International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Real-Time Fire & Smoke Detection from CCTV Footage: A Comprehensive Case Study on AI-Driven Visual Safety Systems

Kunal
Guru Gobind Singh Indraprastha University
Artificial Intelligence & Data Science

Abstract. This paper presents a study of AI-driven, real-time fire and smoke detection using CCTV footage.It positions this approach as a visual-first safety solutionthat is proactive. The analysis shows a shift from traditional, reactive point-sensors to an "Internet of Eyes," whichuses existing surveillance systems to provide critical situational awareness. The study looks at the evolution of detection algorithms, moving from earlier flawed models to modern deep learning object detectors like YOLO. Through three case studies—commercial, industrial, and smart city—this paper identifies practical deployment setups, such as Human-in-the-Loop and multi-sensor fusion. A significant contribution is the detailed discussion of the "false alarm problem" and its root cause in data scarcity, especiallyfor "incipient" fires. The paper also offers a technical analysis of key solutions, including Explainable AI, model optimization for edge computing, and the role of Generative AI. It concludes by summarizing the findings and suggesting a hybrid strategy for future development that emphasizes multi-sensor fusion, edge computing, and standardized regulatory frameworks.

Index Terms: AI-Driven Fire Detection, CCTV, Computer Vision, Deep Learning, Explainable AI, Model Optimization

I.INTRODUCTION

Fire hazards are among the most serious threats to human life, property, and the environment. Globally, uncontrolled fires cause billions of dollars in economic losses each year and lead to thousands of preventable deaths and injuries. Early detection has always been vital for fire safety. The time between the initial stage of a fire and its full growth stage is often just minutes. Reducing the "Time-to-Detect" and thus the "Time-to-Respond" is the most important factor in determining the outcome of a fire event.

Traditional fire safety systems mainly rely on smoke and heat sensors. These systems fall under the category of the Internet of Things. They function as point-based, discrete devices that report simple binary states, such as "Smoke: True/False." However, the last ten years have seen a massive increase in high-definition Closed-Circuit Television cameras, which create a widespread "Internet of Eyes." These millions of cameras, already in use for security and operational purposes, are a largely untapped resource for safety. This report focuses on applying Artificial Intelligence to this existing "Internet of Eyes" to create a proactive visual safety network.

Conventional fire detection systems, such as ionization, photoelectric smoke detectors, and heat sensors, are reactive and spatially limited. Their drawbacks are well- recognized:

• Physical Contact: They need smoke or heat particles to drift to the sensor.
• Latency in Large Spaces: In places with high ceilings, smoke can cool and stratify, taking vital minutes to reach ceiling-mounted sensors, if it reaches them at all.
• Inapplicability Outdoors: They completely fail in outdoor environments, like forests, industrial yards, and city streets.
• Lack of Context: Triggered sensors give no visual information. A "Zone 4" alarm doesn't provide details about the threat's scale, source, or nature. Is it a smoldering wastebasket or a full-blown fire? Is it a false alarm triggered by steam? This lack of situational awareness slows down an effective response.

The goals of this report are fourfold:

1. To examine the foundational concepts and algorithmic evolution of vision-based fire detection, from early color-space models to modern deep learning systems.
2. To analyze real-world case studies of these systems in different environments to understand deployment setups and results.
3. To identify and discuss the major technical, operational, and financial challenges associated with this technology, including a detailed look at the false alarm issue.
4. To offer practical recommendations for stakeholders, including engineers, regulators, and end-users, to tackle these challenges and improve the reliability and adoption of AI-driven fire detection.

This report is structured into six chapters. It emphasizes AI, data science, and system-engineering aspects of video- based fire detection. It will not cover the chemical engineering of fire suppression, like sprinkler systems, or the physics of fire dynamics, except where relevant to detection principles.

II. RESEARCH ELABORATIONS: LITERATURE REVIEW AND RESEARCH GAPS

The academic literature shows a clear progression in fire detection technology, transitioning from physical sensors to intelligent, vision-based systems.

A. Overview of Fire Detection System Evolution

• Classical and Sensor-Based Methods (Point Sensors): Early literature focuses on the chemistry and physics of ionization and photoelectric sensors. Key research in this area looks at the fluid dynamics of smoke, optimal sensor placement, and the challenge of smoke stratification in tall spaces, a primary failure point that vision-based systems aim to address.
• Early Vision-Based Algorithms (Non-AI): The initial attempts from the late 1990s and 2000s used traditional image processing techniques.
o Color-Space Models: These algorithms operated under the assumption that fire occupies a specific range in the RGB or HSL color space. The models flagged any large cluster of red, orange, or yellow pixels.
o Motion Analysis: To reduce false positives from static red objects, motion analysis like background subtraction or optical flow was incorporated. The theory held that fire "flickers."

o Failure Point: These non-AI models were often unreliable. They caused a high rate of false alarms triggered by sunsets, red lighting, people in red shirts, or rustling leaves.

• The Deep Learning Revolution: Image Classification: With the rise of deep learning around 2012, research shifted toward Convolutional Neural Networks. Early studies tested architectures like AlexNet, VGG-16, and ResNet. A pre-trained model would be fine-tuned on a fire image dataset. This method treats the entire video frame as a single image and provides a binary classification: "Fire" or "No-Fire." However, it cannot detect small fires in larger scenes, such as a wastebasket in an office. It also struggles to localize the fire or differentiate it from a picture of fire on a TV screen in the background.

• **Real-Time Object Detection (YOLO, SSD):** This is the current best practice in the literature. Models like YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and Faster R-CNN are used. These models not only classify but also localize. They scan the image and draw bounding boxes around every object they recognize, like "person," "car," "fire," and "smoke." This capability allows for detecting multiple, small incipient fires in complex scenes and provides coordinates for the fire. These coordinates can guide PTZ cameras or aim automated suppression systems.B. Key Problems Identified in Literature

- • False Positive vs. False Negative Trade-off: This is the main focus. Literature often points out the problem of "hard negatives," which are non-fire objects that look just like fire or smoke. These include fog, steam, welding sparks, industrial discharge, dust clouds, and sunsets. Models must be trained on millions of these "hard negative" images to learn what to ignore.
-
- • Environmental Robustness: A model trained on clear, daytime video, which makes up most public datasets, will fail in real-world situations. The literature shows significant performance drops in Low-Light & Night, Weather (like rain, snow, fog), and Obscuration (partially blocked views).
-
- • Computational and Latency Constraints: Running a complex model like YOLOv8 on over 100 high-definition (1080p) video streams at 30 frames per second is a huge computational challenge. It requires expensive, power-hungry GPU servers. This creates a dilemma: complex models are more accurate, but simpler models are easier to implement at scale.
-
- C. Theoretical Foundations Supporting Vision-Based Detection
- • Convolutional Neural Networks (CNNs): These are the basis of modern image analysis. The main idea is that the network learns hierarchical features. Early layers detect simple edges, corners, and color blobs. Mid layers combine edges to identify textures (like the "flickering" texture of smoke). Deep layers merge textures to recognize "objects" (for example, "this mix of color, flicker, and wispiness is 'Fire'").
-
- • Temporal Analysis (RNNs, 3D-CNNs): This is a more advanced topic. Analyzing static images is flawed because a photo of a sunset can resemble fire. The behaviors are different. Recurrent Neural Networks (RNNs/LSTMs) look at sequences of frames. They can learn that "fire flickers randomly, smoke rises and spreads," while a red light stays still and fog moves uniformly. 3D-CNNs view video as a 3D space (width x height x time) and perform 3D convolutions, learning spatio-temporal features at the same time. This approach is computationally demanding but very effective.
-
- D. Underlying Problems and Research Gaps
- • **The "Incipient Fire" Data Scarcity Problem:** This is the biggest data issue. For clear safety reasons, researchers can't easily or ethically start thousands of real fires in actual office buildings to record video of the first 60 seconds. Consequently, the majority of datasets focus on large, established fires, making the AI good at spotting big blazes but poor at identifying early threats.

• **Lack of Standardized Benchmarks**: There is no equivalent to "ImageNet" for fire detection. Many studies refer to small, proprietary, or un-curated public datasets. This disparity makes it hard to compare different algorithms fairly.

• **Research Gaps Identified**: The review points out several gaps:

- Robust Multi-Sensor Fusion: There's a lack of proven systems that intelligently combine video, thermal, and gas sensor data in real-time.
- Generative AI for Data Augmentation: There is a significant gap in using GANs or diffusion models to create realistic, synthetic videos of "incipient" fires to tackle the data scarcity problem.
- Edge Model Optimization: More research is needed on "model quantization" and "pruning" to develop accurate, energy-efficient models that can operate on low-power "edge" devices.
- Adversarial Robustness: Very little research exists on thecybersecurity of these systems (for example, "blinding" the AI to a real fire).
- Regulatory Standards: There is a complete absence of aregulatory framework (such as from UL, NFPA, or BIS)to certify an AI algorithm for life safety.
-

III.RESULTS OR FINDING: CASE STUDY ANALYSIS

- AI-powered fire detection has advanced from academic settings to critical real-world applications. This chapter examines three different cases to understand their practical setups, workflows, and outcomes.
- **A. Case 1**: Commercial High-Rise & Public Malls
- • Deployment Context: Large public areas (like malls, lobbies, and parking garages) with over 100 existing CCTV cameras. The issue is that traditional ceiling sensors are slow in high-atrium spaces, and a fire can lead to panic.
-
- • Architecture: Generally, this involves a software-only upgrade. An "AI analytics" server is set up in the central security or operations center (SOC). This server connects to the video feeds from the building's existing Video Management System (VMS), where the AI models operate.
-
- • Operational Framework: The "Human-in-the-Loop" (HITL) approach: The AI doesn't trigger the main fire alarm. This crucial design choice prevents a single false alarm from causing mass evacuations. The AI identifies a fire signature, sends a high-priority alert to the operator's dashboard, and the VMS shows the live feed with a bounding box. The human operator checks the threat in under 10 seconds. If confirmed, the operator triggers the response protocol. If false, the operator marks it as a "false positive."

- •Documented Impact: In pilot tests, this HITL system provided verified alarms 1 to 3 minutes faster than traditional smoke detectors in high-ceiling settings. It also effectively prevents nuisance alarms from sources such as cooking fumes or steam.

- **B. Case 2**: High-Risk Industrial (Lithium-ion Battery Storage)
- • Deployment Context: A "Battery Energy Storage System" (BESS) room. Lithium-ion battery fires are particularly hazardous due to the risk of "thermal runaway," which is an explosive chemical reaction. Detecting the early stage of "off-gassing" or "overheating" is the only safe mitigation.

- • Architecture: This requires a high-cost, high-reliability new system rather than a software upgrade. It integrates various sensor types at the "rack" level.

- • Operational Framework: Multi-Sensor Fusion: This system combines three data streams into a central "safety PLC" (Programmable Logic Controller):
- 1. Thermal Cameras: These continuously monitor the surface temperature of each battery module.
- 2. Visual-AI Cameras: These look for early signs of "off-gassing" smoke.
- 3. IoT Gas Sensors: These test the air for specific chemical compounds (like H_2, CO, VOCs).

- The system uses an automated "AND/OR" logic gate to issue alerts or an automatic response. The response doesn't wait for human input; it de-energizes the rack and activates a specialized suppression system.

- • Documented Impact: In reported instances, this fusion system detected and isolated a single failing battery module during its "off-gassing" phase, preventing a multi-million dollar, unquenchable fire. It effectively identifies fire precursors.

- **C. Case 3**: Smart City and Environmental Monitoring
- • Wildfire Detection in Large-Area Forests: AI-powered, solar-charged, 4G/5G-connected Pan-Tilt-Zoom (PTZ) cameras are mounted on existing fire watchtowers. The AI, often running on an "edge" device, continuously scans a 360-degree view, trained to recognize the "plume" signature of new smoke. When it detects a plume, it calculates the (GPS) location and alerts a central command center. Reported services can detect wildfires hours before a public "911" call.

- • Transportation: Tunnel and Highway Safety: AI analytics operate on current traffic-monitoring cameras. The AI is trained for "multi-threat" detection (like stopped vehicles, debris, and fire/smoke). A vehicle fire in a tunnel is a high-casualty scenario. The AI detects the fire/smoke within seconds and automatically alerts the traffic center, activates ventilation fans, and turns on digital signage to close the tunnel. This cuts the time needed to activate life-saving systems from minutes to seconds.

- D. Comparative Analysis of Case Studies
- Feature Case 1: Commercial Case 2: Industrial (BESS) Case 3: Smart City (Wildfire)
- Primary Goal Life Safety (Panic Prevention) Asset Protection (Catastrophe Prevention) Early Warning & Rapid Response
- Key Technology Software-Only AI on VMS Multi-Sensor Fusion (Thermal, Gas, AI) Edge AI on PTZ Cameras
- Response Human-in-the-Loop (HITL) Fully Automated (PLC Logic) HITL (Command Center)
- Main Challenge False alarms from the public High cost; sensor reliability Vast area; environmental (fog)
- Key Metric Time-to-Human-Verification Time-to-Pre-Fire-Detection Time-to-Plume- Detection

IV. DISCUSSION

While the case studies show great potential, widespread and reliable use faces significant challenges. This chapter combines insights from the literature review and case studies to analyze these key issues in detail, focusing on problems that need to be addressed for this technology to evolve from a "promising add-on" to a "life-safety- certified" standard.

- A. Detailed Analysis of the Major Problems
- • The False Alarm Problem: Sensitivity vs. Specificity: This is the most critical barrier to adoption. It represents a classic trade-off between sensitivity and specificity.
- o High Sensitivity (Good): A model with high sensitivity can pick up the slightest wisp of smoke (a true positive).
- o Low Specificity (Bad): To achieve this, the model becomes less specific and flags fog, steam, welding sparks, and dust clouds as well (false positives).
- o The Consequence ("Alarm Fatigue"): A system that raises false alarms 20 times a day is worse than having no system at all. Operators will start to ignore, silence, or distrust the alerts, leading to a psychological phenomenon called alarm fatigue. This makes the system a dangerous liability.

- •ModeElnvironmental Variability and Generalization: An AI model is only as good as the data it was trained on. Most academic datasets are clean (daytime, high-resolution). The messy reality includes low-light/night conditions (video noise), blocked views, dynamic lights, and weather challenges (like rain, fog). A model trained on clean data will likely fail in this messy reality.

- • The "Black Box" Problem and Explainability (XAI): This issue is crucial for regulatory approval. When a traditional detector activates, it's clear why. But when a deep learning model triggers an alarm, it's a "black box." If a model produces a false alert, engineers may be unable to figure out why it failed. This drives research into Explainable AI (XAI), focused on making models easier to understand.
- o Key Technique: Grad-CAM (Gradient-weighted Class Activation Mapping): Grad-CAM is an important XAI technique that generates a "heat map" overlay, showing the specific areas of the image that influenced the model's decision.
- o How it Works: The model analyzes an image and makes a prediction (like "Fire"). Grad-CAM examines the last convolutional layer (where the most detailed features are found) and calculates

the "gradient" (importance) of each feature related to the "Fire" prediction. It uses these weights to create the heat map.

- o Impact: This is crucial for debugging and building trust. If the model incorrectly flags "fog" as "smoke," the heat map reveals what the AI was focusing on (like the texture of the fog). An auditable record of these heat maps would be vital for certification.

• **Data Scarcity** and the High Cost of "Ground Truth": The "Incipient Fire" data shortage is a serious problem. You cannot train an AI to find something without examples. To create a useful dataset, you must manually draw bounding boxes, known as "annotation," frame by frame. This process is slow, costly, and subjective. Because good data is hard to come by, many systems rely on limited, proprietary datasets. This makes their performance uncertain and may introduce bias.

B. Alternative Solutions to the Identified Problems

• Solution A: Multi-Sensor Fusion: As shown in Case 2, this approach combines data from video, thermal imaging, and Internet of Things sensors like gas and temperature sensors.
- Pros: It offers the highest reliability. It significantly reduces false alarms; for instance, fog may look like smoke but is actually cold and chemically inert. It can identify "hot spots" before a fire starts.

- Cons: It involves high costs and complexity. It is not just a simple software upgrade. A new challenge is temporal synchronization, as it involves fusing 30-FPS video with a gas sensor report that takes 10 seconds.

• **Solution B**: Advanced AI (Temporal Analysis & Generative AI): This approach makes one sensor, the camera, much more intelligent.

- Method 1: Temporal Analysis: It uses models like LSTMs and 3D-CNNs to analyze video, not just images, to learn behaviors like fire flickers and smoke rising.
- Method 2: Generative AI (GANs & Diffusion): This tackles the "Incipient Fire" data scarcity issue. Instead of capturing real data, we can create it. A Generative Adversarial Network, or GAN, consists of two competitive models: a Generator that creates fake fire and a Discriminator that identifies the fakes. They continue to improve until the Generator produces realistic fire. Diffusion Models, like DALL-E, can also be prompted to create specific images, such as "a small wisp of smoke rising from a wastebasket in a dark office."

- Pros: It can use existing hardware. Generative AI could potentially create an unlimited and diverse dataset of "incipient fire" and "hard negatives" that are impossible to film in real life.
- Cons: It has a high computational cost. Generative AI is promising but is still in research and development. Models that are only trained on synthetic data often struggle to generalize.

• **Solution C**: Edge Computing Architecture (Edge-to-Cloud):

This architectural solution addresses the scalability issue. Instead of sending 1,000 video streams to a central server, processing is distributed. A small, low-power "Edge" computer is installed with each camera. This requires model optimization.

- Key Technique: Pruning & Quantization: These two techniques help shrink a large "lab" model to fit on an "edge" device:
1. Pruning: This involves removing unnecessary neural connections from the model, making it sparser and faster.
2. Quantization: This reduces the precision of the model's weights from 32-bit floating-point numbers to smaller 8-bit integers. This change makes the model four times smaller and much faster on modern edge chips.

- Pros: It is highly scalable and resolves bandwidth and processing bottlenecks. It offers low latency and preserves privacy.
- Cons: There is hardware cost at each camera. The optimization process can slightly lower accuracy if not done carefully.

• **Solution D:** Human-in-the-Loop (HITL) Verification: As seen in Case 1, this is a procedural solution. It recognizes that AI can make mistakes and includes a human operator as the final decision-maker.

- Pros: It is 100% effective in preventing AI false positives from triggering evacuations. It has low operational costs and is the standard for commercial deployments.
- Cons: It introduces a delay of 10 to 30 seconds for the human operator. It serves as a temporary fix that does not address the root problem and adds cognitive load on the operator.

C. Comparative Evaluation and Identification of Most Suitable Approach

Solution | Solves False Alarms? | Solves Scalability? | Cost/Complexity | Best For...
A: Multi-Sensor Fusion | Yes (Best) | No (Worst) | Very High | High-Risk Industrial
B: Advanced AI (Temporal) | Yes (High Potential) | No (High Compute) | High (Software) | Future-Proofing
C: Edge Computing | No (It's an architecture) | Yes (Best) | Medium (Hardware) | Large-Scale Deployments
D: Human-in-the-Loop | Yes (Procedurally) | Yes (Uses humans) | Low (Operational) | Commercial / Public

No single solution is enough. The analysis shows that a mixed, phased strategy is the most practical and effective approach:
- Baseline (Today): The best option for most deployments is Solution D, Human-in-the-Loop, combined with Solution C, Edge Computing. This combination is scalable, cost-effective, and reliable.
- R&D Focus (Tomorrow): All research investment should go into Solution B, Advanced AI (Temporal & GANs). The aim is to enhance AI intelligence to reduce the number of false positives directed to human operators.
- High-Risk (Specialized): Solution A, Multi-Sensor Fusion,

should be reserved for high-stakes, environments, such as battery energy storage systems, data centers, and chemical plants, where the high cost is justified by the significant risk.V.

CONCLUSION

This seminar report explored the emerging field of autonomous, real-time fire and smoke detection using CCTV footage. The main goal was to close the gap between its theoretical promise and its practical, real-world use. The central idea is that while AI-driven visual detection represents a significant change, it is not always reliable. Its success heavily relies on a tailored approach that considers its current limitations. The study confirms that AI-driven visual detection is a major upgrade from traditional systems. Traditional point sensors provide a simple "Smoke: True/False" answer. In contrast, visual AI offers detailed information and context. The ability to see the threat's location, size, and characteristics instantly changes the response from reactive to proactive. The technology has advanced significantly, evolving from early, flawed color- space models to effective deep learning frameworks. Current state-of-the-art includes object detection models like YOLO. The latest research focuses on Temporal Analysis using 3D- CNNs and LSTMs to understand fire and smoke behavior. A key finding is the principle that there is no one-size-fits-all solution. The case studies demonstrated that the ideal system design depends entirely on the specific risk and environment:

• Commercial (Low-Risk, High-Traffic): The Human-in-the-Loop (HITL) model is the only practical option.

• Industrial (High-Risk, High-Consequence): A high-cost, high-reliability Fully Automated Multi-Sensor Fusion system is the only logical choice, as it targets pre-fire conditions.

• Smart City (Vast-Area, Low-Power): Edge Computing serves as the supporting structure.

The biggest technical and commercial challenge is the problem of "false positives." This issue acts as the technology's weak point. It stems from the difficulty of "hard negatives," such as fog, steam, dust, and welding sparks. This technical failure contributes to the human-factors issue of "alarm fatigue," where operators become desensitized to alerts, making an expensive system ineffective.

The root cause of false alarms is the lack of high-quality, diverse training data. Specifically, the issue of "Incipient Fire" data scarcity creates a cycle: 1) We lack data on the first 60 seconds of a fire; 2) Consequently, models are under- trained to detect them; 3) To adjust, engineers increase model sensitivity; 4) This results in many false alarms; 5) Trust erodes. Breaking this cycle is essential.

The most reasonable and realistic conclusion from this analysis is that a hybrid strategy is necessary. The best approach combines:

1. Procedural Safety: Use HITL as a baseline.
2. Architectural Scalability: Use Edge Computing.
3. R&D Focus: Invest in Advanced AI, particularly in Temporal and GANs.
4. Specialized Solution: Reserve Multi-Sensor Fusion for high-risk situations.

While this report is thorough, it faces significant limitations. Most commercial systems keep their true field performance data as a trade secret. The AI field is growing rapidly, which means any technology-specific analysis is just a snapshot in time. Finally, the absence of an independent certification standard means there is no third-party testing and validation of vendor claims.

The future of this technology will likely unfold in three phases:

• Near-Term (1-3 Years): Focus on optimizing and improving robustness, with lighter models and better HITL workflows.

• Mid-Term (3-5 Years): Marked by building trust and integration. The first AI Safety Certified models will appear, and Multi-Sensor Fusion will set the standard. The use of GAN-generated synthetic data will become common.

• Long-Term (5-10+ Years): This phase will involve autonomy and ecosystem integration, including AI-to-AI communication with drones or suppression systems, digital twin integration to model fire spread, and On-Device AI processing on cameras.

AI-driven fire detection is one of the most significant advancements in life-safety technology in decades. It can potentially save thousands of lives and prevent billions in damage. However, this study concludes that it is not a simple "magic bullet" or a plug-and-play solution. It is a complex, powerful tool that requires a deep understanding of its limitations. The key to its widespread and reliable use lies not in a single groundbreaking algorithm but in a collective effort involving better data, smarter systems, multi-sensor fusion, and most importantly, developing strong, independent standards to build trust.

VI. RECOMMENDATIONS

Based on the literature review, case study analysis, and discussions about the challenges, this chapter presents a set of actionable recommendations. These aim to guide stakeholders—including engineers, research institutions, regulators, and end-users—in improving the safety, reliability, and broader adoption of autonomous fire detection systems. The challenges identified are multifaceted, affecting technology, data, and policy. Therefore, a comprehensive strategy is necessary. **A. Recommendation 1: Prioritize Multi-Sensor Fusion** The industry should move beyond visual-only systems for all new life-safety-critical installations. A "fusion-first" approach, combining visual AI with thermal sensors and IoT gas/chemical sensors, should become the new standard.

• Actionable Steps: Support research into "fusion logic"—AI models that intelligently weigh and cross-check data from different sensor streams. Manufacturers should be encouraged to create "tri-sensor" devices that combine visual, thermal, and VOC sensors to lower installation costs.

• Key Stakeholders: AI/VMS Companies, Sensor Manufacturers, Industrial End-Users. **B. Recommendation 2: Launch the "Fire-Incipient-Net" Initiative** Academic institutions and industry groups should work together to create a large-scale public and standardized benchmark dataset for fire detection, focusing on "incipient" fires and "hard negatives." • Actionable Steps: Fund and conduct controlled burns in various, equipped environments to gather high-quality, multi-sensor "incipient" data. Initiate a "grand challenge" to use GANs or diffusion models to generate one million photorealistic, annotated video clips of early-stage fires and false-positive sources, like fog and steam. Make this "Fire- Incipient-Net" dataset publicly available, similar to ImageNet, with a standardized scoring system for comparing new algorithms. • Key Stakeholders: Academic Researchers, AI Companies (Google, Meta, OpenAI), Public Fire Safety Institutes. **C. Recommendation 3: Create National Regulatory and Certification Frameworks** National and international safety organizations, such as NFPA, UL, BIS, and ISO, need to quickly develop new standards and certification protocols for AI-based life safety systems. This could involve creating a new standard like "UL 864-AI." • Actionable Steps: The framework must outline what to certify. This includes:

o Time-to-Detect (TTD): A standardized test for detecting fires of various sizes.

o False Alarm Rate (FAR): For example, there should be "no more than 1 false alarm per 1000 operational hours" when tested against the "hard negative" dataset.

o Robustness: Tests should check performance in low-light, occluded, and adverse weather conditions.

o Explainability (XAI): Certified systems must provide logs for "explainability." This includes features like Grad-CAM heatmaps for all alarm events.

• Key Stakeholders: Regulatory Bodies (NFPA, UL, BIS), Government, Insurance Industry.

Recommendation 4: Mandate Cybersecurity & Resilience Standards

Alongside safety certification, regulatory bodies must establish cybersecurity standards to safeguard these systems from attacks. • Actionable Steps: Require testing against known attack methods, such as "patch" attacks meant to conceal a fire from the AI. Require end-to-end encryption for all video feeds. Certified systems should include a "heartbeat" or "watchdog" mechanism. If the AI model crashes, it must trigger a "System Fail" alert. • Key Stakeholders: Cybersecurity Agencies, AI Companies, Regulatory Bodies. **Implementation Plan and Cost Analysis**

• Implementation Plan: The plan should take a phased approach:

o Short-Term (0–12 Months): Launch pilot programs in non-critical settings using a strict HITL model. Form the "Fire-Incipient-Net" consortium. Start data collection.

o Medium-Term (12–36 Months): Release the first version of the dataset and benchmark. Publish the initial draft of the "AI-Safety Certification" for Edge-based HITL systems. Start pilot deployments of Multi-Sensor Fusion systems.

o Long-Term (36+ Months): The first "UL-AI" certified systems will be commercially available. Insurance companies will start offering premium reductions. Deployments will become standard practice. •Estimated Costs: (Based on a medium-sized commercial building, e.g., 50-camera deployment).

REFERENCES

(References are formatted from case study 001.docx to match the IJSRP "Periodical style" and "Book style" examples.) [1] S. Khan, A. Gupta, and R. Mehta, “A deep learning approach for real-time fire and smoke detection in smart city surveillance,” IEEE Transactions on Industrial Informatics, vol. 18, no. 5, pp. 3122–3131, 2022. [2] P. Sharma, V. Singh, and M. Kumar, “A review of traditional vs. AI-based fire detection systems: Challenges and opportunities,” Fire Safety Journal, vol. 129, p. 103522, 2022. [3] Y. Li and R. Zhao, “YOLO-Fire: A
lightweight and real-time fire detection model for complex environments,” Proc. International Conference on Image Processing (ICIP), pp. 1– 5, 2023. [4] Olea Sensor Networks, “AI-based visual analytics for safety and security,” Olea White Paper, 2023. [5] D. T. Tran, H. Kim, and J. Kim, “Temporal-aware deep neural networks (3D-CNNs) for video-based smoke detection,” Journal of Real-Time Image Processing, vol. 17, no. 4, pp. 789–802, 2021. [6] V. A. M. et al., “Multi-sensor fusion for fire detection: A systematic review of thermal, visual, and gas sensor integration,” Sensors, vol. 20, no. 14, p. 3920, 2020. [7] National Fire Protection Association (NFPA), NFPA 72: National Fire Alarm and Signaling Code. NFPA Publications, 2022. [8] M. G. Kim, B. K. Lee, and S. H. Park, “Robust fire detection in low-light and obscured conditions using generative adversarial networks (GANs),” IEEE Access, vol. 9, pp. 71002–71015, 2021. [9] D. C. C. M. and A. T. S., “A
public dataset for computer- vision based fire and smoke detection,” Fire Technology, vol. 57, pp. 1123–1140, 2021.

[10] Actuate AI, “AI-powered threat detection for commercial properties,” Actuate Inc. Technical Brief, 2023.

[11] S. J. Yoo, H. W. Kim, and D. S. Kim, “Cybersecurity for AI- abnasded surveillance systems: Vulnerabilities countermeasures,” IEEE Security & Privacy, vol. 19, no. 2, pp. 40–47, 2021.

[12] R. J. R. and P. Singh, “Explainable AI (XAI) for safety- critical systems: A review of Grad-CAM in fire-detection,” Journal of AI Ethics, vol. 4, pp. 201–215, 2024.

[13] Pano AI, “AI-based wildfire detection and response,” Pano AI Case Studies, 2023.

[14] M. T. H., L. K. M, and D. S. P, “Model quantization and pruning for efficient fire-detection on edge-TPU,” Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2023.

[15] D. G. L. and S. M. F., “The psychology of alarm fatigue in security operations centers,” Journal of Human-Computer Interaction, vol. 38, no. 3, pp. 245–267, 2022.