scientific reports

OPEN

Research on accurate fire source localization and seconds-level autonomous fire extinguishing technology

Xu Ren[1], Keyu Qu[2], Junlong Guo[1], Hesheng Yin[1]✉ & Bo Huang[1]

With the continuous development of intelligent technology, robots have entered various industries. Firefighting robots have become a hot topic in the field of firefighting and rescue equipment. For firefighting robots, autonomous firefighting technology is the core capability. It includes three steps: flame recognition, fire source location, autonomous firefighting. At present, flame recognition has been widely studied, but the adaptability is poor for different flames. For fire source location technology only rough positioning has been achieved. For autonomous firefighting, the time-consuming water point feedback adjustment method is often used. Those are not suitable for the actual fire rescue. So we proposed to use the visual information, thermal imaging morphological and thermal data of the flame for deep learning, which greatly improves the adaptability of flame recognition. The centimeter-level high-precision positioning of the fire source is achieved. Finally, the proposed water cannon fire source projection method is used to realize rapid water cannon movement instruction generation, and achieve rapid autonomous firefighting. The test results show that the proposed fire source identification algorithm can identify all fire sources up to 15 m at a speed about 15Fps. It can rapid autonomous firefighting within 0.5 s.

Fire is a serious threat to the safety of human life and property. Every year, countries invest a lot of resources in fire rescue. Intelligent equipment for fire rescue has gradually appeared on the rescue scene and has become an indispensable and important force in fire rescue. Fire-fighting robots have become the stand of those equipment. For robot the autonomous firefighting function is an important manifestation of its role in firefighting work[1–5]. To this end, it is necessary to solve the three problems: fire source identification[6–8], fire source positioning[9,10], and rapid water cannon movement[4].

In existing technical research, fire source identification technology has been extensively studied. The use of image processing technology based on visual images to identify fire sources is very popular[8]. Using simple color image information to study image pixel recognition operators is the fastest way to identify flame[4,11–13]. However, the identification method of simple pixel operators often causes a large number of misidentifications. Objects with the same pixel characteristics (RGB pixel characteristics, HSV pixel characteristics) as the fire source are mistakenly recognized as flames. With the development of machine learning, more and more researchers are beginning to use learning to solve this problem[6,7,14]. Such as, Wenxia et al. designed support vector machines to identify flame[15]. It tested in a variety of environments, the recognition success rate reaches about 90%.

In addition, a series of small network learning models like CNN[16–18], RNN[19], UFS-Net[20] were developed to meet the demand for rapid flame identification. For example Muhammad et al. proposed an original, energy-friendly, and computationally efficient CNN architecture for fire detection. It can more accurately identify flame in a variety of scenarios[16]. With the continuous upgrade and iteration of hardware, the use of large networks like YOLO for fire source identification has become a hot topic[21–23]. However, the problem of misjudgment caused by the luminous properties of the flame itself has not been completely solved. A large number of warm-toned and high-brightness objects will be identified as flame. To address this problem, some researchers have begun to use smoke[6,20,24], heat[1] and other flame characteristics to enhance identification. Among them, the addition of thermal images has become the first choice. However, the use of thermal data is more about directly reading high

[1]The State Key Laboratory of Robotics and System, Harbin Institute of Technology, Harbin 150001, China. [2]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore. ✉email: yhs_hit@163.com

temperature value areas after coordinate correction with the depth camera. It is invaluable to feed the thermal imaging data directly into the neural network for flame recognition, rather than processing it as image data. Its advantages are not reflected. Therefore, this paper adds the flame morphological features and temperature values in the thermal image data to the learning network and constructs a lightweight flame recognition network with low resource consumption and high recognition efficiency. It can quickly and accurately recognize flames of any shape and burning material based on color images, thermal images, and temperature information, which greatly improves its adaptability and is no longer limited to recognition in the form of structured fire sources. This is something that traditional technology has not achieved.

In the field of traditional intelligent fire protection technology, accurate positioning of the fire source has never been achieved. After identifying the flame, more research work apply it to alarm and flame spread prediction. There is no further research on the positioning technology of the spatial position of the fire source. At present, the vast majority of autonomous fire-fighting equipment usually adopts the method of interference coverage. Based on the orientation information of the fire source obtained during the flame recognition process, a full coverage extinguishing method is used. This method is relatively simple and crude. In the actual firefighting and rescue process, full coverage cannot be used for extinguishing in many cases. This may lead to explosions and other accidents that seriously affect people's safety. The only way to locate the fire source currently relies on depth cameras or stereo vision cameras for positioning. Its positioning accuracy is low and is not suitable for locating the fire source. It is more about locating pots, buckets and other containers where the fire source is burning. In terms of locating the fire source, the flame has no physical medium and it constantly produces energy and light outwards. It is very unfriendly to non-contact measurement methods such as optical devices[14], which will cause a large number of blind areas and distortion of measurement data. The current existing technologies are more focused on monitoring large-scale fires[11,12,18,21,22]. Monitoring a large fire source can accept a small measurement blind zone. In summary, related technologies for high-precision fire source positioning are very scarce. Most of the positioning information obtained has low accuracy[10] and is not suitable for actual rescue missions. The current intelligent fire-fighting technology has a detection range of less than 100 m, a detection error of about 1 m, and an identification time of about 3–5 s. The inability to obtain high-precision fire source positioning information will directly cause intelligent firefighting equipment such as fire-fighting robots to become blind. Though it can know the fire burning, but it cannot put out the fire quickly and accurately. Therefore, high-precision fire source positioning is very urgent and necessary.

Based on the inaccuracy of fire source positioning, research on rapid motion control of water cannons often uses visual feedback[2,4]. After roughly locating a certain area, spray a water column in the center of the area. Use visual identification of water columns and fire sources to calculate the water column offset. As a result, the water column approaches the fire source little by little. The overall implementation method is complex, time-consuming and inefficient. The average time is in minutes. At the same time, due to the integrated installation method of the water cannon and the monitoring equipment, it is very easy to lose the fire source perspective when spraying the water column. Therefore, the water column alignment cannot be completed. In addition, there is a method of simulating the spray of the water column through the physics simulation of the simulation software. Taking into account gravity, air resistance, and water column pressure, the landing point of the water column is determined using a static formula[25]. However, during the actual fire extinguishing process, the conventional method requires intermediate adjustments and waiting periods, resulting in prolonged operation times and limiting extinguishing effectiveness to short-range targets (approximately 7 m). To address these limitations, we propose a water cannon fire projection method based on pulsating spraying, enabling the rapid generation of motion commands and achieving millisecond-level alignment of the water cannon with the fire source. By dispersing the water column into small water clusters through pulsating spray, interference with the identification system is minimized. Combined with fast motion command generation, this approach enables millimeter-level autonomous fire extinguishing. Notably, this speed performance is unprecedented within the industry.

In summary, to address the poor adaptability of current flame recognition methods, thermal image data are extensively mined, and flame morphological features and thermal characteristics are combined with color imagery to construct a flame recognition network suitable for unstructured environments. This network is capable of recognizing various flame types.

To solve the problem of limited high-precision fire source positioning, a multimodal fusion approach is proposed, wherein a high-energy–density LiDAR sensor is employed to integrate deep visual information. This enables the development of a highly accurate fire source positioning method, effectively replacing traditional direction-based identification methods lacking precision.

Experimental results demonstrate that the system can accurately identify fire sources located approximately 15 m away, achieving a leading level of performance in current technology. The maximum longitudinal positioning error is 0.5 m, and the maximum horizontal positioning error is 0.25 m.

Finally, the water cannon fire projection method is utilized to enable the rapid generation of movement commands. Experimental validation shows that centimeter-level positioning accuracy is achieved, with motion command generation requiring only 0.002 s. Consequently, high precision, high accuracy, and efficient autonomous firefighting are successfully realized.

Experimental systems design
Experimental system composition
To validate the efficient autonomous fire extinguishing technology proposed in this study, field experiments were conducted using the firefighting robot illustrated in Fig. 1. The robot is equipped with a thermal imaging device capable of capturing both the morphological and thermal characteristics of flames, which are then fused with color image features to verify the proposed flame recognition network with strong adaptability to unstructured

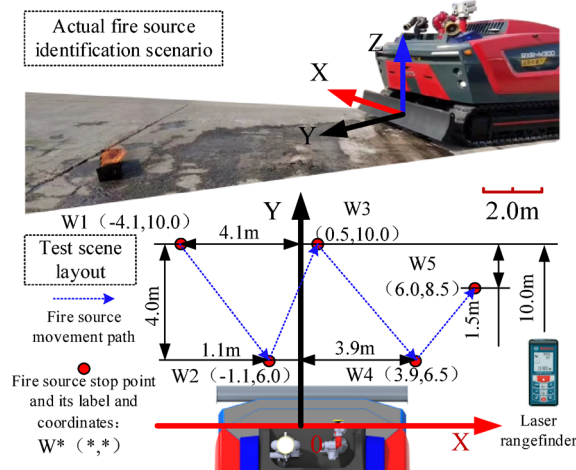**Fig. 1**. Schematic diagram of firefighting robot system.

Fig. 2. Schematic diagram of the autonomous fire test arrangement. W* (*, *) represents the five fixed test points, and their coordinates are the position under the robot coordinate system with the origin at the center of the robot and the front end of the track. The laser rangefinder in the picture is BOSCH GLM70.

environments based on three-dimensional data. Unprecedentedly, in order to achieve high-precision spatial localization of fire sources, a multimodal fusion approach was employed, integrating data from depth cameras and LiDAR sensors. This approach addresses the existing gap in high-precision fire source localization within academic research. Based on the acquired high-accuracy spatial data, rapid, millisecond-level water cannon alignment was performed to enable autonomous fire suppression.

The infrared thermal imaging module utilized was the HM-TD2037T-15/X (HIKMICRO), with a temperature measurement range of 0–550 °C. The depth camera employed was the ZED 2i (STEREOLABS), the LiDAR unit was the M1 (RoboSense), the deputy water cannon was the EXM2 (Elkhart Brass), and the control system was powered by an NVIDIA Jetson Orin.

Test plan

The overall test scene is shown in the Fig. 2. The fire-fighting robot identifies the fire source while moving. In order to ensure the scientificity and reliability of the test, a fire source fixation test between fixed points and a motion test in which the fire source moves between fixed points were chosen. There are two types of flames in the experiment: a torch-shaped flame with a height of 1.5 m and a long-shaped flame with a height of 1 m and a length of 1.5 m.

The five selected fixed points together form a w-shape and are numbered as W1(−4.1,10.0), W2(−1.1,6.0), W3(0.5,10.0), W4(3.9,6.5), W5(6.0,8.5) respectively. The detailed coordinate values are the average value obtained after three ranging measurements with a laser rangefinder (Bosch LM70). The origin of the coordinate system is located at the installation depth of the lidar at the front of the robot in the Y direction and the center of the robot in the X direction. The burning flame used in the test is placed on a large tray and moved at five points. During the movement, the firefighting robot system shown in Fig. 1 is used for the test. The firefighting robot identifies the moving fire source. It locates the fire source and then generates motion control instructions. It controls the water cannon to spray and extinguish the fire source. During the test, in order to eliminate the influence of the size of the fire source, when the fire source moved to the fixed point position, it paused slightly to ensure the stability of the fixed point measurement, and then quickly passed through the moving path. And the details of the test as shown in Table 1.

The overall workflow of the firefighting robot is shown in Fig. 3. From the start of the fire, the system identifies and locks the fire source. Based on the locked fire source, the fire source is located with high precision to obtain

| No	Test plan	
	Test type	Test details
1	L-shape	Ensure that the fire source gradually moves away from the fire source without changing in the X direction relative to the fire robot
2	W-shaped test	The fire source moves in a W shape relative to the fire-fighting robot, and its trajectory forms a W. All sensor data such as thermal infrared imaging, depth cameras, lidar, etc. are obtained in the process

Table 1. Test detail plan.
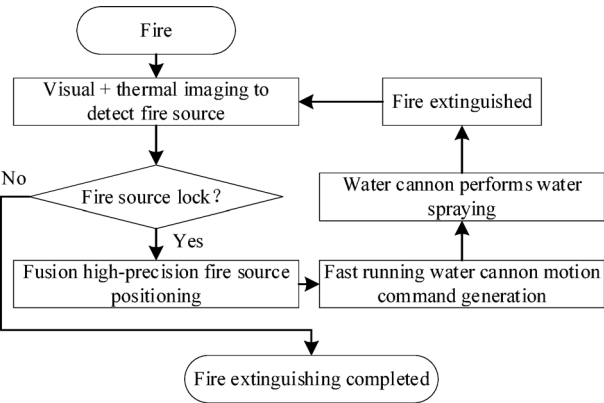
Fig. 3. Overall flow chart of the experiment.

the centimeter-level spatial position coordinates of the fire source. Based on the obtained spatial coordinates of the fire source, the motion control instructions for the water cannon are generated. The generated motion control instructions are executed to aim the water cannon at the fire source and spray the firefighting medium. Then the fire source identification and locking are performed again to determine whether there is a fire source. If no fire source is found, the fire is extinguished.

Key method for efficient and autonomous fire extinguishing

Efficient and autonomous fire extinguishing technology encompasses three key aspects: multi-form flame recognition, high-precision fire source localization, and rapid water cannon maneuvering.

Traditional research has primarily focused on the recognition of standardized, torch-shaped orange flames, where the target characteristics are highly uniform. However, in real fire scenarios, flame behavior varies significantly depending on the environment and the materials involved, leading to considerable differences in flame color and morphology. Consequently, conventional recognition methods cannot be directly applied to diverse real-world flame conditions. Furthermore, most existing flame recognition studies are designed for alarm systems or early flame spread prediction rather than for determining the precise spatial location of fire sources. Typically, only rough directional information (e.g., identifying that a fire is located in the upper-left field of view) is provided, without any spatial accuracy. Such imprecise localization is inadequate for practical firefighting and rescue operations.

According to operational experience from fire departments, effective firefighting requires the precise targeting of fire sources rather than broad-coverage extinguishing strategies. Therefore, high-precision spatial localization of fire sources is both critical and urgent for the development of intelligent firefighting and rescue equipment. In current research on autonomous fire extinguishing, feedback-based approaches are commonly employed, wherein the water cannon's operating angle is continuously adjusted based on the observed deviation between the water column landing point and the fire source location. However, these methods inherently lack real-time precision and operational efficiency. This method has a long operation cycle and a limited range of action within ten meters[25]. It is not suitable for actual firefighting and rescue. Based on the above problems, through the study of the actual mechanism, a targeted solution is proposed.

Identification network based on flame formation mechanism

For the flame formation, the essence of flame is the phenomenon that the air molecules around the reaction zone are heated and move at high speed in the exothermic reaction, thereby emitting light. The energy released from the reaction zone gradually increases from the flame core to the outer flame, and then drops sharply, making the flame have a clearer outline. The boundary between the flame and the surrounding air is where the reaction energy drops sharply. The range from the flame center to the outer flame boundary of the flame is gaseous combustibles or vaporized combustibles, which are undergoing a violent or relatively violent oxidation reaction with the combustion-supporting materials. In the process of gaseous molecules combining, energy waves of different frequencies are released, thus emitting light of different colors in the medium.

4

Therefore, the difference between the combustion material and the combustion-supporting material will lead to different colors of flames. The common orange flame is produced by the full combustion of carbon materials in oxygen. The typical torch-shaped flame is caused by the rise of air flow due to heat. The air flow flows smoothly around the flame and gathers it into a point. When the heat material of the flame burns and the periodic cold air exchanges heat and mass, the conflict between the two forms a clear boundary, thus forming the visual shape of the flame. Therefore, the shape and color of the flame at the fire rescue site are not fixed and are affected by the environment. Pure image deep learning networks are not suitable for flame recognition. After studying the formation mechanism of the above flames, it was found that the most fundamental factor in the formation of flame color and shape is heat. Based on this discovery, the thermal information of the flame is obtained using infrared thermal imaging data, including the morphological information of the flame and the thermal information of the flame. A deep learning network that integrates the flame morphological information, thermal information and color image information is proposed, as shown in Fig. 4.

Targets with flame morphological structural characteristics are set as potential fire source targets from depth camera's color image. The thermal imaging image of the potential target is obtained using data association technology based on morphological characteristics. The identified fire source is matched with the pixel features of the thermal infrared imaging image. The image information of thermal infrared imaging can provide heat information on the basis of providing the morphological characteristics of the fire source. A temperature above ambient temperature is determined as a flame. Perform pixel-level data association between the image information of the depth camera and the image information of thermal infrared imaging. Obtain the correspondence between color image information and heat information. After obtaining the undetermined fire source point identified by the dual network, the heat information of the corresponding area is determined to determine whether it is a fire source. Only when the heat information shows that it belongs to the fire source level, the current undetermined fire source point is determined to be the real fire source.

In principle, any state-of-the-art deep learning based object detector can be used for flame recognition. However, consider the need to run the instrumentation in an embedded system (Orin) with limited resources. Choose the lightweight YOLOv4 tiny[21–23]. Because it provides rapid object detection with guaranteed accuracy. Modify it based on it to make it suitable for three-dimensional data input. The data set used for training uses color image data, infrared thermal imaging data and thermal information data. The acquisition equipment is depth camera and thermal imaging cameras. The collection equipment captures data within a range of 2 m ~ 20 m from the fire source and under various lighting conditions. This dataset contains 15,378 images. The data set is randomly divided into training set (90%) and test set (10%). In the first stage of identification, a large number of pending flame targets are obtained based on color image information. This includes a large number of high-brightness and red-yellow non-flame targets. Based on these undetermined targets, second-level network identification is performed to obtain the flame morphological structure and heat information of the corresponding thermal imaging. Determine the final source of fire. The detection network is trained in Ubuntu 20.04 (64-bit) operating system, which contains Intel(R) Core i7 7700 CPU and Nvidia GeForce RTX 1050Ti GPU. The models were trained with a batch size of 16, a subdivision of 1, and 10,000 iterations. Each input image was resized to a fixed resolution of 416×416 before being fed into the network. The training process took about 3.8 h in total.

High-precision fire source spatial location determination based on energy characteristics

Most of the recognition results obtained by traditional flame recognition are used for alarm and prediction. Even when firefighting is required, the method adopted involves spraying excessive firefighting media in a full coverage manner, following a rough orientation determination. However, such a method is not advisable in
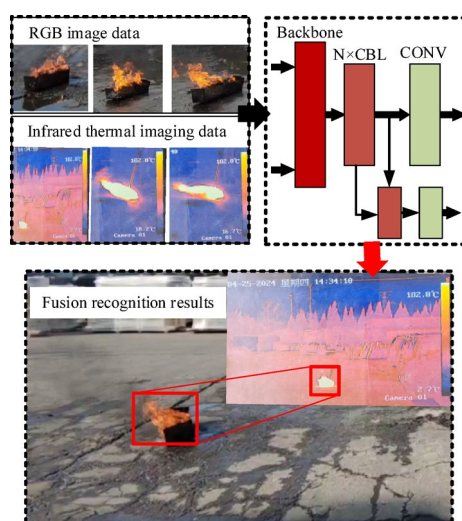

Fig. 4. Schematic diagram of the fire source identification algorithm using thermal imaging fusion enhanced color images.

actual rescue. The full coverage method can only be used when safety conditions are determined. In other cases, more precise fire source strikes are required. To this end, this paper proposes a high-precision fire source spatial positioning algorithm using multimodal sensor fusion. This is an innovative exploration in the field of intelligent firefighting and rescue.

The essence of flame is the phenomenon that the air molecules around the reaction zone in the exothermic reaction are heated and move at high speed, thereby emitting light. Energy scattering brings great trouble to the method of optically determining the location of the fire source. It will disrupt the sensor's perception data and bring a large number of blind spots to the optical ranging method (such as visual ranging). As a result, the method of directly using vision for ranging will cause the depth information of the fire source location to fail. At the same time, the fire source is an energy body and does not have a real entity, so reflection measurement methods are greatly limited. Innovation introduces the concept of multimodal sensor fusion used in the field of autonomous driving into the field of intelligent fire rescue equipment. This paper adopts the lidar enhancement and cooperates with a binocular vision depth camera to perform multi-sensor fusion fire source positioning as shown in Fig. 5. Through preliminary experiments, it can be found that high-energy LiDAR can detect particle phenomena caused by flame combustion. This verifies the effectiveness and feasibility of our scheme. High-density LiDAR point cloud can greatly solve the blind spot problem caused by visual positioning of fire sources.

After obtaining the image information for fire source identification, the fire source area is framed to obtain its pixel coordinate system in the image information as (1). Using the camera's own internal parameters as (2) that obtained by measurement, the pixel coordinates of the fire source are converted into camera coordinates, and the coordinate values of the fire source in the camera coordinate system are obtained. Among them, $K_{3\times3}$ is the internal parameters of the camera. After obtaining the coordinate position of the fire source in the camera coordinate system, the structural characteristics of the robot body are used to convert it into the lidar coordinate system also the word coordinate. The coordinate relationship of the fire source in the lidar coordinate system is obtained as (3). Among them, $E_{4\times4}^{-1}$ is the conversion matrix from the lidar coordinate system to the camera coordinate system. This matrix is the external parameter matrix from the lidar to the camera, which is obtained through measurement and feature point matching. As shown in (4), it is composed of a homogenized rotation matrix and a translation matrix. Finally, the projection relationship between the camera and the lidar coordinate point is obtained (5). Due to the existence of the measurement blind zone, the fire source depth value determination of the camera is relatively unstable. Therefore, the visual positioning information obtained using the above correspondence relationship has uncertainty in the Y direction. Based on the fire source information in the YZ direction obtained by conversion, the fire source depth information of the corresponding LiDAR point cloud is obtained in this area. In this regard, the depth information obtained by vision and the depth information obtained by lidar are weighted fused, and finally high-reliability fire source positioning and identification information is obtained.

$$p_1 = \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix}. \tag{1}$$

$$P_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = z_1 \begin{bmatrix} f_x & 0 & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ v_1 \\ z_1 \end{bmatrix} = z_1 K_{3\times3} \begin{bmatrix} u_1 \\ v_1 \\ z_1 \end{bmatrix}. \tag{2}$$

$$\begin{bmatrix} P_{word} \\ 1 \end{bmatrix} = E_{4\times4}^{-1} \times \begin{bmatrix} P_1 \\ 1 \end{bmatrix} = \begin{bmatrix} P_{1L} \\ 1 \end{bmatrix}. \tag{3}$$

$$E_{4\times4}^{-1} = \begin{bmatrix} r_{x1} & r_{y1} & r_{z1} & t_x \\ r_{x2} & r_{y2} & r_{z2} & t_y \\ r_{x3} & r_{y3} & r_{z3} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R_{3\times3}^{-1} & T \\ 0 & 1 \end{bmatrix}. \tag{4}$$

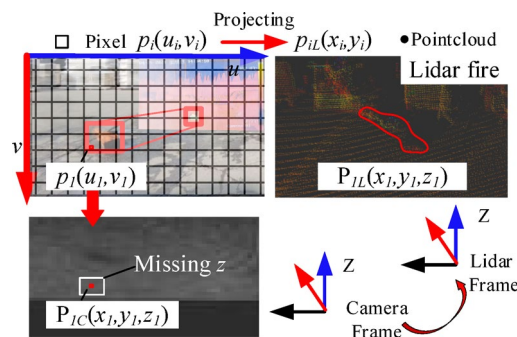$$P_{word} = R_{3\times3}^{-1} \cdot P_1 + T. \tag{5}$$

Fig. 5. Lidar enhanced visual positioning flame algorithm process.

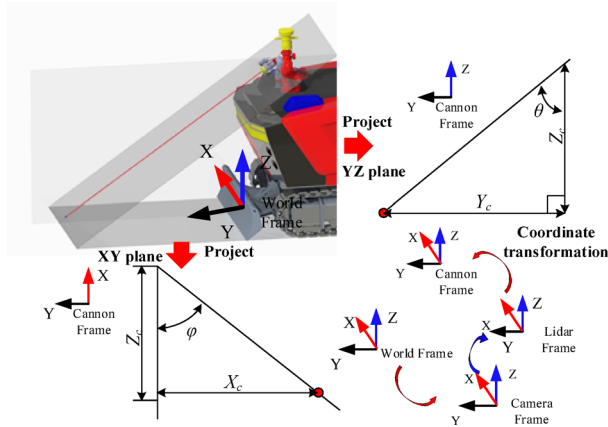*C. Millisecond-Level Water Cannon Fire Source Projection Positioning*

Fig. 6. Water cannon fire source projection and positioning method.

Method	Torch flame (%)	Long line flame (%)	Transparent flame (%)	Blue flame (%)
YOLOv8 (accuracy)	97.6	96.5	53.4	62.1
Our proposed (accuracy)	95.7	95.3	95.9	94.8

Table 2. Recognition network comparison results.

Millisecond-level water cannon fire source projection positioning

After obtaining the high-precision fire source coordinate system, aligning the drop point of the water cannon to the fire source is necessary to achieve autonomous firefighting. Currently, the water column physics model is used to construct complex water cannon control instructions, or the water cannon's movement angle is continuously adjusted by feedback recognition to direct the water column to extinguish the fire source. However, these methods are time-consuming and often take seconds, which is impractical for fast-moving firefighting robots. A faster method is needed to quickly extinguish the fire source during the robot's movement. To address this, complex control networks and slow visual feedback adjustments were abandoned. A new method, the Water Cannon Fire Source Projection Positioning, was proposed to achieve high-frequency and rapid water column drop point alignment, as shown in Fig. 6.

The water column generated by the water cannon takes the form of a columnar ray, which can be equivalent to a ray in a three-dimensional space determined by the water cannon outlet as the origin and the pitch and yaw of the water cannon. In the Fig. 6, the red ray in the figure is equivalent to the projection ray of the water column. Based on this, the fire source projection positioning of the water cannon is carried out. The obtained high-reliability positioning data after camera LiDAR fusion of the fire source is converted into the water monitor coordinate system through a transformation matrix. Project it into the YZ plane to obtain the YZ plane positioning triangle. Its height is Z_c and its length is Y_c, from which the pitch angle θ of the water cannon can be obtained. Project it into the XY plane to obtain the XY plane positioning triangle. Its height is Z_c and its length is X_c, from which the yaw angle φ of the water cannon can be obtained. On the premise of ensuring positioning accuracy at the control level, the highest speed of control command generation is achieved.

Results and discussion
Highly adaptable flame recognition

A full range of experiments were conducted on the recognition network under the flame formation mechanism based on flame characteristics mentioned in the previous chapter. The recognition results of the recognition network constructed using traditional flame color image information were compared with the three-dimensional data learning network model constructed based on the flame formation mechanism proposed by us. A comparison chart of the test results is formed as shown in Table 2. In table , transparent flames represent flames that are difficult to observe with the naked eye in the macroscopic world. Blue flames represent flames that appear blue overall in the macroscopic world. Since there is currently no model that uses data of the same dimension for network training, a simple comparison with the YOLOv8 network is chosen. This shows that the improvement of the proposed network does not reduce the success rate of recognition. On the contrary, the proposed method expands the application scope of the recognition network.

As shown in Table 2, the latest YOLOv8 was used for flame recognition under color images as a comparison test. The test results show that for torch-shaped and long-line flames, the YOLOv8 network achieved a higher recognition success rate than the network we proposed. However, for special transparent flames and blue flames, which are similar to common flames, its performance is poor, with a recognition success rate of only about 60%. Although our proposed recognition network does not achieve as high a recognition success rate as YOLOv8, it

excels in handling various flame conditions, with an average accuracy rate of around 95%. This demonstrates the strong adaptability of our network.

The flame recognition network of common color images has poor adaptability to variable flames. From the perspective of the mechanism of flame formation, the shape and color of the flame seen by the human eye can only represent a small part of the flame. The essence of flame combustion is the change of energy, and the color and shape of the flame only represent the intensity of the heat exchange between its energy and the external atmosphere. The orange flame is only a small part of the flame family. Therefore, the color image neural network based on orange flame learning cannot accurately identify non-orange flames. If all colors of flames are included in the learning data, it will lead to the generalization of the entire learning network. Flame color basically includes most of the colors produced by the three primary colors. This is fully reflected in the poor recognition rate of transparent and blue flames. Therefore, the color image neural network based on orange flame learning cannot accurately identify non-orange flames. If all colors of flames are included in the learning data, it will lead to the generalization of the entire learning network. Flame color basically includes most of the colors produced by the three primary colors. This is fully reflected in the poor recognition rate of transparent and blue flames. Compared with traditional methods, the addition of thermal imaging data makes the network we proposed have a wider range of applications and can cope with more complex situations.

High-precision fire source spatial positioning

After obtaining the flame recognition result, the multimodal sensor fusion technology was used to achieve high-precision fire source positioning at the centimeter level. This was based on high-energy density LiDAR and depth images, preparing for precise fire source attacks in fire rescue operations. High-precision fire source positioning is rare in the field of intelligent fire rescue equipment. To verify the feasibility of the proposed method and algorithm, a performance test experiment was conducted on the fire-fighting robot system, as shown in Fig. 1. The fire-fighting robot starts from a position two meters away from the fire source and moves away from the fire source in a straight line as Table 1 L-shape.

The fire source location distance information and time result image identified by the algorithm are shown in Fig. 7. The fire source identification network finally deployed in the embedded system can achieve a speed of 15FPS fire source identification. The positioning method proposed in this paper can identify the fire source distance up to about 13.5 m within the range of centimeter-level accuracy. This is a new breakthrough in the field of fire source identification. In order to verify the specific accuracy of fire source positioning under the proposed fusion algorithm, the experiment shown in Fig. 2 was used to conduct three ratios (1/9 fusion, 2/8 fusion, 3/7 fusion) of LiDAR camera fusion test experiments, and the results are shown in Fig. 8.

The horizontal and vertical coordinates in the figure represent the distance of the fire source based on the X and Y directions of the robot respectively. It can be seen from the figure that the distribution of the four groups of identification points presents a W-shape as a whole, which shows the scientific nature and effectiveness of the experiment. The four groups of points basically overlap, which shows the effectiveness of the fusion algorithm test. However, it can be found that the traditional depth visual positioning (TIC) results represented by the black triangles are more divergent than the positioning method we proposed, indicating that the fusion lidar method mentioned in this paper is effective. Based on this, the experimental data is further analyzed. The Z direction of the TIC data and the fusion data are extracted for comparison, and the results shown in Fig. 9 are obtained.

The horizontal and vertical coordinates respectively represent the distance between the fire source and the fire-fighting robot in the X and Z directions. It can be seen in the Fig. 9 that there are obvious differences between TIC data and fused data. TIC data is discretely distributed in two areas, and its values are concentrated in the range of -0.3 to 0.3. In theory, the fire source positioning coordinate values in this experiment should be continuous in the X and Y direction and in the Z direction concentrated at the height of the ground (the fire source must be horizontal on the ground). The data after LiDAR fusion is obviously more consistent with this characteristic. This result fully proves the strong interference of flames on optical ranging elements mentioned
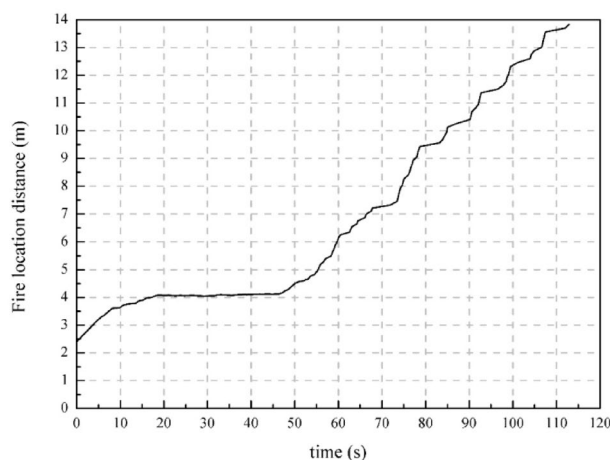

Fig. 7. Fire source positioning extreme position performance test chart.
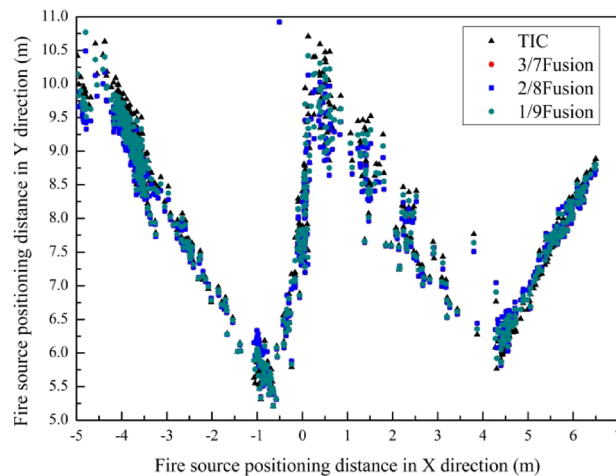
Fig. 8. Fire source identification and positioning map under the fusion ratio of infrared thermal imaging fusion enhancement and lidar multi-sensory fusion with different weights: "a/bFusion" represents the test data when the lidar detection value weight ratio is $a/(a+b)$ and the thermal infrared imaging fusion enhancement ratio is $b/(a+b)$.
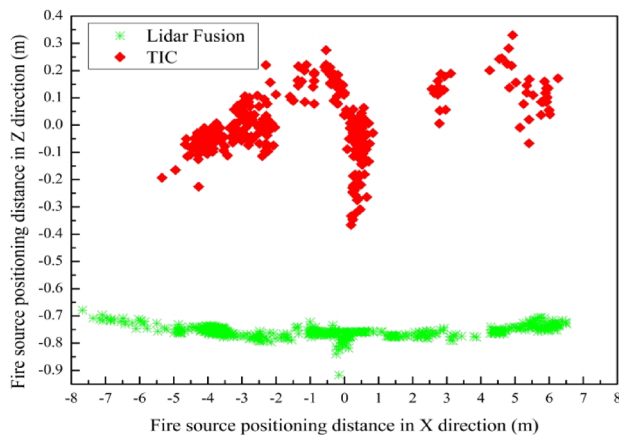

Fig. 9. Z-direction comparison chart of fire source identification and positioning results under infrared thermal image fusion enhancement and lidar multi-sensory fusion.

in the method section. The traditional widely used method of using depth vision to locate the fire source is not feasible. This fully proves the necessity of our proposed method.

In order to further accurately determine the optimal fusion ratio of LiDAR, perform fixed-point error analysis on the data at five points W1-5. The error distribution in the X and Y directions is shown in Fig. 10. The left half of the abscissa in the figure represents the error in the Y direction, and the right half represents the error in the X direction. It can be seen from the figure that the error in the Y direction of the left coordinate is obviously larger than the error in the Z direction on the right, and the difference is obvious. This is due to the difference in measurement range. The Y-direction measurement range is 6.0–10.0 m, and the X-direction measurement range is −4.1–6 m. The difference in measurement base causes a large difference in error. Comparing the errors under three different fusion ratios, it can be found that the error under the 1/9 fusion ratio is significantly smaller than the other two fusion ratios. The error of 1/9 in the X direction of point W4(3.9,6.5) is larger than the other two fusion ratios, but the larger ratio is about 0.04 m. Overall, 1/9 is the best fusion ratio. Considering that the water column cross-section diameter of the water cannon is 0.5–1.5 m, a 1/9 ratio is used to fuse the lidar data and TIC data.

Therefore, the fusion fire source positioning results at 1/9 scale were quantitatively analyzed. The positioning mean and standard deviation within the range of five points are shown in Fig. 11. It can be seen that the maximum error occurs in the Y direction at W3 (3.9, 10.0), which is 48 cm, accounting for 4.8%. The maximum standard deviation appears in the Y direction of W2 (−1.1, 6.0) and is 0.3. The standard deviation of about 0.3 means that the data is relatively concentrated and there is not much deviation. The designed positioning algorithm has good stability. In summary, its positioning accuracy can achieve a horizontal (X direction) deviation of no more than 7 cm, an error of 1.79%, and a longitudinal (Y direction) deviation of no more than 48 cm, an accuracy of
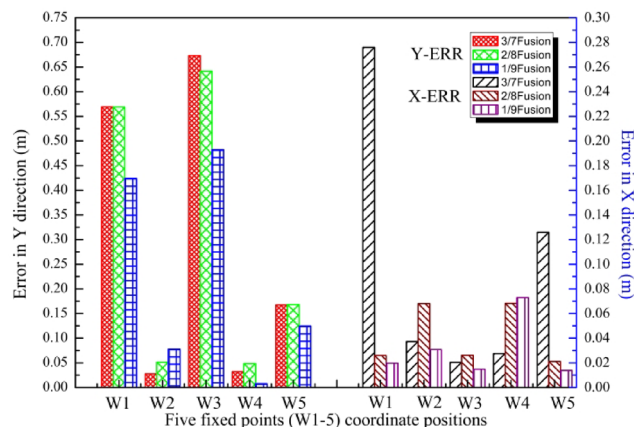
Fig. 10. Comparison of fire source positioning errors under different fusion ratios: Y-ERR is the positioning error in the Y direction; X-ERR is the positioning error in the X direction.
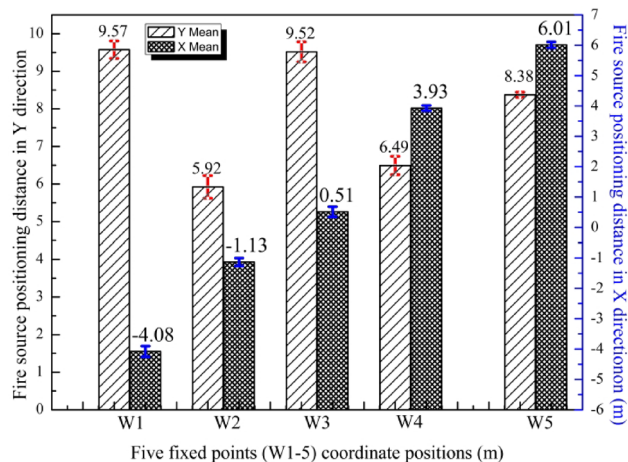

Fig. 11. Positioning mean results and corresponding standard deviation under the optimal fusion ratio of 1/9.

4.8%. The high-precision fire source positioning algorithm proposed by us can achieve fire source positioning of about 13.5 m within the range of centimeter-level accuracy, and the positioning error accuracy can be guaranteed at the centimeter level. It far exceeds most existing related works. Compared with the traditional fire source positioning technology (detection error of about 1 m, identification time of about 3-5 s), the accuracy of our multi-sensor fusion positioning technology is not much different, but the identification time has been qualitatively improved, making a huge contribution to the timeliness requirements of fire rescue. It is expected to be applied in actual fire rescue work.

Millisecond-level efficient autonomous fire extinguishing execution

After obtaining the high-precision fire source spatial position, it is necessary to quickly generate water cannon motion control instructions. Accurately aim the water cannon at the fire source before the mobile robot produces a large displacement, and accurately extinguish the fire source. During the experiment, the relative speed between the fire source and the firefighting robot was maintained at 1 m/s. At this speed, the water cannon needs to achieve a millisecond-level control speed and finally accurately aim at the fire source. Therefore, the traditional feedback control method and the method of generating using physical models are not applicable. They often take tens of seconds to complete the alignment work. This is completely unacceptable in firefighting robots. Based on this, we proposed the millisecond-level water cannon fire source projection positioning method, which can efficiently realize the high-precision motion control of the water cannon. The experimental results are shown in Fig. 12.

In the Fig. 12, the landing point of the water cannon water column is determined by the above-mentioned water cannon fire source projection positioning method, which is recorded as (WCL) and the fire source positioning data is recorded as (TICL). Image processing technology is used to process TICL data and WCL data. Overlap two image data composed of points of the same size and type under the same coordinate system at the same scale to calculate the coverage area. The final measurement of the coverage rate of drop points reached 95%. The specific performance of the entire autonomous fire extinguishing operation is shown in Table
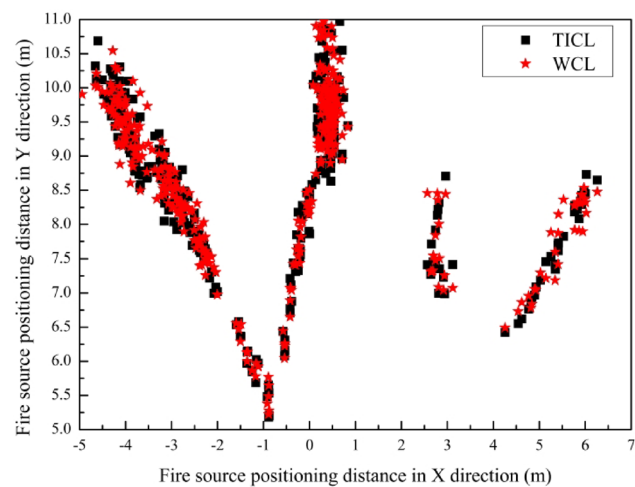
Fig. 12. The fire source identification and positioning results after multi-sensor fusion are compared with the water cannon water column impact point results derived from the water cannon fire source projection principle.

Identification	15Fps
Positioning	0.006 s
Command generation	0.002 s
Motion execution	<0.5 s
Success rate	98%

Table 3. Autonomous fire extinguishing operation.

Fig. 13. Record of actual test results.

3 and Fig. 13. Among them, the time and success rate of identification, positioning, and instruction generation are the average values of 400 tests as shown in Fig. 12. The motion execution time is the limit time within the maximum range of water cannon movement. The maximum distance is the maximum distance at which fire can be extinguished under the current success rate. The entire autonomous fire extinguishing execution time can be controlled within the range of 0.5 s, which is limited by the movement speed of the water cannon itself. This is limited by its mechanical structure. For the entire software system, the time taken from identification, positioning, and motion command generation is completely within the millimeter range. This speed is rare in the entire intelligent fire rescue equipment.

As shown in Fig. 13, this is a recording of the actual experiment. At time Fig. 13a, the fire source is identified, at time Fig. 13b, the auxiliary water cannon is controlled to start spraying water, and at time Fig. 13c, the fire source is extinguished. Through frame-by-frame analysis of the actual recorded video, the average value of stages a-c is about 0.5 s. The method we proposed shortens the firefighting execution time from minutes to less than 1 s using the traditional feedback control method.

Conclusion

Autonomous fire extinguishing is a systematic project. This article takes efficient and autonomous fire extinguishing as the core and completely upgrades and enhances the technical theory from three aspects. Autonomous extinguishing of fire is an indispensable function of intelligent firefighting and rescue equipment represented by firefighting robots. In-depth research has been carried out on the basis of the traditional use of color images for fire source identification and alarm. Based on the needs of actual firefighting and rescue, the fire source positioning method has been expanded, and the current rough positioning strategy has been upgraded to centimeter-level high-precision positioning. The water cannon fire source projection positioning method is used to realize the generation of fast water cannon movement instructions. The millisecond-level speed of the entire autonomous firefighting is achieved. In summary, the specific contributions of this article are as follows:

Based on the flame formation mechanism, combined with the flame morphological information, flame thermal information and color image information. A deep learning network that integrates flame morphological information, thermal information and color image information is proposed, which can achieve highly adaptable flame recognition. Compared with the latest YOLOv8 network, a wider range of flame recognition is obtained while ensuring a small reduction in accuracy.

Based on the concept of multimodal sensor fusion, the high-energy density LiDAR fusion deep vision is used to achieve high-precision centimeter-level positioning of the fire source for the first time. This is a qualitative leap compared to the current rough positioning of the fire source or direct positioning using deep vision. It solves the problem of blind spots when deep vision is used to locate the fire source. This is an innovative exploration in the field of intelligent firefighting and rescue.

Based on the proposed water cannon fire source projection positioning method, combined with the recognition of flames and the positioning of fire sources, millisecond-level water cannon motion control is achieved, and the entire autonomous fire extinguishing operation can be completed within 0.5 s, and the maximum fire extinguishing distance reaches 13.5 m. This is a brand-new breakthrough in the performance of autonomous fire extinguishing systems.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

1. Chaoxia, C., Shang, W., Zhang, F. & Cong, S. Weakly aligned multimodal flame detection for fire-fighting robots. *IEEE Trans. Industr. Inf.* **19**(3), 2866–2875 (2023).
2. McNeil, J. G. & Lattimer, B. Y. Robotic Fire Suppression Through Autonomous Feedback Control. *Fire Technol.* **53**(3), 1171–1199 (2016).
3. Chao Ching, H., Ming Chen, C. & ChihHao, L. Machine vision-based intelligent firefighting robot. *Key Eng. Mater.* **450**, 312–315 (2010).
4. Jinsong, Z., Wei, L., Da, L., Hengyu, C. & Ge, Z. Intelligent fire monitor for fire robot based on infrared image feedback control. *Fire Technol.* **56**(5), 2089–2109 (2020).
5. Shuo, Z. et al. Design of intelligent fire-fighting robot based on multi-sensor fusion and experimental study on fire scene patrol. *Robot. Auton. Syst.* **154**, 21 (2022).
6. Gaur, A., Singh, A., Kumar, A., Kumar, A. & Kapoor, K. Video flame and smoke based fire detection algorithms: A literature review. *Fire Technol.* **56**(5), 1943–1980 (2020).
7. Chengtuo, J. et al. Video fire detection methods based on deep learning: Datasets, methods, and future directions. *Fire* **6**(8), 27 (2023).
8. Çetin, A. E. et al. Video fire detection—review. *Digit. Signal Process.* **23**(6), 1827–1843 (2013).
9. Tao, S., Baoping, T., Minghang, Z. & Lei, D. An accurate 3-D fire location method based on sub-pixel edge detection and non-parametric stereo matching. *Measurement* **50**, 160–171 (2014).
10. Dongyao, J., Wang, Z. & Jiayang, L. Research on flame location based on adaptive window and weight stereo matching algorithm. *Multimed. Tools Applicat.* **79**(11–12), 13 (2020).
11. Sudhakar, S. et al. Unmanned aerial vehicle (UAV) based forest fire detection and monitoring for reducing false alarms in forest-fires. *Comput. Commun.* **149**, 1–16 (2020).
12. Xiong, D. & Jinding, G. A new intelligent fire color space approach for forest fire detection. *J. Intell. Fuzzy Syst.* **42**(6), 18 (2022).
13. Yang, L., Zhang, D. & Wang, Y.-h. A new flame segmentation algorithm based color space model. *Presented at the In 2017 29th Chinese Control And Decision Conference* 2708–2713 (2017).
14. Haojun, Z., Xue, D. & Zhiwei, S. Object detection through fires using violet illumination coupled with deep learning. *Fire* **6**(6), 15 (2023).
15. Bao, W., Zhong, Z., Zhu, M. & Liang, D. Image flame recognition algorithm based on M-DTCWT. *J. Phys: Conf. Ser.* **1187**(4), 6 (2019).
16. Muhammad, K. et al. Efficient deep CNN-based fire detection and localization in video surveillance applications. *IEEE Trans. Syst., Man, Cybern.: Syst.* **49**(7), 1419–1434 (2019).
17. Latif, A. & Chung, H. Fire detection and spatial localization approach for autonomous suppression systems based on artificial intelligence. *Fire Technol.* **59**(5), 2621–2644 (2023).
18. Shaoxiong, Z. et al. A forest fire recognition method based on modified deep CNN model. *Forests* **15**(1), 18 (2024).

19. Wu, X. et al. Smart detection of fire source in tunnel based on the numerical database and artificial intelligence. *Fire Technol.* **57**(2), 657–682 (2020).
20. Hosseini, A., Hashemzadeh, M. & Farajzadeh, N. UFS-Net: A unified flame and smoke detection method for early detection of fire in video surveillance applications using CNNs. *J. Comput. Sci.* **61**, 25 (2022).
21. Gong, C. et al. Multi-scale forest fire recognition model based on improved YOLOv5s. *Forests* **14**(2), 16 (2023).
22. Qian, J. & Lin, H. A forest fire identification system based on weighted fusion algorithm. *Forests* **13**(8), 1301 (2022).
23. Sen, L. et al. An indoor autonomous inspection and firefighting robot based on SLAM and flame image recognition. *Fire* **6**(3), 21 (2023).
24. Yehan, S., Lijun, J., Jun, P., Shiting, S. & Libo, H. A satellite imagery smoke detection framework based on the Mahalanobis distance for early fire identification and positioning. *Int. J. Appl. Earth Obs. Geoinf.* **118**, 13 (2023).
25. Gao, X. et al. Design and Experimental Verification of an Intelligent Fire-fighting Robot. *Presented at the 2021 6th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM)* 943–948 (2021).

Author contributions

R.X. wrote the main manuscript text. H.B. was responsible for the provision and guarantee of experimental equipment. G.J. was responsible for the derivation and verification of the formula. Y.H. was responsible for the overall structural planning of the article, and Q.K. was responsible for the data collation. All authors reviewed the manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.