International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 5, Sep-Oct 2025
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Intelligent Fire Detection Systems Using Deep Learning and Multi-Sensor Data Fusion: A Comprehensive Review

Gwamzhi Ponsah Emmanuel*, Jibril Danladi Jiya**, Ya'u Haruna Shuaibu***, Joseph Timothy Abraham****, Rebecca Ogwu*****, Lynda Elesa Bitrus******.
*(Zonal Advanced Space Technology Applications Laboratory (ZASTAL), P.M.B 47, Pajat, Langtang North, Plateau State, Nigeria. Email: eponsah@gmail.com)
** (Department of Mechatronics, Abubakar Tafawa Balewa University, P.M.B. 0248, Bauchi, Nigeria. Email: jdjiya@atbu.edu.ng)
*** (Department of Electrical and Electronics, Abubakar Tafawa Balewa University, P.M.B. 0248, Bauchi, Nigeria. Email: jdjiya@atbu.edu.ng)
**** (Department of Electrical and Electronics, Abubakar Tafawa Balewa University, P.M.B. 0248, Bauchi, Nigeria. Email: Timothyabrahamjoseph@gmail.com)
***** (National Centre for Remote Sensing, Rizek Village, Jos East, Nigeria. Email: beckymbaya@gmail.com)
****** (National Productivity Centre, No 7 Cape Town, Wuse, Abuja, Nigeria. Email: lyndabitrus@gmail.com)

--*************************----------------------------------

Abstract:

Fire detection technology has evolved significantly from basic mechanical systems to sophisticated intelligent platforms integrating artificial intelligence and multi-sensor fusion. This comprehensive review examines the state-of-the-art in fire detection systems, with particular emphasis on Convolutional Neural Network (CNN) approaches, multi-sensor data fusion methodologies, and Internet of Things (IoT) integration. Traditional single-sensor fire alarm systems suffer from high false alarm rates (up to 95% in some jurisdictions), detection latency issues, and limited discriminative capability between genuine fire events and benign environmental conditions. Recent advances in deep learning, particularly CNN architectures such as You Only Look Once (YOLO), Residual Networks (ResNet), MobileNet, and specialized fire detection models, demonstrate substantial improvements in detection accuracy while maintaining real-time performance capabilities. Multi-sensor data fusion techniques combining temperature, smoke, flame, and visual sensors through decision-level, feature-level, and signal-level integration strategies offer enhanced robustness and reduced false alarm rates compared to single-sensor systems. Kalman filtering and Bayesian fusion methods provide mathematical frameworks for optimal sensor integration under uncertainty. However, significant challenges remain in deploying sophisticated machine learning models on resource-constrained embedded platforms, necessitating model optimization through quantization, pruning, and efficient architecture design. This review synthesizes current research across fire detection methodologies, critically analyses their strengths and limitations, identifies persistent research gaps including standardized evaluation frameworks and real-world validation needs, and proposes directions for future investigation. The integration of embedded machine learning, IoT connectivity, and intelligent sensor fusion represents a promising pathway toward next-generation fire safety infrastructure capable of reliable, autonomous operation across diverse building environments.

Keywords — **Fire Detection, Convolutional Neural Networks, Data Fusion, Embedded Systems, False Alarm Reduction.**

--*************************----------------------------------

I. INTRODUCTION

Fire incidents constitute a persistent global threat with devastating consequences for human safety, property, and economic stability. Global fire statistics indicate approximately 180,000 fire-related deaths annually, accompanied by billions of dollars in economic losses worldwide [1]. The fundamental challenge in fire detection lies in achieving rapid, accurate identification of genuine fire events while minimizing false alarms that undermine system credibility and emergency response effectiveness. The National Fire Protection Association reports that approximately 95% of fire alarm activations in Germany represent false alarms rather than genuine fire incidents [2], reflecting a systemic problem undermining public confidence in fire safety infrastructure.

Traditional fire detection systems rely predominantly on single-sensor modalities—ionization detectors, photoelectric smoke detectors, or heat detectors—each exhibiting inherent limitations and trade-offs. Ionization detectors respond effectively to fast-flaming fires but perform poorly with smouldering fires, while photoelectric detectors exhibit the inverse performance characteristics. Heat detectors avoid many false alarm triggers but may activate too late to enable effective early warning and occupant evacuation. These fundamental limitations have driven research toward more sophisticated detection methodologies integrating multiple complementary information sources.

Recent technological advances in artificial intelligence, particularly deep learning and convolutional neural networks, have demonstrated transformative potential for visual fire detection applications. CNN-based approaches achieve substantially higher accuracy than traditional image processing techniques while maintaining feasible computational requirements for real-time deployment [3]. The original YOLO architecture introduced by Redmon et al. [4] revolutionized object detection by enabling single-pass real-time detection at 45 frames per second. Subsequent architectures including ResNet [5] with residual learning frameworks and MobileNet [6] with depthwise separable convolutions have further advanced efficient deep learning for resource-constrained applications.

Concurrently, the proliferation of Internet of Things (IoT) technologies enables networked, intelligent fire detection systems capable of remote monitoring, coordinated response, and data-driven optimization. Multi-sensor data fusion techniques, grounded in mathematical frameworks such as Kalman filtering and Bayesian inference, provide optimal methods for combining uncertain measurements from heterogeneous sensors [7]. This review examines the current state of fire detection technology across these multiple dimensions, critically evaluates competing approaches, identifies persistent research gaps, and proposes directions for future investigation toward increasingly intelligent, reliable, and effective fire safety systems.

II. TRADITIONAL FIRE DETECTION SYSTEMS AND LIMITATIONS

Fire alarm systems comprise multiple integrated components including detection sensors (smoke, heat, flame), control panels, notification appliances, and communication systems. The historical evolution progressed from basic mechanical systems exploiting thermal expansion principles to electronic sensors utilizing ionization, photoelectric, and thermal detection mechanisms [1].

Ionization smoke detectors employ radioactive material (typically Americium-241) to ionize air within a detection chamber, establishing a small electric current between electrodes [1]. When smoke particles enter, they disrupt the current flow, triggering alarm activation. These detectors exhibit high sensitivity to small smoke particles from fast-flaming fires but demonstrate poor response to smoldering fires and susceptibility to false alarms from cooking aerosols, humidity, and dust [2].

Photoelectric smoke detectors operate on light scattering principles, projecting a light beam within a sensing chamber configured such that light does not normally reach a photoelectric sensor [1]. When smoke particles scatter light toward the sensor, alarm activation occurs. These excel at detecting large smoke particles from smoldering fires but exhibit reduced sensitivity to fast-flaming fires and remain vulnerable to false alarms from dust accumulation and high humidity [2].

Heat detectors respond to ambient temperature increases through either fixed-temperature activation (typically 57-68°C) or rate-of-rise activation (typically 6-8°C per minute) [1]. They offer substantially lower false alarm rates, making them suitable for environments where smoke detector false alarms are problematic. However, they inherently respond more slowly because significant temperature elevation requires substantial fire development, potentially delaying critical early warning [2].

Flame detectors identify fires through electromagnetic radiation emitted by flames, operating in ultraviolet (190-260 nm) or infrared (4.3-4.4 µm) spectral regions [1]. They provide extremely rapid response to open flames but require direct line-of-sight, perform poorly with smoldering fires, and may experience false alarms from sunlight or welding operations [2].

The fundamental challenge with conventional single-sensor systems lies in their susceptibility to environmental interference and false alarms triggered by non-fire phenomena [2]. Each sensor type optimizes detection for specific fire characteristics while exhibiting reduced sensitivity to others, creating inherent trade-offs between detection speed, sensitivity, and false alarm rates. This limitation has motivated extensive research into multi-sensor approaches providing cross-validation of fire indicators through integration of complementary detection modalities.

III. CONVOLUTIONAL NEURAL NETWORKS FOR FIRE DETECTION

A. Theoretical Foundations and Key Architectures

Convolutional Neural Networks represent the dominant paradigm for visual pattern recognition tasks, including fire detection in images and video streams [8]. CNNs automatically learn hierarchical feature representations through successive convolutional layers, progressively extracting increasingly abstract features from low-level edges and textures to high-level semantic concepts enabling accurate object classification and localization, eliminating the need for manual feature engineering that characterized traditional computer vision approaches [3].

YOLO (You Only Look Once) architecture introduced by Redmon et al. [4] fundamentally transformed object detection by framing it as a regression problem rather than classification, dividing input images into a grid and predicting bounding boxes and class probabilities directly in a single forward pass, achieving 45 frames per second on the base model. Research comparing multiple CNN architectures found that YOLOv3 provided superior balance between detection accuracy and computational efficiency, achieving 83.7% average precision at 28 frames per second [3]. YOLOv5 with CSP-Darknet53 backbone demonstrates mean Average Precision (mAP@50) of 0.88 with inference speed of 73 frames per second [1]. Recent work on YOLOv7 and YOLOv8 shows continued evolution, with YOLOv8 achieving approximately 89% accuracy for fire object detection [9].

ResNet (Deep Residual Learning) introduced by He et al. [5] addressed the fundamental problem of training very deep neural networks through residual connections or skip connections, reformulating layers as learning residual functions with reference to layer inputs rather than learning unreferenced functions, enabling successful training of networks with 152 layers—8x deeper than VGG nets—while maintaining lower complexity. ResNet-152 achieved 3.57% error on ImageNet, surpassing human-level performance. By enabling substantially deeper networks that can learn more complex feature hierarchies, ResNet architectures capture subtle visual patterns distinguishing genuine fires from fire-like phenomena such as sunsets, orange lighting, or reflective surfaces.

MobileNet introduced by Howard et al. [6] represents a class of efficient models specifically designed for mobile and embedded vision applications, employing depthwise separable convolutions—factorizing standard convolutions into depthwise and pointwise convolutions—dramatically reducing computational requirements and parameters while maintaining classification accuracy. MobileNet introduces two hyperparameters (width multiplier α and resolution multiplier ρ) enabling flexible trade-offs between latency and accuracy based on application constraints. Subsequent variants including MobileNetV2 [10] with inverted residual structures and MobileNetV3 [11] incorporating neural architecture search further improve efficiency. Quantized MobileNetV2 models maintain adequate accuracy while achieving dramatic computational requirement reductions suitable for microcontroller deployment [12], [13].

Other notable architectures include Modified Deep CNN models (MDCNN) integrating transfer learning and feature fusion algorithms achieving enhanced flame detection accuracy [14], VGG16 and VGG19 architectures providing high accuracy though at substantial computational cost [15], Inception architectures enabling effective learning of both local and global features [16], and DenseNet architectures facilitating feature reuse and gradient propagation [8].

B. Comparative Analysis and Performance

Recent comprehensive surveys examining deep learning methods for fire detection provide systematic comparisons across architectures [8], [17]. Three main approaches dominate: image classification using CNNs to distinguish fire-affected from normal regions, object detection localizing fires with bounding boxes, and semantic segmentation providing pixel-level fire masks. Classification approaches achieve accuracies typically ranging from 85-98% depending on dataset complexity and model size [8]. Object detection methods using YOLO variants, Faster R-CNN, and SSD achieve mean average precisions of 75-90% with inference speeds from 15-100 fps [3]. A comprehensive review of 37 research articles implementing deep learning for forest fire detection published between 2018-2023 reveals that YOLO-based methods dominate recent research due to their balance of accuracy and speed, though persistent challenges include limited generalization to unseen environmental conditions, sensitivity to smoke occlusion, and difficulty with small or distant fires [18].

C. Advantages and Limitations

CNN-based approaches offer compelling advantages: automatic feature extraction adapting to diverse fire appearances, translation invariance enabling detection regardless of position, robustness to variability across fire size and environmental conditions, temporal integration identifying characteristic fire development patterns, semantic understanding discriminating genuine fires from fire-like visual phenomena, and multi-scale detection across varying sizes and distances [8], [3].

However, significant challenges remain: substantial computational requirements particularly challenging for resource-constrained embedded platforms, large diverse training data requirements, struggles with extreme lighting conditions or heavy smoke obscuring visual features, interpretability difficulties making failure mode diagnosis challenging, extensive model optimization needs for embedded deployment, and environmental generalization issues when models encounter conditions not well-represented in training data [18], [8].

IV. MULTI-SENSOR DATA FUSION METHODOLOGIES

Multi-sensor data fusion integrates information from multiple heterogeneous sensors to achieve more accurate, reliable, and comprehensive situational awareness than any individual sensor alone [7]. The fundamental principle is that different sensors respond to different fire aspects, and intelligent combination of their outputs provides more robust detection than relying on any single modality.

A. Fusion Levels and Approaches

Signal-level fusion combines raw sensor measurements before feature extraction, preserving maximum information content but requiring careful handling of different sampling rates, measurement units, and noise characteristics, typically employing weighted averaging, Kalman filtering, or particle filtering [7]. Feature-level fusion extracts relevant features from individual sensors before combination, including statistical moments, frequency domain characteristics, or learned representations, reducing dimensionality while maintaining richer information than decision-level fusion [7]. Decision-level fusion combines independent classification decisions from multiple sensors, offering maximum modularity and computational efficiency, typically employing voting schemes, Dempster-Shafer theory, or Bayesian inference [19].

B. Key Fusion Techniques

Kalman filtering represents one of the most popular fusion methods due to its simplicity, ease of implementation, and optimality in a mean-squared error sense [7]. The Kalman filter provides a recursive solution to the linear filtering problem, making it ideal for estimating states in dynamic systems by reducing noise. For multi-sensor fire detection, Kalman filtering fuses measurements from temperature, smoke, and flame sensors operating at different sampling rates with different noise characteristics, computing optimal weightings based on measurement noise covariances, automatically emphasizing more reliable sensors while downweighting those experiencing interference [20]. Extended Kalman Filters (EKF) handle nonlinear sensor models through linearization via Jacobian matrices [7].

Dempster-Shafer evidential reasoning provides a mathematical framework for reasoning with uncertain and potentially conflicting evidence from multiple sources, assigning belief masses to different hypotheses and providing combination rules for updating beliefs as new evidence arrives [19]. This framework proves particularly valuable for fire detection because sensors may provide ambiguous or contradictory evidence, enabling principled reasoning about these conflicts and assigning appropriate uncertainty to final fire assessment.

Neural network-based fusion utilizing temperature, smoke, and carbon monoxide sensor data with trend extraction methods has demonstrated ability to correctly identify fires while reducing detection time by 32% compared to single-sensor approaches [21]. Neural network fusion learns optimal weighting of different sensor inputs, adapting weights based on environmental context and historical performance data, treating fusion as a supervised learning problem and capturing complex nonlinear relationships between sensor readings and fire conditions.

Temporal Convolutional Networks (TCN) for multi-sensor fusion achieve classification accuracy improvements exceeding 2.5% and detection speed improvements over 15% compared with standard TCN, back propagation neural networks, and LSTM approaches [22]. The temporal dimension proves particularly important because fires exhibit distinctive temporal signatures—progressive temperature increase, accelerating smoke production, and persistent rather than transient sensor responses—distinguishing them from benign events such as cooking activity.

Ensemble learning approaches such as LogitBoost, integrating multiple classifiers including Naive Bayes, backpropagation neural networks, support vector machines, and k-nearest neighbor algorithms, substantially improve detection accuracy and robustness [23]. The diversity principle in ensemble learning provides robustness against classifier-specific failure modes.

Bayesian fusion methods provide probabilistic frameworks for combining evidence under uncertainty [7]. Bayesian networks represent joint probability distributions over sensor measurements and fire states, enabling principled inference about fire presence. Hierarchical Bayesian models enable reasoning at multiple abstraction levels, from individual sensor measurements to room-level and building-level fire assessments [19].

Different fusion approaches exhibit distinct strengths: signal-level fusion preserves maximum information but imposes computational burden; feature-level fusion balances information preservation with computational efficiency; decision-level fusion offers maximum modularity and lowest computational requirements; Kalman filtering provides optimal fusion for linear systems with Gaussian noise; Dempster-Shafer theory excels when explicit uncertainty representation is valuable; neural network fusion provides adaptive learning but requires substantial training data; and ensemble approaches enhance robustness through classifier diversity [7], [19], [23].

V. INTERNET OF THINGS INTEGRATION IN FIRE ALARM SYSTEMS

Internet of Things technologies have transformed fire detection from isolated sensing devices to networked, intelligent systems capable of coordinated response and remote monitoring [24]. IoT-based fire alarm systems integrate fire detection sensors with microcontrollers, wireless communication modules, and cloud-based or edge computing platforms, enabling capabilities impossible with standalone devices.

A. Architecture and Communication

Core architectural components include distributed sensor nodes performing local fire detection, wireless communication infrastructure (Wi-Fi, cellular, LoRaWAN, or other protocols) for data transmission, edge computing devices for local intelligence and response coordination, cloud platforms for centralized monitoring and data analytics, and user interfaces (mobile applications, web dashboards) for remote monitoring and control [24]. Wireless sensor network implementations provide real-time monitoring and control through mobile applications and web interfaces, enabling security personnel to assess building conditions from centralized control rooms [25]. The distributed architecture enables monitoring of large building areas using multiple networked sensor nodes, providing comprehensive coverage while maintaining scalability and cost-effectiveness.

IoT fire detection systems employ various wireless communication technologies with distinct trade-offs [24]: **Wi-Fi (IEEE 802.11)** provides high bandwidth and ubiquitous infrastructure availability but consumes substantial power [25]; **Cellular (GSM/3G/4G/5G)** offers wide-area coverage and infrastructure independence but incurs subscription costs, with SIM900 GSM modules enabling SMS notifications even when internet connectivity is unavailable; **LoRaWAN** provides extremely long range (up to 15 km) and low power consumption suitable for battery-powered sensors but offers limited bandwidth inadequate for video transmission [26]; **Zigbee (IEEE 802.15.4)** provides low-power mesh networking suitable for building-scale deployments with self-healing topology [25]; **Bluetooth Low Energy (BLE)** enables low-power communication for proximity-based applications; and **NB-IoT** offers cellular connectivity optimized for IoT applications with extended coverage and ultra-low power consumption [24].

B. Edge Computing and Capabilities

Modern IoT fire detection systems increasingly employ edge computing—processing data locally on sensor nodes or gateway devices rather than transmitting all data to cloud platforms—offering reduced latency enabling faster response, decreased bandwidth consumption, enhanced privacy by processing sensitive data locally, and improved reliability through operation during network connectivity failures. Edge devices equipped with embedded AI accelerators such as NVIDIA Jetson Nano enable real-time CNN inference for vision-based fire detection [9]. Hierarchical edge-cloud architectures distribute intelligence across multiple levels: sensor nodes perform local preprocessing and threshold-based detection, edge gateways aggregate data from multiple sensors and apply fusion algorithms and CNN inference, and cloud platforms provide long-term storage, analytics, and system-wide coordination.

IoT technologies enable capabilities unattainable with traditional standalone detection systems: remote monitoring enabling continuous oversight of multiple facilities, automated alerting providing immediate notification through SMS, email, or mobile push notifications with geo-tagged alerts, data logging and analytics enabling post-incident analysis and predictive maintenance, coordinated response enabling building-wide responses such as automated fire suppression activation and HVAC control, predictive maintenance through analysis of sensor performance trends, and system self-testing through automated periodic testing verifying sensor functionality and communication links [24], [25], [26].

VI. MODEL OPTIMIZATION FOR EMBEDDED DEPLOYMENT

Deployment of sophisticated machine learning models on resource-constrained embedded platforms presents substantial challenges [27]. Typical microcontrollers feature limited processing capability (16-80 MHz clock speeds), minimal RAM (2-256 KB), and restricted flash storage (32-512 KB), rendering direct deployment of full-precision deep neural networks infeasible. The Arduino ATmega328P microcontroller features an 8-bit AVR processor running at 16 MHz with 32 KB flash memory and 2 KB SRAM [25], necessitating aggressive model optimization.

A. Optimization Techniques

Quantization reduces model memory requirements and computational costs by representing weights and activations using lower-precision numerical formats than standard 32-bit floating-point representations. Post-training quantization converts trained floating-point models to 8-bit integer representations, reducing model size by approximately 75% without substantial accuracy loss [28]. Quantization-friendly MobileNet architectures demonstrate that quantized models maintain inference accuracy within 4% of floating-point equivalents while achieving 33% reduction in time complexity [12]. Quantization of MobileNetV2 for resource-constrained microcontrollers demonstrates effectiveness for achieving efficient, low-power, low-latency execution on low-power MCUs, with successful deployment on STM32 boards validating the approach [13]. Quantization-Aware Training (QAT) simulates quantization effects during training, achieving better accuracy than post-training quantization [28]. Integer-Only Quantization converts both weights and activations to 8-bit integers, eliminating floating-point operations entirely [13].

Network pruning removes redundant connections and parameters contributing minimally to model output. Through three-stage pipeline compression involving pruning, quantization, and Huffman encoding, neural networks can achieve 35x to 49x reduction in storage requirements, with pruning alone reducing connections 9x to 13x, while maintaining accuracy comparable to original models [28]. Structured pruning removes entire filters, channels, or layers, resulting in models with regular structure amenable to efficient implementation on hardware without specialized sparse computation support [28]. The Lottery Ticket Hypothesis [29] suggests that dense networks contain sparse subnetworks that, when trained in isolation from initialization, can match the performance of the original network.

Knowledge distillation trains compact "student" models to mimic predictions of larger "teacher" models, transferring learned knowledge from complex models to simpler ones suitable for embedded deployment [30]. For fire detection applications, a large ResNet or ensemble model trained on comprehensive datasets can serve as the teacher, with a compact MobileNet serving as the student. Research demonstrates that well-distilled student models often outperform equivalent architectures trained from scratch.

Neural Architecture Search (NAS) automates the discovery of efficient network architectures optimized for specific hardware constraints and performance targets [31]. MobileNetV3 employed NAS combined with novel architecture modifications, achieving state-of-the-art accuracy-efficiency trade-offs, demonstrating 3-5% better accuracy than MobileNetV2 with equivalent latency [11]. EfficientNet demonstrated that systematically scaling network depth, width, and resolution according to principled compound scaling rules discovered through NAS yields substantial improvements [31].

B. Deployment Frameworks and Hardware

TensorFlow Lite for Microcontrollers represents Google's framework specifically designed for deploying machine learning models on microcontrollers with only kilobytes of memory, with core runtime fitting in 16 KB on ARM Cortex M3 processors, requiring no operating system support or dynamic memory allocation [32]. Ultra-low-power embedded AI fire detection systems employing TensorFlow Lite Micro and quantized deep neural networks demonstrate practical feasibility of deploying machine learning on resource-constrained microcontrollers [27]. Alternative frameworks include PyTorch Mobile, ARM CMSIS-NN offering optimized neural network kernels for ARM Cortex-M processors [13], X-CUBE-AI from STMicroelectronics generating optimized C code for STM32 microcontrollers [13], and Edge Impulse providing end-to-end development platform for embedded ML.

Specialized **hardware accelerators** enhance embedded ML performance: Neural Processing Units (NPUs) provide dedicated silicon for neural network inference, with processors such as Kendryte K210 integrating RISC-V cores with CNN accelerators capable of YOLO inference [27]; Edge TPUs from Google provide high-throughput ML inference accelerators offering 4 TOPS performance; and NVIDIA Jetson family provides GPU-accelerated computing platforms enabling deployment of sophisticated models [9].

Model optimization for embedded deployment involves careful navigation of trade-offs between accuracy, latency, memory consumption, and power consumption [12]. Power consumption becomes critical for battery-powered sensors in distributed fire detection networks [24]. Robust embedded ML systems require careful attention to model update mechanisms, diagnostic capabilities for detecting model degradation, and fail-safe behaviors when resource constraints are exceeded [27].

VII. CRITICAL ANALYSIS AND RESEARCH GAPS

Despite substantial progress in individual component technologies, comprehensive systems integrating CNNs, multi-sensor fusion, IoT connectivity, and embedded deployment remain underdeveloped [8]. Existing literature predominantly addresses isolated components without demonstrating end-to-end integration addressing practical deployment challenges including electromagnetic interference in industrial environments, sensor calibration drift over months or years, maintenance accessibility in hazardous areas, integration with legacy building management systems, and compliance with fire safety codes and standards that may not anticipate AI-based detection approaches [2].

Most fire alarm literature addresses detection accuracy but provides limited guidance on implementing intelligent reset functionality without manual intervention [2]. Intelligent reset mechanisms must balance avoiding prolonged alarm states after fire extinction versus ensuring adequate verification that fire conditions have genuinely ceased [21]. Limited exploration exists of decision-level fusion strategies combining CNN predictions with traditional sensor threshold detection [22]. Research should investigate adaptive fusion strategies that dynamically adjust weights assigned to different modalities based on current environmental conditions and sensor reliability indicators [19].

Absence of comprehensive evaluation frameworks for intelligent fire alarm systems across diverse fire scenarios, environmental conditions, and building types hinders systematic comparison of competing approaches [8], [18]. Development of standardized benchmarks encompassing diverse fire types, environmental conditions, and performance metrics would substantially advance the field. Such frameworks should include standardized datasets representing diverse fire scenarios [8], performance metrics capturing accuracy and computational requirements [18], test protocols specifying evaluation procedures, baseline implementations providing reference performance levels [3], and adversarial test cases assessing robustness. The computer vision community's ImageNet competition and COCO dataset demonstrate the value of standardized benchmarks [5].

Much existing research relies on laboratory testing or simulation rather than extensive field validation in operational building environments [2]. Real-world deployment introduces environmental variability, sensor degradation over time, electromagnetic interference, installation variations, and maintenance issues. Long-term field studies evaluating system performance over months or years in diverse building types remain scarce but essential for establishing practical viability [25].

Limited economic analysis exists comparing intelligent fire detection systems to conventional alternatives [2], though comprehensive cost-benefit analyses would inform adoption decisions.

CNN-based systems function as "black boxes," making it difficult for fire safety professionals to understand detection logic or diagnose failure modes [8]. Techniques such as attention visualization, saliency mapping, or rule extraction could address this gap, with uncertainty quantification helping operators assess detection reliability [18]. Vision-based fire detection systems raise privacy concerns [24], while security vulnerabilities in networked fire detection systems could enable malicious actors to disable protection or access building surveillance data [25]. Fire alarm systems must comply with numerous regulations and standards such as NFPA 72 and EN 54, though AI-based detection systems introduce challenges for regulatory frameworks designed around conventional sensor technologies.

VIII. FUTURE RESEARCH DIRECTIONS

Integration of intelligent fire detection systems with building automation platforms could enable coordinated responses extending beyond simple alarming, including HVAC control adjusting ventilation to limit smoke spread, elevator control recalling elevators to designated floors, access control unlocking exit doors along egress paths, lighting control activating emergency lighting and providing dynamic wayfinding illumination, and fire suppression automatic activation of sprinkler or gaseous suppression systems based on fire characteristics and location. Research demonstrating effective integration protocols and control strategies would substantially enhance overall fire safety effectiveness.

Investigation of emerging CNN architectures specifically optimized for embedded deployment could yield further performance improvements: EfficientNet family systematically optimizing network scaling achieved state-of-the-art accuracy-efficiency trade-offs [31]; Vision Transformers (ViT) demonstrated that transformer architectures can match or exceed CNN performance on vision tasks; Neural Architecture Search specifically targeting fire detection rather than general object recognition could discover specialized architectures incorporating domain-specific inductive biases; and Spiking Neural Networks (SNNs) offer extreme energy efficiency through event-driven computation.

Federated learning enables distributed machine learning across multiple deployed systems while preserving privacy by keeping training data localized. In fire detection applications, federated learning would allow models to improve through aggregated experience across numerous deployed systems without transmitting sensitive building surveillance data to centralized servers, enabling continuous improvement, privacy preservation, personalization, and rare event learning. Investigation of distributed intelligence approaches enabling collaborative decision-making across networked sensor nodes could enhance detection accuracy through spatial correlation of fire indicators, with edge computing architectures enabling distributed inference, spatial-temporal correlation, graceful degradation, and bandwidth optimization.

Most current research focuses on visual and environmental sensor fusion but overlooks audio modalities. Fire generates characteristic acoustic signatures including crackling sounds, structural material collapse, and glass breaking. Audio processing offers advantages complementary to visual and traditional sensors: operates effectively when visual observation is obscured by smoke, provides information about fire intensity, enables detection of events indicating fire progression, and works in darkness or poor lighting conditions. Research on multimodal fusion incorporating audio, visual, and environmental sensors remains largely unexplored.

As intelligent fire detection systems become increasingly reliant on machine learning, ensuring robustness against adversarial conditions and rare edge cases becomes critical. Research directions include adversarial training improving robustness to unexpected conditions, certified robustness providing mathematical guarantees that models maintain correct behavior, out-of-distribution detection triggering fallback behaviors, robustness testing across extreme environmental conditions, and safety cases documenting hazard analysis and verification evidence. Rather than fully automated detection, research on human-AI collaborative approaches where intelligent systems provide fire probability assessments, evidence visualization, and decision support to human operators could combine machine learning capabilities with human judgment.

Progress in machine learning is often limited by available training data. Development of large-scale, comprehensive fire detection datasets would substantially benefit the research community. Desired characteristics include diverse fire scenarios, temporal annotations with frame-level labels tracking fire development, multi-modal data with synchronized visual, thermal, audio, and environmental sensor data, extensive non-fire scenarios easily confused with fires, real-world conditions from actual buildings, and standardized formats facilitating cross-study comparisons. Collaborative efforts involving research institutions, fire departments, and industry partners could pool resources for dataset creation, with simulation environments generating synthetic fire data supplementing real-world data collection.

IX. CONCLUSIONS

This comprehensive review has examined the current state of intelligent fire detection technology, encompassing traditional detection methodologies, CNN-based visual fire detection, multi-sensor data fusion approaches, IoT integration, and embedded deployment optimization. Significant progress has been achieved in each domain, with CNN architectures such as YOLO, ResNet, and MobileNet demonstrating superior accuracy compared to traditional approaches, multi-sensor fusion substantially reducing false alarm rates while improving detection speed, and model optimization techniques including quantization, pruning, and knowledge distillation enabling deployment on resource-constrained embedded platforms.

The evolution from single-sensor detection to integrated multi-modal systems incorporating deep learning represents a paradigm shift in fire safety technology. Modern approaches leverage complementary strengths of diverse sensing modalities—visual pattern recognition through CNNs, direct physical measurements via temperature and smoke sensors, and infrared flame detection—synthesized through sophisticated fusion algorithms grounded in mathematical frameworks including Kalman filtering, Dempster-Shafer theory, and neural network-based integration.

However, substantial research gaps remain, particularly regarding integrated systems combining these component technologies into comprehensive, deployable fire safety infrastructure. The field would benefit from standardized evaluation frameworks enabling systematic comparison of competing approaches across diverse fire scenarios and building types, extensive field validation establishing long-term operational reliability under real-world conditions, and investigation of advanced techniques such as federated learning enabling continuous model improvement while preserving privacy, distributed edge intelligence for collaborative detection across sensor networks, and multimodal fusion incorporating audio analysis alongside visual and environmental sensing.

The convergence of artificial intelligence, Internet of Things, and embedded systems technologies presents unprecedented opportunities for advancing fire safety infrastructure. Future intelligent fire detection systems integrating these technologies promise to substantially reduce false alarm rates that currently undermine system credibility, decrease detection latency enabling earlier occupant warning and evacuation, provide autonomous operation reducing dependence on manual intervention, enable coordinated building-wide response through integration with automation systems, and continuously improve through learning from aggregated experience across deployed systems.

Realizing this vision requires collaborative efforts spanning multiple disciplines—fire safety engineering bringing domain expertise on fire dynamics and detection requirements, computer science contributing machine learning and distributed systems expertise, embedded systems engineering addressing deployment on resource-constrained platforms, building automation integrating fire detection with coordinated response capabilities, and regulatory engagement ensuring compliance with safety standards while enabling innovation. Through such interdisciplinary collaboration, the research community can advance fire detection technology toward increasingly intelligent, reliable, and effective protection of lives and property worldwide.

REFERENCES

[1] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient deep CNN-based fire detection and localization in video surveillance applications," *IEEE Transactions on Systems, Man, and Cybernetics, Syst.*, vol. 50, no. 6, pp. 2160–2169, Jun. 2020.

[2] S. Festag, "False alarm ratio of fire detection and fire alarm systems in Germany—A meta-analysis," *Fire Safety Journal.*, vol. 79, pp. 119–126, Jan. 2016.

[3] C. Li and X. Zhao, "Image fire detection algorithms based on convolutional neural networks," *Case Studies in Thermal Engineering.*, vol. 19, p. 100625, Jun. 2020.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc.* In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770-778), Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[6] A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, Apr. 2017.

[7] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, Jan. 2013.

[8] C. Jin, T. Wang, N. Alhusaini, S. Zhao, H. Liu, K. Xu, and J. Zhang, "Video fire detection methods based on deep learning: Datasets, methods, and future directions," *Fire*, vol. 6, no. 8, p. 315, Aug. 2023.

[9] A. Shamta and K. A. Demir, "Forest fire detection using YOLOv7 and YOLOv8," *International Journal of Engineering and Innovative Research.*, vol. 6, no. 1, pp. 26–38, 2024.

[10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510–4520.

[11] Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. Searching for MobileNetV3. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 1314-1324) 2019.

[12] U. Kulkarni, S. M. Meena, S. V. Gurlahosur, and G. Bhogar, "Quantization friendly MobileNet (QF-MobileNet) architecture for vision based applications on embedded platforms," *Neural Networks.*, vol. 136, pp. 28–39, Apr. 2021.

[13] L. Lazzaroni, F. Bellotti, A. Dabbous, A. Pighetti, and R. Berta, "Quantization of MobileNetV2 for resource-constrained microcontrollers," in *Applications in Electronics Pervading Industry, Environment and Society*. Cham, Switzerland: Springer, 2025, pp. 685–695.

[14] G. Chen, R. Cheng, X. Lin, W. Jiao, D. Bai, and H. Lin, "MDCNN: An improved deep convolutional neural network for flame detection," *Journal of Safety Science and Resilience*, vol. 5, no. 1, pp. 1–10, Mar. 2024.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," In *International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, May 2015.

[16] A. K. Vishwakarma and M. Deshmukh, "CNNM-FDI: Novel convolutional neural network model for fire detection in images," *IETE Journal of Research.*, vol. 71, no. 4, pp. 2245–2258, Apr. 2025.

[17] V. E. Sathishkumar, J. Park, and Y. Cho, "Seoul bike trip duration prediction using data mining techniques," *IET Intelligent Transport Systems.*, vol. 14, no. 11, pp. 1465–1474, Sep. 2023.

[18] R. Ghali, M. A. Akhloufi, and M. Jmal, "Deep learning and transformer approaches for UAV-based wildfire detection and segmentation," *Sensors*, vol. 24, no. 5, p. 1374, Feb. 2024.

[19] S. Hadjiefthymiades, E. Zervas, and C. Anagnostopoulos, "Multisensor data fusion for fire detection," *Information Fusion*, vol. 11, no. 4, pp. 300–312, Oct. 2010.

[20] M. Veysi, H. Sazgar, and M. Aghaei, "Multi-sensor data fusion based on the Kalman filter for in-situ state-of-charge estimation of lithium-ion batteries," *Renewable Energy Focus*, vol. 48, p. 100525, Mar. 2024.

[21] J. Liu, R. Bao, S. Yang, X. Zeng, and H. Liu, "Multi-sensor data fusion algorithm for indoor fire early warning based on BP neural network," *Information*, vol. 12, no. 2, p. 59, Jan. 2021.

[22] C. Wang, J. Liu, and H. Liu, "Research on multi-sensor fusion indoor fire perception algorithm based on improved TCN," *Sensors*, vol. 22, no. 12, p. 4721, Jun. 2022.

[23] L. Wang and J. Zhang, "Multi-sensor data fusion algorithm for indoor fire detection based on ensemble learning," in *Advanced Computational Intelligence and Intelligent Informatics*. Singapore: Springer, 2024, pp. 45–58.

[24] K. Al Mamari, S. Al Mamari, S. A. U. D. Kazmi, A. K. Pandey, and A. Al Hinai, "Smart fire alarm system using IOT," In *Proceedings of the 7th International Conference on Soft Computing & Machine Intelligence.*, Stockholm, Sweden, Nov. 2020, pp. 112–116.

[25] F. S. Perilla, G. R. Villanueva, N. M. Cacanindin, and T. D. Palaoag, "Fire safety and alert system using Arduino sensors with IoT integration," In *Proceedings of the 2018 7th International Conference on Software and Computer Applications.*, Kuantan, Malaysia, Feb. 2018, pp. 180–185.

[26] M. R. Rani, B. Siva, P. Rajesh, S. Purnima, and M. H. Kousar, "Smart fire detection system with early notifications using IoT," *Journal of Advances in Developmental Research (IJAIDR)*, vol. 16, no. 1, pp. 1–9, Jan.–Jun. 2025

[27] A. Nazir, H. Mosleh, M. Takruri, A. H. Jallad, and H. Alhebsi, "An ultra-low-power embedded AI fire detection and crowd counting system for indoor areas," *ACM Transactions on Embedded Computing Systems.*, vol. 21, no. 6, pp. 1–27, Nov. 2022.

[28] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," *arXiv preprint arXiv:1510.00149*, Feb. 2016.

[29] J. Frankle and M. Carbin, "The lottery ticket hypothesis: Finding sparse, trainable neural networks," In *International Conference on Learning Representations (ICLR)*, New Orleans, LA, USA, May 2019.

[30] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv preprint arXiv:1503.02531*, Mar. 2015.

[31] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, Long Beach, CA, USA, Jun. 2019, pp. 6105–6114.

[32] "TensorFlow Lite for Microcontrollers," TensorFlow, 2023. [Online]. Available: https://www.tensorflow.org/lite/microcontrollers

.