

Internet of Things-Based Physical Security Systems for Smart Cities in Saudi Arabia: A Vision 2030 Framework

Abdulrahman Alshaikhali

0009-0005-9819-5823

Published October 2026

Abstract

The rapid digitalisation of cities in Saudi Arabia is leading to a greater number of connected cameras, access-control devices, environmental sensors, edge gateways and city platforms being used for the protection of people, buildings and critical infrastructure. While these systems can enhance situational awareness, they also turn physical security into a cyber-physical issue in which compromised identities, networks, analytics or actuators can have real-world effects. This review draws on peer-reviewed evidence published between 2020 and 2025 to look at how Internet of Things technologies can support physical security in Saudi smart cities while remaining resilient, privacy-conscious and operationally governable. Thirty studies verified by DOI were examined using a structured integrative approach that included topics such as smart-city security, IoT architectures, video analytics, intrusion detection, distributed trust, federated learning, digital twins, emergency response and Saudi urban governance. The synthesis highlights five interdependent capability areas: trusted sensing, edge intelligence, resilient connectivity, secure city platforms and coordinated response. It concludes that the technical effectiveness of these systems is more a matter of managing device identity, software provenance, data minimisation, model assurance, interoperability and multi-agency operating procedures than of adding more sensors. A framework aligned with Vision 2030 is suggested, focusing on risk-based zoning, zero-trust device governance, privacy-preserving analytics, resilient edge-cloud orchestration, integrated command workflows and lifecycle assurance. The framework views physical security as an adaptive public-service capability rather than as a set of separate surveillance products. The research priorities are Saudi field trials, climate-resilient sensing, common evaluation benchmarks, measuring privacy impacts and carrying out a longitudinal assessment of human-centred security outcomes.

1 Introduction

The development of smart cities alters the concept of urban physical security. In the past, protection has been based on the use of barriers, guards, closed-circuit television and separately managed alarm systems. Smart urban environments now include sensing, computing and communication alongside these functions: cameras are able to classify events, access readers can swap identity data, gateways can correlate alarms, and city platforms can combine information from transport, utilities and emergency services. While this enhanced capability can reduce the time taken to detect and respond to situations, it also introduces dependencies which were not present in standalone systems. A faulty camera, a compromised gateway credential or a flawed analytics model could now affect decisions at a number of different physical locations. Security must therefore be designed as an end-to-end cyber-physical system rather than being obtained by buying separate devices.

This is especially pertinent to Saudi Arabia since the urban changes underway involve the incorporation of existing metropolitan areas with new districts, digitally enabled infrastructure, and the country's broader goals for improving the quality of urban life. Recent research into smart cities in Saudi Arabia stresses that technology should be combined with governance, sustainability, and the local institutional situation rather than be seen as a standalone solution [1]. Similarly, studies on citizen participation indicate that the results of smart city initiatives rely on communication among stakeholders, legitimacy, and inclusive urban management [2-4]. These findings are important with regard to physical security since urban sensing takes place in public and semi-public areas, often includes sensitive behavioural data, and relies on agencies which have different mandates, operating procedures, and levels of risk tolerance.

Laufs, Borrión and Bradford demonstrate that security measures in smart cities range from the fitting of new sensors to existing response mechanisms to the introduction of completely new security capabilities which are mediated digitally [5]. Rizi and Seno also point out that privacy and security protection are a cross-cutting design issue for citizens in smart cities [6]. At the level of the infrastructure, IoT systems inherit the weaknesses of constrained devices, use a variety of protocols, experience problems with software updates and have extensive attack surfaces. Approaches such as machine-learning based intrusion detection, distributed trust and improved device authentication can help to reduce some of these risks, but their effectiveness depends on the deployment conditions and lifecycle management [7]. The literature thus highlights a persistent tension in that the same connectivity which allows for faster decisions in physical security also opens up additional avenues for manipulation, surveillance overreach and cascading failure.

The surveys of IoT-enabled smart cities show that the technology involved is extremely varied, including sensors, wireless networks, edge computing, cloud platforms, data analytics and application services [8-11]. More up-to-date synthesised studies point out that security, privacy, interoperability, scalability and integration with city planning remain ongoing challenges [30]. Research into physical security is usually spread out among different fields such as video surveillance, access control, cyber intrusion, emergency response and privacy-preserving analytics. It follows that city officials might be given detailed technical solutions without receiving a clear model of how these solutions should function as a comprehensive citywide security system.

The review fills this gap by concentrating on IoT-based physical security in Saudi smart cities. Physical security, as defined in this context, means preventing, detecting, assessing and responding to events which could endanger people, property, facilities, transport assets or essential urban services. This definition also covers cyber events when they are able to affect the sensing, decisions or response aspects of physical security. The paper does not suppose that increased surveillance will lead to greater safety; rather, it examines how sensing and analytics can be constrained by a trustworthy architecture, privacy controls and accountable operational procedures.

First, the review brings together the latest evidence concerning sensing, edge analytics, cyber resilience, privacy-preserving learning and incident management into one physical-security architecture. Second, it converts this architecture into a set of requirements appropriate to the conditions in Saudi Arabia, such as large-scale developments, severe environmental exposure, the presence of diverse legacy infrastructure, a large amount of public space and cooperation among multiple agencies. Third, it suggests an implementation framework which connects the technical controls to tangible operational results. By doing this, the approach moves the focus away from individual smart-

security products and instead addresses the assurance of the entire process from sensing through to human decision and coordinated response.

2 Aim, objectives and review questions

This study seeks to combine the latest evidence regarding IoT-enabled physical-security systems and to create a framework that is suitable in different contexts for achieving secure, privacy-conscious and resilient deployment in Saudi smart cities, in line with the wider transformation goals linked to Vision 2030. It has four objectives. The first of these is to identify the main physical-security functions made possible by IoT, such as perimeter protection, access and identity management, intelligent video, sensing of the environment and infrastructure, anomaly detection and emergency coordination. The second objective is to look at the cyber-physical vulnerabilities that could compromise these functions and to find controls that are suitable for resource-constrained, distributed devices. The third objective is to assess various architectural approaches—such as edge intelligence, distributed trust, privacy-preserving learning and digital twins—in terms of latency, resilience, privacy and practicality in operation. The fourth objective is to produce an implementation model which links together technology, governance, human decision-making and measurable assurance.

The synthesis is guided by four review questions: which IoT capabilities have the greatest impact on improving the detection and response to physical security threats in urban areas; which cyber, privacy and operational risks could offset those benefits; which architectural patterns offer the best balance between responsiveness and the proper handling of trustworthy data; and which contextual requirements should influence the implementation in Saudi cities? For instance, a camera is not considered a security outcome in itself; rather, its value is determined by the extent of its coverage, the quality of the images, the performance of the model, the secure transmission of the data, the process of triaging alerts, the handling of evidence, and the presence of an authorised response procedure.

3 Review methodology

A structured integrative review was carried out since the subject area involves computer networking, cybersecurity, computer vision, urban governance and emergency management. In order to ensure that the synthesis covers current IoT architectures and security methods, the time frame for the evidence was limited to the period 2020 to 2025. The searches were based on combinations of the terms 'Internet of Things', 'IoT', 'smart city', 'physical security', 'public safety', 'surveillance', 'video analytics', 'access control', 'intrusion detection', 'edge computing', 'privacy', 'federated learning', 'blockchain', 'digital twin', 'emergency response' and 'Saudi Arabia'. Peer-reviewed journal articles were found using academic indexing services and publisher databases, and related studies were identified through backward and forward citation checking.

The study was accepted if it dealt with at least one of the following four areas: an IoT architecture or a security mechanism suitable for use in urban protection; sensing or analytics that have a direct physical-security or public-safety function; privacy and trust controls applicable to city-wide IoT; or evidence regarding governance and implementation that is relevant to Saudi smart cities. Descriptive papers about smart cities were only included if they contained a Saudi-specific requirement that could influence the design of security systems. Those studies which focused solely on applications unrelated to the topic, unpeer-reviewed papers, duplicate entries and items outside the date range

were excluded. In order to ensure the reliability of the references, each source included had to have a resolvable DOI and bibliographic information that matched the scholarly record.

The evidence base that was finally assembled consisted of thirty studies. Reviews were employed in order to identify recurring patterns relating to technology and risk; empirical or prototype studies were used to look at feasibility and performance mechanisms; and studies focusing on Saudi Arabia's urban environment were used to establish the relevant contextual and governance requirements. Each paper was assigned codes in five analytical areas: sensing and access, edge intelligence, cyber resilience, privacy and trust, and command or response integration. Further codes were included to cover interoperability, latency, scalability, energy constraints, human oversight, explainability, citizen acceptance and lifecycle maintenance.

The evidence was first assigned to one of the stages in the security chain: sense, verify, communicate, interpret, decide, respond or learn. Risks were likewise assigned to this same chain so that the controls could be evaluated in relation to the function they were intended to protect. For instance, encryption was not regarded as a general benefit; instead it was connected with the protection of data transmitted from device to gateway or from gateway to platform. Federated learning was looked at as a way of reducing centralised data exposure, while separate consideration was given to its own risks relating to poisoning, inference and aggregation [21-24,28,29].

The quality assessment took into account five criteria: relevance to a physical-security function, the transparency of the technical or review method, evidence of deployment or experimental evaluation, an explicit consideration of security or privacy, and the extent to which the findings can be applied on a city scale. Since the studies involve incompatible outcomes, datasets, and architectures, the approach does not assert that it has carried out a statistical meta-analysis. Rather, it involves producing a design-oriented synthesis in which agreement among the different types of evidence is used to spot robust requirements and unresolved tensions.

4 Architecture of IoT-enabled physical security

The relevant literature points to a layered architecture in which security value results from controlled information flow rather than from the number of sensors. At the perception level, physical events are monitored using fixed and mobile cameras, thermal imagers, access readers, door and gate contacts, intrusion sensors, environmental detectors, acoustic sensors, vehicle identifiers and condition-monitoring devices. As shown by IoT surveys, these diverse types of endpoints are a key characteristic of smart-city systems [10,11]. The required design approach therefore involves sensor fusion coupled with clear confidence levels and provenance, not the unselective accumulation of data.

By carrying out object detection, tracking or event classification close to the camera, edge-enabled video analytics is able to decrease backhaul demand and alert latency [17]. Moreover, region-based methods have shown that selective processing can reduce the amount of data while still keeping the information relevant to the task [18]. Yet performance obtained on benchmark datasets does not necessarily carry over to outdoor conditions in Saudi Arabia. Factors such as heat, glare, dust, night-time contrast, camera vibration, variations in crowd density and clothing that differs according to culture can all affect detection performance. It follows that a deployable architecture should therefore treat model drift and image-quality degradation as operational security indicators, not merely as maintenance problems.

Putting analytics at the edge enhances responsiveness in situations where connectivity to the cloud is limited and allows for data minimisation since the raw video or biometric data need not always leave the local area [17,25]. However, it also increases the potential impact because a breach of a gateway could cause multiple sensor feeds to be compromised. It is therefore necessary to have device identity, secure boot, signed firmware, hardware-backed keys where possible, services running with the principle of least privilege and remote attestation before carrying out advanced analytics. Machine-learning based intrusion detection can provide behavioural monitoring, especially in cases where normal traffic patterns are stable enough to enable anomaly detection models [7], but it should be used to supplement rather than to replace sound configuration and patch management.

A physical-security network can transmit low-bandwidth alarm signals, high-bandwidth video, and time-critical control messages at the same time. By separating the network according to function and risk it is possible to restrict lateral movement and ensure that traffic which is essential for the response is protected from less trusted endpoints. Mutual authentication together with encrypted transmission help to prevent interception and spoofing, and redundant paths enable the network to keep running in the event of a failure of the carrier or of the local network. IoT security reviews regularly point out weak authentication, insecure interfaces, and resource limitations as common areas for attack [9,12,13]. As a result, when carrying out procurement the requirement should be for interoperable identity and update mechanisms, not just protocol compatibility.

At the city-platform level, events are normalised, correlated and made available to operators. This level can involve security information and event management, physical-security information management, geospatial services, evidence stores, analytics engines and digital-twin components. Smart-city data science can enhance the interpretation across different domains, but the quality of the decisions is determined by the quality of the data, the context and governance [26]. Digital twins are especially important in the case where they show the present state of assets, the geography of incidents and the response resources. Research into multi-agency incident management demonstrates how a shared digital representation can support common situational awareness among organisations [19]. The practical benefit is most evident when the twin is linked to verified live data and tested operating procedures rather than being used merely as a visual display.

The response phase involves making decisions in the control room, dispatching guards, revoking access, managing traffic, issuing public warnings, mobilising emergency services and carrying out recovery measures. Research into intelligent transportation and connected vehicles shows the way in which digital information can be used to support routing and co-ordination of emergency responses [27]. The main architectural principle is that the degree of automation should correspond to the severity of the consequence. Automating low-risk actions, for example by improving the quality of recordings following a confirmed alarm, is acceptable. In the case of high-consequence actions—such as blocking public egress routes or triggering a major emergency response—human approval, redundant verification or carefully limited fail-safe logic is required. Figure 1 combines these different layers into an architecture which is oriented towards assurance.

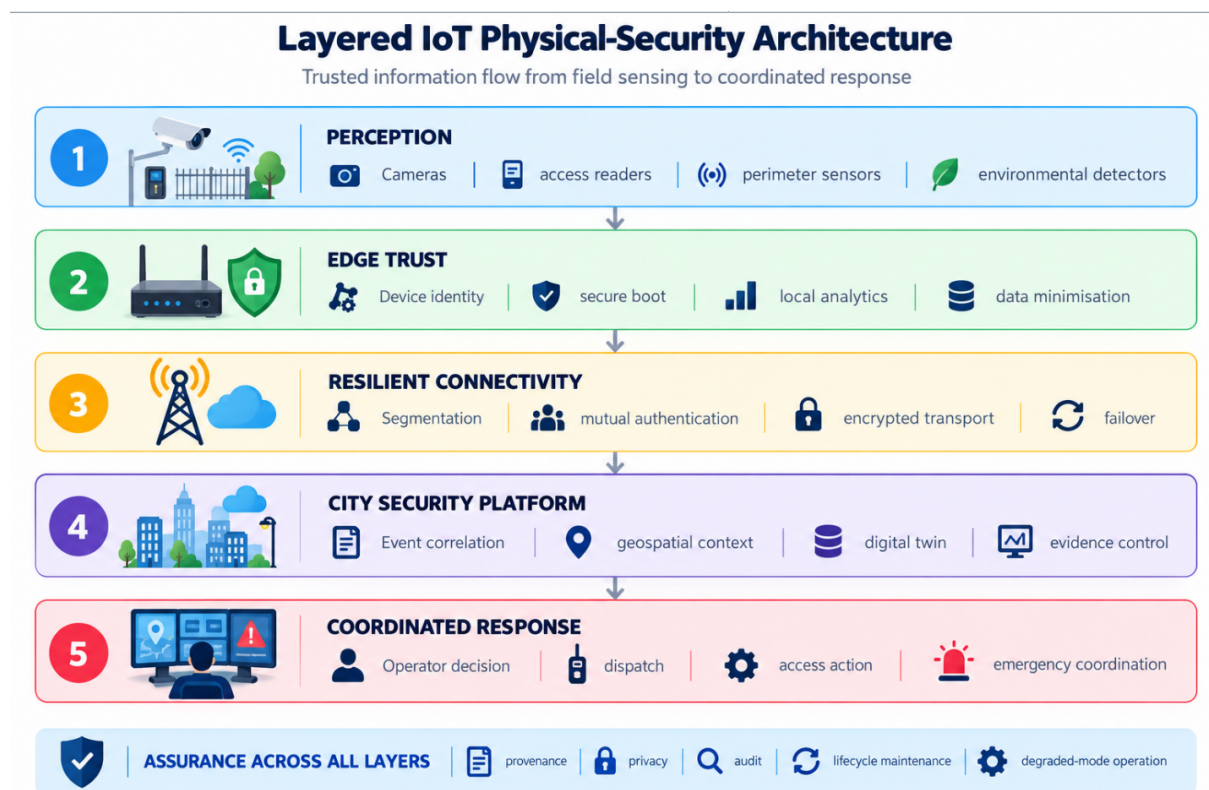


Figure 1. Layered IoT physical-security architecture with assurance controls across sensing, edge, connectivity, platform and response functions.

5 Cyber-physical security, privacy and trust

The deployment of IoT technology in physical security extends the attack surface from just the protected site to include a range of devices, identities, models, and service dependencies. It is not necessary for a malicious person to overcome a physical barrier if it is possible to clone a credential, suppress a camera feed, or alter an alarm message. On the other hand, cyber safeguards which are acceptable for information systems in offices may not be appropriate for safety-critical field devices since events such as a reboot, a certificate failure, or an update error could disrupt protection. Threat modelling should therefore start by considering the physical consequences and then move backwards to the digital conditions that could lead to them.

Because large IoT estates make use of a variety of chipsets, operating systems, and support lifecycles, there is a greater likelihood that outdated firmware or default configurations will continue to exist. Nowadays, IoT cybersecurity assessments focus on authentication flaws, malware, denial-of-service attacks, insecure communications, and software vulnerabilities [12,13]. As a result, in the case of physical-security installations, an inventory should include the device owner, its location, the firmware, the credentials, the certificate status, the network segment, the supported update period, and the criticality of the device. A device which cannot be included in such an inventory or which cannot be securely updated must not be regarded as a reliable source just because it is physically located within a controlled facility.

Research into blockchain-based smart cities suggests the use of tamper-evident event histories, decentralised identity relationships and auditable transactions for resource-constrained IoT [14-16]. However, it does not eliminate the problems of endpoint compromise, inadequate identity verification or false sensor data; an immutable record can still hold an incorrect observation. As a result,

distributed ledgers should be used selectively when it comes to ensuring the integrity of evidence across agencies, for device-registration events or for high-value configuration changes, while high-rate video streams and routine telemetry should stay in systems that are optimised for those types of workload.

Video models are able to filter out large numbers of streams and give priority to operator attention, but they can cause alarm fatigue through false positives and lead to misplaced confidence through false negatives. The review of edge analytics identifies compute limitations, model optimisation and data heterogeneity as constraints when it comes to deployment [17]. Physical-security governance should include the management of thresholds, the display of confidence levels, rules for human review and scenario-based validation. The models employed for detecting crowd anomalies, abandoned objects or intrusion into restricted areas should be assessed in light of the local operating conditions and kept under observation following any updates. When a model's output has a material impact on a decision, its version, training provenance and validation status should be included in the incident record.

The over-retention of data, the unnecessary identification of individuals or the use of the data for secondary purposes can damage public confidence and result in the creation of further data assets which in turn need to be protected. In the literature on privacy in smart cities, security and privacy are seen as interacting factors rather than as competing issues that are merely addressed afterwards [6,8,9,20]. Measures based on privacy by design involve purpose limitation, control of the field of view, filtering directly on the device, retention that is triggered by an event, pseudonymisation, role-based access, immutable access logging and scheduled deletion. If identification is not required, the system should detect the event without trying to identify the person concerned.

Federated learning provides a means of reducing reliance on a central system by allowing models to be trained across a number of distributed locations without having to pool the raw data. Surveys indicate a high level of interest in federated methods for use in smart-city IoT, together with still unresolved problems relating to non-identical local data, communication costs, poisoning and privacy leaks [21-23]. Secure aggregation and homomorphic encryption can help to improve confidentiality in federated environments [24], but do so at the expense of greater computational and operational complexity. For security purposes in Saudi cities, federated learning could be especially useful in cases where airports, transport hubs, districts or organisations want to improve shared models while at the same time keeping control of their own data. It must only be adopted if the governing body is able to determine who takes part, how the updates are verified and how contributors who have been compromised can be isolated.

Privacy-preserving visual analytics also shows a valuable design approach consisting of reducing sensitive visual details while still keeping the information required for machine analysis [25]. It is not necessary for operators, investigators, machine-learning services and maintenance teams to have identical access to the raw imagery.

Security systems have to keep providing protection even if there is network congestion, a cloud outage, a cyberattack, or a local power failure. The architectures should specify degraded modes in advance by stating which analytics will remain at the edge, for how long devices can store evidence, whether each door fails safe or fails secure, how local operators are informed of alarms, and how the clocks are synchronised after reconnection. Table 1 gives a summary of the resulting threat-control logic.

Table 1. Threat-control matrix for IoT-enabled physical security.

Security domain	Typical assets	Dominant threat/failure	Physical consequence	Recommended assurance controls	Evaluation indicators
Video surveillance and edge analytics	Cameras; edge nodes	Spoofed feeds, credential compromise, model error	Missed events, false alarms, loss of evidential confidence	Mutual authentication; signed firmware; source attestation; local buffering; scenario validation	Feed integrity; false-alarm rate; recovery time
Access and identity	Readers; controllers; locks; credentials	Cloning, replay, API compromise, unauthorised override	Unauthorised entry or unsafe lockout	Protected credentials; privileged-access control; anti-replay; defined fail-safe/fail-secure behaviour	Rejected spoof attempts; override-log completeness; authorisation latency
Perimeter and environmental sensing	Contacts; radar; acoustic; smoke; environmental sensors	Tampering, jamming, false measurements, sensor loss	Undetected intrusion or delayed hazard recognition	Tamper detection; redundant modalities; sensor fusion; heart-beat monitoring	Availability; corroboration rate; unresolved sensor faults
Edge gateways	IoT gateways; local compute	Malware, insecure update, key theft	Compromise of multiple field devices or local analytics	Secure boot; signed updates; segmentation; least privilege; attestation	Compliant firmware coverage; attestation failures; patch latency
Communications	LAN; wireless; backhaul	Denial-of-service, interception, routing failure	Loss or delay of alarms, video and control messages	Mutual authentication; encryption; quality-of-service; redundant paths	Failover time; critical-traffic loss; link availability
City platform and digital twin	Event bus; PSIM/SIEM; geospatial and evidence services	API abuse, poisoned data, time mismatch, excessive privilege	Incorrect common operating picture; evidence integrity failure	Role-based access; provenance; audit; time synchronisation; schema validation	Audit completeness; correlation accuracy; time-sync health
Response and actuators	Control room; dispatch; barriers; public warning	Automation error, unauthorised command, ambiguous alert	Unsafe physical action or delayed response	Human verification for high-consequence actions; bounded automation; rehearsed playbooks; manual override	Acknowledgement time; exercise success; override traceability

6 A Vision 2030 framework for Saudi smart cities

Research into smart cities in Saudi Arabia shows that there is quick urban development, a need for governance that is tailored to the specific context, and the fact that it is important to combine technology with social and institutional aspects [1-4]. A physical-security framework should therefore be applicable to established city areas, to new mixed-use developments, to transport infrastructure and to major public facilities rather than suppose that a single greenfield architecture is sufficient. The framework that is suggested consists of six mutually supportive pillars: risk-based zoning, trusted device identity, edge-first intelligence, resilient interoperability, privacy-accountable governance and integrated response assurance.

The first of the pillars, risk-based zoning, begins with the physical mission rather than with a list of technologies. A city should allocate its assets and spaces according to the likely harm, the required response time, the sensitivity concerning privacy and the degree of dependence on other services. Thus, for example, a public park, a metro control room, a utility substation and the entrance to a stadium all have different threat profiles and acceptable ways of failing. The zoning decides on the density of sensors, the level of redundancy, the amount of data retained, the strength of authentication and the need for human supervision. It also avoids the usual mistake of applying the most intrusive capability everywhere just because a platform is capable of doing so.

The second aspect concerns trusted device and identity governance. Each sensor, reader, gateway, operator and service account must have a unique identity associated with an accountable owner and a lifecycle state. This principle is in line with the general evidence on IoT security that shows unmanaged identities and insecure endpoints weaken higher-level controls [9,12,13]. When onboarding devices, approved firmware, cryptographic credentials, configuration baselines and a network assignment should be required. During continuous operation, certificate rotation, vulnerability monitoring and defined support periods should be carried out. Procurement contracts should treat security-update obligations and vulnerability disclosure as measurable deliverables rather than as optional services provided by the vendor.

The third pillar involves edge-first intelligence with a controlled escalation to the central platforms. In Saudi Arabia, such deployments can be used in densely populated indoor areas, along road corridors, on remote utility assets and in outdoor locations subjected to heat and dust. By processing locally the demand on the bandwidth is reduced and core detection can be maintained during outages when data is being sent up to the central system. It also allows for privacy filtering to take place before the data leaves the site. However, city platforms are still required for correlation between different sites, for long-term pattern analysis and for achieving multi-agency situational awareness. The aim of the design is to locate each computation in a position where a balance can be achieved between latency, confidentiality, resilience and the ease of maintenance. The critical rules should still be capable of being carried out locally in the event that the central platform is not available.

The fourth pillar, resilient interoperability, regards integration as a security feature. Since proprietary interfaces can lead to operational blind spots, make it difficult to combine incident data and result in a dependency on a single supplier, interoperability is repeatedly cited in the smart-city literature as a long-standing obstacle [10,11,30]. Security platforms should therefore make use of documented schemas, time synchronisation, event identifiers and authenticated application interfaces. Even if one subsystem is temporarily disconnected, a camera alert, an access event and a dispatch record must still keep the same time and location context.

The fifth pillar integrates privacy, accountability and public legitimacy. Research into urban areas in Saudi Arabia stresses stakeholder involvement and communication as essential elements of achieving sustainable smart city development [2-4]. Even though operations relating to physical security cannot reveal sensitive tactics, governance may nonetheless specify the categories of data that are collected, the reasons for their necessity, who is allowed to access them, how long they are kept and how any misuse is investigated. Whenever systems are designed to identify or track individuals, a privacy impact assessment should form part of the approval process for those systems. Audit logs should record access relating to viewing, exporting and analysis, and not just administrative changes.

The sixth pillar involves integrated response assurance. Sensors only generate value if alerts get sent to an authorised responder together with adequate context and a procedure that has been tested. The example of multi-agency digital-twin collaboration shows the potential of a common information model for coordinating incidents [19], and research into connected emergency response highlights the benefits of timely data exchange for routing and taking action [27]. Saudi city platforms should therefore establish event taxonomies, severity levels, ownership, acknowledgement times, escalation rules and the handing off of evidence among municipal security, facility teams, transport operators and emergency services.

The pillars function by going through a continuous cycle of govern, sense, verify, interpret, respond and learn. When sensing is carried out, only the minimum amount of information necessary is captured. Verification ensures that the identity of the device, the origin of the data and the quality of the signal are confirmed. Interpretation involves the use of rules, analytics and human judgement. The response includes the application of appropriate action together with escalation and fail-safe controls. Learning makes use of the results of incidents, near misses and maintenance data in order to improve the thresholds, models and procedures. Figure 2 shows the six pillars and the continuous assurance lifecycle.

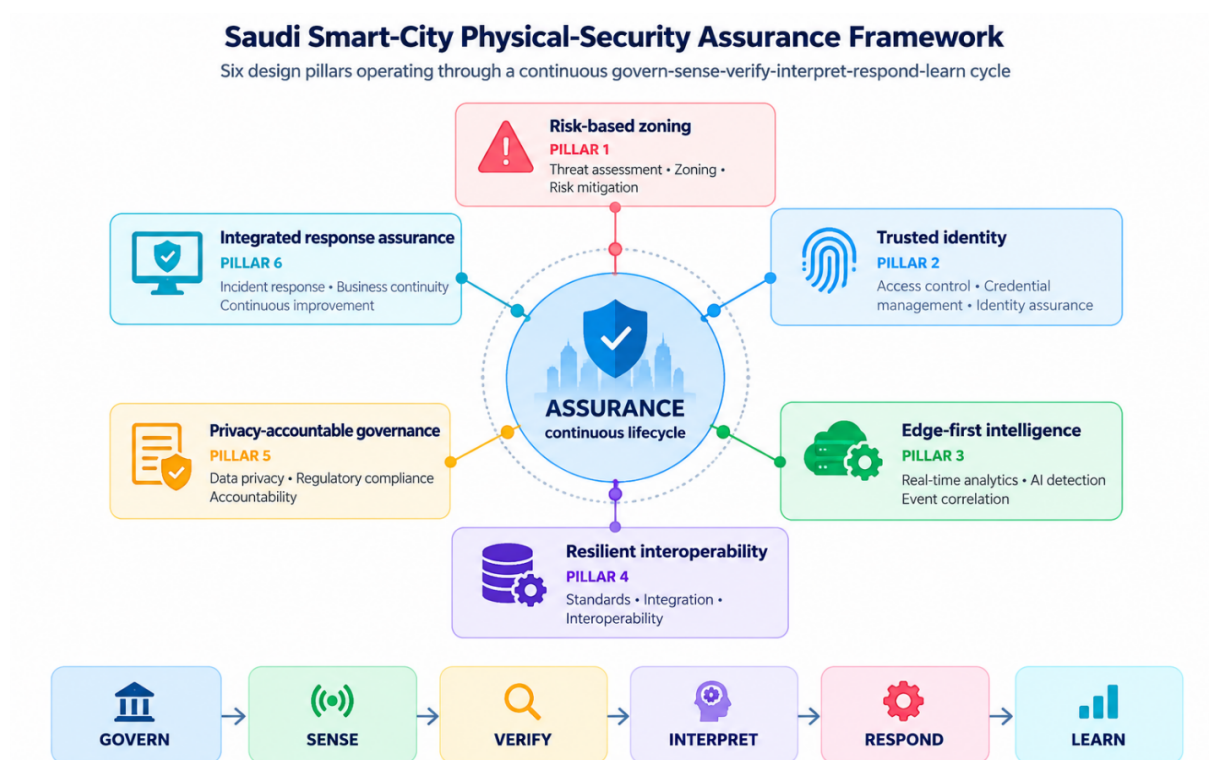


Figure 2. Saudi smart-city physical-security assurance framework: six design pillars operating through a continuous govern-sense-verify-interpret-respond-learn cycle.

During the foundation stage, cities should set up asset inventories, define risk areas, establish standards for device identity, formulate retention rules and put in place baseline network segmentation. Instead of simply replacing existing cameras and access systems because they predated the smart-city platform, they can be evaluated for secure integration. In the integration stage, edge analytics, event correlation, shared geospatial context and cross-agency workflows can be incorporated into particular high-value applications. The adaptive stage includes federated model improve-

ment, advanced anomaly detection, digital-twin simulation and automated assurance, provided that evidence of value has been obtained. Table 2 converts the framework into a staged implementation roadmap.

Table 2. Staged implementation roadmap for the proposed framework.

Stage	Primary focus	Core actions	Governance/assurance gate	Representative measures
Phase 1 - Foundation	Establish trusted baseline	Risk zoning; asset and firmware inventory; device identity; network segmentation; retention rules; vendor update obligations	Approved purpose and ownership for each security data flow; degraded-mode definition	Inventory coverage; valid credentials; supported firmware; recovery readiness
Phase 2 - Integration	Connect high-value use cases	Edge-analytics pilots; event correlation; common location/time context; evidence workflows; cross-agency escalation	Scenario validation; privacy impact review; failover and interoperability testing	Alarm verification time; false-alarm rate; evidence completeness; responder acknowledgement time
Phase 3 - Adaptive assurance	Improve models and coordination	Federated model improvement where justified; digital-twin simulation; advanced anomaly detection; continuous configuration assurance	Model-drift monitoring; adversarial testing; periodic benefit and proportionality review	Drift detection; incident recovery time; cross-agency exercise performance; privacy exceptions

More meaningful examples are the percentage of devices that have the current firmware and valid credentials, the alarm verification time, the false-alarm rate according to use case, the median time between a verified event and the responder's acknowledgement, evidence completeness, the percentage of critical areas that have degraded-mode capability, privacy-policy exceptions, the number of unauthorised-access attempts, and the recovery time following a communications loss. The model's metrics should also be broken down by environmental condition in order that good average performance may not mask failures that occur at night, during dust events or in dense crowds.

Procurement should have secure update support, documented interfaces, auditability, local failover, data minimisation, model monitoring, evidence portability and the right to exit.

7 Discussion

The report shows that IoT can improve physical security mainly by reducing the uncertainty that exists between an event and the response. Because sensors are connected, visibility is increased, edge analytics can focus attention on the most important issues, and shared platforms are able to link up signals that would otherwise stay separate. However, the more sensing there is, the greater the impact on privacy and the higher the risk of attack; the more automation is introduced, the greater the reliance on the quality of the models; and the more integration takes place, the wider the potential blast radius of a compromise. The right design objective is therefore not to aim for maximum connectivity but rather for controlled connectivity, with evidence of trust at each point of transition.

Even if a city has advanced video analytics it can still be vulnerable whenever camera logins are shared, the network sections are not segmented, or the dispatch procedures are unclear. The framework thus places a priority on strong basic controls before introducing advanced intelligence. This sequence is especially crucial in the case of large-scale programmes, since feature-driven procurement can exceed the organisation's ability to maintain the equipment and examine the alerts.

It is a key architectural decision because it achieves a number of conflicting aims: it lowers latency and central bandwidth usage, enables local continuity and allows sensitive data to be minimised in terms of transmission [17,25]. On the other hand, it spreads the software and models over a significantly wider maintenance area. In the case of Saudi Arabia's deployments, the field hardware is placed in harsh environmental conditions, which means that enclosure, thermal design, cleaning access and the replacement of components become part of cyber-physical resilience. When carrying out future procurements, consideration should be given to both the serviceability of an edge node and its secure-update procedure; a device that is secure under laboratory conditions but hard to maintain in the field will build up risk over time.

Distributed ledgers are able to enhance tamper evidence and enable audits across organisations, yet they cannot verify the truthfulness of data from a compromised sensor [14-16]. While federated learning can cut down on the need to share raw data with a central location, the training process is still susceptible to malicious updates and distribution shift [21-24,28,29]. Therefore, emerging technologies should only be made into city standards if their particular trust assumptions correspond to the operational problem and if a stated failure mode, a measurable benefit, and a fallback procedure have been established as part of the architecture review.

Human judgement is still essential. It is necessary to provide operators with interfaces that show the source, the level of confidence, the supporting signals and the recommended actions without hiding the uncertainty. Over-automation can cause errors to be propagated more quickly, whereas too much manual checking can cancel out the speed advantages offered by IoT. A useful design criterion is the quality of decisions made under a realistic workload, specifically relating to how quickly an operator can identify a false alarm, ask for further evidence and pass the incident on to another agency.

The studies into Saudi smart cities show that institutional coordination and citizen involvement are essential for achieving sustainable implementation [1-4]. In terms of physical security, this means establishing clear rules regarding data ownership, the legal purpose for which data is used, access permissions, retention periods, auditing and accountability in the event of an incident. Privacy should not be seen as a constraint that is introduced after the engineering phase; by reducing the amount of unnecessary data, storage costs are lowered, exposure to insiders is decreased and the impact of a breach is reduced, while at the same time the purpose of the system becomes clearer.

8 Research agenda and limitations

The biggest gap in terms of evidence is the lack of long-term, city-wide assessments that connect technical performance with physical-security results. Saudi research should go beyond achievements in the laboratory and brief demonstrations and instead carry out multi-season field tests which measure false alarms, device failure, communications degradation, operator workload, and response time. Particularly important is validation under harsh climate conditions for cameras, edge computers, batteries, and outdoor access equipment. The datasets and benchmarks should reflect local lighting, dust, building architecture, crowd density, and transport conditions while at the same time protecting personal data.

Future research ought to include in its reports not only detection accuracy but also latency, energy consumption, bandwidth usage, update success, model drift, privacy exposure, recovery behaviour, and operator decision performance. Tests comparing centralised, edge and federated methods should show when the benefits regarding privacy or resilience are sufficient to warrant the extra complexity. Furthermore, digital-twin research should investigate whether shared models actually lead to measurable improvements in multi-agency coordination during exercises and in real incidents rather than merely enhancing visualisation.

The studies referred to also make use of diverse methods, which makes it impossible to combine the results quantitatively. Lastly, the idea that the findings can be applied to cities in Saudi Arabia is based on evidence relating to the architecture and the context rather than being shown by means of a single national dataset.

9 Conclusion

When IoT-based physical-security systems are designed as cyber-physical service chains rather than simply as a set of connected devices, they are able to contribute to safer and more responsive smart cities. The evidence examined indicates that cameras, access control systems, environmental sensors, edge analytics and city platforms can enhance detection and situational awareness, while distributed trust, federated learning and privacy-preserving analytics can deal with certain integrity and data-governance issues. The systems' value, however, relies on secure device identity, software assurance, resilient communications, limited data use, model monitoring and well-tested response procedures.

The framework suggested for Saudi smart cities connects six key areas: risk-based zoning, trusted identity, edge-first intelligence, resilient interoperability, privacy-accountable governance and integrated response assurance. It is in line with the wider Vision 2030 context since it focuses on reliable urban services, scalable digital infrastructure and accountable implementation rather than on the adoption of technology for its own sake. Furthermore, the framework allows for a step-by-step implementation approach by first setting up an inventory and establishing trust, then integrating selected use cases, and only afterwards introducing advanced adaptive analytics once the basic controls can be measured.

The main takeaway is that physical security improves when uncertainty is eliminated without resulting in unmanaged dependence. A reliable system must be able to identify which device caused an event, assess whether the data are trustworthy, determine what analytical processes affected the interpretation, know who authorised the response, and also understand how the system acts when connectivity or the models fail. Future field research in Saudi Arabia should turn these principles into validated engineering limits and governance standards. In the meantime, the best design approach is disciplined integration—this means connecting only those components that can be identified, maintained, audited, minimised, and safely degraded.

References

- [1] Alqahtany AM. Smart cities as a pathway to sustainable urbanism in the Arab world: A case analysis of Saudi cities. *Sustainability*. 2025;17:1525. <https://doi.org/10.3390/su17041525>
- [2] Alamoudi AK, Abidoye RB, Lam TYM. Implementing smart sustainable cities in Saudi Arabia: A framework for citizens' participation towards Saudi Vision 2030. *Sustainability*. 2023;15:6648. <https://doi.org/10.3390/su15086648>

-
- [3] Alamoudi AK, Abidoye RB, Lam TYM. An evaluation of stakeholders' participation process in developing smart sustainable cities in Saudi Arabia. *Smart and Sustainable Built Environment*. 2024;13(5):1074–1095. <https://doi.org/10.1108/SASBE-08-2022-0170>
- [4] Mutambik I, Almuqrin A, Alharbi F, Abusharhah M. How to encourage public engagement in smart city development: Learning from Saudi Arabia. *Land*. 2023;12:1851. <https://doi.org/10.3390/land12101851>
- [5] Laufs J, Borrión H, Bradford B. Security and the smart city: A systematic review. *Sustainable Cities and Society*. 2020;55:102023. <https://doi.org/10.1016/j.scs.2020.102023>
- [6] Rizi MHP, Seno SAH. A systematic review of technologies and solutions to improve security and privacy protection of citizens in the smart city. *Internet of Things*. 2022;20:100584. <https://doi.org/10.1016/j.iot.2022.100584>
- [7] Rahman MA, Asyhari AT, Leong LS, Satria GB, Hai Tao M, Zolkipli MF. Scalable machine learning-based intrusion detection system for IoT-enabled smart cities. *Sustainable Cities and Society*. 2020;61:102324. <https://doi.org/10.1016/j.scs.2020.102324>
- [8] Ismagilova E, Hughes L, Rana NP, Dwivedi YK. Security, privacy and risks within smart cities: Literature review and development of a smart city interaction framework. *Information Systems Frontiers*. 2022;24(2):393–414. <https://doi.org/10.1007/s10796-020-10044-1>
- [9] Hernandez-Ramos JL, Martinez JA, Savarino V, Angelini M, et al. Security and privacy in Internet of Things-enabled smart cities: Challenges and future directions. *IEEE Security & Privacy*. 2021;19(1):12–23. <https://doi.org/10.1109/MSEC.2020.3012353>
- [10] Syed AS, Sierra-Sosa D, Kumar A, Elmaghraby A. IoT in smart cities: A survey of technologies, practices and challenges. *Smart Cities*. 2021;4(2):429–475. <https://doi.org/10.3390/smartcities4020024>
- [11] Bauer M, Sanchez L, Song J. IoT-enabled smart cities: Evolution and outlook. *Sensors*. 2021;21:4511. <https://doi.org/10.3390/s21134511>
- [12] AlSalem TS, Almaiah MA, Lutfi A. Cybersecurity risk analysis in the IoT: A systematic review. *Electronics*. 2023;12:3958. <https://doi.org/10.3390/electronics12183958>
- [13] Altulaihan E, Almaiah MA, Aljughaiman A. Cybersecurity threats, countermeasures and mitigation techniques on the IoT: Future research directions. *Electronics*. 2022;11:3330. <https://doi.org/10.3390/electronics11203330>
- [14] Paul R, Ghosh N, Sau S, Chakrabarti A, Mohapatra P. Blockchain based secure smart city architecture using low resource IoTs. *Computer Networks*. 2021;196:108234. <https://doi.org/10.1016/j.comnet.2021.108234>
- [15] Majeed U, Khan LU, Yaqoob I, Kazmi SMA, Salah K, Hong CS. Blockchain for IoT-based smart cities: Recent advances, requirements, and future challenges. *Journal of Network and Computer Applications*. 2021;181:103007. <https://doi.org/10.1016/j.jnca.2021.103007>
- [16] Ferreira CMS, Garrocho CTB, Oliveira RAR, Silva JS, Cavalcanti CFMDC. IoT registration and authentication in smart city applications with blockchain. *Sensors*. 2021;21:1323. <https://doi.org/10.3390/s21041323>
- [17] Badidi E, Moumane K, El Ghazi F. Opportunities, applications, and challenges of edge-AI enabled video analytics in smart cities: A systematic review. *IEEE Access*. 2023;11:80543–80572. <https://doi.org/10.1109/ACCESS.2023.3300658>
- [18] Zahra A, Ghafoor M, Munir K, Ullah A, Abideen ZU. Application of region-based video surveillance in smart cities using deep learning. *Multimedia Tools and Applications*. 2024;83:15313–15338. <https://doi.org/10.1007/s11042-021-11468-w>
- [19] Wolf K, Dawson R, Mills J, et al. Towards a digital twin for supporting multi-agency incident management in a smart city. *Scientific Reports*. 2022;12:16221. <https://doi.org/10.1038/s41598-022-20178-8>
- [20] Shorfuzzaman M, Hossain MS, Alhamid MF. Towards the sustainable development of smart cities through mass video surveillance: A response to the COVID-19 pandemic. *Sustainable Cities and Society*. 2021;64:102582. <https://doi.org/10.1016/j.scs.2020.102582>
- [21] Al-Huthaifi R, Li T, Huang W, Gu J, Li C. Federated learning in smart cities: Privacy and security survey. *Information Sciences*. 2023;632:833–857. <https://doi.org/10.1016/j.ins.2023.03.033>
- [22] Pandya S, Srivastava G, Jhaveri R, et al. Federated learning for smart cities: A comprehensive survey. *Sustainable Energy Technologies and Assessments*. 2023;55:102987. <https://doi.org/10.1016/j.seta.2022.102987>
-

-
- [23] Ghadi YY, Mazhar T, Shah SFA, et al. Integration of federated learning with IoT for smart cities applications, challenges, and solutions. *PeerJ Computer Science*. 2023;9:e1657. <https://doi.org/10.7717/peerj-cs.1657>
- [24] Hijazi NM, Aloqaily M, Guizani M, Ouni B, Karray F. Secure federated learning with fully homomorphic encryption for IoT communications. *IEEE Internet of Things Journal*. 2024;11(3):4289–4300. <https://doi.org/10.1109/JIOT.2023.3302065>
- [25] De Coninck S, Wang WC, Leroux S, Simoens P. Privacy-preserving visual analysis: training video obfuscation models without sensitive labels. *Applied Intelligence*. 2024;54:6041–6052. <https://doi.org/10.1007/s10489-024-05489-9>
- [26] Sarker IH. Smart city data science: Towards data-driven smart cities with open research issues. *Internet of Things*. 2022;19:100528. <https://doi.org/10.1016/j.iot.2022.100528>
- [27] Capodiecì N, Cavicchioli R, Muzzini F, Montagna L. Improving emergency response in the era of ADAS vehicles in the smart city. *ICT Express*. 2021;7(4):481–486. <https://doi.org/10.1016/j.icte.2021.03.005>
- [28] Yaacoub JPA, Noura HN, Salman O. Security of federated learning with IoT systems: Issues, limitations, challenges, and solutions. *Internet of Things and Cyber-Physical Systems*. 2023;3:155–179. <https://doi.org/10.1016/j.iotcps.2023.04.001>
- [29] Gugueoth V, Safavat S, Shetty S. Security of Internet of Things (IoT) using federated learning and deep learning-recent advancements, issues and prospects. *ICT Express*. 2023;9(5):941–960. <https://doi.org/10.1016/j.icte.2023.03.006>
- [30] Omrany H, Al-Obaidi KM, Hossain M, Alduais NAM, Al-Duais HS, Ghaffarianhoseini A. IoT-enabled smart cities: a hybrid systematic analysis of key research areas, challenges, and recommendations for future direction. *Discover Cities*. 2024;1:2. <https://doi.org/10.1007/s44327-024-00002-w>