A Smoke Source Location Method Based on Deep Learning Smoke Segmentation

Yuanpan ZHENG*, Zeyuan HUANG, Hui WANG, Binbin CHEN, Chao WANG, Yu ZHANG
School of Computer Science and Technology
Zhengzhou University of Light Industry, Zhengzhou 450000, Henan, China

***Abstract*—The generation of smoke is an early warning sign of a fire, and fast, accurate detection of smoke sources is crucial for fire prevention. However, due to the strong diffusivity of smoke, its morphology is easily influenced by environmental factors, and in complex real-world scenarios, smoke sources are often obscured. Current methods lack precision, generalization ability, and robustness in complex environments. With the advancement of deep learning-based smoke segmentation technology, new approaches to smoke source localization have emerged. Smoke segmentation, driven by deep learning models, can accurately capture the morphological characteristics of smoke. This paper proposes a precise and robust smoke source localization method based on deep learning-enabled smoke segmentation. We first conducted experimental evaluations of commonly used deep learning segmentation models and selected the best-performing model as input. Based on the segmentation results, we analyzed the diffusion characteristics and transmittance of smoke, constructed a concentration model, and used it to accurately locate the smoke source. Experimental results demonstrate that, compared with existing methods, this approach maintains high localization accuracy in multi-target smoke scenarios and complex environments, with superior generalization ability and robustness.**

Keywords—Smoke segmentation; smoke source detection; deep learning; instance segmentation; mathematical modeling

I. Introduction

Disaster monitoring and prevention is a broad research area directly related to the safety of people's lives and property. The key to fire monitoring and prevention lies in early detection and timely response. In the field of computer vision, early fire detection and prevention essentially focuses on the detection and warning of fire smoke, as smoke is a significant indicator of fire occurrence. Research shows that during the early stages of a fire, the combustion intensity is weak, the temperature and radiant heat of the fire scene are low, and flames are not produced when the temperature of combustible materials has not yet reached the ignition point. Usually, smoke is generated on the surface of combustible materials at this stage [1]. Therefore, the appearance of smoke is a precursor to fire outbreaks [2].

Accurately locating the source of fire smoke is crucial in fire monitoring and prevention. Pinpointing the smoke source helps rapidly identify the location of a fire, providing key information for emergency response, reducing response time, and improving firefighting efficiency. Additionally, smoke source localization helps understand the fire’s development and spread direction, which supports the formulation of evacuation routes and protective measures to minimize casualties and property losses. Furthermore, accurately determining the smoke source can prevent false alarms caused by similar phenomena, thereby enhancing the reliability and effectiveness of fire warning systems. Thus, smoke source localization is not only a critical technology for early fire warning but also an essential part of ensuring fire response and disaster mitigation capabilities.

With the continuous development of related technologies in the field of computer vision, methods for detecting smoke using visual sensors have gradually matured. However, due to smoke strong diffusion and the multiple reflections and occlusions it experiences in everyday environments, accurately determining the smoke source remains a challenge. In recent years, significant progress has been made in deep learning-based smoke segmentation methods. Smoke segmentation technology can visually reflect the spatial distribution of smoke, making it easier to analyze its diffusion and indirectly infer the smoke source location, providing new insights for smoke source localization.

This paper proposes a precise smoke's source localization method that combines instance segmentation and mathematical modeling. First, the smoke area was extracted from the background using segmentation techniques. Then, mathematical modeling was applied to the smoke region to extract pixel coordinates and grayscale values, and the smoke diffusion characteristics and direction were derived by curve fitting. Based on this, a smoke concentration model is constructed using the dark channel algorithm to locate the area with the highest concentration, thereby estimating the exact position of the smoke source. The main contributions of this paper are as follows:

1) We constructed a fire smoke segmentation dataset with multi-scale, complex scenes and varying attributes.

2) We proposed a smoke source detection method. By analyzing the smoke segmentation results, we established a diffusion model and transmittance model based on smoke contours and static characteristics, combined these two models to construct a smoke concentration model, and inferred the smoke source location using the concentration model.

3) We analyzed several instance segmentation models, including the U-Net series [4, 15], DeepLab series [7-10], and others like FCN [3], SegNet [5], and PSPNet [6], for fire smoke segmentation tasks. After evaluating their performance and accuracy, we selected the optimal model to support smoke source detection.

The structure of this paper is as follows: Section II introduces related work on fire smoke source localization and smoke segmentation; Section III describes the algorithm proposed for locating fire smoke sources; Section IV presents a comparative analysis of several commonly used smoke segmentation models, selects the most suitable model for fire smoke segmentation, and tests the proposed smoke source localization scheme; Section V concludes the paper and discusses future research directions.

II. Related Work

A. Smoke Segmentation

Research in recent years has shown that scholars are more inclined to use deep learning methods to segment smoke, but most of the methods rely on existing semantic segmentation networks. Researchers achieve smoke segmentation by improving the semantic segmentation model with better performance, Wang et al. [13] improved the Deeplab V3 network model by embedding channel attention and deformable pyramid in the model, and feature refinement of the input image to improve the model segmentation accuracy, although the enhancement is not obvious, it provides a direction for the model improvement. Liu et al. [14] improved the Deeplab v3+ network model by improving ASPP structure through heterogeneous sensory wild fusion and incorporating a channel attention module. Salman Khan et al. [11] added an improved multi-scale separable convolutional encoder and decoder to the Deeplab v3+ model, along with a per-pixel classifier to improve the model. Taoyang Wang et al. [12] improved on Unet++ by introducing a convolutional attention module CBAM in the original coding layer making the network adaptively selectable and feature-enhanced to be able to focus more on smoke targets.

In addition to improving the model, some scholars have also implemented smoke segmentation by designing new model structures. Yuan et al. [16] designed a W-Net network model to segment smoke, which uses asymmetric encoders and decoders stacked in multiple layers, and segments the smoke region through the use of smoke density estimation. Yuming Li et al. [17] proposed a two-path smoke semantic segmentation network, although the architecture is different from conventional segmentation models, the core idea is still using pyramid pooling and channel attention. After proving that the attention mechanism can effectively improve the accuracy of the model, Yuan et al. [18] proposed a cross-scale hybrid attention network, a kind of fusion of 3D attention, multi-scale channel attention, and hybrid cross-enhancement, which is experimentally proved to have a certain degree of smoke segmentation ability.

B. Smoke Source Localization

Research on smoke source localization is relatively limited, especially regarding automated or intelligent methods for locating smoke sources. Traditionally, smoke source localization has relied on conventional visual monitoring and manual intervention techniques [19]. While these methods have been used in practical applications, their efficiency and accuracy largely depend on experience. In recent years, with the rapid development of computer vision technology, some image processing and machine learning methods have been applied in fire monitoring systems, but specialized techniques for smoke source localization remain relatively scarce.

Most existing smoke source localization methods are based on the physical properties of smoke diffusion, sensor data fusion, and multi-angle visual information. Cao et al. [20] proposed a feature foreground network that models the complex variation patterns of smoke features, enhancing smoke recognition capabilities. By using feature extraction and separation techniques, and integrating environmental conditions such as temperature and wind speed provided by sensors, they effectively separated smoke foreground from complex backgrounds, enabling smoke source prediction in dynamic environments. Zhou et al. [21] developed a vision-based localization method that analyzes the dynamic diffusion patterns of smoke captured by cameras using the U-Net algorithm to estimate the smoke source. However, the high cost of sensors and their susceptibility to environmental conditions, along with the highly random nature of smoke diffusion in the air [22], limit the widespread application of these methods in complex environments.

III. Methods

Smoke detection and segmentation can roughly determine the overall area where the smoke is located, but due to the diffusive nature of smoke, accurately pinpointing the smoke source is often difficult. Therefore, when predicting the smoke source, it is necessary to incorporate data related to smoke morphology, color, transmittance, and its diffusion characteristics into the calculations to model the smoke region. By constructing a fitted curve, we can solve for the pixel diffusion rate and the diffusion direction within the smoke region. Additionally, we apply the dark channel algorithm [23] to process the image, reducing the impact of light sources on the smoke and analyzing the grayscale values presented by the smoke. This allows us to build a smoke transmittance model. Finally, by combining the diffusion rate model and the transmittance model, we construct a smoke concentration model, identifying the point of highest smoke concentration. Using the diffusion direction, we can then predict the precise area of the smoke source. The flowchart for smoke source localization is shown in Fig. 1.

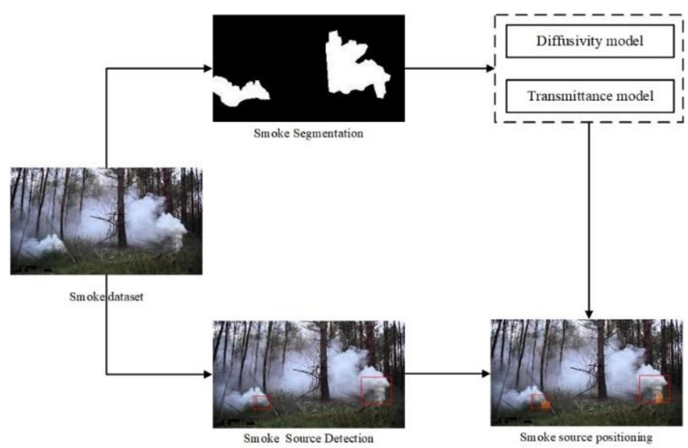


Fig. 1. Smoke source localization flowchart.

A. Smoke Information Extraction

First, the grayscale value of each pixel in the smoke region is obtained, as the grayscale value can reflect the foreground degree of the current smoke region. The foreground degree can be inferred from the relative thickness of the smoke, indirectly reflecting the transmittance of the smoke. Before converting the image to grayscale, the dark channel algorithm is applied to reduce the impact of bright fog on the smoke region, ensuring that the grayscale values more accurately reflect the smoke's foreground intensity. Then, by traversing the coordinates of all smoke pixels, a set of coordinate points for the smoke region is constructed, allowing for binarization of the smoke region.

B. Diffusivity and Transmittance Modeling

The diffusion direction of smoke generally spreads from bottom to top and from inside to outside. In terms of computer vision, the construction of the diffusion model is primarily based on the contour information of the smoke. The transmittance model is constructed based on the grayscale values of the smoke processed using the dark channel algorithm. First, a cubic polynomial regression curve is constructed based on the coordinate data from the set of coordinate points, and the coefficients are solved using the least squares method. Next, after constructing the fitted curve, the data from the set of coordinate points is grouped into segments of 10 pixels each, and the average variance between the pixel points in each group and the fitted curve is calculated. The size of the average variance is considered the value of the diffusion rate, with lighter colors having smaller diffusion rates and darker colors having larger diffusion rates. Additionally, the sets of points with the largest and smallest diffusion rates are obtained, and the direction from the area with the smallest diffusion rate to the area with the largest diffusion rate is regarded as the direction of smoke diffusion.

$$Y = AX^3 + BX^2 + CX + D \tag{1}$$

The transmittance model is based on the grayscale value of each pixel. After processing the image using the dark channel algorithm, part of the bright fog effect can be eliminated, allowing the transmittance of the smoke to be judged based on the grayscale values. Regions with larger grayscale values indicate that they are closer to the foreground, meaning the smoke is denser and the transmittance is relatively low. Conversely, regions with smaller grayscale values are closer to the background, where the smoke is less dense than in the foreground, resulting in higher transmittance. In the field of computer vision, the image model is defined as follows:

$$I(x) = J(x)t(x) + A(1 - t(x)) \tag{2}$$

where, I represents the image with haze, J represents the image without haze, t represents the transmittance, and A represents the global atmospheric light.

The formula for the dark channel algorithm is as follows:

$$J_{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in (r,g,b)} J_c(y) \right) \to 0 \tag{3}$$

where J_{dark} represents the dark channel, Ω(x) represents a square region centered at x, and Jc represents the image processed through the dark channel in the c-channel.

$$t(x) = 1 - \min_{y \in \Omega(x)} \left(\min_{c \in (r,g,b)} \frac{I_c(y)}{A_c} \right) \tag{4}$$

The diffusivity model and the transmittance model are illustrated in Fig. 2.

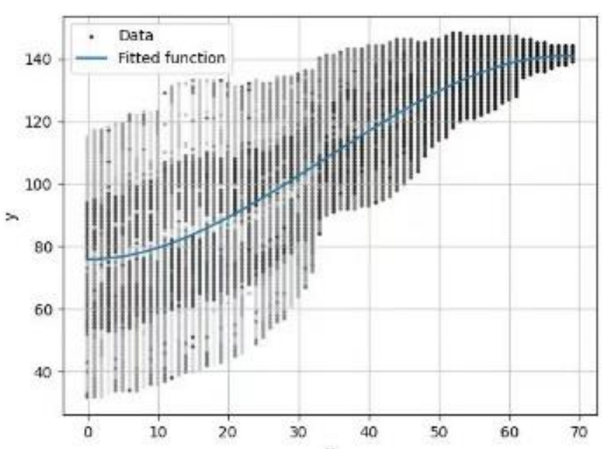
C. Smoke Concentration Modeling

The smoke concentration model is based on the diffusivity and transmittance models, where the diffusivity and transmittance values are added at different ratios to form the concentration value. After modeling the concentration, extreme concentration areas are marked, with red indicating areas of highest smoke concentration and blue indicating areas of lowest smoke concentration.

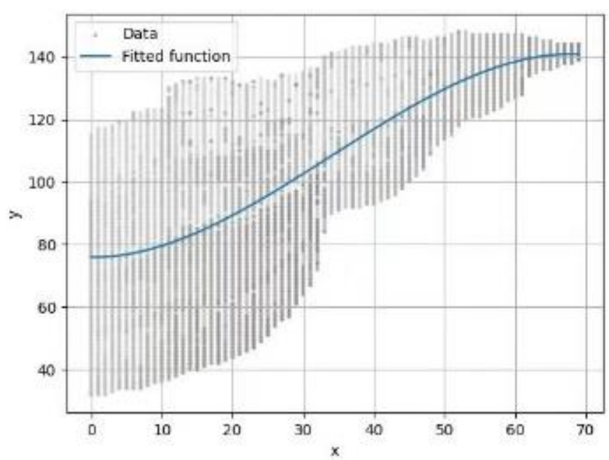
The smoke concentration can be expressed as:

$$N = yK + (1 - y)(1 - T) \tag{5}$$

where N represents the concentration value, K represents the diffusivity value, T represents the transmittance, and y represents the ratio.



Diffusivity model.



Transmittance model.

Fig. 2. Diffusivity and transmittance model.

In addition to single-target smoke, there is also multi-target smoke. The multi-target smoke concentration model is constructed by dividing the smoke based on enclosed regions, where the smoke in each enclosed region is modeled separately as a concentration. After completing the smoke model construction, a set of pixels with the highest smoke concentration is obtained. The highest concentration area is tangential to the fitted curve, and the area formed by the extension of the tangent line and the extension line in the opposite direction of diffusion is used as the predicted smoke source area. The size of the predicted smoke source region is set to the size of the outer bounding box of the area with the highest concentration.

The predicted smoke source region is illustrated in Fig. 3. To further enhance the effectiveness of smoke source localization, we also selected the YOLOv8n [24] smoke source region detection model as a supplementary method to assist in limiting the location of the smoke source. Fig. 4 illustrates the schematic diagram of the smoke source localization method.

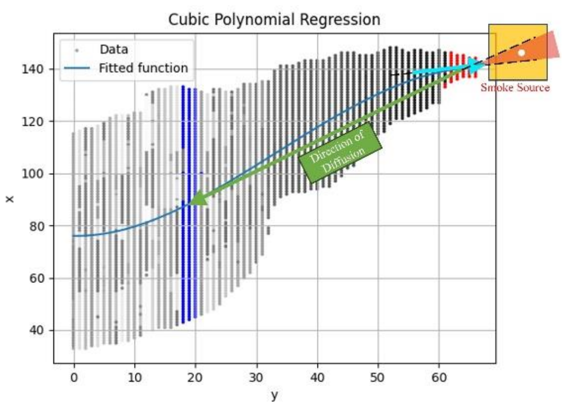


Fig. 3. Smoke source prediction area.

Fig. 4. Schematic diagram of smoke source localization method.

IV. Experimental Results

A. Segmentation Model Selection

In this section, our goal is to select a segmentation model that strikes a balance between segmentation accuracy and speed, meeting the needs of smoke source localization. Specifically, the model must maintain high accuracy while maximizing segmentation speed to ensure the timeliness of smoke source localization. Fig. 5 shows the proposed model testing flowchart.

Building on previous work, we have selected the nine most representative and high-performing segmentation models from commonly used models. These models were tested and evaluated on a self-constructed smoke segmentation dataset to ensure that the selected model possesses good adaptability and stability in practical applications.

1) Dataset and Evaluation Metrics: Next, we will introduce the dataset and evaluation metrics used to analyze the strengths and weaknesses of the segmentation models. The experimental dataset was selected from publicly available sources that could serve as valid samples. For video files, we used frame segmentation to divide the entire video data frame by frame and sampled up to 20 frames at intervals to avoid having overly similar samples. Additionally, we used real fire videos and images from scenarios such as urban areas, forests, and grasslands, selecting relevant fire smoke footage for frame-by-frame segmentation and interval sampling.

Furthermore, we simulated fire smoke generation in open areas, using video tools placed at different locations relative to the smoke source to record the smoke formation. These recordings were then segmented frame by frame, and interval samples were collected to capture fire smoke samples of various scales. Lower-resolution samples were also included to address extreme conditions. In total, the dataset collected 1,511 valid data points, including samples with multiple scales and targets, samples with complex backgrounds, and those exhibiting complex static attributes such as color, concentration, shape, texture, and expansion.

We used image processing software to manually annotate each collected sample. Since this study focuses on the detailed edges of smoke, most of the annotations were made at the pixel level. However, to minimize issues such as overfitting, generalization errors, and information loss, the edges of the labels were appropriately smoothed. We also applied symmetric processing to all samples and their corresponding labels on the left and right sides, and increased the size of the training set through data augmentation to address the challenges of sample refinement.

After these processes, our effective dataset contained 3,022 fire smoke sample images and 3,022 corresponding label annotations. To verify the validity of our custom fire smoke segmentation dataset, we randomly selected 1,000 samples and their labels to conduct comparative experiments using a basic semantic segmentation model on the training set. From the model's training results, our custom fire smoke segmentation dataset generally yielded better training outcomes compared to publicly available fire smoke segmentation datasets. Part of the dataset is shown in Fig. 6.

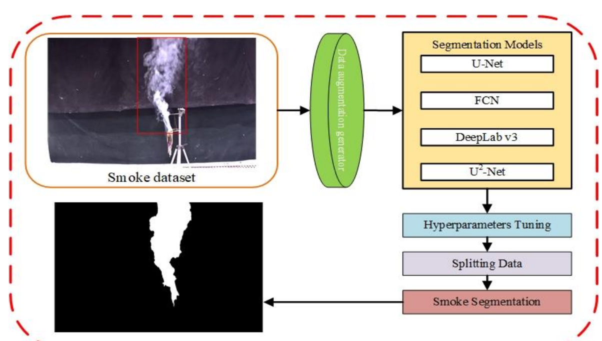


Fig. 5. Model testing flowchart.

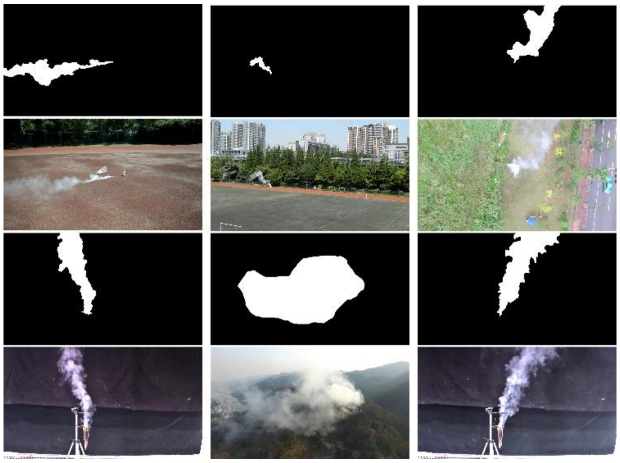


Fig. 6. Partial dataset.

2) Experimental Environment and Evaluation Metrics

The hardware environment used for the experiments in this paper is shown in Table I.

TABLE I. EXPERIMENTAL ENVIRONMENT

CPU	AMD EPYC 7773X @ 3.50GHz
GPU	GeForce RTX 4090
RAM	80G
Operating System	Ubuntu 20.04
Programming Language	Python 3.8
Deep Learning Framework	Pytorch 2.0.0
GPU Acceleration Library	Cuda 11.8

For the smoke segmentation model, we primarily focus on its segmentation accuracy and speed. Therefore, we selected mIoU, mPA, MaxF1, and FPS as the evaluation metrics for the segmentation model.

*a)*mIoU measures the overlap between the predicted segmentation and the ground truth. It is the average IoU for each class. The calculation formulas are shown in Eq. (6) and Eq. (7).

$$IoU = \frac{TP}{TP + FP + FN} \tag{6}$$

$$mIoU = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FP_i + FN_i} \tag{7}$$

where N represents the number of classes, and i represents the total number of categories.

*b)*mPA measures the pixel classification accuracy for each class, averaging the pixel accuracy for each category. The calculation formula is shown in Eq. (8).

$$mPA = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FN_i} \tag{8}$$

F1 Score is the harmonic mean of precision and recall, balancing both metrics. MaxF1 is the maximum F1 score obtained during multi-threshold testing. The formulas are shown in Eq. (9) and Eq. (10).

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{9}$$

$$MaxF1 = \max\left(F1\right) \tag{10}$$

The formulas for Precision and Recall are provided in Eq. (11) and Eq. (12).

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

FPS measures the processing speed of the model, indicating how many image frames the model can process per second. It is used to evaluate the real-time performance of the model.

3) Experimental Results and Analysis: We trained and tested several models, including U-Net, FCN, SegNet, PSPNet, DeepLab V3+, U2-Net, U2-Net+, DPESS-Net+ (64 channels), and DPESS-Net, using our custom fire smoke segmentation dataset. The final test results are shown in Table II, and the mIoU training process is illustrated in Fig. 7. The experimental data in Table II indicate that DPESS-Net performed the best overall, achieving mIoU scores of 91.09% under AVG20 and 91.18% under TOP1, an mPA of 96.96%, and a MaxF1 score of 94.37%. However, its FPS was 15.60, placing it in the medium range for processing speed.

TABLE II. THE PERFORMANCE OF THE MODEL IN COMPARATIVE EXPERIMENTS AND BOLD FONT INDICATES BEST GRADE.

Model	mIoU (%)		mPA (%)	MaxF1 (%)	FPS
	AVG[20]	*TOP[1]*			
U-Net	86.07	87.70	96.09	93.46	26.39
FCN	84.59	85.97	95.31	91.95	**27.37**
SegNet	85.47	86.45	95.70	92.48	21.22
PSPNet	84.23	85.16	95.20	91.61	17.54
DeepLab V3+	85.50	87.31	95.28	92.29	11.50
U^2-Net	86.83	87.80	96.01	93.01	13.37
U^2-Net+	85.98	86.68	95.90	92.67	12.91
DPESS-Net+	89.08	89.73	96.75	93.45	9.34
DPESS-Net (Ours)	**91.09**	**91.18**	**96.96**	**94.37**	15.60

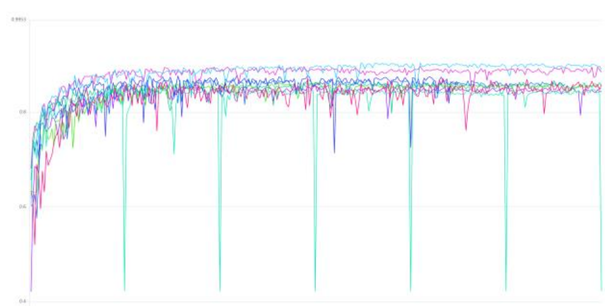


Fig. 7. The results of mIoU of the models in the comparative experiment.

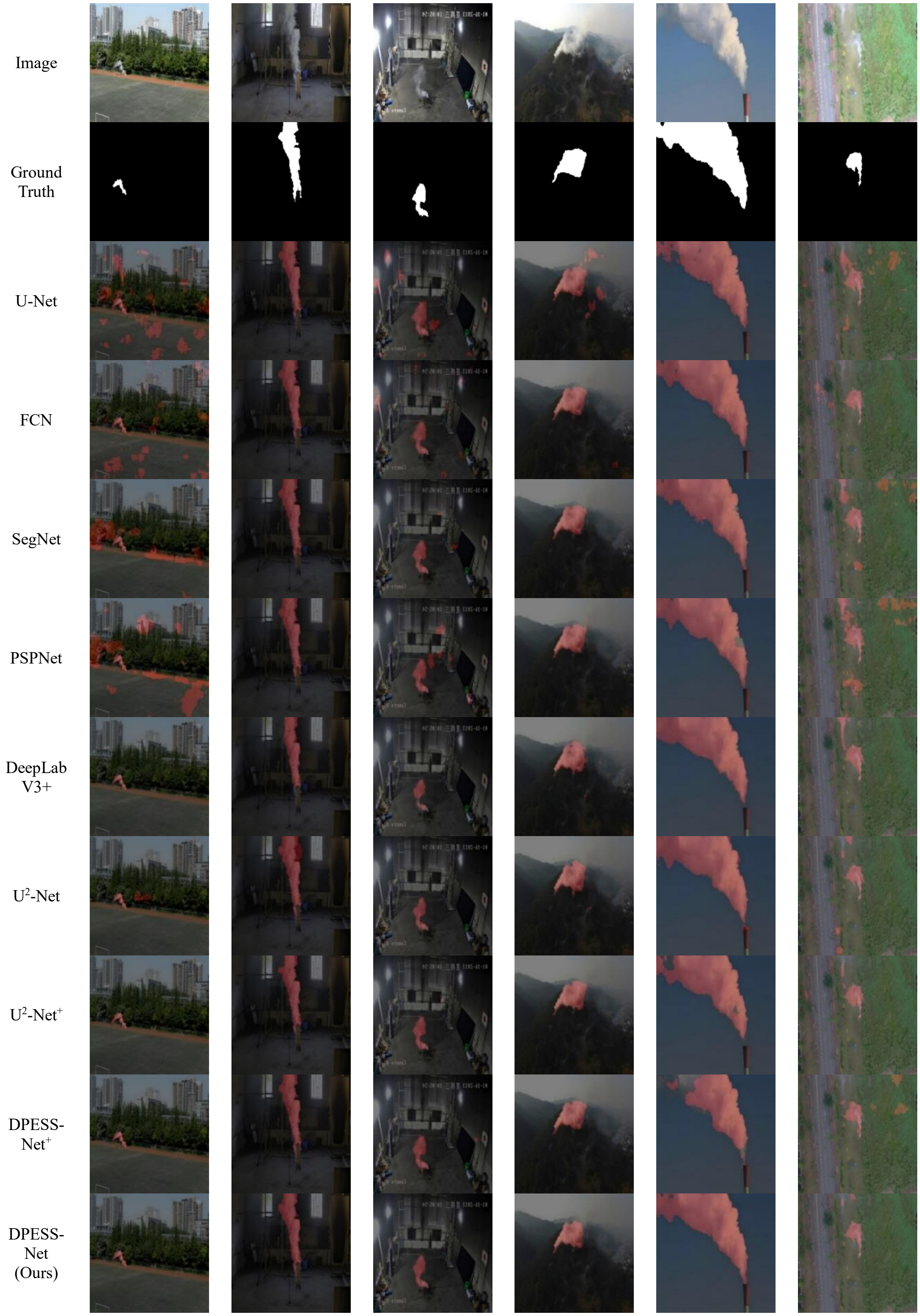


Fig. 8. Comparison of segmentation performance between our proposed method and the other eight models.

DPESS-Net's ability to achieve these results is largely attributed to its dual-channel encoder design, which enables effective information exchange between the channels. The fusion of multi-scale body features and edge-enhanced features further improves the model's ability to capture smoke targets, resulting in higher segmentation accuracy while focusing on both smoke mass and edge details.

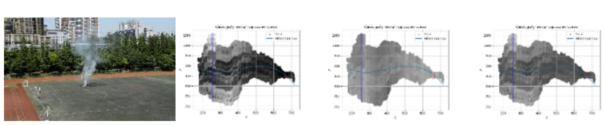
As observed in the training process shown in Fig. 7, U-Net maintains a moderate level of performance, while U2-Net slightly outperforms U-Net. FCN, PSPNet, and SegNet exhibit relatively lower segmentation accuracy but higher FPS because their architectures are relatively simple, allowing for faster segmentation speeds. DeepLab V3+, due to its more complex architecture, remains stable with relatively slower segmentation speed. DPESS-Net+, a version of DPESS-Net with the number of channels reduced to 64, decreases the parameter count and model size, improving efficiency at the cost of slightly lower accuracy.

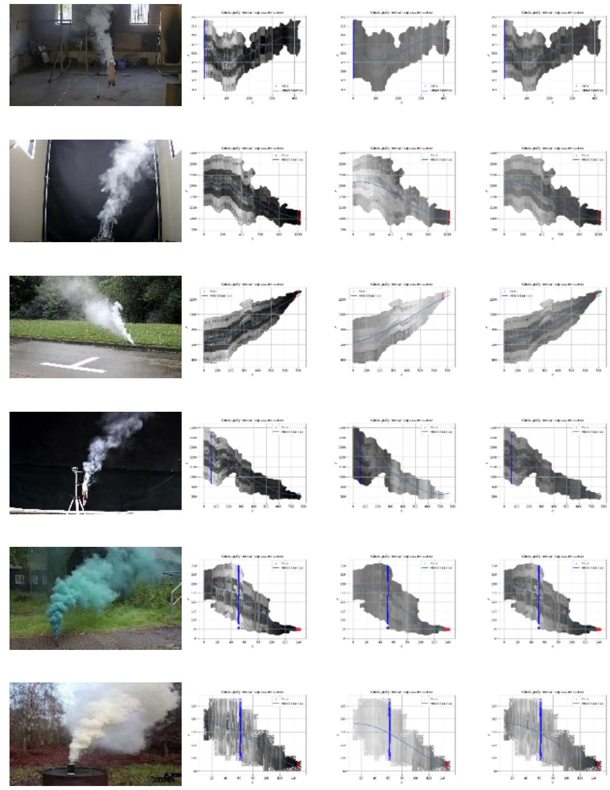
To more intuitively display the segmentation accuracy and performance of the models, we present the segmentation results of each model in Fig. 8. From the comparison images, it can be seen that in scenarios with complex backgrounds and small smoke bodies, most segmentation models experience errors and fail to segment the smoke completely. In scenarios with relatively simple backgrounds and clear main subjects, most models are able to segment the main body of the smoke; however, the details along the edges are not segmented as clearly.

In many scenes, DPESS-Net outperforms the other models, showing minimal segmentation errors. It achieves complete smoke body segmentation with clear edge details, meeting the segmentation accuracy requirements for smoke semantic segmentation as discussed in this paper. Therefore, DPESS-Net is selected as the smoke segmentation solution for the smoke source localization method.

B. Experimental Results of Smoke Source Localization

During the modeling process of the concentration model, it was found that the concentration modeling for single-target smoke is relatively stable and can accurately find the area with the highest concentration. However, in case of multi-target smoke, the results of the smoke segmentation model often show multiple smoke targets, which may be connected if the distance between the smoke targets is too small, and this has led to leakage detection in the concentration model when inferring the smoke source area. For this situation, we output the prediction results of YOLOv8n smoke region detection at the same time to maximize the accuracy of the algorithm. The results of the concentration model construction and the results of the smoke source localization algorithm are shown in Fig. 9 and Fig. 10.





Images Diffusivity Transmittance Concentration

Fig. 9. Model building results.

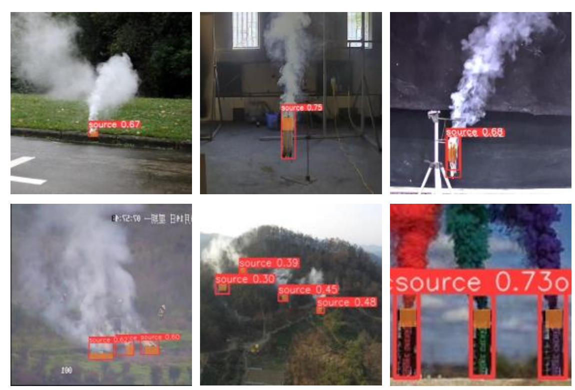


Fig. 10. Smoke source localization results.

V. CONCLUSIONS

This paper presents a smoke source localization method based on smoke segmentation, aimed at improving the precision of smoke source localization. Firstly, the paper compares and analyzes various mainstream smoke segmentation models, evaluating their performance in practical applications, and ultimately selects the most suitable segmentation model as the core algorithm. This model effectively extracts the morphological and concentration distribution characteristics of smoke, which are then used for reverse inference of the smoke diffusion process. Combined with a smoke source detection algorithm, the paper achieves high-precision localization of the smoke source. The core of this method lies in utilizing the results of smoke segmentation to correlate the dynamic diffusion process of smoke with the source point, thereby enabling accurate trace-back localization.

From a practical perspective, this method has significant potential in enhancing early fire detection systems and industrial safety monitoring by providing accurate smoke source localization in complex environments. However, the proposed method still relies heavily on high-quality segmentation results, and its performance in extremely noisy or occluded scenarios requires further investigation. Additionally, the computational cost of the current algorithm may pose challenges for large-scale or real-time applications, which calls for future optimizations.

Future research will focus on two key areas of optimization: first, further improving the processing speed of the segmentation model to meet real-time application requirements; and second, enhancing localization accuracy in multi-target scenarios to ensure efficient and precise smoke source identification in complex environments. Additionally, efforts will be made to simplify the algorithmic process of source localization to reduce computational complexity and improve the system's applicability.

Looking ahead, integrating this method with other data sources, such as thermal imaging or environmental sensors, could further enhance its robustness and adaptability. The continued advancement of AI and IoT technologies also holds promise for creating more intelligent and efficient smoke detection systems. By addressing these challenges and leveraging these opportunities, the proposed method can serve as a foundation for more reliable and effective safety warning systems in the future.

DATA AVAILABILITY

Data used for this article were collected by the research team and will be given to other researchers upon request.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGMENT

This study was supported by the Key Scientific Research Project Plan for Higher Education Institutions of Henan Province, China（No.25A520033)

REFERENCES

[1] M. A. Finney, "The wildland fire system and challenges for engineering," Fire Safety Journal, vol. 120, p. 103085, 2021.

[2] S. Chaturvedi, P. Khanna, and A. Ojha, "Comparative Analysis of Traditional and Deep Learning Techniques for Industrial and Wildfire Smoke Segmentation," in 2021 Sixth International Conference on Image Information Processing (ICIIP), 2021, vol. 6, pp. 326-331.

[3] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431-3440.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, 2015, pp. 234-241: Springer.

[5] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," vol. 39, no. 12, pp. 2481-2495, 2017.

[6] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2881-2890.

[7] L. C. Chen, G. Papandreou, and I. Kokkinos, "Semantic image segmentation with deep convolutional nets and fully connected crfs," in International conference on learning representations, 2015.

[8] L.C. Chen, G. Papandreou, and I. Kokkinos, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," vol. 40, no. 4, pp. 834-848, 2017.

[9] L.C. Chen, G. Papandreou, and Schroff F, "Rethinking atrous convolution for semantic image segmentation," arXiv preprint arXiv:1706.05587, 2017.

[10] L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 801-818.

[11] S. Khan et al., "Deepsmoke: Deep learning model for smoke detection and segmentation in outdoor environments," Expert Systems with Applications, vol. 182, p. 115125, 2021.

[12] T. Wang et al., "AOSVSSNet: Attention-guided optical satellite video smoke segmentation network," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 8552-8566, 2022.

[13] Z. Y. Wang, Y. T. Su, Y. Y. Liu, and W. Zhang, "An improved smoke segmentation algorithm for DeeplabV3 network," Journal of Xidian University, vol. 46, no. 6, pp. 52-59, 2019.

[14] Z. Liu, C. Xie, J. Li, and Y. Sang, "Smoke area segmentation and recognition algorithm based on improved Deeplabv3+," System Engineering and Electronic Technology, vol. 43, no.2, pp. 328-335, 2021.

[15] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, 2018, pp. 3-11.

[16] F. Yuan, L. Zhang, X. Xia, Q. Huang and X. Li, "A Wave-Shaped Deep Neural Network for Smoke Density Estimation," in IEEE Transactions on Image Processing, vol. 29, pp. 2301-2313, 2019.

[17] Y. Li, W. Zhang, Y. Liu, and X. Shao, "A lightweight network for real-time smoke semantic segmentation based on dual paths," Neurocomputing, vol. 501, pp. 258-269, 2022.

[18] F. Yuan, Y. Shi, L. Zhang, and Y. Fang, "A cross-scale mixed attention network for smoke segmentation," Digital Signal Processing, vol. 134, p. 103924, 2023.

[19] M. Park, J. Bak, and S. Park, "Advanced wildfire detection using generative adversarial network-based augmented datasets and weakly supervised object localization," International Journal of Applied Earth Observation and Geoinformation, vol. 114, p. 103052, 2022.

[20] Y. Cao, Q. Tang, X. Wu, and X. Lu, "EFFNet: Enhanced feature foreground network for video smoke source prediction and detection," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 4, pp. 1820-1833, 2021.

[21] H. Zhou, H. Cong, Y. Wang, and Z. Dou, "A computer-vision-based deep learning model of smoke diffusion," Process Safety and Environmental Protection, vol. 187, pp. 721-735, 2024.

[22] C.Y. Xu, B.T. Zha, J.Q. Bao, H. Zhang, and H.X. Li, "Analysis of temporal and spatial distribution characteristics of ammonium chloride smoke particles in confined spaces," Defence Technology, vol. 18, no. 7, pp. 1269-1280, 2022.

[23] X. Zhuang, F Tan, Z Li, L Li, " Image Defogging Algorithm Based On Dark Channel Priorand Optimized Auto-Color," Computer Applications and Software, vol. 38, no. 7, pp. 190-195, 2021.

[24] A. C. G. Jocher, and J. Qiu, “YOLO by Ultralytics.” https://github.com/ultralytics/ ultralytics, 2023. Accessed: February 30, 2023.