DEVELOPMENT OF AN EARLY FIRE AND SMOKE DETECTION SYSTEM USING RASPBERRY PI AND MULTIMODAL ARTIFICIAL INTELLIGENCE

Safarov Rustambek Sunnatillo-o'g'li
Bukhara State University, Faculty of Physics, Mathematics
and Information Technologies, 3rd year student
E-mail: rustam0062006@gmail.com |
Tel: (+998) 91 444-55-82
https://doi.org/10.5281/zenodo.20769505

Abstract

This paper investigates the development of an early fire and smoke detection system based on the Raspberry Pi microcomputer and multimodal artificial intelligence technologies. The system is based on the simultaneous processing of visual data video stream via Raspberry Pi Camera Module and chemical indicators . YOLOv8 and MobileNet-SSD neural networks are applied for image analysis, while a weighted decision-making algorithm is used for sensor data evaluation. The research results demonstrate that combining two channels allows fire detection in less than 1.8 seconds on average with 96.4% accuracy, which is 4–6 times faster than traditional signaling systems. Remote real-time notification capabilities via IoT platforms are also discussed.

Keywords: *fire detection, smoke, Raspberry Pi, artificial intelligence, YOLOv8, multimodal sensors, neural networks, IoT, MQ-2, open flame.*

Introduction

Fires are disasters that cause enormous economic damage and pose threats to human lives on a global scale. According to data from international fire safety organizations, millions of fire incidents are recorded worldwide every year. The most critical factor in fire prevention is detecting it at the earliest stage - immediately upon the appearance of smoke and heat - and providing rapid notification. Traditional alarm systems, as a rule, depend on only one parameter , have low sensitivity, and the widespread problem of false alarms is well known.

In recent years, the rapid development of artificial intelligence and computer vision technologies has significantly expanded the possibilities of fire detection based on images. However, systems based solely on visual data also have limitations in cases such as changes in lighting, invisible smoke, or hidden flames. Therefore, the most effective approach is combining multimodal data - that is, jointly analyzing visual and physicochemical indicators. The purpose of this article is to design a fast, affordable, and reliable fire detection system based on the Raspberry Pi platform and multimodal artificial intelligence, and to evaluate its effectiveness.

Literature review and methodology

Existing fire detection systems can be divided into the following groups: traditional sensor-based systems, image-based systems, and hybrid systems. Table 1 below presents their main characteristics.

Table 1. Comparative analysis of fire detection systems.

System Type	Primary Method	Sensitivity / Accuracy	Disadvantages
Traditional sensor (ionization, optical)	Smoke particles / temperature	Average (50–70%)	False alarms, delays
Image-based (classical CV)	Color and contour filter (HSV)	Average (70–85%)	Light dependency
Deep learning (YOLO/CNN)	Neural networks	High (90–95%)	Missing hidden smoke
Multimodal (AI + sensors)	Image + gas + temperature	Very high (95%+)	Complex integration

The analysis shows that systems relying on only one data source continue to have shortcomings in real-world conditions. For this reason, in this work a multimodal approach was chosen - an architecture combining visual (camera) and chemical (gas/smoke) channels. The hardware and software components of the system are described in detail in the following section.

The system uses a Raspberry Pi 4 Model B (4 GB RAM) microcomputer as the central control unit, running a Broadcom BCM2711 Quad-core Cortex-A72 64-bit processor. The following sensors are connected for environmental data collection: MQ-2 for smoke and combustible gases, MQ-7 for carbon monoxide , DHT22 for temperature and humidity, and a KY-026 infrared (IR) flame sensor for detecting open flames. A Raspberry Pi Camera Module V2 (8 MP, 1080p) is used for visual data acquisition. An MCP3008 ADC chip serves to convert analog signals to digital values.

Table 2. Hardware components and parameters of the system.

Component	Function	Key Parameter
Raspberry Pi 4 B	Central control	4 GB RAM, 1.5 GHz
Pi Camera V2	Image capture	8 MP, 1080p@30fps

MQ-2	Smoke and combustible gas	300–10,000 ppm
MQ-7	Carbon monoxide (CO)	20–2,000 ppm
DHT22	Temperature and humidity	±0.5°C, ±2% RH
KY-026	Open flame (IR)	760–1,100 nm
MCP3008	Analog → Digital	10-bit, 8 channel
Buzzer / LED	Local signal	85 dB

The software part of the system is written in Python 3, using OpenCV, PyTorch and RPi.GPIO as the main libraries. In the visual channel, after a frame is captured, smoke and flame objects are first detected using the pre-trained YOLOv8n model; then sensor data is combined based on weighted indicators, and a final decision is made. This decision synthesis constitutes the central idea of the multimodal approach.

Mathematical and algorithmic model

Temperature Dynamics Formula:

In the sensor channel, temperature, smoke, and gas concentration are evaluated. The following relation is used to express the dynamics of temperature increase:

$$T(t) = T_0 + k \cdot (t - t_0)$$

where:

1. $T(t)$ - current temperature at time t
2. T_0 - initial temperature
3. k - rate of change coefficient
4. t - current time
5. t_0 - initial time

If $k > 0$, the temperature is rising, which indicates a fire hazard.

Accuracy Formula:

The accuracy value of sensors is expressed through the following formula:

$$A = \frac{TP}{TP + FP + FN} \times 100\%$$

1. TP - true positives (correctly detected cases)
2. FP - false positives (incorrect signals)
3. FN - false negatives (missed cases)

Weighted Decision-Making Model:

Since the reliability level of each detection channel differs, the following weighted sum model is applied for multimodal synthesis:

$$S = \omega_1 \cdot P_{cam} + \omega_2 \cdot P_{smoke} + \omega_3 \cdot P_{gas} + \omega_4 \cdot P_{temp}$$

where:

1. P_{cam}- probability of flame/smoke detection by the YOLOv8 model (range: [0, 1])
2. $P_{cam}, P_{smoke}, P_{gas}, P_{temp}$ - normalized values of the corresponding sensors
3. $\omega_1, \omega_2, \omega_3, \omega_4$ - weight coefficients for each channel

The optimal weights were determined experimentally as follows:

$$\omega_1 = 0.45, \qquad \omega_2 = 0.25, \qquad \omega_3 = 0.20, \qquad \omega_4 = 0.10$$

subject to: $\sum_{i=1}^{4} \omega_i = 1$

The decision rule is defined as:

Fire Detected = $\begin{cases} True, & if\ S \geq 0.65 \\ False, & if\ S < 0.65 \end{cases}$

If the total score $if\ S \geq 0.65$, the system considers that a fire is in its initial stage and triggers an alert.

F1-Score Metric:

The F1-score (harmonic mean between precision and recall) is used to evaluate the overall quality of the model:

$$F_1 = 2 \cdot \frac{Precision \times Recall}{Precision + Recal}$$

where:

$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}$

Therefore, the complete expression is:

$F_1 = \frac{2 \cdot TP}{2 \cdot TP+FP+FN}$

Experiments and results

Experimental tests were conducted in laboratory and forest-area conditions. The model was trained on a dataset of 2,500 images (with fire and smoke samples). Table 3 compares the performance of various systems.

Table 3. Performance comparison of different fire detection systems.

Metric	Traditional Sensor	Camera Only (YOLOv8)	Multimodal (Ours)
Accuracy	62%	91.2%	**96.4%**
Precision	55%	89.5%	**95.1%**
Recall	70%	87.8%	**94.6%**

F1-score	0.615	0.886	**0.948**
Average Response Time	12.4 s	2.6 s	**1.8 s**
False Positives	High	Medium	**Very Low**

The results demonstrate that the multimodal approach has significant advantages over camera-only or sensor-only systems: accuracy reached 96.4%, and response time was reduced to 1.8 seconds. The number of false alarms was also greatly reduced compared to other systems, since image and sensor data mutually confirm each other.

Conclusion

In this research, an early fire and smoke detection system was developed and tested based on the Raspberry Pi microcomputer and multimodal artificial intelligence.

Main results:

1. The multimodal approach (camera + gas + temperature sensors) raised accuracy to the level of 96.4% and significantly reduced the number of false alarms.
2. Image analysis based on the YOLOv8 neural network enabled real-time detection of fire and smoke within 1.8 seconds.
3. The weighted decision-making model ($\omega_1 = 0.45$, $\omega_2 = 0.25$, $\omega_3 = 0.20$, $\omega_4 = 0.10$) ensured stable system operation by accounting for the reliability of each channel.
4. IoT integration enabled rapid remote notification about fire conditions.
5. The overall cost of the system is considerably low, making it promising for wide application in forestry, warehouses, and residential areas.

In future research, it would be appropriate to investigate adding a thermal camera , further optimizing the model for edge AI devices, and exploring the possibility of autonomous operation with solar panels

References:

1. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767, 2018.
2. Jocher G. et al. Ultralytics YOLOv8. GitHub repository, 2023. URL: https://github.com/ultralytics/ultralytics
3. Howard A. G. et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv:1704.04861, 2017.

4. Muhammad K., Ahmad J., Baik S. W. Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing, 2018.
5. Raspberry Pi Foundation. Raspberry Pi 4 Model B Documentation. Cambridge, UK, 2023. URL: https://www.raspberrypi.com/documentation/
6. Hanqing W., Yong C. Fire detection with image processing and deep learning: A comprehensive review. Journal of Building Engineering, 2022.
7. MQ-2 Gas Sensor Datasheet. Hanwei Electronics Co., Ltd., 2022.
8. MQ-7 Carbon Monoxide Sensor Datasheet. SparkFun Electronics, 2021.
9. DHT22 Humidity and Temperature Sensor Datasheet. Aosong Electronics, 2022.
10. Giyosov U. Sh. Fundamentals of Semiconductor Technologies and Microprocessor System Design. Tashkent, 2024.
11. Sandler M. et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks. IEEE/CVF CVPR, 2018.
12. Foggia P., Saggese A., Vento M. Real-time fire detection for video surveillance applications using a combination of experts based on color, shape and movement. IEEE Transactions on Circuits and Systems for Video Technology, 2015.