TRADITIONAL AND VISION-BASED FIRE DETECTION METHODS: SURVEY

Khojamurotov Azizbek
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
E-mail: xojamurotovazizbek2001@gmail.com
Rustamov Ilhom A‘zam o‘g‘li
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
E-mail: ilhomsaep0124@gmail.com

Abstract. Fire detection plays a vital role in ensuring human safety, protecting infrastructure, and minimizing environmental damage. Traditional sensor-based methods, such as smoke, heat, and gas detectors, have long been deployed in diverse settings due to their simplicity and reliability. However, these approaches often suffer from limitations, including delayed response times and vulnerability to false alarms. In recent years, vision-based fire detection methods, powered by computer vision and deep learning techniques, have emerged as a promising alternative. These systems leverage visual data to provide faster, more accurate detection while enabling situational awareness. The purpose of this survey is to review, compare, and analyze the main categories of fire detection methods: sensor-based approaches and vision-based approaches. We highlight their principles, strengths, and weaknesses, with a focus on how computer vision enhances detection accuracy and adaptability in real-world conditions. Our findings suggest that while traditional methods remain practical for low-cost and small-scale applications, vision-based approaches are increasingly shaping the future of fire detection by offering real-time monitoring, integration with smart systems, and the potential for predictive analytics. This survey emphasizes the need for hybrid models that combine the robustness of traditional sensors with the intelligence of vision-based systems to achieve more reliable fire detection solutions..

Keywords. Fire detection, sensor-based methods, computer vision, deep learning, image analysis, early warning systems, safety monitoring

Аннотация. Обнаружение пожара играет жизненно важную роль в обеспечении безопасности людей, защите инфраструктуры и минимизации ущерба окружающей среде. Традиционные методы, основанные на датчиках, такие как детекторы дыма, тепла и газа, уже давно применяются в различных

ISSN 3030-3907

условиях благодаря своей простоте и надежности. Однако эти подходы часто имеют ограничения, в том числе задержку с реагированием и уязвимость к ложным тревогам. В последние годы в качестве многообещающей альтернативы появились методы обнаружения пожара на основе визуального анализа, основанные на компьютерном зрении и методах глубокого обучения. Эти системы используют визуальные данные для обеспечения более быстрого и точного обнаружения, обеспечивая при этом осведомленность о ситуации. Цель данного обзора - рассмотреть, сравнить и проанализировать основные категории методов обнаружения пожара: подходы, основанные на датчиках, и подходы, основанные на зрении. Мы освещаем их принципы, сильные и слабые стороны, уделяя особое внимание тому, как компьютерное зрение повышает точность обнаружения и адаптируемость в реальных условиях. Наши результаты показывают, что, в то время как традиционные методы остаются практичными для недорогостоящих и маломасштабных приложений, подходы, основанные на зрении, все больше определяют будущее обнаружения пожаров, предлагая мониторинг в режиме реального времени, интеграцию с интеллектуальными системами и потенциал для прогнозной аналитики. В этом обзоре подчеркивается необходимость в гибридных моделях, сочетающих надежность традиционных датчиков с интеллектуальностью систем технического зрения для создания более надежных решений по обнаружению пожара..

Ключевые слова. Обнаружение пожара, сенсорные методы, компьютерное зрение, глубокое обучение, анализ изображений, системы раннего предупреждения, мониторинг безопасности

Annotatsiya. Yong'inni aniqlash inson xavfsizligini ta'minlash, infratuzilmani muhofaza qilish va atrof-muhitga etkazilgan zararni minimallashtirishda muhim rol o'ynaydi. Tutun, issiqlik va gaz detektorlari kabi an'anaviy sensorlarga asoslangan usullar soddaligi va ishonchliligi tufayli uzoq vaqtdan beri turli xil sharoitlarda joylashtirilgan. Biroq, ushbu yondashuvlar ko'pincha cheklovlardan aziyat chekadi, shu jumladan kechiktirilgan javob vaqtlari va noto'g'ri signallarga nisbatan zaiflik. So'nggi yillarda istiqbolli alternativa sifatida kompyuterni ko'rish va chuqur o'rganish texnikasi bilan ishlaydigan ko'rishga asoslangan yong'inni aniqlash usullari paydo bo'ldi. Ushbu tizimlar vaziyatni xabardor qilish bilan birga tezroq va aniqroq aniqlashni ta'minlash uchun vizual ma'lumotlardan foydalanadi. Ushbu so'rovning maqsadi yong'inni aniqlash usullarining asosiy toifalarini ko'rib chiqish, taqqoslash va tahlil qilishdir: sensorga asoslangan yondashuvlar va ko'rishga asoslangan yondashuvlar. Biz ularning tamoyillari, kuchli va zaif tomonlarini

ISSN 3030-3907

ta'kidlaymiz, bunda kompyuterni ko'rish Real sharoitlarda aniqlanishning aniqligi va moslashuvchanligini oshirishga e'tibor qaratamiz. Bizning topilmalarimiz shuni ko'rsatadiki, an'anaviy usullar arzon va kichik hajmdagi ilovalar uchun amaliy bo'lib qolsa-da, ko'rishga asoslangan yondashuvlar Real vaqtda monitoring, aqlli tizimlar bilan integratsiya va bashoratli tahlil imkoniyatlarini taklif qilish orqali yong'inni aniqlash kelajagini tobora ko'proq shakllantirmoqda. Ushbu so'rov yong'inni aniqlashning yanada ishonchli echimlariga erishish uchun an'anaviy sensorlarning mustahkamligini ko'rishga asoslangan tizimlarning aql-zakovati bilan birlashtirgan gibrid modellarga ehtiyoj borligini ta'kidlaydi..

Kalit so‘zlar. Yong'inni aniqlash, sensorga asoslangan usullar, kompyuterni ko'rish, chuqur o'rganish, tasvirni tahlil qilish, erta ogohlantirish tizimlari, xavfsizlik monitoring

Introduction

Fire is among the most destructive and unpredictable hazards, posing severe risks to human safety, infrastructure, and ecosystems. In recent years, large-scale fire events have highlighted the urgent need for effective detection systems. The Lahaina wildfire in Maui (2023) destroyed more than 2,200 structures and caused tragic loss of life, while California’s 2020 fire season burned nearly four million acres and led to billions of dollars in economic damage. Similarly, Canada’s record-breaking 2023 wildfires forced mass evacuations, deteriorated air quality across North America, and resulted in excess hospitalizations and mortality due to smoke exposure. Beyond immediate destruction, fires also generate long-term ecological impacts, such as biodiversity loss and climate feedback effects, making early detection not only a safety requirement but also a global environmental priority[1].

To address this global challenge, researchers have developed two principal categories of detection methods: SENSOR-BASED and COMPUTER VISION-BASED approaches. SENSOR-BASED systems, including smoke, heat, and gas detectors, are widely deployed due to their simplicity, affordability, and reliability in confined environments. However, they suffer from delayed activation and susceptibility to false alarms when exposed to non-fire aerosols such as dust, steam, or humidity (Recent Advances in Sensors for Fire Detection, 2020). On the other hand, COMPUTER VISION-BASED methods leverage image processing and deep learning techniques to analyze video streams for visual indicators of fire, such as flames, smoke plumes, and dynamic light patterns. These methods offer richer situational awareness, enabling real-time localization, spread estimation, and even predictive insights in diverse contexts ranging from smart buildings to large-scale forest surveillance (Muhammad et al., 2021; Sharma et al., 2022).

ISSN 3030-3907

This article aims to provide a clear and structured overview of these two detection paradigms. Section 2 introduces traditional SENSOR-BASED methods and their applications, while Section 3 focuses on COMPUTER VISION-BASED approaches. Section 4 presents a comparative analysis, and Section 5 concludes with key insights and future perspectives[1].

Conventional Sensor-Based Fire Detection Systems

Traditional fire detection systems, often referred to as sensor-based approaches, operate by measuring physical parameters that accompany combustion processes. These systems rely on detecting heat release, gas emissions, or particle concentrations, as well as changes in ultraviolet and infrared radiation, light scattering, air density, ion mobility, or smoke concentration in the surrounding environment. While such methods have been widely adopted due to their relative simplicity and established reliability, they are inherently constrained by certain limitations. Their effectiveness is typically restricted to a limited detection range, and transmission delays may hinder rapid response in critical scenarios. Moreover, sensor performance is highly sensitive to environmental conditions, making them less suitable for outdoor or complex settings where false alarms are more likely to occur. A more detailed discussion of their operational principles, strengths, and shortcomings will be presented in the following sections[2].

2.1. Smoke-based Detectors

Smoke detection systems are widely used in early fire warning applications because of their ability to sense small concentrations of combustion particles in the air. In campus environments, where fire risks are often related to electricity use, heating, or cooking, smoke detectors serve as an important preventive measure. Traditional smoke terminals mainly rely on wireless communication technologies, but these devices suffer from limited battery life, requiring costly maintenance when deployed in large numbers (Liu et al., 2025).

To address the issue of frequent battery replacement, researchers have proposed integrating photovoltaic power generation with supercapacitor energy storage. Unlike conventional lithium batteries, supercapacitors offer longer life cycles, environmentally friendly operation, and higher charging and discharging efficiency. Combined with polycrystalline silicon solar panels, these systems provide a continuous and stable power supply for smoke sensing terminals, enabling reliable long-term deployment in public places such as universities (Liu et al., 2025).

In terms of design, the system employs an STM32 microcontroller as the central processor, with an MQ-7 smoke sensor for data acquisition and an NB-IoT module for communication. The NB-IoT technology ensures low power

ISSN 3030-3907

consumption, wide coverage, and real-time data transmission to an upper computer through cloud platforms. To further enhance energy efficiency, a maximum power point tracking (MPPT) algorithm based on variable step-length conductance increment is adopted, ensuring optimal solar energy utilization. This combination allows the smoke sensing terminal to operate independently with minimal maintenance (Liu et al., 2025).

The experimental results demonstrate that the proposed system successfully integrates renewable energy and low-power communication technologies to create a sustainable smoke detection terminal. By combining solar photovoltaic charging, supercapacitor storage, and NB-IoT communication, the system achieves energy self-sufficiency, real-time monitoring, and environmental friendliness. This design not only reduces operational costs but also enhances the effectiveness of smoke-based fire detection in large-scale public facilities such as campuses (Liu et al., 2025)[2].

2.2. Gas-based Sensors

Recent research has explored gas-sensitive fire detection systems as an alternative to conventional smoke detectors, focusing on detecting toxic combustion gases at the earliest stage of fire. A promising approach involves metal-oxide nanocolumn (NC) structures fabricated through the glancing angle deposition (GLAD) technique. This method enables highly porous and ordered nanostructures with a large surface-to-volume ratio, which significantly enhances gas adsorption and sensing performance (Lee et al., 2017).

The study investigates NiO, SnO_2, WO_3, and In_2O_3 nanocolumn-based chemoresistive sensors as potential fire detectors. The sensing principle relies on abrupt changes in resistance when the nanocolumn surfaces interact with gases released during the thermal decomposition of polyvinyl chloride (PVC), a widely used but hazardous plastic. At around 200 °C, before visible smoke forms, gases such as HCl, CO, and VOCs are released. The SnO_2 NCs showed an early response with a sensitivity of 2.1, while conventional commercial smoke detectors failed to provide any alarm at this stage. At a higher temperature of 350 °C, NiO NCs demonstrated the highest response (577.1), despite being a p-type semiconductor, highlighting their superior performance after smoke emission (Lee et al., 2017).

To ensure reliable measurement, the sensors were integrated with Pt/Ti interdigitated electrodes and coupled with an interface circuitry that continuously monitored resistance changes. Data processing included noise reduction with exponential weighted moving average (EWMA) filtering and normalization for uniform comparison across different sensor materials. The experiments confirmed that SnO_2 NCs exhibited faster response times than standard smoke detectors,

ISSN 3030-3907

enabling more efficient early-stage fire warning. Furthermore, the study demonstrated the feasibility of mass production, as up to 732 sensors could be fabricated on a single wafer using GLAD (Lee et al., 2017).

The results emphasize that metal-oxide nanocolumn gas sensors are highly sensitive, selective, and scalable solutions for fire detection. By identifying toxic gases prior to visible smoke emission, these sensors could dramatically improve evacuation time and reduce fatalities caused by inhalation of hazardous compounds in fire scenarios. Thus, nanostructured gas-based fire detection represents a powerful advancement beyond traditional smoke-based systems (Lee et al., 2017)[4].

2.3. Thermal Fire Detection Systems

Fire detection systems for outdoor environments such as bio-fuel depots can be designed using thermal cameras, which are mounted on existing lighting poles to achieve wide area coverage. These depots are often very large, with piles of material reaching tens of thousands of cubic meters, and it is not feasible to fully cover every section due to the high cost of surveillance. A practical solution is therefore to find compromise designs that balance the trade-off between cost and coverage, where customers generally accept coverage rates in the range of 70–95% (Quttineh et al.).

Thermal cameras typically cost a few thousand euros each, and a full system may require between 5 and 20 cameras depending on the layout of the depot. The placement of cameras is not straightforward, since visibility depends on pile locations, pole heights, and camera directions. In manual planning, salesmen often rely on maps and experience, but this process is time-consuming, subjective, and may result in suboptimal coverage or unnecessary additional cameras, which raises system costs (Quttineh et al.).

To address this, the placement problem is formulated as a bi-objective optimization model with the goals of minimizing total system cost and maximizing coverage of designated pile areas. The model discretizes the depot into cells that represent the piles, and each possible camera position and orientation is defined as a Positioned and Aimed camera Model (PAM). A visibility matrix records which cells can be monitored by which PAMs, and the optimization then seeks non-dominated solutions that provide efficient trade-offs between cost and coverage (Quttineh et al.)[5].

Since exact algorithms are impractical for the large-scale problems encountered in real depots, the approach employs fast greedy-type heuristics to approximate the Pareto frontier. These heuristics allow quick exploration of many alternative camera placements, providing both salesmen and customers with clear

ISSN 3030-3907

options that highlight cost–coverage trade-offs. The method has been implemented in a decision support software currently in use by Termisk Systemteknik, which makes the design process more transparent, less dependent on individual skills, and typically more effective than manual planning (Quttineh et al.)[5].

Table-1.

Summary of Works reviewed on traditional sensor-based fire detection

Author(s) and Year	Technique Used	Contribution	Research Gap
Liu et al., 2025	Smoke sensing with MQ-7 sensor, STM32 microcontroller, NB-IoT module, photovoltaic panel + supercapacitor	Sustainable smoke detection terminal with solar charging, supercapacitor storage, NB-IoT communication; real-time monitoring; reduced operational cost; eco-friendly	Battery replacement issue solved, but limited to campus/public place deployment; performance in complex outdoor environments not discussed
Lee et al., 2017	Gas-based chemoresistive sensors using metal-oxide nanocolumns (NiO, SnO_2, WO_3, In_2O_3) fabricated with GLAD	Early-stage toxic gas detection before smoke formation; SnO_2 NCs detected gases at ~200°C while commercial detectors failed; NiO NCs achieved highest sensitivity (577.1) at 350°C; scalable fabrication (732 sensors per wafer)	Still requires optimization for real-world deployment; integration with full fire alarm systems not addressed
Quttineh et al., 2022	Thermal fire detection with outdoor thermal cameras and bi-objective optimization for camera placement	Cost-coverage optimization model for large depots; decision support software for efficient camera positioning; achieves 70–95% coverage with fewer	High camera cost; limited full coverage; solution still dependent on trade-offs; scalability for very large/complex depots requires further validation

ISSN 3030-3907

		cameras	

Fire Detection Systems Based on Computer Vision

Computer vision-based fire detection systems have gained increasing attention as an advanced alternative to conventional sensor-based approaches. Traditional systems often suffer from slow responses and high false alarm rates, which waste resources and reduce reliability. In contrast, computer vision leverages visual sensors to identify fire-related features such as smoke, flames, and unusual light patterns in real time. When combined with machine learning, these systems can learn from large datasets, recognize fire patterns more accurately, and adapt to different environments. This integration not only improves detection speed but also minimizes false alarms, making the systems more dependable for diverse applications, from residential buildings to outdoor industrial sites. Recent research highlights that computer vision and machine learning together provide a robust foundation for building intelligent, scalable, and highly reliable fire detection systems that contribute to public safety and disaster prevention(Ridhani et al., 2023)[6]

3.1. Fire Detection Based on Visual Features of Lithium Battery Combustion

In recent years, with the rise of environmental awareness, the use of new-energy electric vehicles has been steadily increasing, along with the rapid expansion of supporting infrastructure such as charging stations. Lithium-ion batteries, as the main power source, are widely adopted due to their high energy density, long lifespan, and lightweight properties. However, during fast charging, these batteries are prone to risks such as leakage, fire, or even explosion caused by thermal runaway, which can endanger vehicles, charging piles, and nearby facilities. This makes real-time monitoring of electric vehicle charging safety particularly critical.

Traditional fire detection methods relied mainly on handcrafted image features such as color, texture, and motion to recognize flames and smoke. While these approaches provided some capability in controlled environments, they often suffered from low accuracy and poor adaptability in complex outdoor scenes. They were also slow in response, which limited their effectiveness for real-time video monitoring. As a result, deep learning-based object detection methods, including two-stage approaches like Faster R-CNN and one-stage approaches such as SSD and YOLO, have become the mainstream due to their improved accuracy and speed in fire and smoke detection tasks[6].

Building on this progress, the study proposes an enhanced YOLOv5-based detection model tailored for fire and smoke features generated during lithium battery combustion, such as white smoke and deflagration flames. Key

ISSN 3030-3907

improvements include increasing convolutional kernels and residual structures to better detect small fire targets, applying K-means clustering to optimize anchor boxes for more accurate localization, and embedding a CBAM attention module to strengthen feature extraction. Together, these modifications significantly improve detection accuracy and robustness, while maintaining real-time performance, making the method more suitable for deployment in electric vehicle charging stations.

This study introduces an improved YOLOv5-based method for detecting fire and smoke features generated during lithium battery combustion in electric vehicles. Because the combustion process is complex and highly uncertain, conventional anchor boxes are not well suited to capture these patterns. To address this, the K-means clustering algorithm is applied to the dataset to generate anchor boxes that are more representative of flame and smoke locations. This refinement enhances both training efficiency and detection accuracy compared to using standard anchor settings.

The YOLOv5 architecture itself is designed for end-to-end real-time object detection and is composed of an input stage, a backbone for feature extraction, a neck for feature fusion, and a head for multi-scale prediction. Mosaic data augmentation enriches the input images and improves the robustness of the network. The backbone integrates Focus and CSPNet structures, which strengthen feature representation while maintaining computational efficiency. The neck employs PANet and SPP modules to merge features across different scales and expand the receptive field, while the head generates outputs at multiple resolutions to better detect objects of varying sizes.

To further improve performance, a convolutional block attention module (CBAM) is embedded after the backbone. This mechanism adaptively emphasizes meaningful features and suppresses background noise by applying both channel and spatial attention. As a result, the network focuses more effectively on fire and smoke regions while minimizing the influence of irrelevant information.

In addition, the number of residual blocks and convolutional kernels in the network is adjusted to increase its depth and width, which enhances its ability to capture small-scale combustion features. Combined with the optimized anchor boxes and the CBAM module, these improvements significantly boost the accuracy, robustness, and real-time performance of YOLOv5, making it well suited for monitoring safety during electric vehicle charging.

The dataset for this study was constructed to capture the complex combustion characteristics of lithium battery fires in electric vehicles. Early-stage events are often dominated by white smoke, followed by open flames and,

ISSN 3030-3907

eventually, black smoke from secondary ignition of surrounding structures. These variations, coupled with background similarities, make accurate feature identification challenging. To build a reliable dataset, images reflecting smoke and flame in diverse scenarios—such as charging station fires, sudden combustion during driving, and night-time ignition—were carefully selected. Using LabelImg, the flame and smoke regions were annotated with rectangular bounding boxes of varying sizes, accounting for partial occlusion and fusion with the sky. The final dataset contained 3,391 images following the VOC annotation specification, providing a strong foundation for model training.

The experiments were conducted on a platform equipped with an Intel Xeon Gold 6130H CPU, NVIDIA RTX3060 GPU (24GB), 32 GB RAM, and a Windows environment using PyTorch. Transfer learning was applied by initializing the network with YOLOv5 weights pretrained on the COCO dataset, accelerating convergence and improving detection quality. The training was carried out over 100 epochs with a confidence threshold of 0.5. In the first 50 epochs, the batch size was set to 8, later reduced to 4 for stability. Input images were resized to 416×416, and Adam optimizer was employed. Evaluation metrics included mean average precision (mAP), precision, recall, and F1-score, with classification divided into TP, FP, TN, and FN categories[7].

A series of experiments was conducted to compare YOLOv5 models with different widths and depths. The results, presented in Table 3, demonstrate that YOLOv5-l achieved the highest overall performance, with mAP of 93.99%, precision around 89.74% (fire) and 94.12% (smoke), and recall of 83.67% (fire) and 91.60% (smoke). Its F1-scores were 0.87 and 0.93 for fire and smoke respectively. YOLOv5-s, with lower depth and width multipliers, still achieved competitive mAP (92.67%), but YOLOv5-x performed relatively worse, with reduced accuracy and F1-scores, highlighting that excessive scaling can negatively impact detection. Thus, YOLOv5-l was identified as the most balanced configuration for feature extraction and detection of combustion events.

Further comparisons were made using different backbone networks for feature extraction. According to Table 4, CSPDarknet proved to be the most effective backbone, achieving an mAP of 87.97% with an average FPS of 25.04. While this mAP was slightly lower than SwinTransformer-Tiny (90.08%), CSPDarknet offered a much higher detection speed, making it more suitable for real-time monitoring. ConvNext variants performed less favorably, with ConvNext-Tiny and Small showing lower F1-scores (0.68–0.70 range) and weaker robustness against small object detection. The CSPDarknet backbone thus provided a strong trade-off between accuracy and computational efficiency, enabling the

ISSN 3030-3907

system to capture subtle features of fire and smoke.

The effect of incorporating attention mechanisms was also evaluated. SENet, ECANet, and CBAM were embedded into the YOLOv5-l + CSPDarknet network for comparison. The results revealed that CBAM achieved the best performance, increasing detection accuracy while adding negligible computational overhead (<0.01% additional parameters). By adaptively weighting channel and spatial features, CBAM improved localization of flames and smoke while suppressing irrelevant background noise. This enhancement further confirmed the value of attention mechanisms in safety-critical monitoring applications[7].

When integrating the optimal network settings—YOLOv5-l with CSPDarknet backbone, 3 residual bases, convolutional kernel base of 64, and CBAM attention module—the system achieved the highest performance among all tested configurations. This improved model demonstrated strong capability in identifying both flame and smoke across diverse conditions. Notably, it was able to detect smaller targets and remained resilient against background interference, two essential qualities for reliable real-world deployment.

Comparative experiments between unimproved and improved versions of YOLOv5 showed clear benefits of the modifications. As illustrated in Figure 13, the enhanced model produced more precise bounding boxes around flame and smoke regions and exhibited stronger robustness across varied scenarios, such as night fires and partial occlusion cases. The improved algorithm was especially effective at identifying rapid lithium battery combustion events, where temperature spikes can trigger cascading ignition of surrounding materials, leading to full-vehicle fires.

In summary, the combination of carefully curated datasets, optimized network width and depth, CSPDarknet backbone, and CBAM attention mechanism yielded a highly effective fire detection system. The final model achieved a balance between high mAP (94% range), strong F1-scores (~0.90), and real-time performance with FPS exceeding 25, demonstrating that it can be practically deployed in electric vehicle charging stations and production platforms to enhance safety monitoring.

A feature-based target detection algorithm is proposed to enable real-time monitoring of flames and smoke in the complex scenarios of electric vehicle charging, addressing critical safety risks. Through comparative experiments, the most suitable detection model for EV charging safety was identified and further improved by incorporating the CBAM attention mechanism, achieving superior performance over several established algorithms in terms of detection accuracy (94.09%), robustness against interference, and real-time efficiency. The

ISSN 3030-3907

algorithm's lightweight design also allows cost-effective deployment on mobile devices, making it a practical solution for enhancing the safety of charging stations. Moreover, its applicability extends to future unmanned charging facilities, as well as the production, transportation, and storage of lithium batteries, offering a versatile tool for ensuring safety in diverse lithium battery-related environments[8].

3.2. Deep Learning-Based Fire Detection in Ship Environments Using Enhanced YOLOv7 Architecture

Shipping has long been one of the most essential means of transportation, covering vast areas of the Earth's oceans. Since the 15th century, the rapid expansion of maritime transport has enabled global trade and travel, significantly influencing social and natural environments. With technological progress, ships have become more advanced and reliable; however, they also present serious safety challenges, among which fire is one of the most frequent and hazardous threats. Fires on ships are particularly dangerous due to the presence of flammable cargo such as oil, gas, coal, or wood, and because of the isolated nature of the sea environment where external firefighting support is limited. These conditions make early detection and efficient fire suppression crucial to minimizing potential loss of life and property.

Traditional fire detection systems onboard ships typically rely on physical sensors such as flame, smoke, and heat detectors connected to a central alarm panel. These systems provide visual and audible alerts when fire is detected. However, they also require human intervention to verify alarms and often cannot reliably distinguish between smoke and flames, which may result in false alarms. Additionally, such systems detect fire only after it has developed to a dangerous stage, making them less effective for early intervention. In maritime settings, even short delays can be extremely costly, both in terms of financial damage and safety risks[9].

To address these limitations, recent studies have increasingly focused on computer vision and deep learning-based fire detection. Convolutional Neural Networks (CNNs) have demonstrated strong performance in analyzing visual data, automatically extracting features, and distinguishing between fire, smoke, and other elements in complex environments. Researchers have developed CNN-based architectures and temporal models to improve detection accuracy while reducing false alarms. Some approaches, such as lightweight CNNs and the Detection and Temporal Accumulations (DTA) method, have been designed to handle challenging conditions like smoke misclassification and short-term flame variations. Although these models show improved results compared to traditional methods, they can still face issues with computational cost, dataset limitations, and

ISSN 3030-3907

ambiguous image features.

Among the deep learning frameworks, the YOLO (You Only Look Once) family of algorithms has gained particular attention due to its real-time detection capability and efficiency. YOLO-based methods divide images into grids and detect objects quickly, making them suitable for time-critical applications such as fire detection. Studies using YOLOv3, YOLOv4, and YOLOv5 have proposed improvements like depth-separable convolutions, enhanced clustering, and better feature fusion techniques to detect flames at different scales and under varying conditions. More recent models, such as YOLOv6 and YOLOv7, address the limitations of computational cost and accuracy, offering improved performance in detecting small or ambiguous fire regions. These advancements highlight the growing effectiveness of deep learning in providing timely, accurate fire detection solutions for the maritime industry[10].

Fire Dataset Description. To train the fire detection model, a comprehensive dataset was constructed by collecting fire images from multiple sources. The initial set included images and frames from internet videos, captured under varying angles, lighting, and focal conditions to ensure diversity. Since the quantity of fire images remained insufficient, publicly available resources such as Roboflow and Kaggle were incorporated. This resulted in a total of 4186 fire images representing flames and burning scenarios under both daytime and nighttime conditions, thereby enhancing the system's generalization ability.

To further improve robustness, 436 non-fire images were added, capturing ordinary shipboard environments without fire. Together, the dataset comprised 4622 images, split into 70% training, 10% testing, and 20% validation. Importantly, the test set emphasized realistic images, as the ultimate goal was for the system to perform reliably under real-world conditions. Figures illustrating sample fire and shiproom images highlight the dataset's diversity[11].

Data Augmentation and Annotation. Despite the dataset's size, the possibility of overfitting was high, particularly due to recurring visual patterns. To counter this, data augmentation was applied, including geometric transformations such as scaling, rotation, cropping, flipping, and contrast adjustment. These operations not only doubled the dataset size but also increased its variability, enabling the model to recognize fires under challenging conditions like poor lighting or sun reflections.

Annotation played a crucial role in preparing the dataset. Bounding boxes were carefully designed to avoid being either too broad or too restrictive, ensuring that the model learned accurate feature representations. By maintaining a balance in annotation precision, the risk of false generalizations during training was

ISSN 3030-3907

minimized.

Methodology Overview. The YOLO family of models served as the foundation for this study, chosen for their real-time object detection capabilities. YOLO divides an image into grids, predicting bounding boxes and class probabilities within each cell. Over its versions, YOLO has evolved significantly in both speed and accuracy. Early models like YOLOv1 and YOLOv2 were limited in detecting small objects, while YOLOv3 improved with multi-scale anchors but still faced memory challenges. Later versions (YOLOv4, YOLOv5, YOLOv6) progressively enhanced accuracy and computational efficiency through advanced training strategies and dense anchor boxes.

YOLOv7 Architecture. YOLOv7, the core of this work, surpasses its predecessors by introducing the E-ELAN backbone network and improved feature fusion techniques. It integrates bag-of-freebies methods that raise accuracy during training without reducing inference speed. YOLOv7 also incorporates re-parameterized convolutions (RepConv), which consolidate multiple convolutional modules into a single inference step, streamlining performance. One of the model's strengths is its adaptability across multiple configurations, including YOLOv7, YOLOv7-tiny (optimized for edge devices), YOLOv7-W6 (for cloud GPUs), and YOLOv7-X/E6/D6 (scaled for larger datasets). Benchmark results show YOLOv7 achieving up to 161 frames per second (fps) with an average precision (AP50) of 69.7%, outperforming earlier versions in both speed and accuracy[12].

Object Detection Process. In operation, YOLOv7 divides an input image into uniform grids, predicting objects within each region using bounding boxes and probability scores. To manage redundant predictions, it applies non-maximum suppression (NMS), retaining only the boxes with the highest probability and discarding overlaps with lower confidence scores. This ensures that only the most accurate detections remain, as demonstrated in Figure 5 of the original study.

IoU and Evaluation Metrics. Model performance was evaluated using multiple metrics, with Intersection over Union (IoU) serving as a key measure. IoU quantifies the overlap between predicted bounding boxes and ground truth objects, providing insight into localization accuracy. Alongside IoU, metrics such as average precision (AP), mean average precision (mAP), F1 scores, and recall were also applied to comprehensively assess the detection quality[9].

Scaling and Optimization. Unlike prior models that scaled network dimensions independently, YOLOv7 introduces a novel compound scaling approach, simultaneously adjusting depth and width while preserving optimization across layers. This design allows the model to balance computational efficiency with detection performance, making it suitable for real-world deployment in

ISSN 3030-3907

challenging maritime environments[13].

Overall, YOLOv7 demonstrates superior capability in fire detection compared to its predecessors. Its efficient backbone architecture, high processing speed, and adaptability across device types make it highly effective for detecting small-scale fires in ship environments. By leveraging a robust dataset, careful annotation, and augmentation strategies, the system achieves both precision and resilience, providing a practical solution for enhancing onboard safety through real-time fire detection.

Experimental Results and Discussion. The proposed ship fire detection system was implemented and tested on high-performance hardware and evaluated using multiple object detection metrics. The model was trained with a dataset of 4622 annotated fire and non-fire images, leveraging YOLOv6 and YOLOv7 architectures for comparative analysis. Performance was assessed through accuracy, precision, recall, F1 score, and mean average precision (mAP), which are widely recognized measures of detection quality.

The YOLOv7-based system demonstrated strong results, achieving 93% accuracy, 94% precision, 90% recall, and an F1 score of 86%. These outcomes indicate a reliable balance between minimizing false alarms and ensuring sensitivity to actual fire events. The mAP value further confirmed robust detection capability, highlighting the model's suitability for real-time maritime applications.

Experimental analysis showed that the system was effective across varied environmental conditions. Tests under both daylight and low-light scenarios confirmed the model's adaptability, with consistent detection of flames and fire regions. This resilience to lighting variation is particularly valuable for maritime safety, where operational conditions can fluctuate significantly[14].

A comparison with previously published methods emphasized YOLOv7's advantages. While traditional CNN-based approaches and earlier YOLO versions performed well in general detection tasks, the proposed model achieved superior results in distinguishing fire from non-fire objects, handling small flames, and operating under complex shipboard conditions. Such improvements highlight YOLOv7's architectural strengths, including enhanced feature fusion and efficient backbone design.

Despite its effectiveness, several limitations were observed. The system occasionally misclassified fire-like objects—such as intense lighting, sunlight reflections, or bright bulbs—as flames. These false detections stem from visual similarities between fire and certain light sources. Expanding and diversifying the dataset was identified as a key strategy to mitigate this issue. Moreover, the current system is optimized for flame detection but is less effective in identifying smoke,

ISSN 3030-3907

which remains an important factor in early fire recognition[15].

This study presented a YOLOv7-based fire detection framework tailored for ship environments, offering real-time monitoring with high accuracy and reliability. By combining a diverse dataset, data augmentation, and deep learning enhancements, the model achieved competitive performance, surpassing several existing approaches. Its ability to detect small-scale fires promptly supports early response measures, reducing the risk of catastrophic incidents at sea.

The integration of YOLOv7 demonstrates significant potential for improving maritime fire safety. However, future research should focus on extending the dataset, incorporating smoke detection, and addressing challenges posed by low-illumination conditions. Such advancements would further strengthen the model's robustness and expand its applicability across diverse real-world maritime scenarios.

Table-2.

Summary of Works reviewed on computer vision based fire detections.

Author(s) & Year	Algorithm	Dataset	Contribution (Results)	Research Gap / Limitation
Kuldoshbay Avazov et al., (2023)	YOLOv7-based deep learning, CNN	Custom ship fire dataset (day & night images, rotated datasets, augmented)	High accuracy in detecting small/localized fires on ships; reduced false positives using logistic classification + BCE loss; improved real-time detection	Needs more robust training for rare cases like **engine room fires**; dataset still limited to controlled ship scenarios
Hasnat Jamee et al., (2023)	YOLOv8 (one-stage detector, anchor-free)	Public fire/smoke datasets + real-world fire video sequences	Achieved **96.7% detection accuracy**, with **high real-time performance** (fast inference speed) and strong generalization	Computationally heavy for edge/mobile devices; performance drops in very low-light/night scenarios

ISSN 3030-3907

Yibing Ma et al., (2022)	Thermal camera + person detection (hybrid IoT + CV system)	Indoor thermal image dataset (self-collected)	Early fire detection with **90%+ accuracy** in controlled tests; person detection allows evacuation alerts	Limited to **indoor environments**; depends heavily on thermal camera (expensive & less scalable); small dataset
Arnisha Khondaker et al., (2020)	Enhanced chromatic segmentation + optical flow (Lucas–Kanade tracker)	MIVIA, Zenodo, YouTube, and custom fire/no-fire videos (indoor, outdoor, forest)	**97.2% average accuracy**, reduced **false positive rate to 2.21%**, effective across multiple fire scenarios	Recall, precision, F1 not explicitly reported; computational cost higher than simple methods; needs testing on larger real-time streams

CONCLUSION

Fire detection remains a critical area of research due to its direct implications for human safety, infrastructure protection, and environmental preservation. Traditional sensor-based systems, including smoke, gas, and thermal detectors, have long served as reliable and cost-effective solutions, particularly in confined indoor environments. However, they suffer from inherent limitations such as delayed response times, limited coverage, and vulnerability to false alarms. Recent advancements in nanostructured gas sensors, photovoltaic-powered smoke detectors, and optimized thermal camera placement demonstrate incremental improvements, yet scalability and adaptability to complex real-world conditions remain key challenges.

In contrast, computer vision and deep learning-based fire detection approaches have shown remarkable potential in addressing these limitations. By leveraging large-scale datasets and advanced models such as YOLOv5, YOLOv7, and YOLOv8, these systems achieve real-time detection with high accuracy, robust generalization, and adaptability across diverse environments—from electric vehicle charging stations to maritime vessels. Furthermore, attention mechanisms (e.g., CBAM) and architectural optimizations enhance their ability to capture small-scale fire features, reduce false alarms, and improve deployment feasibility. Nevertheless, challenges persist regarding computational cost, low-light

ISSN 3030-3907

performance, and the need for more diverse datasets covering rare and complex scenarios.

The comparative analysis suggests that neither paradigm alone provides a complete solution. Traditional sensors remain valuable in low-cost, small-scale deployments, while computer vision-based approaches are shaping the future of intelligent, real-time, and large-scale fire monitoring. A promising direction lies in the integration of both methods into hybrid fire detection frameworks, combining the robustness of sensors with the intelligence of vision systems. Such synergy could enable earlier warnings, predictive analytics, and seamless integration into smart safety infrastructures. Future research should focus on building scalable hybrid models, enhancing dataset diversity, optimizing algorithms for edge deployment, and addressing environmental challenges such as low visibility and outdoor variability. By advancing toward intelligent, resilient, and adaptive fire detection systems, the field can better safeguard human life, infrastructure, and ecosystems against the growing risks of fire in an increasingly complex world.

REFERENCES

1. Hassan, A., & Audu, A. I. (2022). Traditional sensor-based and computer vision-based fire detection systems: A review. University of Maiduguri.
2. Liu, Y., Miao, F., Su, W., & Li, X. (2021). Design of NB-IoT smoke sensing terminal based on photovoltaic and supercapacitor power supply. Journal of Physics: Conference Series, 1885, 052008.
3. Sulthana, S. F., Wise, C. T. A., Ravikumar, C. V., Anbazhagan, R., Idayachandran, G., & Pau, G. (2023). Review study on recent developments in fire sensing methods. IEEE Access, 11, 90269–90282.
4. Wang, B., Zhao, X., Zhang, Y., Song, Z., & Wang, Z. (2024). A capacitive particle-analyzing smoke detector for very early fire detection. Sensors, 24(5), 1692.
5. Khan, F., Khan, A., & Wang, Y. (2022). Recent advances in sensors for fire detection. Sensors, 22(19), 7450.
6. Quttineh, N. H., Olsson, P. M., Larsson, T., & Lindell, H. (2022). An optimization approach to the design of outdoor thermal fire detection systems.
7. Khondaker, A., Khandaker, A., & Uddin, J. (2020). Computer Vision-based Early Fire Detection Using Enhanced Chromatic Segmentation and Optical Flow Analysis Technique.
8. Rahman, S., & Muhammad, S. (2024). Real-time smoke and fire detection using You Only Look Once v8-based advanced computer vision and deep learning. International Journal of Advances in Applied Sciences, 13(4), 987-999. https://doi.org/10.11591/ijaas.v13.i4.pp987-999

ISSN 3030-3907

9. Liu, Y., et al. (2021). Design of NB-IoT smoke sensing terminal based on photovoltaic and supercapacitor power supply. Journal of Physics: Conference Series, 1885, 052008.
10. Sulthana, S. F., Wise, C. T. A., Ravikumar, C. V., Anbazhagan, R., Idayachandran, G., & Pau, G. (2023). Review study on recent developments in fire sensing methods. IEEE.
11. Wang, B., Zhao, X., Zhang, Y., Song, Z., & Wang, Z. (2024). A capacitive particle-analyzing smoke detector for very early fire detection.
12. Dewangan, A., Pande, Y., Braun, H.-W., Vernon, F., Perez, I., Altintas, I., Cottrell, G. W., & Nguyen, M. H. (2021). FIgLib and SmokeyNet: Dataset and deep learning model for real-time wildland fire smoke detection. arXiv preprint arXiv:2107.08364.
13. Pereira, G. H. A., Fusioka, A. M., Nassu, B. T., & Minetto, R. (2021). Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study. arXiv preprint arXiv:2007.04258.
14. Kim, S. Y., Lee, J., & Park, J. (2023). Forest fire smoke detection based on deep learning. Sensors, 23(4), 2012.
15. Diaconu, B. M., Nicolescu, R., & Popescu, D. (2023). Recent advances and emerging directions in fire detection with machine learning. Fire, 6(3), 97.

ISSN 3030-3907