Fire Technology
https://doi.org/10.1007/s10694-025-01738-6

Fire Vigilance Pocket: An Intelligent APP for Real-Time Fire Hazard Quantification

Zilong Wang[1] · Tianhang Zhang[1] · Xinyan Huang[1]

Received: 10 September 2024 / Accepted: 1 April 2025

Abstract
Real-time fire hazard estimation is an essential step for smart firefighting practice. This paper introduces the Fire Vigilance Pocket Edition application (FV Pocket), which is designed to enable automatic fire identification and quantification using computer vision and deep learning techniques, for real-time fire surveillance. The application comprises four main functions, namely, fire detection, fire segmentation, fire measurement, and fire calorimetry. Fire detection is performed by YOLOv5, which localizes the fire source in the image and marks the location of the flame area. Subsequently, the detected fire area is input into the Swin-Unet model to separate the flame and background, enabling the real-time display of the fire area. Additionally, image-based fire measurement techniques are used to determine the flame height and the flame area according to the estimated reference scales, which also facilitates the rescaling of raw images. Finally, the rescaled images are fed into a pre-trained fire calorimetry model to identify the heat release rate of the fire. The models used in FV Pocket, their design, and main features are discussed, and the application is demonstrated using real fire events under various scenarios. The potential uses and limitations of FV Pocket are also addressed in this work.

Keywords Fire safety · Smart firefighting · Fire calorimetry · Computer vision · Deep learning

1 Introduction

Fire detection and hazard quantification play a crucial role in safeguarding lives, property, and the environment away from the devastating fire disasters. A timely fire detection allows for swift evacuation of occupants, activation of fire suppression systems, and the initiation of appropriate emergency response procedures [1, 2]. The detected fire signal can also trigger the building automatic fire services system to rapidly reduce the spread and intensity of fires, limiting the damage and potential loss of life [3]. Moreover, the local fire

✉ Xinyan Huang
xy.huang@polyu.edu.hk

1 Departmemt of Building Environment and Energy Engineering, Research Centre for Smart Urban Resilience and Firefighting, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Published online: 29 April 2025

brigade will be alerted to prepare the firefighting and rescue operations, which has been standard practice globally.

Current fire detection methods (Fig. 1), including the human surveillance [4], automated systems [5], or the combination of human and sensors [6]. Human surveillance relies on trained individuals to actively monitor an area for signs of fire or the occupants' observation when they are near the fire scene. However, it suffers from limitations such as limited coverage, potential for human error, and high cost associated with employing sufficient human observers [7]. Automated systems utilize sensors and cameras to detect fires and trigger alerts, but they are prone to false alarms, lack adaptability, and require frequent maintenance. Most modern fire detection practices integrate human surveillance with automated systems to leverage their respective strengths. For example, once a fire detector sends an alarm, the property manager can further check the surveillance camera or patrol the area to confirm the fire alarm and reduce the false alarm. Despite the high complexity and cost of the current fire detection system, it can neither provide quantitative fire information to monitor the fire evolution [8, 9] nor reduce the manpower and false fire alarm [10].

To address these drawbacks, ongoing research focuses on improving accuracy and reliability of fire detection system [6], developing advanced algorithms and computer vision methods for automatic fire detection [11], and exploring emerging technologies such as drones and remote sensing [12]. Dampage et al. [13] adopted a wireless sensor network combined with a machine learning regression model to detect forest fires at their early stages, aiming to improve the accuracy of fire detection. Vikram et al. [14] proposed a novel localization technique that employed the Support Vector Machine and wireless sensor network for early prediction of forest fires. Sharma et al. [15] designed an early fire detection system using sensor network and Unmanned Aerial Vehicle technology to mitigate fire incidents. Other researchers looked into the fuel distribution [16] and human behaviors [17] to evaluate the real-time fire risk [18]. Today, deep learning models such as Convolutional Neural Networks (CNN) [19–21], Faster R-CNN [22–24], and YOLO [25–27] have been widely employed to improve the fire detection.

Though these approaches contribute to enhancing fire detection capabilities, most of them can only judge the occurrence of the fire (i.e., Y/N) [28] without more quantitative information for fire hazard evaluation. The measurement of flame height [29], flame area [30], and fire heat release rate (HRR) [31, 32] is of paramount importance in quantifying fire hazard and assessing its potential consequences. To be more specific, flame height provides valuable information about the vertical extent of the fire, indicating the potential fire spread to the ceiling and higher floors. The burning area quantifies the spatial coverage of the fire, which determines the potential impact on surrounding fuels and structures. The fire HRR or its power is considered the most critical parameter of fire, aiding in assessing

Fig. 1 Fire inspection and detection by humans, automated sensors and surveillance camera systems

the destructive potential of fire spread and the associated risk of structural failure. However, only limited studies have attempted to quantify these critical fire parameters. For example, Zhang et al. [8, 33] utilized temperature sensor network and artificial intelligence to track the real-time fire growth and forecast the risk and moment of critical fire events like flashover and backdraft. Cao et al. [34] proposed an LSTM-Kriging neural network model that utilizes temporal data from fire sensors to generate real-time indoor fire threat fields. Zhang et al. [35] developed a CNN-BiLSTM-based fire location detection model to address challenges in firefighting strategies for extrawide immersed tunnels. Reddy et al. [35] proposed an AI-based recurrent neural network optimized with a whale optimization algorithm (AI-RNN-WO) for early fire hazard detection in smart cities. Shadrin et al. [36] utilized remote sensing data and a neural network based on the MA-Net architecture to predict wildfire spread over 1–5 days, focusing on environmental and climate data with spatial distribution features.

Despite the utility of sensor data, particularly time-series data, these methods often suffer from limitations such as low spatial resolution, restricted coverage, and susceptibility to sensor malfunctions in extreme conditions. In contrast, image-based methods provide richer spatial–temporal information, enabling more detailed and interpretable analyses. More recently, we proposed the AI-image fire calorimetry method, which uses smoke and fire images to feed the AI model for measuring the fire HRR [37–40], these approaches demonstrate the potential to overcome the limitations of sensor-based methods and highlights the robustness of using image data. However, such a method has not been integrated into a portable tool to support daily fire safety management and firefighting operations.

This work introduces Fire Vigilance Pocket Edition application (FV Pocket), a web version designed for real-time intelligent fire monitoring and quantification of fire hazards. This smart tool incorporates image-based deep learning models to perform fire identification and quantification, providing crucial information for decision-making processes. Unlike many existing fire models, which are often restricted to internal use by researchers or specific teams, FV Pocket is the first to openly share its advanced fire models with all users. This open-access approach bridges the gap between fire model development and practical applications, making these tools accessible to a wider audience. Additionally, FV Pocket features an intuitive interface and user-friendly design, ensuring that even non-expert users can effectively utilize its functionalities.

We describe the development process, design considerations, and key features of FV Pocket, highlighting its effectiveness through demonstrations using real fire events. Furthermore, the limitations and assumptions associated with the application's usage, as well as potential avenues for future improvements are discussed. By maintaining continuous vigilance for fire hazards, implementing proactive fire prevention measures, and promptly responding to fire incidents, fire vigilance plays a crucial role in ensuring life safety and the preservation of valuable assets.

2 Methodology

In this study, the FV Pocket system employs established deep learning models in computer vision to fulfill distinct objectives. Firstly, YOLOv5 [39] is utilized for fire detection, enabling the identification and localization of fire sources. Secondly, Swin-Unet [40, 41] is employed for fire segmentation, facilitating the precise delineation of flames from the background. Lastly, VGG [37, 38, 42] is employed for fire calorimetry, enabling the estimation of the transient heat release rate (HRR) based on the captured fire images. To establish the correspondence between image dimensions and actual dimensions, a reference scale is incorporated. Furthermore, the flame segmentation process is expedited by leveraging the outcomes of fire detection, as the fire-monitored flame regions are input into the fire segmentation model to reduce the effect of flame reflections. The framework of the FV Pocket is shown in Fig. 2, details on the specific AI models can be found in the reference.

2.1 Fire Detection

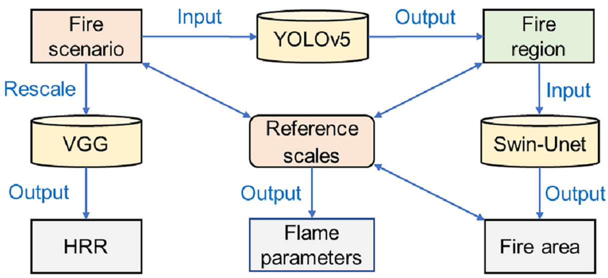
Fire detection plays a crucial role in ensuring the safety of people and property in various settings. It serves as an early warning system that alerts individuals to the presence of a fire, enabling timely evacuation and response actions. By detecting fires at their early stages, potential risks and damages can be minimized.

This study employs YOLOv5 (Fig. 3) as an automated fire detection system to overcome the constraints of human vigilance. YOLOv5 incorporates a deep convolutional neural network as its underlying architecture to extract hierarchical features across various scales. Additionally, it incorporates a neck component to enhance feature representations by aggregating multi-scale features. The detection head of YOLOv5 is responsible for generating predictions, including bounding box coordinates and class probabilities, for the identified objects in the input image.

YOLOv5 offers several advantages over traditional methods, including continuous and consistent operation, thereby eliminating the impact of human factors such as fatigue and distractions. The system is capable of monitoring large areas in real-time, enabling swift response to fire incidents. The algorithm is trained on diverse datasets, enhancing its ability to accurately identify fire patterns and differentiate them from other objects or false alarms.

Furthermore, YOLOv5 demonstrates robust performance in detecting fires even in challenging environments with limited visibility. By analyzing the bounding box of the detected fire instances, the flame height can be calculated when the camera captures an orthogonal view of the flame. This information provides valuable insights into the fire's behavior, intensity, and potential hazards. Overall, integrating YOLOv5 into fire vigilance efforts can significantly improve the overall safety and efficiency of fire detection, facilitating early detection and timely interventions.

Fig. 2 Framework of the FV Pocket



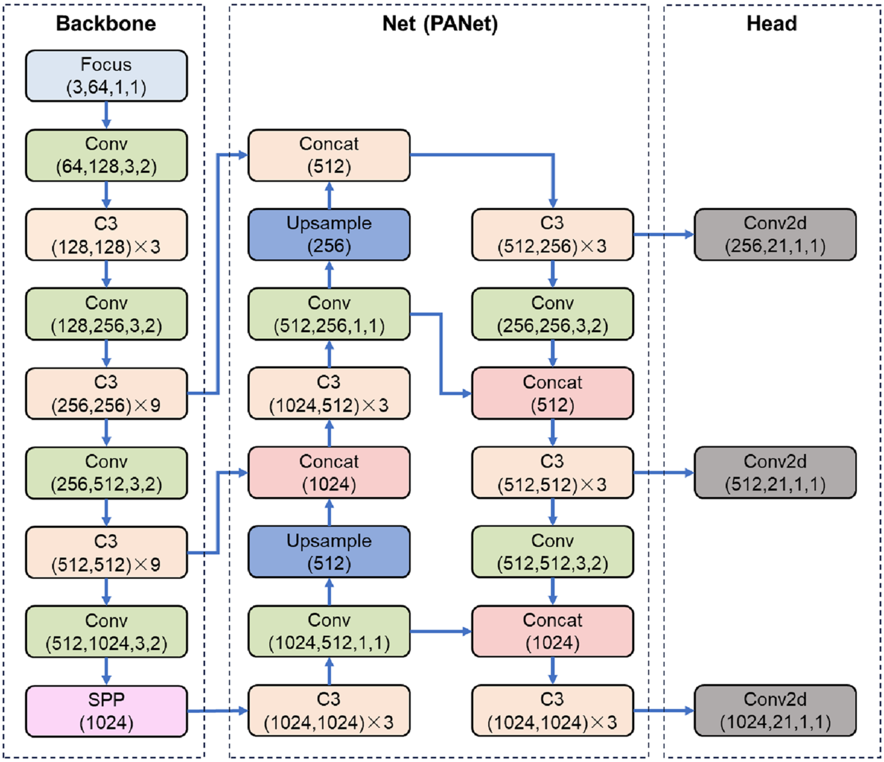


Fig. 3 The architecture of YOLOv5 for fire detection

2.2 Fire Segmentation

When a fire occurs, in addition to promptly raising an alarm, it is crucial to keep track of the development of the fire in real time. This necessitates the analysis of critical fire parameters.

In the context of fire analysis and understanding, accurate fire segmentation plays a critical role in tracking the development of fires and analyzing their parameters. By effectively delineating the boundaries of the fire region in images or videos, fire segmentation enables precise measurement and quantification of various fire-related parameters, such as size, shape, and spatial distribution. These parameters are essential for assessing the intensity, spread, and potential hazards associated with fire.

Swin-Unet (Fig. 4) is adopted for real-time fire segmentation. It is a variant of the U-Net architecture that incorporates the Swin Transformer as its backbone. All images are divided into 224×224 blocks as input to Swin-Unet. The input is passed through the initial layers of the Swin Transformer, which extract high-level spatial representations. These representations are then passed through a series of encoder and decoder blocks, similar to the U-Net architecture, to capture multi-scale features. Skip connections are used to concatenate features from the down-sampling path with the up-sampling path, enabling the integration of both local and global information. The final output is a segmentation map that assigns class labels to each pixel in the input image.

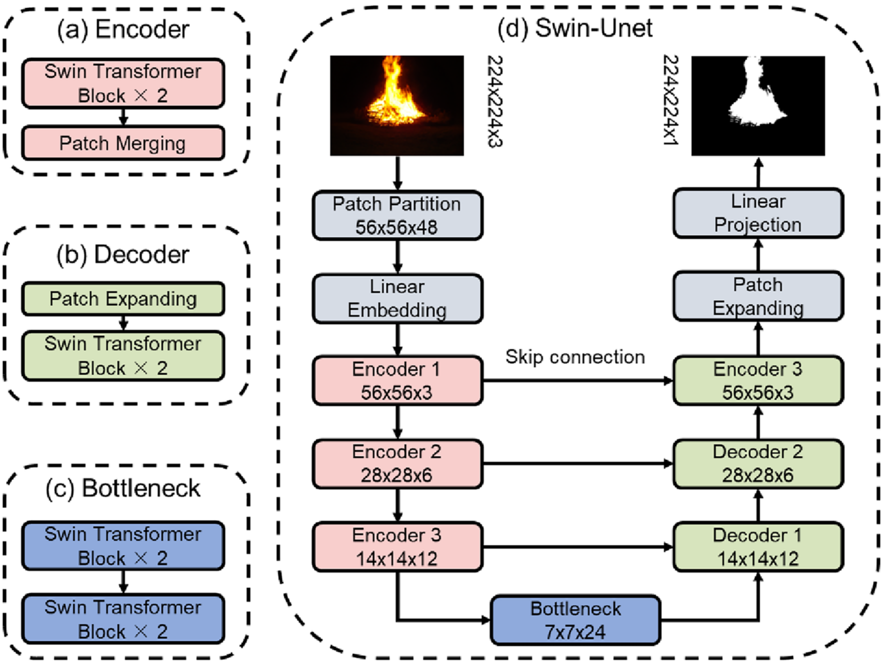


Fig. 4 The architecture of Swin-Unet for fire segmentation

Compared to traditional methods, Swin-Unet eliminates the need for manual threshold setting by training on diverse fire datasets, enabling flexible and adaptive fire segmentation. Additionally, its fast segmentation capability provides real-time fire information, greatly supporting decision-making in fire-related scenarios.

In our FV Pocket, two modes are employed for fire segmentation, namely the global mode and the fast mode (Fig. 5). In the global mode, the entire raw image is utilized as

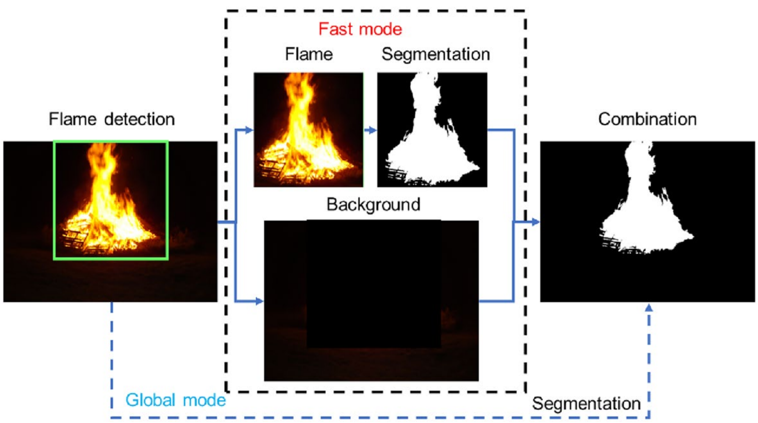


Fig. 5 Two modes of fire segmentation, fast mode and global mode

input. The image is divided into small blocks of size 224 × 224, which are then fed into the Swin-Unet model for segmentation. This mode ensures accurate identification of all regions containing flames; however, its drawback lies in the time-consuming process of segmenting each small block individually. In addition, the global mode may misidentify flame reflection as actual flame area, resulting in overestimation of the flame area. To address these issues, the fast mode is introduced to enhance the segmentation speed and mitigate the impact of flame reflections. In this mode, only the flame area detected by the YOLOv5 model is considered as input to Swin-Unet, while the remaining regions are assumed to be flame-free. This approach significantly accelerates the segmentation process while still maintaining satisfactory accuracy. Moreover, by combining fire detection with flame segmentation, the overestimation of the flame area (e.g., flame reflection or objects with flame-like color) by the model can be effectively reduced.

2.3 Fire Measurement

Following fire detection and segmentation, the pixel values corresponding to flame height and area can be calculated. However, relying solely on raw pixel measurements can be challenging for users to accurately assess fire parameters in real-world contexts. Therefore, it is necessary to convert these pixel values into specific, real-world parameters.

Real-world parameters refer to physical measurements, such as meters for height and square meters for area, which are essential for practical fire analysis and safety assessments. To determine the real fire parameters, a reference scale (R_s) is required, which defines the actual length that each pixel represents in the images. Using this reference scale, the flame height can be determined by multiplying the pixel count of the flame height by the reference scale ($H = H_p \times R_s$, where H is actual flame height and H_p is flame height in pixels), and the flame area can be calculated by summing the flame pixels and then multiplying by the square of the reference scale ($A = A_p \times {R_s}^2$, where A is actual flame area and A_p is flame area in pixels).

There are several methods available to determine the reference scale: pre-calibration, human estimation, and the use of a binocular camera. Table 1 provides a comparison of these methods. The pre-calibration method provides high estimation precision but requires site-specific calibration, and once calibrated, the camera must remain stationary. This method typically involves placing objects of known dimensions (L) in the camera's field of view and establishing the pixel-to-real-world relationship ($R_s = L/L_p$, where L_p is the length in pixels). The human estimation method can be applied in any environment with reference objects, but the accuracy depends on the user's experience. This approach involves using familiar objects in the scene, such as doors, windows, or standardized equipment, as reference points. Additionally, if the camera moves, each frame will need to be manually assessed, leading to a high workload. The binocular camera method combines the advantages of the previous two methods but requires a specialized binocular camera, which also needs to be calibrated prior to use. This method uses stereoscopic vision principles to automatically calculate distances ($D = B \times f/d$, where D is the distance to object, B is the baseline between cameras, f is focal length, and d is disparity) similar to human depth perception.

For general-purpose applications, our system allows the direct input of the reference scale, ensuring that it can be applied to various video types and is accessible to all users regardless of the environment or camera setup. This flexibility allows users to input reference scales obtained through any of the aforementioned methods or other

Table 1 Comparison of reference scale determination methods

Method	Pre-calibration	Human estimation	Binocular camera [39]
Procedure	1. Fix the camera 2. Calculate the value of the reference scale at different locations 3. Make a reference table 4. Determine the reference scale according to the location	1. Find reference objects near the fire source 2. Estimate the size of object and the pixel size in the images 3. Calculate the reference scale	1. Binocular camera calibration 2. Distance measurement using binocular camera 3. Calculate references scale using the distance and the focal length of camera
Advantages	High precision	Easy to use	Easy to use and High precision
Disadvantages	High workload	Low precision and high workload	Special camera is needed

reliable sources, making the system adaptable to different scenarios while maintaining measurement accuracy.

2.4 Fire Calorimetry by Computer Vision

Heat release rate (HRR) is a fundamental parameter that determines the fire's energy output and potential hazard. By quantifying the HRR, fire calorimetry enables the evaluation of fire size and growth, which is important in fire vigilance. However, traditional lab-based methods such as mass balance and oxygen calorimeter are not applicable in real fire scenarios. Therefore, image-based fire calorimetry has emerged as a necessary alternative.

In this process, a modified Visual Geometry Group (VGG) network (Fig. 6), trained on the NIST Fire Calorimetry Database (FCD) [43], is employed for fire size estimation. The VGG network is a deep convolutional neural network commonly used for image classification tasks. Modifications are made to the network architecture to meet the specific requirements of fire calorimetry. These include adjusting the last fully connected layer to a single output layer with linear activation, and reducing the number of hidden nodes in the initial layers to reduce training parameters.

The raw images captured from the videos are resized based on the reference scale to ensure that the input video maintains a consistent scale with the images used for model training. Consequently, the identified HRR can be accurately determined and output using the trained AI model.

2.5 Other Functions

In addition to its primary capabilities of fire detection, segmentation, measurement, and calorimetry, the web application offers several advanced features to enhance user experience and data analysis. The webpage allows users to create images or videos from the processed detection and segmentation results, providing a clear and intuitive representation of the fire analysis that makes it easier for users to understand and interpret the data. Users can also generate graphs depicting the variation of flame measurements and calorimetric data over time. These graphs are crucial for tracking the progression and intensity of the fire, providing insights into the fire's behavior and development.

All processed data, including flame measurements and calorimetric results, are saved in CSV format and can be downloaded for further analysis. For instance, quantified heat release rate data can be imported into Computational Fluid Dynamics (CFD) software for detailed fire reconstruction and modeling. By providing downloadable data, the webpage supports users in conducting in-depth analysis and research. This feature is particularly

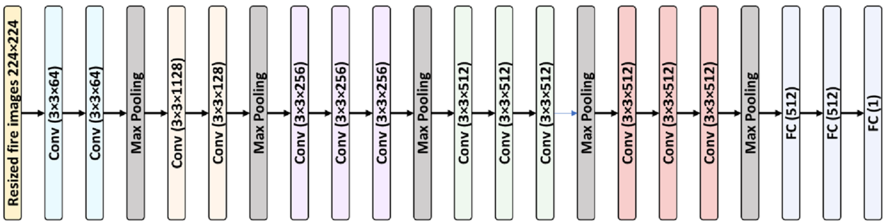


Fig. 6 The architecture of VGG for image-based fire calorimetry

beneficial for fire safety engineers and researchers who need precise data for simulations and predictive modeling.

3 APP Design and Discussion

In order to integrate all the aforementioned functionalities into a comprehensive app, a web application called FV Pocket was developed. This web app is designed to be platform-independent, allowing users to access and utilize it on various devices and operating systems as long as they have a web browser. This eliminates the requirement for users to install and maintain specific software on their individual devices, thereby enhancing convenience and accessibility. Additionally, the web app can be updated and maintained centrally. By hosting the application on a server, updates and bug fixes can be implemented at the server level, ensuring that all users automatically benefit from the latest version without needing to perform individual software updates. Furthermore, this centralized maintenance approach guarantees consistency across all users and devices, enhancing the overall user experience.

3.1 APP Design

The development of the web application involves the utilization of Flask as the web framework and Nginx as the web server. Flask was chosen due to its lightweight and flexible architecture, offering a solid foundation for Python-based web application development. It facilitates the definition of routes, handling of requests and responses, and management of application logic. On the other hand, Nginx serves as an efficient web server, responsible for managing incoming requests and serving static files.

The integration between Flask and Nginx is established by running the Flask application with Gunicorn, which acts as the interface between Flask and Nginx. Gunicorn enables the seamless connection between the Flask application and Nginx, allowing for efficient request forwarding, load balancing, and caching. The combination of Flask and Nginx provides a robust and efficient solution for developing and deploying the FV pocket.

3.2 APP Framework

The developed application is presented as a web-based platform, offering two distinct modes: demo mode and app mode, as illustrated in Fig. 7. In demo mode, users are provided with two example cases to explore the app's capabilities: a box fire from the NIST fire calorimetry database and a bin fire sourced from the internet. After selecting a fire case, users must select the specific functions they wish to perform, including fire detection, segmentation, measurement, and calorimetry. The results of fire measurement are contingent upon fire detection and segmentation. If only fire detection or fire segmentation is selected, the fire measurement will be limited to either flame height or flame area. If neither is selected, no measurement results will be available. Conversely, if both are selected, comprehensive measurement results will be provided.

The output includes both video and text-based results. For video output, the processed videos showing fire detection, fire segmentation, and measurement results are displayed directly on the webpage. Additionally, the system generates temporal plots displaying the variations of flame height, flame area, and heat release rate over time. The text-based output

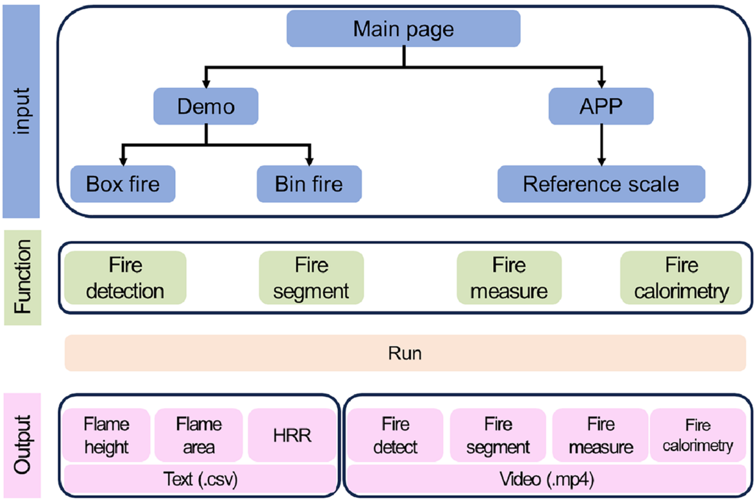


Fig. 7 framework of the developed APP

is saved in CSV format, containing frame numbers along with corresponding flame height, flame area, and heat release rate values for each frame. All generated results, including the processed videos, plot images, and CSV files, can be downloaded through dedicated download buttons on the webpage for further analysis.

In app mode, the process is similar to demo mode, with the key difference being that users can upload their own videos. After entering the reference scale parameter and selecting the desired functions, the webpage processes the uploaded file and generates the results. Users can then download their processed files in the same format as in demo mode, including both the video visualizations and CSV data files containing frame-by-frame measurements of fire parameters.

3.3 Demonstration with a Lab Fire Test

In order to validate the efficacy of the web application, a fire test case sourced from the NIST database was chosen for demonstration purposes [43]. The fire test involved igniting a cardboard box measuring 0.3 m × 0.4 m × 0.3 m, containing crinkled paper as fuel. Four acetone wicks, each soaked in 10 mL of acetone, were strategically placed at the base of each face of the box. The results obtained from the Fire Vigilance (FV) pocket application at various time intervals are depicted in Fig. 8. A more detailed running process of the FV pocket for the box fire test is demonstrated in Video S1.

All the results are rescaled according to the reference scale (the size of the box). The FV pocket yields four distinct outcomes, namely fire detection, fire segmentation, fire measurement, and fire calorimetry. The fire detection results furnish the bounding box encompassing the flame, alongside its precise spatial coordinates within the images. The fire segmentation outcomes quantify the area occupied by the flame within the images. Additionally, the fire measurement results offer pertinent data concerning the flame's height and area, while the fire calorimetry outcomes provide comprehensive information pertaining to the fire's size and thermal characteristics.

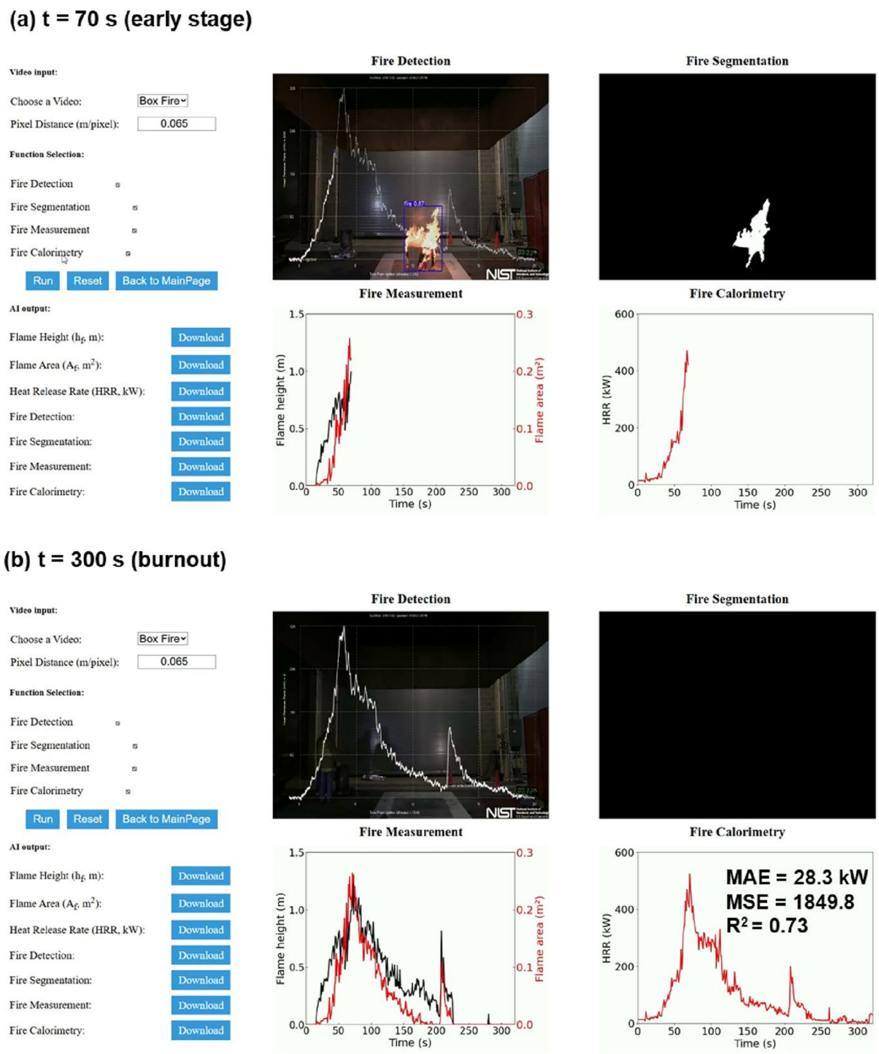


Fig. 8 The results of FV pocket for the box fire. (see Video S1)

Table 2 The overall performance for the paper box fire (Intel Xeon Gold 6126 and Tesla P100)

Model performance	Mode	FPS	Latency (s)
Fire detection	–	103	0.01
Fire segmentation	Global	0.65	1.54
	Fast	18	0.06

The overall performance of the fire test is presented in Table 2 and shown near the burnout in Fig. 8b. Table 2 outlines the key metrics, including frames per second (FPS) and latency, for both fire detection and fire segmentation tasks under different modes of operation. It is important to note that the latency reported here is primarily caused by processing speed, with communication speed excluded, as it can vary depending on the network connection. For fire detection, the model achieves an impressive performance with a high FPS of 103 and minimal latency of 0.01 s, indicating its capability for real-time detection. In contrast, the fire segmentation task demonstrates a trade-off between accuracy and speed, with two distinct operational modes. The "Global" mode, optimized for detailed segmentation, operates at a significantly lower FPS of 0.65 and a higher latency of 1.54 s. This mode is likely suited for scenarios demanding higher precision. Meanwhile, the "Fast" mode offers a balanced approach, achieving an FPS of 18 with a latency of 0.06 s, making it suitable for scenarios where quicker response times are critical. In this work, the “Fast” mode is adopted in the application for real-time fire vigilance.

The FV pocket demonstrates its ability to accurately identify the progression of fire based on fire images. The models effectively locate the fire source on the map and utilize the detected flame area to generate a binary map exclusively containing the flame region. Through fire detection and segmentation, the changes in flame height and area during fire development are easily obtained. Notably, the maximum flame height identified by the fire detection model is 1.29 m, which closely aligns with the maximum flame height of 1.37 m reported by NIST [44], thereby validating the model’s efficacy.

Regarding fire heat release rate (HRR), the identified HRR is compared with the actual HRR measured using an oxygen calorimeter. Overall, the image-based estimation of fire HRR closely corresponds to the measured values, exhibiting a coefficient of determination (R^2) of 0.73. These findings highlight the model’s effectiveness in fire calorimetry, further demonstrating its capability to accurately assess fire power and hazards.

3.4 Demonstration with Real Fire Events

Following the validation of the model through NIST fire tests, real-world fire incidents are subsequently employed to illustrate the practical viability of the application. It is noteworthy that in comparison to the controlled fire experiments within the NIST database, real fire events encompass heightened complexity attributable to diverse fire sources, backgrounds, and lighting conditions. Additionally, disparities in the view angles of video capture, camera configurations, and shooting distances further contribute to the increased difficulty associated with the identification and quantification of parameters when compared to laboratory fire tests.

To substantiate the robustness of the web application’s functionality, a rubbish bin fire has been chosen to serve as a representative exemplar of actual fire scenarios, as depicted in Fig. 9. The processing speeds of different models are shown in Table 3 and the detailed process is shown in Video S2. In the bin fire scenario, the fire detection model achieved a stable FPS of 96 with a latency of 0.01 s. Similarly, the global mode for fire segmentation

Fig. 9 The results of FV pocket for the rubbish bin fire. (see Video S2)

Table 3 The overall performance for the bin fire (Intel Xeon Gold 6126 and Tesla P100)

Model performance	Mode	FPS	Latency (s)
Fire detection	–	96	0.01
Fire segmentation	Global	0.64	1.55
	Fast	13	0.08

maintained a consistent FPS of 0.64 and latency of 1.55 s, comparable to the performance observed in the paper box fire test. Notably, in the fast mode, the FPS increased to 13 with a latency of 0.08 s. This improvement in FPS reflects the efficiency of the fast mode, where only the detected fire regions are segmented. While larger fire regions require more processing time, the system's performance remains well within the requirements for real-time applications.

In order to harmonize with the scale requirements of the deep learning model, all fire scene images have been meticulously rescaled and resized in accordance with a predefined reference scale. The comprehensive scrutiny of genuine fire progression, as presented in Fig. 9b, reveals intricate insights into the identification and quantification of actual fire dynamics. Visual assessments affirm that the web application proficiently aligns with the qualitative patterns observed in the video, thereby effectively discerning and quantifying the fire source in real-world fire events.

In summary, the developed web application can effectively identify and quantify the transient fire development within the proposed fire detection, segmentation, measurement and calorimetry model regardless of the background, fire source, or camera settings.

3.5 Future Improvements

In addition to its primary capabilities of fire detection, segmentation, measurement, and calorimetry, the webpage could offer several advanced features to further enhance user experience and data analysis. Future versions of the webpage might allow users to upload their data to the system, designed to assist in the ongoing optimization and iteration of the fire detection models. By contributing their data, users could help improve the accuracy and reliability of the system, ensuring that it evolves, trains and adapts to new fire scenarios and datasets. The uploaded data would be used by developers to refine and enhance the model, addressing any shortcomings and incorporating new findings. This collaborative approach would ensure that the system remains cutting-edge and effective in various fire detection and analysis contexts.

Another potential enhancement could be an error reporting feature that allows users to manually upload results if they encounter any inaccuracies in flame data processing. This feedback mechanism would be essential for developers to promptly identify and rectify issues, ensuring that the model is continuously improved and updated based on real-world user experiences. By providing a direct channel for error reporting, the webpage would ensure that developers receive timely and accurate feedback, facilitating quick adjustments and improvements to the model, thus enhancing its overall performance and reliability.

These future improvements would not only enhance the user experience but also contribute to the ongoing development and refinement of fire identification and quantification technologies. By offering robust data export capabilities, and mechanisms for user feedback and model optimization, the webpage could stand as a valuable tool for both practical fire management and advanced research.

3.6 Discussion

The FV Pocket system, as an automatic fire vigilance tool, presents certain limitations that necessitate further refinement and integration with advanced technologies to enhance its effectiveness and broaden its application in real-world firefighting scenarios. A critical challenge lies in the accurate measurement of fire parameters, such as flame height, flame area, and heat release rate (HRR), which currently depends on the presence of a reference scale. Without such a scale, these measurements become unreliable, underscoring the need for an automatic distance measurement method. The integration of technologies such as LiDAR or stereo vision could facilitate the automatic acquisition of reference sizes, thereby improving the accuracy of these essential fire parameters.

Another significant limitation is the dependency of flame height and area acquisition on the shooting angle, which requires the camera to be orthogonally aligned with the fire source. This requirement can be impractical in dynamic fire scenarios, suggesting that future developments should focus on either adjusting the camera angle automatically or developing robust angle correction algorithms. Utilizing gimbal-stabilized cameras or AI-based angle correction techniques could ensure accurate measurements even when the camera is not ideally positioned.

Furthermore, the AI model has been primarily trained using data from open fires in open environments. This limiation leads to a reduced accuracy when analyzing indoor fires, where factors such as confined spaces, reflections, and different lighting conditions come into play. Additionally, the model's current training database is limited in terms of fire size, primarily covering fires with heat release rates below 4 MW. This upper limit restricts its ability to accurately analyze larger fire scenarios. To address these limitations, the training dataset should be expanded to include diverse indoor fire scenarios and a broader range of fire sizes, taking into account variations in room size, wall materials, lighting conditions, and fire dynamics.

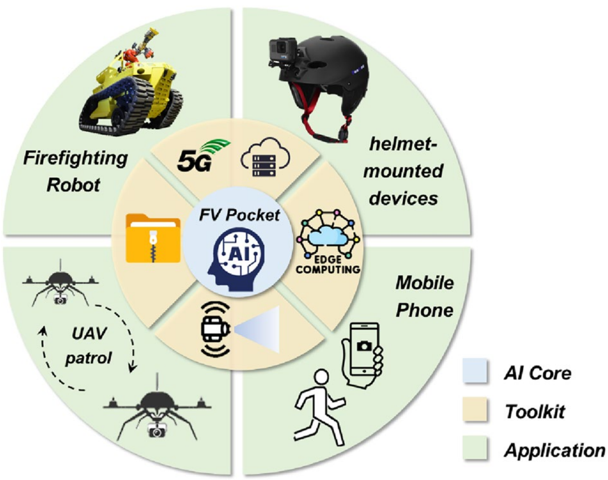


Fig. 10 Smart firefighting ecosystem based on fire vigilance software

The current implementation also assumes that the flame is fully captured within the camera's field of view. However, in real-world scenarios, smoke and obstacles often partially obstruct the flame, leading to incomplete or inaccurate measurements. Future improvements should focus on developing robust algorithms that can estimate fire parameters even when the flame is partially obscured. This could involve incorporating smoke detection and modeling capabilities, as well as methods to compensate for various types of visual obstacles.

When the FV system is deployed for real-time fire scenario assessment, the management of communication and computational resources becomes a critical concern. The system can leverage the advanced communication infrastructure provided by 5G technology to transmit the collected video streams to cloud servers for processing, while the client-side is responsible for the simple rendering and display of the results. Moreover, to further enhance system efficiency and responsiveness, model compression techniques such as knowledge distillation can be employed. By compressing a large, complex model into a more compact and efficient version, the computational burden is significantly reduced, enabling real-time processing at the local level. Additionally, the implementation of edge computing allows data to be processed at the source, proximate to the fire incident, thereby minimizing latency and reducing dependence on continuous cloud connectivity.

Beyond addressing these technical challenges, the FV Pocket system holds considerable potential for integration with emerging technologies and application in various firefighting contexts as shown in Fig. 10. For instance, equipping Unmanned Aerial Vehicles (UAVs) with the FV Pocket system could enable real-time aerial surveillance of large fire scenes, providing critical data from multiple angles. Edge processing capabilities on UAVs would be essential for timely analysis and decision-making, independent of potentially unstable data transmission networks. Cloud processing could complement this by offering detailed post-mission analysis when UAVs return to base stations or connect to reliable networks.

Furthermore, the FV Pocket system is designed to address scenarios with limited computational resources, such as those encountered on mobile devices. All computationally intensive processes, including fire detection, segmentation, measurement, and calorimetry, are performed on cloud servers. Mobile devices serve as display interfaces, showing results and enabling user interactions without the need for significant local processing power. This cloud-based architecture ensures that the system remains lightweight and accessible on a wide range of devices, including smartphones and tablets, regardless of their hardware capabilities.

Moreover, the FV Pocket will also offer a mobile application for use by bystanders, enabling them to report fire risk recognition to the fire service department. This functionality facilitates the provision of supplementary information about the fire following the initial alarm. Given the fact that fires are frequently detected and reported by passersby, this feature constitutes a critical element within the FV smart firefighting ecosystem. Together with the automatic fuel-load quantification by computer vision [45], the real-time fire hazards can be further quantified to support fire emergency response.

Firefighters could also benefit from handheld or helmet-mounted devices running the FV Pocket system, providing immediate feedback on fire dynamics, such as flame height and HRR. Integration with augmented reality (AR) glasses could further enhance situational awareness by overlaying critical fire parameters and safe exit routes in real-time. Furthermore, deploying the FV Pocket system in firefighting robots could enable autonomous navigation and analysis in hazardous environments. As claimed, these robots could leverage edge processing for real-time decisions, such as identifying intense fire areas or locating victims, while cloud processing would support more comprehensive data analysis when network connectivity is available. The incorporation of additional sensors, such as thermal cameras or gas detectors, could further enhance the system's capabilities by providing a multidimensional analysis of the fire scene.

In summary, addressing the current limitations of the FV Pocket system and integrating it with advanced technologies such as 5G connectivity and IoT sensors could significantly enhance its utility in various firefighting applications. These enhancements would improve the accuracy and reliability of fire detection and measurement while expanding the system's applicability to UAVs, firefighters, and firefighting robots, making it a versatile tool for both immediate fire response and in-depth post-incident analysis.

4 Conclusions

This study presents the development and demonstration of an automatic fire vigilance system named FV pocket. The system integrates key functionalities such as fire detection, fire segmentation, fire parameter estimation, and fire calorimetry (power and heat release rate) to enable effective fire vigilance. Experimental evaluations using real fire scenarios demonstrates that the system's outputs are comparable to real measurements.

In future work, we aim to address limitations related to distance measurement and angle correction to further enhance the performance of the FV pocket. We plan to integrate more advanced technologies, such as edge processing, 5G connectivity, and IoT sensors, into the system, making it feasible for real-time fire detection and quantification. With its inherent portability, the FV pocket holds immense potential for utilization in conjunction with portable cameras and unmanned aerial vehicles (UAVs), thereby facilitating the real-time acquisition of vital fire-related information. The proposed approach offers an automated and adaptable solution for the precise measurement of various fire parameters, thus presenting extensive applicability in the domains of firefighting operations and decision-making processes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10694-025-01738-6.

Acknowledgements This work is funded by the Hong Kong Research Grants Council Theme-based Research Scheme (T22-505/19-N). ZW thanks the support from the SFPE Foundation Student Research Grant. TZ thanks the support from the Hong Kong PhD Fellowship Scheme.

Author Contributions Zilong Wang: investigation, methodology, writing—original draft, formal analysis. Tianhang Zhang: writing—review & editing, resources. Xinyan Huang: conceptualization, methodology, supervision, writing—review & editing.

Funding Open access funding provided by The Hong Kong Polytechnic University.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Qureshi WS, Ekpanyapong M, Dailey MN, Rinsurongkawong S, Malenichev A, Krasotkina O (2016) QuickBlaze: early fire detection using a combined video processing approach. Fire Technol 52:1293–1317
2. Saeed F, Paul A, Karthigaikumar P, Nayyar A. Convolutional neural network based early fire detection 2019.
3. Huang X, Tam WC (2024) Intelligent building fire safety and smart firefighting. Springer, New York
4. Sulman N, Sanocki T, Goldgof D, Kasturi R. How effective is human video surveillance performance? 2008 19th International Conference on Pattern Recognition, IEEE; 2008, p. 1–3.
5. Sen K, Sarkar J, Saha S, Roy A, Dey D, Baitalik S, Nandi CS (2015) Automated fire detection and controlling system. Int Adv Res J Sci Eng Technol 2(5):34–37
6. Di Lascio R, Greco A, Saggese A, Vento M (2014) Improving fire detection reliability by a combination of videoanalytics. Lec Note Comp Sci (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8814:477–484
7. Seo J, Han S, Lee S, Kim H (2015) Computer vision techniques for construction safety and health monitoring. Adv Eng Inform 29:239–251
8. Zhang T, Wang Z, Zeng Y, Wu X, Huang X, Xiao F (2022) Building artificial-intelligence digital fire (AID-Fire) system: a real-scale demonstration. J Build Eng 62:105363
9. Zhang X, Jiang Y, Wu X, Nan Z, Jiang Y, Shi J et al (2024) AIoT-enabled digital twin system for smart tunnel fire safety management. Dev Built Environ 18:100381
10. Yu M, Yuan H, Li K, Wang J (2023) Research on multi-detector real-time fire alarm technology based on signal similarity. Fire Saf J 136:103724
11. Xu Z, Guo Y, Saleh JH (2021) Advances toward the next generation fire detection: deep LSTM variational autoencoder for improved sensitivity and reliability. IEEE Access 9:30636–30653
12. Yuan C, Liu Z, Zhang Y (2017) Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles. J Intell Robot Syst Theor Appl 88:635–654
13. Dampage U, Bandaranayake L, Wanasinghe R, Kottahachchi K, Jayasanka B (2022) Forest fire detection system using wireless sensor networks and machine learning. Sci Rep. https://doi.org/10.1038/s41598-021-03882-9
14. Vikram R, Sinha D, De D, Das AK (2020) EEFFL: energy efficient data forwarding for forest fire detection using localization technique in wireless sensor network. Wireless Netw 26:5177–5205
15. Sharma A, Singh PK, Kumar Y (2020) An integrated fire detection system using IoT and image processing technique for smart cities. Sustain Cities Soc 61:102332
16. Zhang X, Chen X, Ding Y, Zhang Y, Wang Z, Shi J et al (2024) Smart real-time evaluation of tunnel fire risk and evacuation safety via computer vision. Saf Sci. https://doi.org/10.1016/j.ssci.2024.106563
17. Ding Y, Zhang Y, Huang X (2023) Intelligent emergency digital twin system for monitoring building fire evacuation. J Build Eng 77:107416
18. Hosseini O, Maghrebi M (2021) Risk of fire emergency evacuation in complex construction sites: integration of 4D-BIM, social force modeling, and fire quantitative risk assessment. Adv Eng Inform 50:101378
19. Li P, Zhao W (2020) Image fire detection algorithms based on convolutional neural networks. Case Stud Therm Eng. https://doi.org/10.1016/j.csite.2020.100625
20. Majid S, Alenezi F, Masood S, Ahmad M, Selda Gündüz E, Polat K (2022) Attention based CNN model for fire detection and localization in real-world images. Expert Syst Appl 189:957–4174
21. Khan ZA, Hussain T, Ullah FUM, Gupta SK, Lee MY, Baik SW (2022) Randomly initialized CNN with densely connected stacked autoencoder for efficient fire detection. Eng Appl Artif Intell 116:105403
22. Barmpoutis P, Dimitropoulos K, Kaza K, Grammalidis N. Fire detection from images using faster R-CNN and multidimensional texture analysis. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* 2019;2019-May:8301–5.

23. Guo J, Hou Z, Xie X, Yao S, Wang Q, Faster JX, Based indoor flame detection R-CNN, for firefighting robot. (2019) IEEE Symposium Series on Computational Intelligence. SSCI 2019(2019):1390–1395
24. Chandana VS, Vasavi S (2022) Autonomous drones based forest surveillance using faster R-CNN. Proc Int Conf Electron Renew Syst ICEARS 2022:1718–1723
25. Wu H, Wu D, Zhao J (2019) An intelligent fire detection approach through cameras based on computer vision methods. Process Saf Environ Prot 127:245–256
26. Chen C, Yu J, Lin Y, Lai F, Zheng G, Lin Y (2022) Fire detection based on improved PP-YOLO. SIViP 17:1061–1067
27. Wu Z, Xue R, Li H (2022) Real-time video fire detection via modified YOLOv5 network model. Fire Technol. https://doi.org/10.1007/s10694-022-01260-z
28. Ma J, Cheng JCP, Jiang F, Gan VJL, Wang M, Zhai C (2020) Real-time detection of wildfire risk caused by powerline vegetation faults using advanced machine learning techniques. Adv Eng Inf. https://doi.org/10.1016/j.aei.2020.101070
29. Shi C, Deng L, Ren F, Tang F (2023) Experimental study on the flame height evolution of two adjacent hydrocarbon pool fires under transverse air flow. Energy 262:125520
30. Xiong C, Fan H, Huang X, Fernandez-Pello C (2022) Evaluation of burning rate in microgravity based on the fuel regression, flame area, and spread rate. Combust Flame 237:111846
31. Babrauskas V, Peacock RD (1992) Heat release rate: the single most important variable in fire hazard. Fire Saf J 18:255–272
32. Sun Q, Turkan Y (2020) A BIM-based simulation framework for fire safety management and investigation of the critical factors affecting human evacuation performance. Adv Eng Inform 44:101093
33. Zhang T, Ding F, Wang Z, Xiao F, Lu CX, Huang X (2024) Forecasting backdraft with multimodal method: fusion of fire image and sensor data. Eng Appl Artif Intell 132:107939
34. Cao X, Wu K, Geng X, Guan Q (2024) Field detection of indoor fire threat situation based on LSTM-Kriging network. J Build Eng 84:108686
35. Zhang Z, Wang L, Liu S, Yin Y (2024) Intelligent fire location detection approach for extrawide immersed tunnels. Expert Syst Appl 239:122251
36. Shadrin D, Illarionova S, Gubanov F, Evteeva K, Mironenko M, Levchunets I et al (2024) Wildfire spreading prediction using multimodal data and deep neural network approach. Sci Rep 14:1–17
37. Wang Z, Zhang T, Wu X, Huang X (2022) Predicting transient building fire based on external smoke images and deep learning. J Build Eng 47:103823
38. Wang Z, Zhang T, Huang X (2023) Predicting real-time fire heat release rate by flame images and deep learning. Proc Combust Inst 39:4115–4123
39. Wang Z, Ding Y, Zhang T, Huang X (2023) Automatic real-time fire distance, size and power measurement driven by stereo camera and deep learning. Fire Saf J 140:103891
40. Wang Z, Zhang T, Huang X (2023) Explainable deep learning for image-driven fire calorimetry. Appl Intell 54:1047–1062
41. Cao H, Wang Y, Chen J, Jiang D, Zhang X, Tian Q, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 2023;13803 LNCS:205–18.
42. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* 2015:1–14.
43. NIST Fire Calorimetry Database 2020.
44. Lindeman A, Randelovic M, Stroup D, Floyd J, DiDomizio M, McGrattan K, et al. Heat release rate and fire characteristics of fuels representative of typical transient fire events in nuclear power plants. U.S. Nuclear Regulatory Commission, Washington, D.C., USA: 2020.
45. Ding D, Deng R, Zhang Y, Huang X, Elhami-Khorasani N, Gernay T (2025) Automatic Assessment of Fuel Load and Fire Risk via Digitized Database and Intelligent Computer Vision, Process Safety and Environmental Protection, 107031.