. <https://doi.org/10.1016/j.firesaf.2022.103541>

10. Zhang, T., Wang, Z., Wong, H. Y., Tam, W. C., Huang, X., & Xiao, F. (2022a). Real-time forecast of compartment fire and flashover based on deep learning. *Fire Safety Journal*, 130, 103579. <https://doi.org/10.1016/j.firesaf.2022.103579>
11. Hosseini, A., Hashemzadeh, M., & Farajzadeh, N. (2022). UFS-Net: A unified flame and smoke detection method for early detection of fire in video surveillance applications using CNNs. *Journal of Computational Science*, 61, 101638. <https://doi.org/10.1016/j.jocs.2022.101638>
12. Li, Y., Zhang, W., Liu, Y., Jing, R., & Liu, C. (2022). An efficient fire and smoke detection algorithm based on an end-to-end structured network. *Engineering Applications of Artificial Intelligence*, 116, 105492. <https://doi.org/10.1016/j.engappai.2022.105492>
13. Hashemzadeh, M., Farajzadeh, N., & Heydari, M. (2022). Smoke detection in video using convolutional neural networks and efficient spatio-temporal features. *Applied Soft Computing*, 128, 109496. <https://doi.org/10.1016/j.asoc.2022.109496>
14. Xavier, K. L. B. L., & Nanayakkara, V. K. (2022). Development of an Early Fire Detection Technique Using a Passive Infrared Sensor and Deep Neural Networks. *Fire Technology*, 58, 3529–3552. <https://doi.org/10.1007/s10694-022-01319-x>
15. Norkobil Saydirasulovich, S., Abdusalomov, A., Jamil, M. K., Nasimov, R., Kozhamzharova, D., & Cho, Y.-I. (2023). A YOLOv6-Based Improved Fire Detection Approach for Smart City Environments. *Sensors*, 23(6), 3161. <https://doi.org/10.3390/s23063161>
16. Xie, W., Zeng, Y., Zhang, X., Wong, H. Y., Zhang, T., Wang, Z., Wu, X., Shi, J., Huang, X., Xiao, F., & Usmani, A. (2025). AIoT-powered building digital twin for smart firefighting and super real-time fire forecast. *Advanced Engineering Informatics*, 65, 103117. <https://doi.org/10.1016/j.aei.2025.103117>
17. Kim, Y.-J., Kim, H., Ha, B., & Kim, W.-T. (2024). Advanced fire emergency management based on potential fire risk assessment with informative digital twins. *Automation in Construction*, 167, 105722. <https://doi.org/10.1016/j.autcon.2024.105722>
18. Chen, X., Ding, Y., Zhang, Y., Wang, Z., Shi, J., Johansson, N., & Huang, X. (2024). Smart real-time evaluation of tunnel fire risk and evacuation safety via computer vision. *Safety Science*, 177, 106563. <https://doi.org/10.1016/j.ssci.2024.106563>

19. Zhang, X., Jiang, Y., Wu, X., Nan, Z., Jiang, Y., Shi, J., Zhang, Y., Huang, X., & Huang, G. Q. (2024). AIoT-enabled digital twin system for smart tunnel fire safety management. *Developments in the Built Environment*, 18, 100381. <https://doi.org/10.1016/j.dibe.2024.100381>
20. <https://doi.org/10.1016/j.ress.2023.109917> Ding, Z., Xu, S., Xie, X., Zheng, K., Wang, D., Fan, J., Li, H., & Liao, L. (2024). A building information modelling-based fire emergency evacuation simulation system for large infrastructures. *Reliability Engineering & System Safety*, 244, 109917. <https://doi.org/10.1016/j.ress.2023.109917>
21. Pan, Y., Braun, A., Brilakis, I., & Borrmann, A. (2022). Enriching geometric digital twins of buildings with small objects by fusing laser scanning and AI-based image recognition. *Automation in Construction*, 140, 104375. <https://doi.org/10.1016/j.autcon.2022.104375>
22. Moon, J., Noh, Y., Jung, S., Lee, J., & Hwang, E. (2023). Anomaly detection using a model-agnostic meta-learning-based variational auto-encoder for facility management. *Journal of Building Engineering*, 68, 106099. <https://doi.org/10.1016/j.jobe.2023.106099>
23. Tariq, S., Loy-Benitez, J., & Yoo, C. K. (2024). Multi-sensor fault detection and correction for automated IAQ monitoring in smart buildings through attention-aware autoencoders with spatial prediction module. *Journal of Building Engineering*, 96, 110573. <https://doi.org/10.1016/j.jobe.2024.110573>
24. Ketzler, B., Naserentin, V., Latino, F., Zangelidis, C., Thuvander, L., & Logg, A. (2020). Digital Twins for Cities: A State of the Art Review. *Built Environment*, 46(4), 547–573. <https://doi.org/10.2148/benv.46.4.547>
25. Yigitcanlar, T., Desouza, K. C., Butler, L., & Roozkhosh, F. (2020). Contributions and Risks of Artificial Intelligence (AI) in Building Smarter Cities: Insights from a Systematic Review of the Literature. *Energies*, 13(6), 1473. <https://doi.org/10.3390/en13061473>
26. Shahat, E., Hyun, C. T., & Yeom, C. (2021). City Digital Twin Potentials: A Review and Research Agenda. *Sustainability*, 13(6), 3386. <https://doi.org/10.3390/su13063386>

- 511 27. Alaboud, N., & Alshahrani, A. (2023). Adoption of Building Information Modelling in
512 the Saudi Construction Industry: An Interpretive Structural Modelling. Sustainability,
513 15(7), 6130. <https://doi.org/10.3390/su15076130>
- 514 28. Hasanain, F. A., & Nawari, N. O. (2022). BIM-based model for sustainable built
515 environment in Saudi Arabia. Frontiers in Built Environment, 8, 950484.
516 <https://doi.org/10.3389/fbuil.2022.950484>
- 517 29. Yusuf, N., & Abdulmohsen, D. (2023). Saudi Arabia's NEOM Project as a Testing
518 Ground for Economically Feasible Planned Cities: Case Study. Sustainability, 15(1), 608.
519 <https://doi.org/10.3390/su15010608>
- 520 30. Iqbal, M., Ullah, I., Abdou, H., Alzara, M., & Yosri, A. M. (2025). Blueprint for
521 progress: Understanding the driving forces of BIM adoption in Kingdom of Saudi Arabia
522 (KSA) construction industry. PLOS ONE, 20(2), e0313135.
523 <https://doi.org/10.1371/journal.pone.0313135>
- 524