fire

Article

An Optimized Smoke Segmentation Method for Forest and Grassland Fire Based on the UNet Framework

Xinyu Hu [1,2], Feng Jiang [1,2], Xianlin Qin [1,2,*], Shuisheng Huang [1,2], Xinyuan Yang [1,2] and Fangxin Meng [1,2]

[1] Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China; hxyuinu@163.com (X.H.); camellia_0601@163.com (F.J.); huangss@caf.ac.cn (S.H.); xinyuan15100637582@163.com (X.Y.); mengfx0208@163.com (F.M.)
[2] Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and Grassland Administration, Beijing 100091, China
* Correspondence: qxl@ifrit.ac.cn; Tel.: +86-10-62888847

Abstract: Smoke, a byproduct of forest and grassland combustion, holds the key to precise and rapid identification—an essential breakthrough in early wildfire detection, critical for forest and grassland fire monitoring and early warning. To address the scarcity of middle–high-resolution satellite datasets for forest and grassland fire smoke, and the associated challenges in identifying smoke, the CAF_SmokeSEG dataset was constructed for smoke segmentation. The dataset was created based on GF-6 WFV smoke images of forest and grassland fire globally from 2019 to 2022. Then, an optimized segmentation algorithm, GFUNet, was proposed based on the UNet framework. Through comprehensive analysis, including method comparison, module ablation, band combination, and data transferability experiments, this study revealed that GF-6 WFV data effectively represent information related to forest and grassland fire smoke. The CAF_SmokeSEG dataset was found to be valuable for pixel-level smoke segmentation tasks. GFUNet exhibited robust smoke feature learning capability and segmentation stability. It demonstrated clear smoke area delineation, significantly outperforming UNet and other optimized methods, with an F1-Score and Jaccard coefficient of 85.50% and 75.76%, respectively. Additionally, augmenting the common spectral bands with additional bands improved the smoke segmentation accuracy, particularly shorter-wavelength bands like the coastal blue band, outperforming longer-wavelength bands such as the red-edge band. GFUNet was trained on the combination of red, green, blue, and NIR bands from common multispectral sensors. The method showed promising transferability and enabled the segmentation of smoke areas in GF-1 WFV and HJ-2A/B CCD images with comparable spatial resolution and similar bands. The integration of high spatiotemporal multispectral data like GF-6 WFV with the advanced information extraction capabilities of deep learning algorithms effectively meets the practical needs for pixel-level identification of smoke areas in forest and grassland fire scenarios. It shows promise in improving and optimizing existing forest and grassland fire monitoring systems, providing valuable decision-making support for fire monitoring and early warning systems.

Keywords: forest and grassland fire; smoke segmentation; GF-6 WFV; deep learning

Citation: Hu, X.; Jiang, F.; Qin, X.; Huang, S.; Yang, X.; Meng, F. An Optimized Smoke Segmentation Method for Forest and Grassland Fire Based on the UNet Framework. *Fire* **2024**, *7*, 68. https://doi.org/10.3390/fire7030068

Academic Editor: Alistair M. S. Smith

Received: 19 January 2024
Revised: 18 February 2024
Accepted: 23 February 2024
Published: 26 February 2024

1. Introduction

Fire is a typical driving factor in terrestrial ecosystems [1–3]. Small-area (<1 ha) and low-intensity forest and grassland fire plays an active role in promoting the stability of forest ecosystems and species evolution [4], for instance by facilitating the removal of surface litter to accelerate nutrient cycling, while large-area and high-intensity forest and grassland fire is the main cause of destruction to natural landscapes and ecosystems [5,6]. For example, it poses a serious threat to wildlife survival through the pollution of water quality. Studies have shown that small-area forest and grassland fire will likely evolve into uncontrollable extreme fire if not detected and intervened with in time [6,7]. Therefore, the prevention and control of forest and grassland fire are more important than governance [8], and early detection of early small fire is crucial [7].

Smoke and thermal anomalies are two important features in the process of forest and grass burning [5]. They are also key breakthroughs in the early warning and monitoring of forest and grassland fire [8]. Thermal infrared data are primarily used to detect heat anomalies due to their wide detection range and high sensitivity [9–11], but it is challenging to detect early small fire and smoldering fire due to tree canopy occlusion and other factors [12]. However, the smoke produced by the burning of forest and grass usually begins in the early stage of the fire, and the spreading area is often much larger than the fire area [13,14]. Due to its fluidity [15], smoke can be used as a good indicator of early small fire and smoldering fire [16].

Videos or images obtained from ground-based platform sensors have commonly been used for the smoke detection of forest and grassland fire [17,18]. Research on smoke feature analysis and detection methods also primarily focuses on them [19,20]. There are a few studies which used multispectral images acquired by space-based platform sensors [21]. However, smoke is better reflected in the visible spectrum with higher spatial resolution satellite images. Making full use of the high temporality and spatiality of satellite images, along with the early appearance of smoke during burnings, can aid in the early detection of forest and grassland fire, which can be helpful to prevent the occurrence of forest and grassland fire [22], thereby reducing the possibility of major forest and grassland fire. As a result, smoke identification based on multispectral satellite images is a hotspot in the research field of wildfire monitoring [21].

The current methods of smoke area identification based on satellite remote sensing mainly include the following: (1) The visual interpretation method. Different wavebands are given red, green, and blue colors to form true (false) color images, and then, smoke areas are identified by visual interpretation [23,24]. (2) The time-series difference method. The smoke area is identified by monitoring the difference between pre-fire and on-fire images of the same area at different times [25]. (3) The empirical threshold method. Thresholds are established based on the differences in physical characteristics between smoke and other surface types. A series of thresholds are used to progressively filter out non-smoke image elements to achieve smoke detection [9,26–28]. (4) Artificial intelligence methods. The early use of traditional machine learning algorithms, such as Random Forests [29], neural networks [30], and Support Vector Machines [31], and the recent successful application of deep learning in the smoke area recognition based on video images has attracted widespread attention from scholars worldwide [12,32–36]. They have tried to conduct research related to satellite image smoke area recognition based on deep learning algorithms, but it is still under exploration and development because of poor interpretability and the need for a large number of training samples. Few studies based on deep learning methods for satellite image smoke segmentation predominantly use low–medium-resolution satellite images [12,35]. This limitation often results in the ineffective segmentation of smoke from small forest and grassland fire.

The visual interpretation method relies entirely on manual work and cannot achieve automatic identification, which is unrealistic for large-scale remote sensing applications; the time-series difference method needs pre-fire data as the background value, which may increase the identification error; the empirical threshold method has poor universality due to the variable threshold value in different images. In contrast, deep learning, as an emerging artificial intelligence methodology, has shown excellent performance in disaster monitoring tasks such as thunderstorms [37], floods [38], droughts [39], landslides [40,41], and water pollution [42,43] due to its stability and transferability. The integration of deep learning with high spatiotemporal remote sensing satellite data is an inevitable trend.

In this study, utilizing middle–high-resolution satellite multispectral imagery, we established a dataset of forest and grassland fire smoke segmentation and developed an optimized smoke segmentation method. This research endeavored to improve the precision of smoke identification by increasing the spatial resolution in satellite data and refining deep learning methods. This systematic approach aimed to enhance the ability to detect small-area fires in forests and grasslands at an early stage. The overarching objective was to provide a methodological reference conducive to advancing technical frameworks for the early monitoring and preemptive warning of forest and grassland fire within the realm of wildfire prevention and management.

2. Materials and Methods

This study utilizes Gaofen-6 satellite Wide-Field-of-View (GF-6 WFV) images and an optimized method based on the UNet framework to carry out smoke segmentation of forest and grassland fire. It primarily involves three aspects: dataset construction, model optimization, and method evaluation, as illustrated in Figure 1. Initially, a global-scale smoke segmentation dataset (CAF_SmokeSEG) was constructed by collecting smoke samples from regions highly susceptible to forest and grassland fire across the world. These samples were collected over multiple years (2019–2022), from diverse locations, during various seasons, and under different fire scenarios. Subsequently, DSASPP (Depthwise Separable Atrous Spatial Pyramid Pooling), SCSE (Spatial and Channel Squeeze-and-Excitation), and HardRes (residual structure with HardSwish activation) modules were incorporated into the traditional UNet framework, resulting in a novel smoke segmentation method named GFUNet. Comparative experiments were conducted against FCN, SegNet, UNet, and other optimization algorithms to confirm the superiority of GFUNet. Furthermore, ablation experiments were performed to demonstrate the effectiveness of module integration. Then, an analysis of band combinations was conducted by adding various bands to the common multispectral sensor settings, exploring the potential of different band combinations in smoke segmentation. Finally, the transferability of the proposed method was assessed by applying it to other similar middle–high-resolution multispectral satellite images.

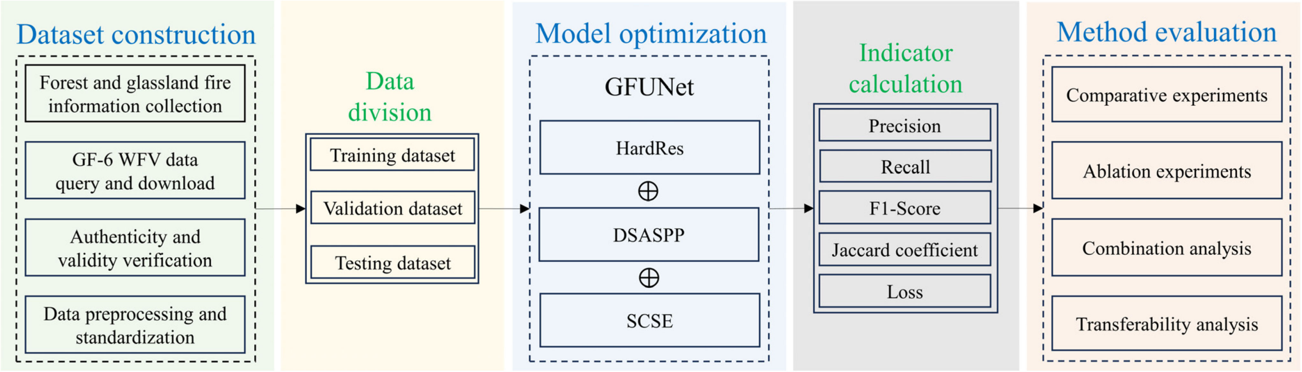

Figure 1. The overall research technical framework.

2.1. Dataset Preparation

2.1.1. Data and Preprocessing

GF-6 WFV images with a spatial resolution of 16 m per ground pixel, 8 spectral bands, and an ultra-wide imaging capability that exceeds 800 km [44] were obtained from the High-Resolution Earth Observation Satellite Project (High-Resolution Project) in China. These images represent the largest optical payload globally in terms of imaging width at an equivalent resolution, which is about three times wider than the Sentinel-2 satellite of the European Space Agency with a similar resolution satellite payload. The specific data parameters are detailed in Table 1. GF-6 can conduct coordinated operations for the first time with the Gaofen-1 satellite (GF-1) and the Environment Disaster Mitigation 2 satellite (HJ-2A/B), which will improve satellite utilization and enhance the efficiency of remote sensing observation in China. GF-6 WFV was used to construct the dataset and to train the smoke segmentation models in this study, while GF-1 WFV and HJ-2A/B CCD were used for model transfer applications.

Table 1. Technical parameters of GF-6 WFV data.

NO.	Band Name	Spectral Range/μm	Spatial Resolution/m	Revisit Cycle/d	Width/km
1	Blue	0.45–0.52	16	4	≥800
2	Green	0.52–0.59			
3	Red	0.63–0.69			
4	NIR	0.77–0.89			
5	Red-Edge1	0.69–0.73			
6	Red-Edge2	0.73–0.77			
7	Coastal Blue	0.40–0.45			
8	Yellow	0.59–0.63			

Global forest and grassland fire data from 2019 to 2022 were extensively collected across search engines. Additionally, the validity of the data was ensured by referencing sources such as NASA FIRMS (https://firms.modaps.eosdis.nasa.gov/, accessed on 19 March 2023). Images of high-incidence forest and grassland fire worldwide were accessed through the China Centre for Resources Satellite Data and Application (https://data.cresda.cn/, accessed on 10 October 2023), totaling 93 images of GF-6 WFV 1A-level data. Subsequently, the raw data were subjected to preprocessing using ENVI 5.6.2. Radiometric calibration was performed to produce apparent reflectance data. Geometric correction was then applied to transform the images into the Universal Transverse Mercator (UTM) projection. Data storage followed the Band Sequential (BSQ) format, and the saved format was TIF. Next, smoke areas within all images were visually interpreted and cropped into sample images of 512 × 512 pixels in size and saved separately. Finally, the regions of smoke areas in the sample images were vectorized using ArcGIS 10.6. Smoke areas were labeled with an attribute value of 255, while non-smoke areas were labeled with an attribute value of 0. These labeled samples were saved and exported as sample labels, following the same data storage and format as previously mentioned. Through the aforementioned comprehensive process, the construction of a global-scale dataset for forest and grassland fire smoke was completed and named CAF_SmokeSEG.

2.1.2. Dataset Overview

This study employed a diverse sampling approach to collect data from various regions, seasons, and smoke scenes worldwide to enhance data diversity, thereby improving model robustness and ensuring application transferability. The spatiotemporal distribution of the data is illustrated in Figure 2. In terms of spatial scale, smoke samples were collected across six major continents (excluding Antarctica). In Asia, the primary sources included southern China, northeastern Russia, and southeastern Myanmar. The focus in Europe was primarily on the Balkan Peninsula region. African samples were drawn mainly from the Central African region, including the Democratic Republic of Congo, Angola, and South Sudan. Oceania's data primarily originated from the northern regions of Australia. The focus in North America was centered around the northwestern United States, while in South America, it was primarily concentrated in the southwestern parts of Brazil and northeastern Paraguay. Smoke samples were collected throughout the year, across all four seasons, from 2019 to 2022. A total of 51 samples were collected from January to March, 43 samples from April to June, 112 samples from July to September, and 34 samples from October to December. Due to significant environmental differences across regions globally, there were variations in combustible materials and differences in the timing of forest and grassland fire; the sampling times also varied. For instance, smoke samples from the Asia were predominantly sampled between January and March, while South America's samples were primarily collected from July to September. In Africa, samples were obtained throughout the year, with representation in each quarter. In terms of smoke scenes, there were transitional areas between agriculture and forestry, as well as forest-urban interfaces, in addition to forest and grassland scenes.

Smoke sample images and labels of forest and grassland fire across various continents under different occurrence scenes are illustrated in Figure 3.

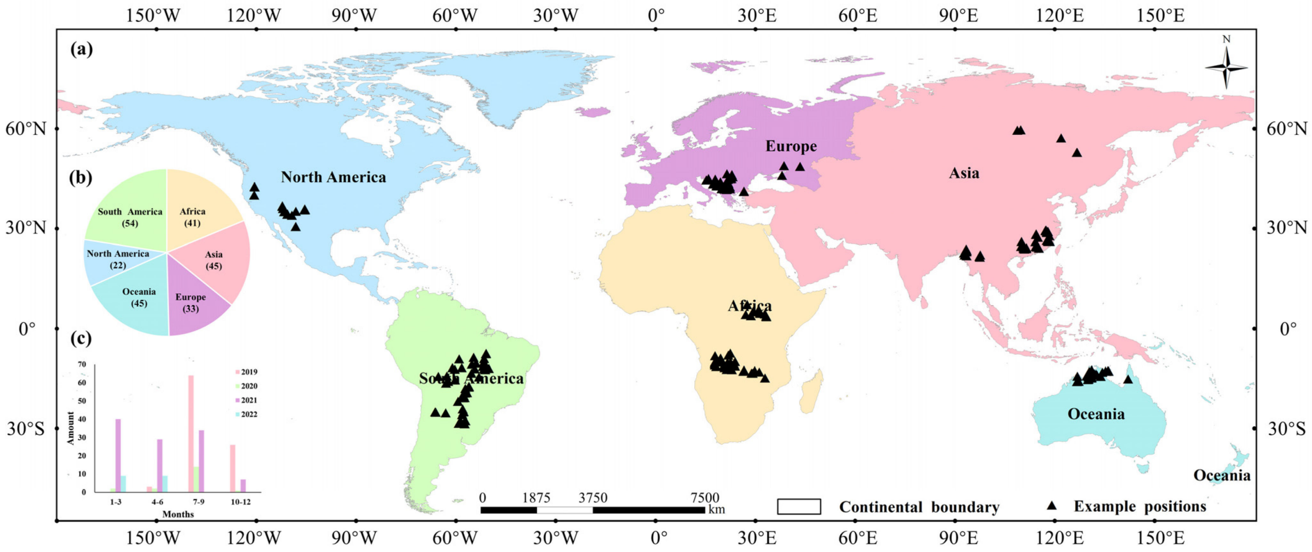

Figure 2. The spatiotemporal distribution of CAF_SmokeSEG. (**a**) Spatial distribution of the data; (**b**) Different intercontinental data distribution; (**c**) Temporal distribution of the data.

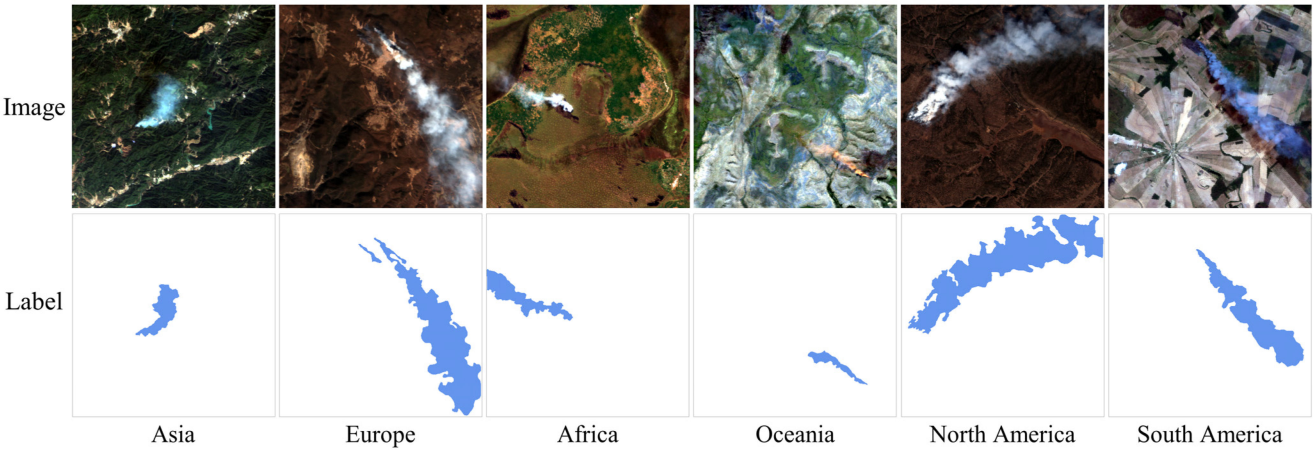

Figure 3. Examples of global samples and labels in CAF_SmokeSEG.

2.2. Methodology

2.2.1. Module Introduction

1. HardRes

As neural network architectures become increasingly deep, challenges related to abnormal gradients and performance degradation ensue. The former refers to issues such as vanishing or exploding gradients, which can severely hinder the training process. In many cases, these problems can be alleviated by employing techniques like data normalization during initialization and batch normalization. However, another effective approach to addressing both gradient issues and enhancing model performance is the incorporation of residual structures. Residual networks have demonstrated the ability to mitigate gradient-related problems and stabilize network training. This architectural innovation helps prevent training errors and ultimately leads to improved model recognition performance. Compared to traditional convolutional structures (Figure 4a), the residual blocks consist of two paths (Figure 4b): the “shortcut path”, which performs a simple identity mapping, and the “main path”, which contains the learnable convolutional layers. The output of these two paths is combined element-wise through addition. This additive combination of the shortcut path and the main path allows the network to learn residual mappings, effectively capturing the deviations from the identity mapping.

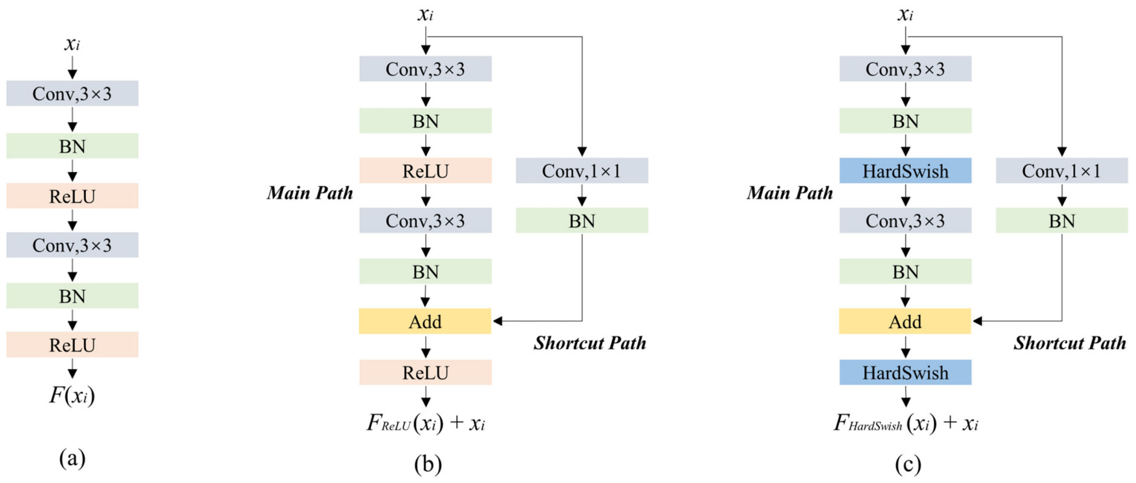

Figure 4. Schematic diagrams of different structures. (**a**) Traditional double convolution; (**b**) Traditional residual convolution; (**c**) HardRes modules.

While the conventional residual convolution modules, utilizing the ReLU [45] activation function, effectively addressed the vanishing gradient problem and facilitated neural network training, a limitation known as the "dying ReLU" emerged, where neurons could become inactive during training, thus limiting the network's representational capacity. To address this limitation, ReLU6 [46] was introduced, which mitigated the "dying ReLU" problem by setting an upper bound on neuron saturation while preserving the core ReLU characteristics. Swish [47] introduced a smooth, non-monotonic activation, potentially enhancing generalization, albeit at the cost of increased computational complexity. To reduce the computational overhead of Swish, HardSwish [48] was introduced, inheriting its advantageous features while optimizing for computational efficiency. HardSwish was introduced to optimize the residual modules, aiming to construct a more computationally efficient and robust network architecture, designated as "HardRes". The curves of various activation functions are illustrated in Figure 5, with their specific mathematical expressions detailed as follows. The structure of the HardRes module is illustrated in Figure 4c.

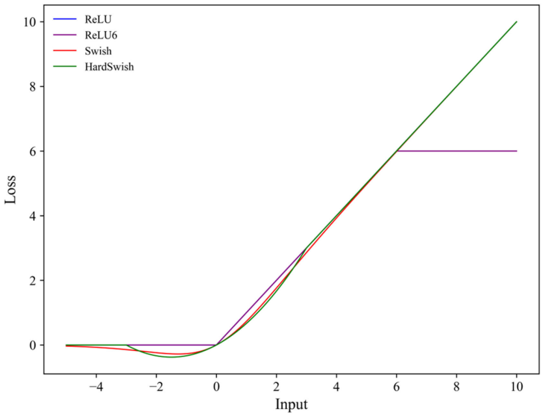

Figure 5. The curves of various activation functions.

2. DSASPP

DSASPP is an improved version of the ASPP [49], incorporating deep separable convolutions into the original dilated convolution operations (Figure 6). It effectively captures multiscale information while maintaining efficient computational performance by decoupling spatial and channel processing. At its core, the module comprises several key elements:

(1) Depthwise separable convolution: DSASPP leverages depthwise separable convolutions, which consist of depthwise convolutions followed by pointwise convolutions. This separation of convolutional operations reduces computational complexity while maintaining strong representational power. It allows the module to extract features at multiscale effectively.

(2) Atrous convolution: Atrous convolution, also known as dilated convolution, is employed within the DSASPP module to expand the receptive field without increasing the number of parameters. By applying convolutional filters with varying dilation rates, the module captures contextual information at different scales, enabling it to perceive both local and global features.

(3) Spatial pyramid pooling: The spatial pyramid pooling component within DSASPP divides the input feature map into grids of different sizes and computes pooling operations independently within each grid. This process effectively captures features at multiple spatial resolutions, enabling the module to handle objects of various sizes within an image.

DSASPP's architectural design allows it to adapt to the specific requirements of the task at hand, capturing contextual information across different scales. This versatility makes it particularly well suited for image segmentation, object detection, and other computer vision tasks where a robust representation of multiscale features is essential.

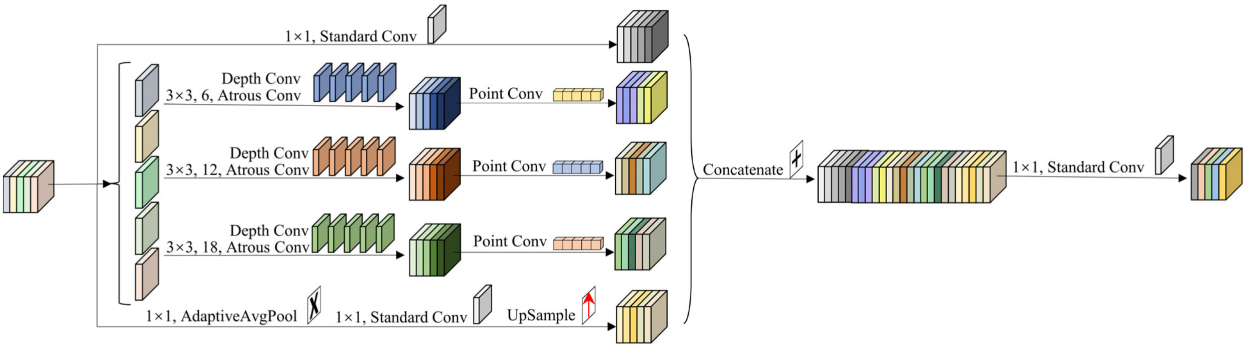

Figure 6. Schematic diagram of the DSASPP structure.

3. SCSE

SCSE [50] seamlessly integrates the strengths of both Channel Squeeze-and-Excitation (CSE) and Spatial Squeeze-and-Excitation (SCE) into convolutional neural networks (Figure 7). The SCSE mechanism is composed of two essential components:

(1) CSE: CSE focuses on recalibrating channel-wise features by capturing global information for each channel through Global Average Pooling (GAP). This process eliminates redundant information while emphasizing crucial features within each channel.

(2) SCE: In addition to CSE, SCE operates at the spatial level, dynamically adjusting the importance of spatial regions within an image. This adaptation is facilitated by learned weights that modulate the significance of spatial features, allowing the network to prioritize areas of interest while de-emphasizing irrelevant regions.

The harmonious integration of CSE and SCE within SCSE enables the network to simultaneously highlight relevant channels and spatial regions during both the forward and backward passes. This adaptability enhances the network's ability to capture critical information across multiple hierarchical levels, ultimately resulting in improved performance across a variety of tasks.

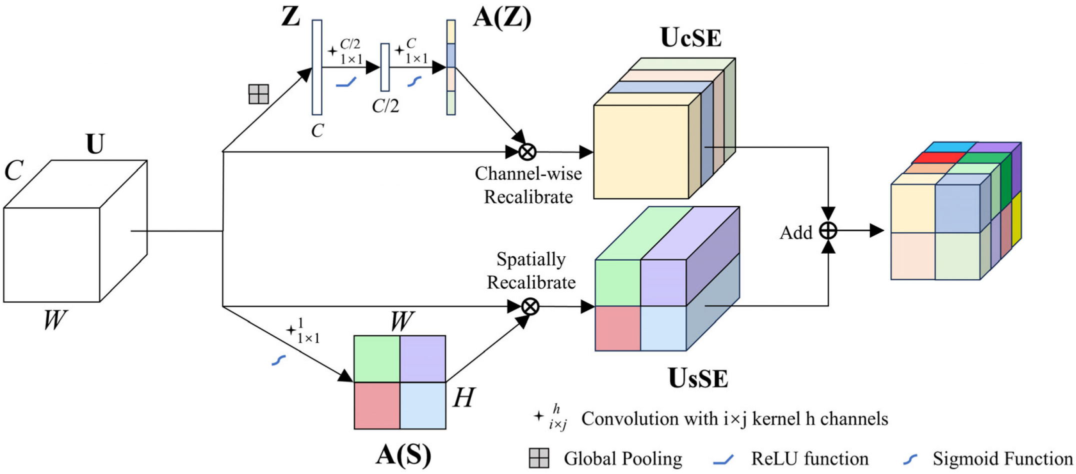

Figure 7. Schematic diagram of the SCSE structure.

2.2.2. Model Optimization

The segmentation method for the smoke of forest and grassland fire proposed in this paper, GFUNet, is illustrated in Figure 8. It is primarily built upon the foundational UNet framework and incorporates three key optimizations mentioned earlier. Firstly, standard convolutional layers were replaced with HardRes. This change allowed for the fusion of features from the previous layer while learning deep features, preventing the loss of smoke-related information as the network depth increases. Secondly, after each HardRes, SCSE was introduced to dynamically update weights in both spatial and channel domains during the learning process, assisting the model in focusing on essential smoke-related information. Thirdly, DSASPP was introduced as a replacement for the traditional skip connection to gather smoke information from different scales and enhances its feature representation capability. These three optimizations collectively contributed to the effectiveness of GFUNet in the segmentation of smoke from forest and grassland fire, enhancing both its feature extraction capabilities and learning efficiency.

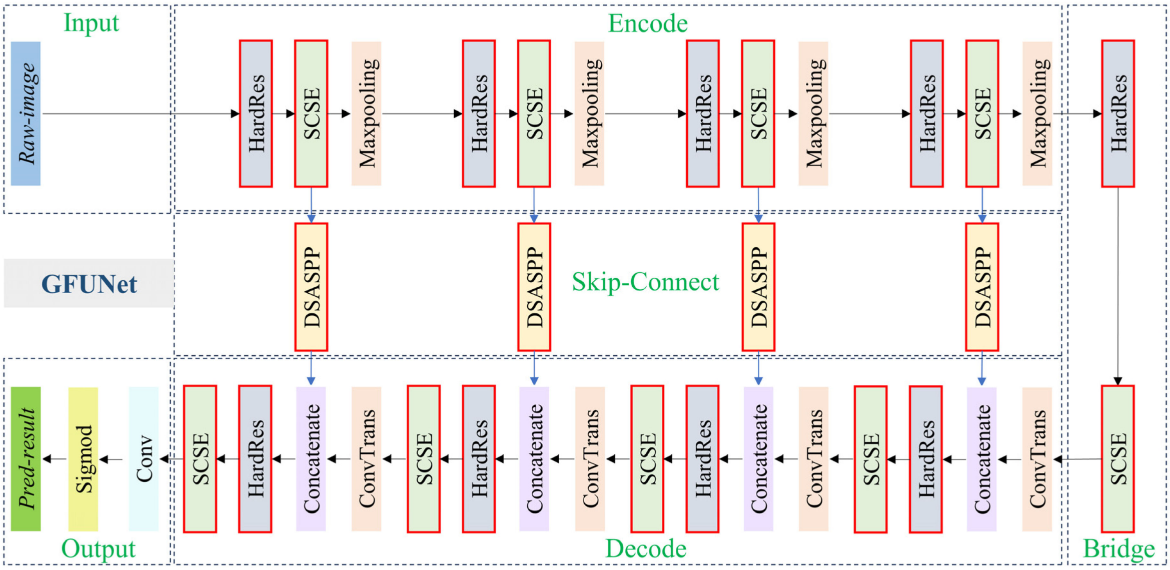

Figure 8. Overview of GFUNet network. (The modules with a red border are the main optimized modules based on the UNet framework).

2.3. Experimental Environment

The experimental platform for this study consisted of the following components: an Intel(R) Xeon(R) E5-2680 V4 2.40GHz processor (CPU) with 64GB of RAM and an NVIDIA GeForce Tesla V100 graphics card (GPU) with 16GB of GPU memory. The image segmentation network was constructed using PyTorch 1.10 and related image processing packages. Data preprocessing was carried out using software tools such as ENVI 5.6.2 and ArcGIS 10.6. The hyperparameters for model training were set as follows: initial learning rate (Learn Rate): 0.0002; batch size: 4; training epochs: 300. At the end of each epoch, the model was validated using a validation dataset, and the evaluation results were recorded. The best-performing segmentation model was saved after training.

2.4. Accuracy Assessment

Four commonly employed evaluation metrics in the field of deep learning, namely precision, recall, F1-Score, and the Jaccard coefficient, were utilized to comprehensively assess the accuracy and completeness of various segmentation results. Precision assesses the proportion of true positive samples in the results returned by a classifier, providing an indication of the accuracy of classification results. Recall quantifies the ratio of correctly retrieved positive samples to all positive samples, reflecting the completeness of classification results. The F1-Score, as the harmonic mean of precision and recall, offers a balanced assessment of both accuracy and completeness. The Jaccard coefficient measures the similarity between the predicted and ground truth regions by calculating the intersection of the two regions divided by their union. Their values range from 0 to 1, with higher values indicating a greater overlap between the prediction and label regions, thus signifying better smoke segmentation performance. The relationships of different accuracy indicators are depicted in Figure 9. The calculation formulas for the relevant accuracy metrics are as follows:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \tag{3}$$

$$Jaccard\ coefficient = \frac{TP}{TP + FP + FN} \tag{4}$$

where these metrics, based on indicators such as true positive (TP), true negative (TN), false negative (FN), and false positive (FP), provide a comprehensive evaluation of the model's ability to accurately delineate and encompass the regions of smoke.

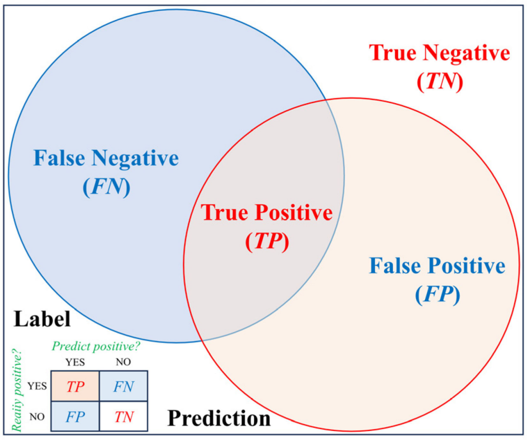

Figure 9. Schematic diagram of accuracy calculation metrics relationship.

3. Results

3.1. Comparative Experiment

To validate the performance advantages of GFUNet, comparative experiments were conducted with UNet [51], FCN [52], SegNet [53], and SmokeUNet [12]. The CAF_SmokeSEG dataset, which included a total of 240 images and their corresponding labels, was split as follows: 80% of the data were allocated as the training set for constructing the smoke segmentation model, 10% served as the validation set for fine-tuning model hyperparameters and saving the best model, and the remaining portion was reserved as the test set to assess the generalization capabilities of the model. To augment the sample size and mitigate overfitting, rotation, transpose, translation, scaling, brightness adjustments, and other enhancement methods were randomly applied during the training stage. Smoke samples of different sizes were selected to enhance the robustness of the model in segmenting different smoke while trying to ensure that the variability in the distribution of smoke samples was similar at each stage. The results of different methods are compared in Table 2 and Figure 10.

Table 2. Accuracy evaluation of smoke segmentation results of different methods.

Methods	Stage	Precision/%	Recall/%	F1-Score/%	Jaccard/%
SegNet	train	82.39	77.70	78.39	66.61
	test	86.77	70.66	72.38	60.70
FCN	train	87.21	83.75	84.84	74.41
	test	92.59	69.35	75.28	63.41
UNet	train	84.08	84.22	83.36	72.44
	test	89.01	73.34	74.62	63.56
SmokeUNet	train	86.54	83.70	84.43	73.76
	test	94.94	70.14	76.87	66.01
GFUNet	train	**89.15**	**89.41**	**88.95**	**80.55**
	test	**94.00**	**80.25**	**85.50**	**75.76**

In Figure 10, examples of smoke segmentation results from the test set across different continents are presented. While all five methods can effectively segment smoke, their segmentation performance varies. GFUNet stood out with the following key advantages: (1) Enhanced smoke region contour description: GFUNet demonstrated a stronger capability to describe the contours of smoke areas, as evident in the smoke samples from Asia and North America. (2) Improved small smoke area detection: GFUNet exceled at capturing small smoke areas, as seen in samples from Oceania and Europe, particularly in the case of smaller smoke portions. (3) Robust thin smoke pixel identification: GFUNet exhibited robust identification of thin smoke pixel regions. This strength was evident not only in the tails of the smoke in samples from Africa and South America but also in the interior of smoke areas in European samples. Compared to other methods, GFUNet tended to have fewer hollow or unsegmented areas. These observations collectively highlighted the superior segmentation performance of GFUNet compared to other methods, making it an effective choice for smoke segmentation in various geographic regions.

In Table 2, a comparison of smoke segmentation accuracy results between the training and testing phases for different methods is presented. It was evident that all five methods exhibit consistent accuracy metrics on the training and testing sets, aligning with the segmentation results discussed earlier. GFUNet outperformed other methods in both phases. In the testing stage, it demonstrated improvements of 4.99% in precision, 6.91% in recall, 10.88% in F1-Score, and 12.20% in the Jaccard coefficient compared to UNet, along with various improvements compared to other methods in terms of different accuracy metrics.

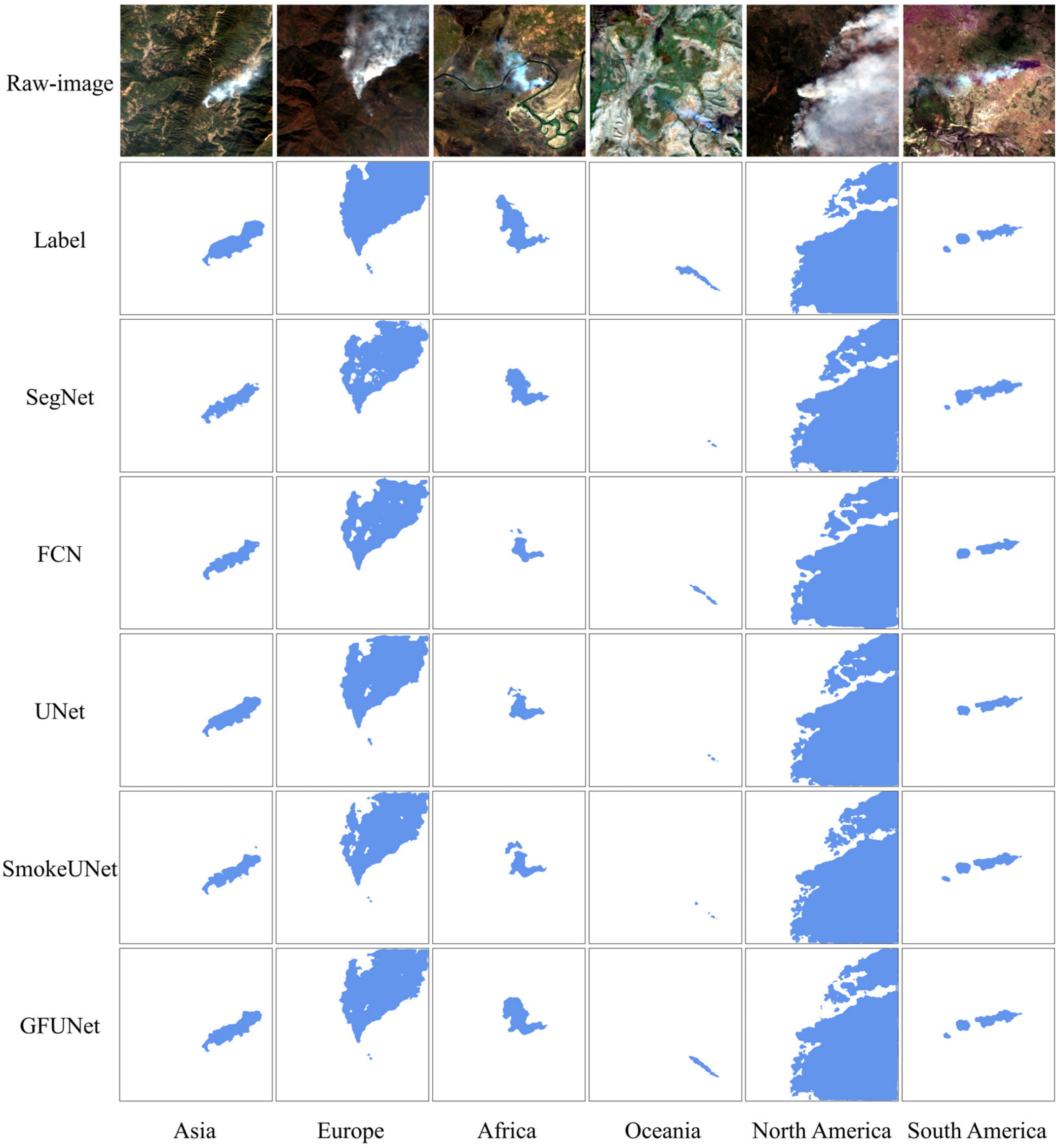

Figure 10. Comparison of smoke segmentation results of different methods.

3.2. Ablation Experiment

To further validate the effectiveness of the added modules, ablation experiments were conducted on three modules, HardRes, DSASPP, and SCSE, each trained with the same parameter settings. Throughout the training process, various alterations in the loss function and multiple accuracy metrics were observed, as illustrated in Figure 11. Compared to the complete GFUNet, module ablation led to varying degrees of performance degradation in model learning. Notably, GFUNet-SCSE resulted in significant changes, with precision showing substantial decreases, and the loss function exhibited a stair-like convergence trend, stabilizing after 50 epochs, and others achieved convergence before 50 epochs. Furthermore, GFUNet-DSASPP resulted in a decrease in recall, while GFUNet-HardRes led to slight reductions in both precision and recall. Based on the changes in the F1-Score and Jaccard coefficient resulting from the factors mentioned above, it was evident that the module ablation had a detrimental impact on the model's ability to learn smoke features, with SCSE having the most significant effect, followed by HardRes and, finally, DSASPP.

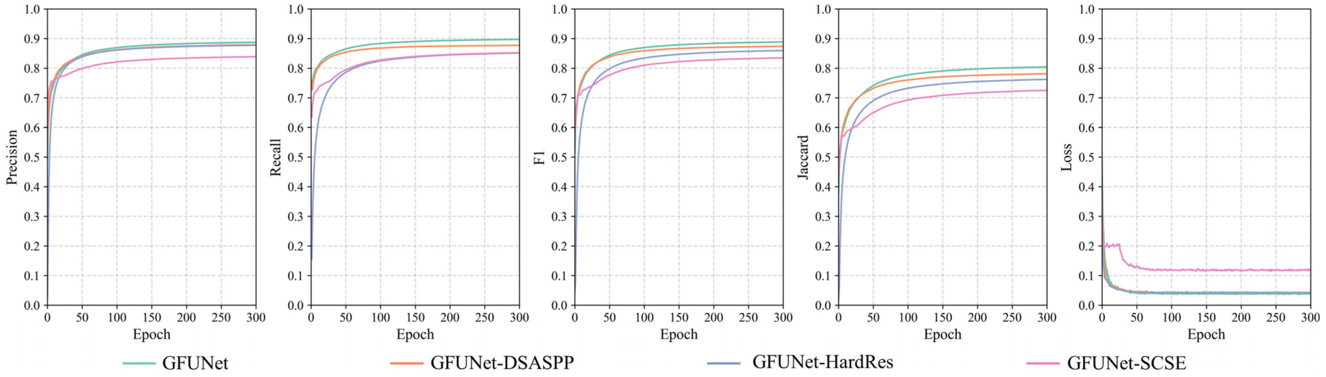

Figure 11. Accuracy changes in module ablation experiments in the training stage.

The accuracy metrics during the testing process with the ablation of different modules are presented in Table 3. The significant impact of GFUNet-SCSE on the precision, leading to a 9.36% decrease, was clearly reflected in the subsequent 7.80% reduction in the F1-Score and a 9.11% decrease in the Jaccard coefficient, affecting the comprehensive accuracy metrics. GFUNet-DSASPP resulted in a marginal 2.78% decrease in recall, consequently causing a 2.32% and 2.91% drop in the F1-Score and Jaccard coefficient, respectively. GFUNet-HardRes led to a slight reduction of 2.86% in precision and 2.05% in recall, resulting in corresponding decreases of 3.90% in the F1-Score and 4.98% in the Jaccard coefficient. The comprehensive impact of module ablations during the testing stage was consistent with the observations made during the training stage. Hence, the incorporation of SCSE contributed to improving the model's smoke segmentation precision, while DSASPP enhanced recall. The addition of HardRes yielded improvements in both the precision and recall. The module ablation experiments demonstrated the effectiveness of these module additions in optimizing the model's capacity for smoke feature learning and generalization, as evidenced by the performance comparison during the training and testing stages.

Table 3. Accuracy metrics of module ablation experiments in the testing stage.

Methods	Precision/%	Recall/%	F1-Score/%	Jaccard/%
GFUNet	**94.00**	**80.25**	**85.50**	**75.76**
GFUNet-DSASPP	94.00	77.47	83.18	72.85
GFUNet-HardRes	91.14	78.20	81.60	70.78
GFUNet-SCSE	84.64	80.58	77.70	66.65

3.3. Combination Analysis

Building upon the foundation of the smoke segmentation method, further training of GFUNet was conducted using nine different band combinations as inputs, with detailed segmentation accuracy evaluations provided in Table 4. Among these combinations, C1 and C2 consisted of two commonly used spectral bands from remote sensing satellite multispectral sensors. C1 comprised the red, green, and blue bands, while C2, building upon C1, added the near-infrared band. C3~5 and C6~8 extend C1 and C2 by introducing additional bands unique to the GF-6 WFV sensor, including the red-edge, coastal blue, and yellow-edge bands. C9 was formed by using all eight bands of the GF-6 WFV sensor.

As shown in Table 4, C1 achieved good results using only the red, green, and blue bands, with an F1-Score and Jaccard coefficient of 80.58% and 70.06%, respectively. The addition of the NIR band in C2 led to an improvement in segmentation accuracy, with a 0.93% increase in the F1-Score and a 0.72% increase in the Jaccard coefficient. C3~5 showed an F1-Score and Jaccard coefficient increase of 0.46% and 0.73%, 2.81% and 3.12%, and 2.72% and 3.00%, compared to C1. For C6~8, the F1-Score and Jaccard coefficient exhibited improvements of 0.48% and 0.61%, 2.32% and 2.41%, and 2.74% and 3.17%, respectively, in comparison to C2. The best segmentation accuracy was achieved with the full-band

Combination 9 of GF-6 WFV data, with an F1-Score and Jaccard coefficient as high as 85.50% and 75.76%, respectively. C9 demonstrated favorable performance with both precision and recall, and it maintained a relative lead over other combinations.

Table 4. Accuracy metrics of different band combinations.

Combinations	Band Composition	Precision/%	Recall/%	F1-Score/%	Jaccard/%
C1	b1, b2, b3	91.29	77.51	80.58	70.06
C2	b1, b2, b3, b4	91.03	78.20	81.51	70.78
C3	b1, b2, b3, b5	92.48	76.93	81.04	70.79
C4	b1, b2, b3, b7	91.38	80.47	83.39	73.18
C5	b1, b2, b3, b8	92.35	79.44	83.30	73.06
C6	b1, b2, b3, b4, b5	91.27	78.46	81.99	71.39
C7	b1, b2, b3, b4, b7	94.84	77.21	83.83	73.19
C8	b1, b2, b3, b4, b8	92.33	79.85	84.25	73.95
C9	All bands	**94.00**	**80.25**	**85.50**	**75.76**

It is worth noting that the precision for different band combinations remains consistently above 90% and exhibits minor variations. In contrast, the recall demonstrated more substantial changes when additional bands were introduced. This observation indicated that adding spectral bands can enhance the model's ability to recall smoke regions, ultimately leading to improved overall accuracy metrics such as the F1-Score and the Jaccard coefficient. The introduction of shorter-wavelength coastal blue and yellow bands demonstrated a more pronounced improvement in accuracy metrics than that observed with longer-wavelength red-edge bands. Furthermore, the GFUNet network can autonomously learn effective features for distinguishing smoke from non-smoke within the original features. Without considering computational costs, this capability allowed for the direct input of all features, enabling the segmentation network to automatically and effectively delineate smoke regions.

3.4. Transferability Analysis

In order to assess the transferability of GFUNet in smoke segmentation across different datasets, a model trained with the common spectral bands (red, green, blue, and NIR) of multispectral satellite data from GF-6 WFV was directly applied to smoke segmentation tests using satellite data from GF-1 WFV and HJ-2A/B CCD datasets, which have the same spatial resolution and similar band configurations. The smoke from forest and grassland fire occurring in Asia, Africa, and North America in 2023 was selected as a sample for conducting transferability analysis. The specific occurrence times and locations are detailed in Table 5. These datasets were used to test the transferability of the GFUNet model to different regions and satellite sources, further assessing its adaptability and performance.

Table 5. Sample information and accuracy metrics of transferability analysis.

Data	Location	Date	Precision/%	Recall/%	F1-Score/%	Jaccard/%
GF-1 WFV	Alberta, Canada	7 June 2023	85.78	89.10	87.41	77.63
	Kigoma, Tanzania	25 June 2023	99.31	82.44	90.09	81.97
	Amurskaya, Russia	9 July 2023	98.44	93.33	95.82	91.97
Average Accuracy Indicator			**94.51**	**88.29**	**91.11**	**83.86**
HJ-2A/B CCD	Alberta, Canada	4 July 2023	99.83	57.90	73.30	57.85
	Tanganyika, Congo	15 July 2023	83.67	90.90	87.13	77.20
	Çanakkale, Turkey	23 August 2023	98.45	84.04	90.67	82.94
Average Accuracy Indicator			**93.98**	**77.61**	**83.70**	**72.66**

The smoke segmentation results using GFUNet on selected GF-1 WFV and HJ-2A/B CCD datasets, as shown in Figure 12, demonstrated that GFUNet model trained with GF-6 WFV data was also applicable to segmenting smoke regions in other datasets. The model exhibited a good ability to distinguish between smoke and non-smoke areas, as seen in Figure 12c,e, where the smoke segmentation was quite complete, and the boundaries were well defined. However, there were still instances of partial under-segmentation in thin smoke areas, which were more pronounced in Figure 12a,d, also prominently observed in the small smoke areas.

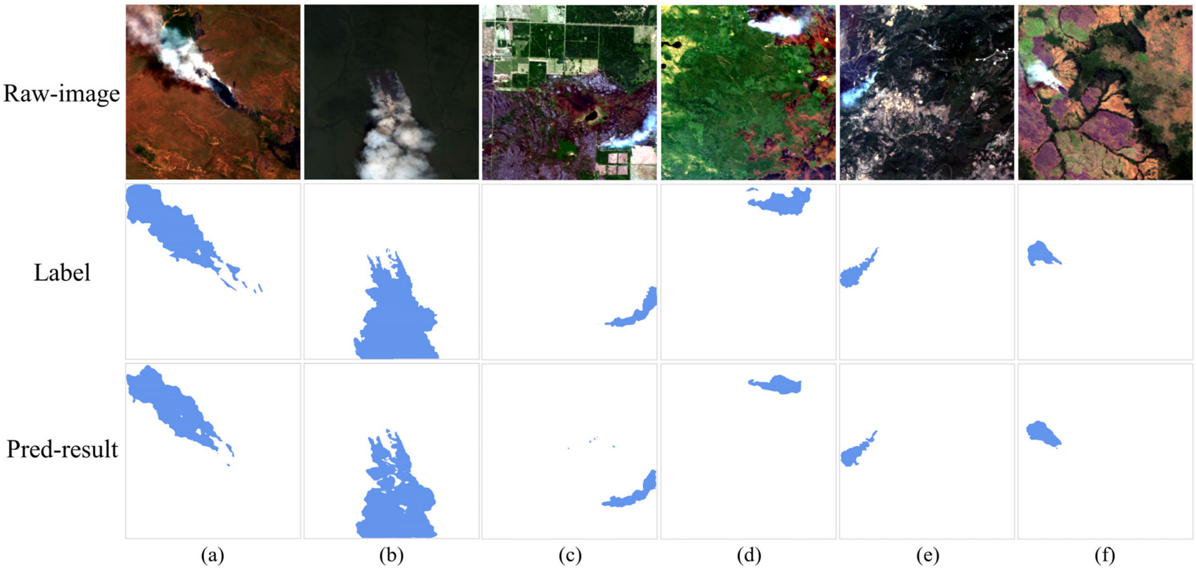

Figure 12. Comparison of smoke segmentation results of different data. Subfigures (**a**–**c**) are GF-1 WFV data; subfigures (**d**–**f**) are HJ-2A/B CCD data.

The transfer accuracy evaluation results in Table 5 reveal that the transfer segmentation accuracy of GF-1 WFV and HJ-2A CCD was similar to the segmentation results of the same band combinations as those in the GF-6 WFV band combination experiment (Table 4). For GF-1 WFV, the precision, recall, F1-Score, and Jaccard coefficient varied from a minimum of 85.78%, 82.44%, 87.41%, and 77.63% to a maximum of 99.31%, 93.33%, 95.82%, and 91.97%, with average values of 94.51%, 88.29%, 91.11%, and 83.86%. Similarly, for HJ-2A CCD, the precision, recall, F1-Score, and Jaccard coefficient ranged from a minimum of 83.67%, 57.90%, 73.30%, and 57.85% to a maximum of 99.83%, 90.90%, 90.67%, and 82.94%, with average values of 93.98%, 77.61%, 83.70%, and 72.66%. In summary, GFUNet exhibited minimal mis-segmentation and some under-segmentation in medium-to-high-resolution multispectral data for smoke segmentation tasks, indicating good transferability across various regions and data sources.

4. Discussion

Smoke segmentation based on remote sensing images enables us to detect early-stage fire with high precision and recall rates. These two metrics are vital for accurately distinguishing between smoke and non-smoke areas, effectively segmenting as many smoke pixels as possible. In comparative experiments, the SegNet, FCN, and UNet methods struggled to balance both metrics. In contrast, SmokeUNet provided a better trade-off, but its limitation may be attributed to performing only three down-sampling operations on the original image. GFUNet extended the original UNet architecture by incorporating four down-sampling operations, thus delving deeper into the feature differences between smoke and non-smoke areas. Furthermore, the skip connections were replaced with the DSASPP module between down-sampling and up-sampling operations, enabling the incremental stacking of multiscale feature differences. The experimental results indicated that recall was enhanced by these optimization strategies while preserving precision, resulting in optimal smoke segmentation accuracy. Additionally, ablation experiments reaffirmed that different optimization modules contributed to improving both the precision and recall in smoke segmentation tasks, ultimately achieving efficient smoke segmentation. Furthermore, timeliness is important for smoke segmentation as well; yet, GFUNet does not exhibit significant improvements compared to other methods. This will also be one of the optimization areas for future research.

The addition of spectral bands to commonly used spectral bands in multispectral sensors of remote sensing satellites can improve the accuracy of smoke segmentation. This suggests that smoke pixels and non-smoke pixels exhibit spatial feature differences not only in the same band but also across different bands. These distinctions are often challenging to visualize. As evidenced in experiments of band combinations, the full GF-6 WFV band combination yielded the best segmentation results. However, this study has not thoroughly investigated the potential interactions, whether supportive or antagonistic, between different spectral bands. Future research will focus on exploiting these interactions to establish effective remote sensing indices for smoke.

This study conducted smoke segmentation experiments using a relatively mature and reliable algorithm under current hardware conditions and limited data. Despite using limited datasets, this study achieved satisfactory segmentation results for the downstream task of smoke segmentation in forest and grassland fire. Furthermore, the small sample size in this study demanded higher segmentation capabilities of models, indirectly highlighting the advantages of the GFUNet network structure in smoke segmentation tasks. In the future, CAF_SmokeSEG will be slated for gradual expansion and enrichment while being employed to explore the latest deep learning networks in smoke detection-related studies.

5. Conclusions

This study introduced a novel smoke segmentation model, GFUNet, by incorporating the DSASPP, HardRes, and SCSE into the UNet framework. Using the CAF_SmkeSEG dataset established based on GF-6 WFV data, this research conducted a global-scale investigation of smoke segmentation in the context of forest and grassland fire. The results demonstrated that GF-6 WFV data exhibit excellent performance in distinguishing smoke areas from non-smoke areas, indicating their advantage in smoke detection for forest and grassland fire. GFUNet showcased superior feature learning capabilities during training and generalization during testing, outperforming UNet, FCN, SegNet, and SmokeUNet, ultimately achieving the best smoke segmentation results. The precision and recall of GFUNet reached 94.00% and 80.25%, while the F1-Score and Jaccard coefficient achieved 85.50% and 75.76%. The effectiveness of smoke segmentation varied with different band combinations. The addition of specific bands of GF-6 WFV to the two common bands of remote sensing multispectral sensors improved the recall, with the full GF-6 WFV band combination yielding the best segmentation results. GFUNet trained with common spectral bands from multispectral images demonstrated strong transferability and was suitable for detecting smoke in forest and grassland fire with the same spatial resolution and similar wavelength bands as GF-1 WFV and HJ-2A/B CCD data.

This work highlights the potential of smoke as an early warning indicator for forest and grassland fire detection. It leverages the synergy between remote sensing and deep learning to advance disaster monitoring applications and contribute to the establishment of a robust global forest and grassland fire monitoring system. However, there still remain challenges, such as insufficient data diversity and the omission of thin smoke pixels in smoke edge regions. Therefore, the continuous collection of smoke data from global forest and grassland fire and advanced segmentation methods for thin smoke areas are necessary. In the future, we will release the thematic dataset for the smoke segmentation of global forest and grassland fires and collaborate with researchers in deep learning and computer vision to explore smoke detection for forest and grassland fire.

Author Contributions: X.H., X.Q. and S.H. provided the conceptualization and methodology. X.H., F.J., X.Y. and F.M. collected the experimental materials. X.H. and X.Q. performed the data collation and analysis. X.H. wrote the manuscript. X.Q. and S.H. contributed with suggestions. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Non-profit Research Institution of CAF, grant number "CAFYBB2019ZB00304", and the National Science and Technology Major Project of China's High Resolution Earth Observation System, grant number "21-Y30B02-9001-19/22-6".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to restrictions of privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bowman, D.M.; Kolden, C.A.; Abatzoglou, J.T.; Johnston, F.H.; van der Werf, G.R.; Flannigan, M. Vegetation fires in the Anthropocene. *Nat. Rev. Earth Environ.* **2020**, *1*, 500–515. [CrossRef]
2. Keith, D.A.; Ferrer-Paris, J.R.; Nicholson, E.; Bishop, M.J.; Polidoro, B.A.; Ramirez-Llodra, E.; Tozer, M.G.; Nel, J.L.; Mac Nally, R.; Gregr, E.J. A function-based typology for Earth's ecosystems. *Nat. Commun.* **2022**, *610*, 513–518. [CrossRef]
3. Pausas, J.G.; Bond, W.J. On the three major recycling pathways in terrestrial ecosystems. *Trends Ecol. Evol.* **2020**, *35*, 767–775. [CrossRef]
4. Hutto, R.L.; Keane, R.E.; Sherriff, R.L.; Rota, C.T.; Eby, L.A.; Saab, V.A. Toward a more ecologically informed view of severe forest fires. *Ecosphere* **2016**, *7*, e01255. [CrossRef]
5. Chowdary, V.; Gupta, M.K.; Singh, R. A Review on forest fire detection techniques: A decadal perspective. *Networks* **2018**, *4*, 12. [CrossRef]
6. Tedim, F.; Leone, V.; Amraoui, M.; Bouillon, C.; Coughlan, M.R.; Delogu, G.M.; Fernandes, P.M.; Ferreira, C.; McCaffrey, S.; McGee, T.K. Defining extreme wildfire events: Difficulties, challenges, and impacts. *Fire* **2018**, *1*, 9. [CrossRef]
7. Martell, D.L. A review of recent forest and wildland fire management decision support systems research. *Curr. For. Rep.* **2015**, *1*, 128–137. [CrossRef]
8. Xianlin, Q.; Xiaotong, L.; Shuchao, L. Forest fire early warning and monitoring techniques using satellite remote sensing in China. *J. Remote Sens.* **2020**, *5*, 511–520.
9. Dewanti, R.; Lolitasari, I. Detection of Forest Fire, Smoke Source Locations in Kalimantan During the Dry Season for the Year 2015 using Landsat 8 from the Threshold of Brightness Temperature Algorithm. *Int. J. Remote Sens. Earth Sci.* **2017**, *12*, 151–160.
10. Fischer, C.; Halle, W.; Säuberlich, T.; Frauenberger, O.; Hartmann, M.; Oertel, D.; Terzibaschian, T. Small Satellite Tools for High-Resolution Infrared Fire Monitoring. *J. Imaging* **2022**, *8*, 49. [CrossRef]
11. Hua, L.; Shao, G. The progress of operational forest fire monitoring with infrared remote sensing. *J. For. Res.* **2017**, *28*, 215–229. [CrossRef]
12. Wang, Z.; Yang, P.; Liang, H.; Zheng, C.; Yin, J.; Tian, Y.; Cui, W. Semantic segmentation and analysis on sensitive parameters of forest fire smoke using smoke-unet and landsat-8 imagery. *Remote Sens.* **2021**, *14*, 45. [CrossRef]
13. Geetha, S.; Abhishek, C.; Akshayanat, C. Machine vision based fire detection techniques: A survey. *Fire Technol.* **2021**, *57*, 591–623. [CrossRef]
14. Qin, X.-L.; Zhu, X.; Yang, F.; Zhao, K.-R.; Pang, Y.; Li, Z.-Y.; Li, X.-Z.; Zhang, J.-X. Analysis of sensitive spectral bands for burning status detection using hyper-spectral images of Tiangong-01. *Spectrosc. Spectr. Anal.* **2013**, *33*, 1908–1911.
15. Çetin, A.E.; Dimitropoulos, K.; Gouverneur, B.; Grammalidis, N.; Günay, O.; Habiboğlu, Y.H.; Töreyin, B.U.; Verstockt, S. Video fire detection—Review. *Digit. Signal Process.* **2013**, *23*, 1827–1843. [CrossRef]
16. Zhan, J.; Hu, Y.; Cai, W.; Zhou, G.; Li, L. PDAM–STPNNet: A small target detection approach for wildland fire smoke through remote sensing images. *Symmetry* **2021**, *13*, 2260. [CrossRef]
17. Chaturvedi, S.; Khanna, P.; Ojha, A. A survey on vision-based outdoor smoke detection techniques for environmental safety. *ISPRS J. Photogramm. Remote Sens.* **2022**, *185*, 158–187. [CrossRef]
18. Gaur, A.; Singh, A.; Kumar, A.; Kumar, A.; Kapoor, K. Video flame and smoke based fire detection algorithms: A literature review. *Fire Technol.* **2020**, *56*, 1943–1980. [CrossRef]
19. Shi, J.; Yuan, F.; Xia, X. Video smoke detection: A literature survey. *Image Graph.* **2018**, *23*, 303–322.
20. Xia, X.; Yuan, F.; Zhang, L.; Yang, L.; Shi, J. From traditional methods to deep ones: Review of visual smoke recognition, detection, and segmentation. *J. Image Graph.* **2019**, *24*, 1627–1647.
21. Barmpoutis, P.; Papaioannou, P.; Dimitropoulos, K.; Grammalidis, N. A review on early forest fire detection systems using optical remote sensing. *Sensors* **2020**, *20*, 6442. [CrossRef]

22. Sun, X.; Sun, L.; Huang, Y. Forest fire smoke recognition based on convolutional neural network. *J. For. Res.* **2021**, *32*, 1921–1927. [CrossRef]
23. Christopher, S.; Chou, J. The potential for collocated AGLP and ERBE data for fire, smoke, and radiation budget studies. *Int. J. Remote Sens.* **1997**, *18*, 2657–2676. [CrossRef]
24. Chung, Y.-S.; Le, H. Detection of forest-fire smoke plumes by satellite imagery. *Atmos. Environ.* **1984**, *18*, 2143–2151. [CrossRef]
25. Chrysoulakis, N.; Herlin, I.; Prastacos, P.; Yahia, H.; Grazzini, J.; Cartalis, C. An improved algorithm for the detection of plumes caused by natural or technological hazards using AVHRR imagery. *Remote Sens. Environ.* **2007**, *108*, 393–406. [CrossRef]
26. Lu, X.; Zhang, X.; Li, F.; Cochrane, M.A.; Ciren, P. Detection of fire smoke plumes based on aerosol scattering using VIIRS data over global fire-prone regions. *Remote Sens.* **2021**, *13*, 196. [CrossRef]
27. Xie, Y. *Detection of Smoke and Dust Aerosols Using Multi-Sensor Satellite Remote Sensing Measurements*; George Mason University: Fairfax, VA, USA, 2009.
28. Xie, Y.; Qu, J.; Xiong, X.; Hao, X.; Che, N.; Sommers, W. Smoke plume detection in the eastern United States using MODIS. *Int. J. Remote Sens.* **2007**, *28*, 2367–2374. [CrossRef]
29. Ko, B.; Park, J.; Nam, J.-Y. Spatiotemporal bag-of-features for early wildfire smoke detection. *Image Vis. Comput.* **2013**, *31*, 786–795. [CrossRef]
30. Li, Z.; Khananian, A.; Fraser, R.H.; Cihlar, J. Automatic detection of fire smoke using artificial neural networks and threshold approaches applied to AVHRR imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1859–1870.
31. Xiong, D.; Yan, L. Early smoke detection of forest fires based on SVM image segmentation. *J. For. Sci.* **2019**, *65*, 150–159. [CrossRef]
32. Ba, R.; Chen, C.; Yuan, J.; Song, W.; Lo, S. SmokeNet: Satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention. *Remote Sens.* **2019**, *11*, 1702. [CrossRef]
33. Chen, S.; Cao, Y.; Feng, X.; Lu, X.J.N. Global2Salient: Self-adaptive feature aggregation for remote sensing smoke detection. *Neurocomputing* **2021**, *466*, 202–220. [CrossRef]
34. Ismanto, H.; Marfai, M. Classification tree analysis (Gini-Index) smoke detection using Himawari_8 satellite data over Sumatera-Borneo maritime continent Sout East Asia. In *Proceedings of the IOP Conference Series: Earth and Environmental Science: The 2nd International Conference on Environmental Resources Management in Global Region (ICERM 2018), Yogyakarta, Indonesia, 22–23 October 2018*; IOP Publishing: Bristol, UK, 2019; p. 012043.
35. Li, X.; Song, W.; Lian, L.; Wei, X. Forest fire smoke detection using back-propagation neural network based on MODIS data. *Remote Sens.* **2015**, *7*, 4473–4498. [CrossRef]
36. Li, X.; Wang, J.; Song, W.; Ma, J.; Telesca, L.; Zhang, Y. Automatic smoke detection in modis satellite data based on k-means clustering and fisher linear discrimination. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 971–982. [CrossRef]
37. Leinonen, J.; Hamann, U.; Sideris, I.V.; Germann, U. Thunderstorm Nowcasting with Deep Learning: A Multi-Hazard Data Fusion Model. *Geophys. Res. Lett.* **2023**, *50*, e2022GL101626. [CrossRef]
38. Yang, L.; Cervone, G. Analysis of remote sensing imagery for disaster assessment using deep learning: A case study of flooding event. *Soft Comput.* **2019**, *23*, 13393–13408. [CrossRef]
39. Zhang, Y.; Xie, D.; Tian, W.; Zhao, H.; Geng, S.; Lu, H.; Ma, G.; Huang, J.; Choy Lim Kam Sian, K.T. Construction of an Integrated Drought Monitoring Model Based on Deep Learning Algorithms. *Remote Sens.* **2023**, *15*, 667. [CrossRef]
40. Liu, Y.; Wu, L. Geological disaster recognition on optical remote sensing images using deep learning. *Procedia Comput. Sci.* **2016**, *91*, 566–575. [CrossRef]
41. Shafapourtehrany, M.; Rezaie, F.; Jun, C.; Heggy, E.; Bateni, S.M.; Panahi, M.; Özener, H.; Shabani, F.; Moeini, H. Mapping Post-Earthquake Landslide Susceptibility Using U-Net, VGG-16, VGG-19, and Metaheuristic Algorithms. *Remote Sens.* **2023**, *15*, 4501. [CrossRef]
42. Pyo, J.; Park, L.J.; Pachepsky, Y.; Baek, S.-S.; Kim, K.; Cho, K.H. Using convolutional neural network for predicting cyanobacteria concentrations in river water. *Water Res.* **2020**, *186*, 116349. [CrossRef] [PubMed]
43. Shamsudeen, T.Y. Advances in remote sensing technology, machine learning and deep learning for marine oil spill detection, prediction and vulnerability assessment. *Remote Sens.* **2020**, *12*, 3416.
44. Mi, W.; Beibei, G.; Xiaoxiang, L.; Lin, X.; Yufeng, C.; Shuying, J.; Xiao, Z. On-orbit geometric calibration and accuracy verification of GF-6 WFV camera. *Acta Geod. Cartogr. Sin.* **2020**, *49*, 171.
45. Xu, B.; Wang, N.; Chen, T.; Li, M. Empirical evaluation of rectified activations in convolutional network. *arXiv* **2015**, arXiv:1505.00853v2.
46. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
47. Ramachandran, P.; Zoph, B.; Le, Q.V. Searching for activation functions. *arXiv* **2017**, arXiv:1710.05941.
48. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
49. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [CrossRef]

50. Roy, A.G.; Navab, N.; Wachinger, C. Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, 16–20 September 2018; pp. 421–429.
51. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; pp. 234–241.
52. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
53. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]