



Article

Semantic Segmentation and Analysis on Sensitive Parameters of Forest Fire Smoke Using Smoke-Unet and Landsat-8 Imagery

Zewei Wang ^{1,†}, Pengfei Yang ^{1,†}, Haotian Liang ¹, Change Zheng ^{1,*} , Jiyan Yin ² , Ye Tian ¹ and Wenbin Cui ³

¹ School of Technology, Beijing Forestry University, Beijing 100083, China; wangzewei@bjfu.edu.cn (Z.W.); yangpf@sinovel.com (P.Y.); LHTht1998@bjfu.edu.cn (H.L.); tytoemail@bjfu.edu.cn (Y.T.)

² China Fire and Rescue Institute, Beijing 102202, China; jkldora@126.com

³ Ontario Ministry of Northern Development, Mines, Natural Resources and Forestry, Sault Ste Marie, ON 279541, Canada; Wenbin.cui@ontario.ca

* Correspondence: zhengchange@bjfu.edu.cn

† These authors contributed equally to the work.

Abstract: Forest fire is a ubiquitous disaster which has a long-term impact on the local climate as well as the ecological balance and fire products based on remote sensing satellite data have developed rapidly. However, the early forest fire smoke in remote sensing images is small in area and easily confused by clouds and fog, which makes it difficult to be identified. Too many redundant frequency bands and remote sensing index for remote sensing satellite data will have an interference on wildfire smoke detection, resulting in a decline in detection accuracy and detection efficiency for wildfire smoke. To solve these problems, this study analyzed the sensitivity of remote sensing satellite data and remote sensing index used for wildfire detection. First, a high-resolution remote sensing multispectral image dataset of forest fire smoke, containing different years, seasons, regions and land cover, was established. Then Smoke-Unet, a smoke segmentation network model based on an improved Unet combined with the attention mechanism and residual block, was proposed. Furthermore, in order to reduce data redundancy and improve the recognition accuracy of the algorithm, the conclusion was made by experiments that the RGB, SWIR2 and AOD bands are sensitive to smoke recognition in Landsat-8 images. The experimental results show that the smoke pixel accuracy rate using the proposed Smoke-Unet is 3.1% higher than that of Unet, which could effectively segment the smoke pixels in remote sensing images. This proposed method under the RGB, SWIR2 and AOD bands can help to segment smoke by using high-sensitivity band and remote sensing index and makes an early alarm of forest fire smoke.

Keywords: forest fire; remote sensing; smoke segmentation; Smoke-Unet; attention mechanism; residual block; Landsat-8; band sensibility



Citation: 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.* **2022**, *14*, 45. <https://doi.org/10.3390/rs14010045>

Academic Editors: Fahimeh Farahnakian, Jukka Heikkonen and Pouya Jafarzadeh

Received: 4 November 2021

Accepted: 20 December 2021

Published: 23 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The forest system, which occupied almost one third of the total land area, provides a variety of critical ecological services such as natural habitat, water conservation, timber products and maintaining biodiversity [1]. It also plays a central role in global carbon circle and energy balance [2,3]. However, the areas of global forests sharply declined at a rate of roughly 10 million hectares per year [4]. Wildfire is the principal threat in terrestrial ecosystems, and many evidences have proved that recent global warming and precipitation anomalies have made forests more susceptible to burning [5,6]. In the period of 2019–2020, the Amazon and South Australia faced the most severe wildfires, and these events have caused wide public concerns because of their considerable ecological and socioeconomic consequences such as consuming generous quantities of tropical rainforest, emitting great volumes of greenhouse gas and aerosols and altering the composition of the atmosphere.

Because smoke appeared at the earliest phase in wildfires, earlier detection and rapid identification of initial wildfire smoke are crucial for wildfire suppression and management

to avoid the damages and negative impacts of wildfires [7]. Wildfire smoke is usually identified by means of manual observation, patrol of forest rangers, infrared and optical sensors of fire lookout towers and aviation monitoring. However, these techniques have shown ineffective, unsystematic, and geographical limit. Wildfires, caused by natural events (e.g., lightening and spontaneous combustion) or human-forcing activities, occurred in the remote regions, making it difficult and cost-consuming for accessibility and suppression. However, data from remote sensing satellites can provide continuous, frequent, and numerous systematic information with various spatial and temporal resolution at global scales, which may overcome several limitations of the conventional wildfire smoke observation methods [8].

Currently, the widely used remote sensing monitoring algorithms are mostly based on satellite remote sensing data of low and medium resolution (>250 m) [9,10], such as Advanced Very High Resolution Radiometer (AVHRR) [11–13], Moderate Resolution Imaging Spectroradiometer (MODIS) [14–16], etc., which has become an important business method to detect wildfire smoke for daily wildfire disaster monitoring in many countries around the world. However, the satellites with lower spatial resolution are unable to capture relevant information effectively at the early stage of forest fires due to too small initial burning area, and thus would cause the detection of early fire spots to be missed. Therefore, high-resolution satellite data are urgently needed to improve the accuracy of fire detection. Landsat-8 data can be publicly obtained and the resolution has increased by an order of magnitude, reaching 30 m, compared with Suomi National Polar-orbiting Partnership (S-NPP) and Visible Infrared Imaging Radiometer Suite (VIIRS) [17–20]. In addition, Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) mounted on Landsat-8 can provide a new data source and capability allowing as small as 1 m^2 active fire to be observed [21]. Therefore, Landsat-8 data were used for wildfire smoke detection in this paper.

The satellite can carry many multispectral sensors and provide large amounts of multispectral data with more valuable information than RGB. Wildfire smoke presents different characteristics in different spectral ranges of remote sensing data and the choice of bands is crucial to smoke recognition. The wildfire smoke detection algorithms [22,23] of AVHRR mainly derived from band 3 (centered at $3.7\text{ }\mu\text{m}$), band 4 (centered at $10.8\text{ }\mu\text{m}$) and band 5 (centered at $12\text{ }\mu\text{m}$). The family of products [24,25] based on MODIS sensors primarily used two MIR bands (band 21 and band 22, centered at $3.96\text{ }\mu\text{m}$) and TIR band 31 (centered at $11\text{ }\mu\text{m}$). Data from band 4 (centered at $3.55\text{--}3.93\text{ }\mu\text{m}$) and band 5 (centered at $10.5\text{--}12.4\text{ }\mu\text{m}$) of VIIRS are used for tracking active fires [26–28]. Nevertheless the Landsat-8 wildfire smoke detection algorithm was based on the reflectance of band 7 (SWIR, centered at $2.2\text{ }\mu\text{m}$), that is sensitive to thermal abnormality [29]. Therefore, the selection of the spectral range of remote sensing data is very important for smoke identification based on different spectral properties.

Due to the development of machine learning and data mining, several studies focused on the automatic retrieving smoke pixels. Li et al. [30] facilitated a neural network algorithm using AVHRR data to search smoke plumes but it failed when smoke pervades in the downwind area. As a powerful and popular machine learning approach, Support Vector Machine (SVM) is widely used in remote sensing task. The SVM classifiers can take advantage of combination of texture, color and other features of the remote sensing scene, and successfully distinguish the pixels contained smoke from non-smoke pixels [31–33]. Other machine learning techniques, such as K-means clustering, fisher linear classification [34] and BPNN algorithm [35], were used to discriminate smoke pixels. Nevertheless, it is still a challenge to extract smoke areas because of the wide range of shapes, color, texture, luminance and heterogeneous component of aerosol as well as diversity of cover types. In addition, with the development of remote sensing technology, a dramatically increasing satellites archive makes it no longer suitable for hand-crafted features of remote sensing data, and it is urgent to develop more automatic detection algorithms.

Deep learning, in the specific area of Convolutional Neural Networks (CNNs), is inspired by the working way of the human brain and recently has acquired many impressive achievements in many scientific fields such as image classification, object detection, and image segmentation. CNN can automatically extract features from data using a structure of multilayers. They are iteratively learning by forward propagation and backward derivation and updating parameters of kernels through complex nonlinear functions. The accuracies can be further improved by providing great amounts of input data, so it would be the best candidate for remote automated detection tasks. CNNs have successfully been employed in variety remote sensing fields such as road detection [36], cloud detection [37] and smoke classification [38]. Recent Unet-based methods [39] have also made good progress in the field of remote sensing [40,41]. However, remote sensing satellite data have many redundant bands so that too much information causes the wildfire smoke detection accuracy drop after the first rise and the detection efficiency decrease. How to reduce the interference of redundant information and make full use of the correlation of feature channels is a key problem on wildfire smoke detection based on remote sensing data.

The objective of this study was to propose a wildfire smoke detection algorithm of Landsat-8 satellite remote sensing imagery at the scene of a wildfire using multispectral data. First, a multispectral smoke dataset of Landsat-8 satellite at global scale, including the information from visible to TIRS1 infrared bands, was built in this paper. Second, a deep learning model, Smoke-Unet, based on Unet architecture incorporating with residual block [42] and attention mechanism [43], was proposed. Then, the performance of this algorithm on different region and various scale of wildfire smoke was evaluated by the experiments based on the abovementioned multispectral smoke dataset. Finally, to better extract the features of remote sensing smoke and reduce the redundancy of remote sensing data, the sensitivity of multiple bands was analyzed.

The main parts of this paper are structured as follows. Section 2 introduces the establishment of a multispectral smoke dataset of Landsat-8 satellite at a global scale, and a proposed deep learning model, Smoke-Unet, based on the Unet architecture incorporating with Attention mechanism and residual block, is presented in Section 3. To reduce the disturbance of the redundant information, the influence of different band combinations of multispectral data and remote sensing parameters on the accuracy of the algorithm are analyzed and the band sensitivity are evaluated in Section 4, and the conclusion is made in Section 5.

2. Data

2.1. Landsat-8 Multispectral Data

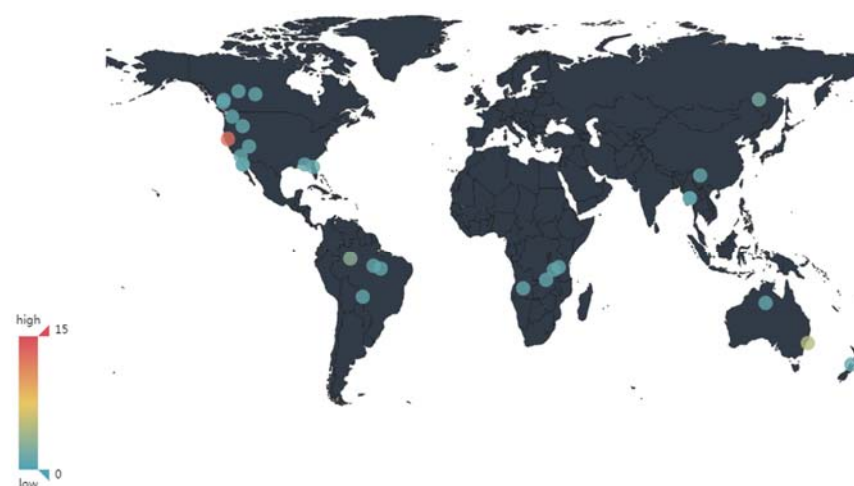
Landsat-8, carrying the OLI and the TIRS, was launched in 2013, and is operated by the US Geological Survey (USGS). As seen in Table 1, OLI is a nine-spectral-band push-broom sensor with spatial resolution of 30 m and 15 m for the panchromatic band, including near-infrared band (NIR) and Panchromatic (Pan). Standard terrain-corrected data (Level 1T) from OLI were used in this study.

2.2. Study Area

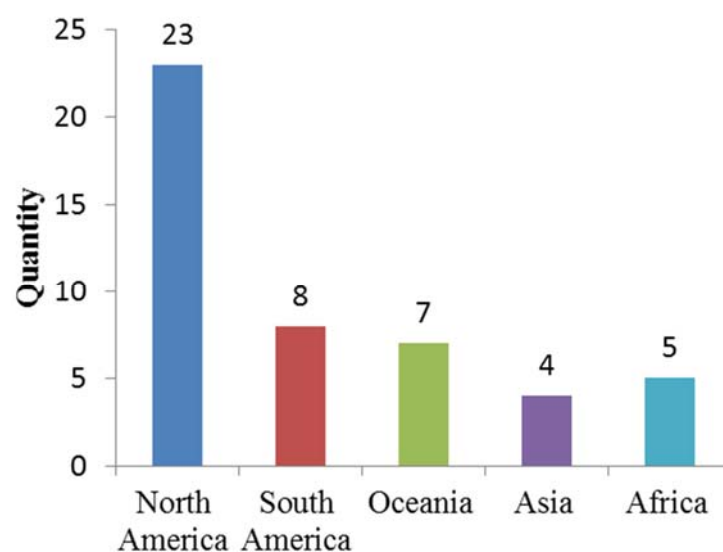
As shown in Figure 1, the various fire-prone ecosystems all over the world were selected as the study areas in this research, containing: (i) needleleaf trees of boreal forests in high latitude regions, such as Canada and Siberia; (ii) subtropical evergreen hard-leaved forest mixed conifer-broadleaf forests in Western America; (iii) dry sclerophyll woodland and open forest in Eastern Australia; (iv) tropical rainforest in the Amazon and Southeastern Asia; (v) tropical grasslands and savannas in Africa.

Table 1. Landsat-8 Satellite Parameters.

Payload Name	Band Number	Band Name	Spectral Range(nm)	Resolution(m)
OLI	1	Coastal	433~453	30
	2	Blue	450~515	30
	3	Green	525~600	30
	4	Red	630~680	30
	5	NIR	845~885	30
	6	SWIR1	560~660	30
	7	SWIR2	100~300	30
	8	Panchromatic	500~680	15
	9	Cirrus	1360~1390	30
TIRS	10	TIRS1	1060~1119	60
	11	TIRS2	1150~1251	60

**Figure 1.** Spatial distribution of study regions in the datasets.

As seen in Figure 2, the study areas are located in Asia, North America, South America, Africa, etc. Considering that the frequent occurrence of wildfires in these areas is representative, the fire-prone regions in the USA, Canada, Brazil and Australia were selected as the primary research areas.

**Figure 2.** Different intercontinental data distribution.

As seen in Figure 3, the land cover data have 4 types, including ocean, city, bare soil and different kinds of vegetation (agricultural land, grassland, forest.)

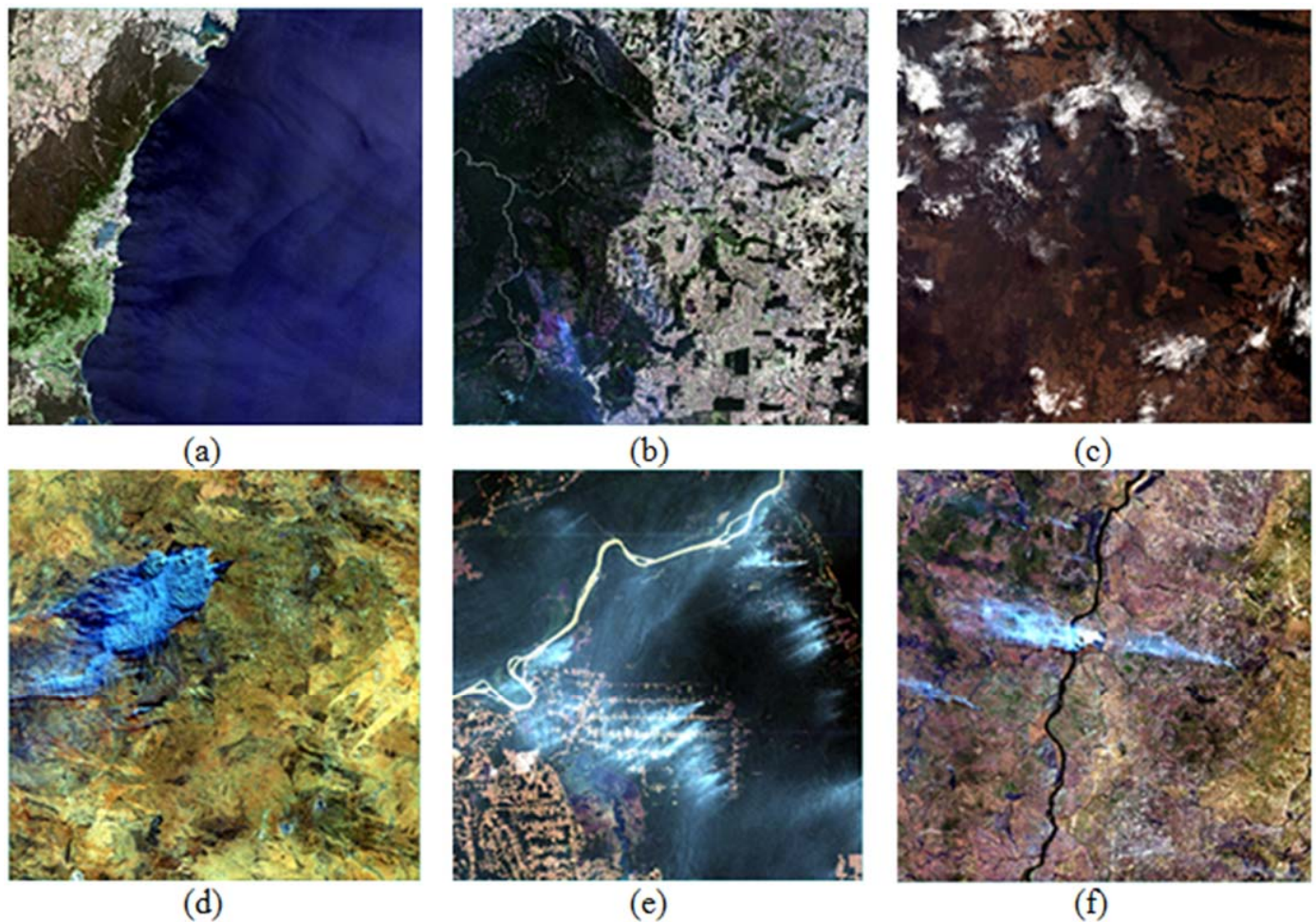


Figure 3. Different land cover types of datasets. (a) Ocean; (b) City; (c) Bare soil; (d) Agricultural land; (e) Grassland; (f) Forest. Different intercontinental data distribution.

2.3. Fire Seasons

Forest fires usually occur in the early stages of springs, autumns and winters due to the influence of climate. As a result of human activity, the wildfire occurrence in summers is dramatically increasing in North America and the Amazon [44,45]. In this study, the period of fire occurrence covered from 2013 to 2019, including different fire seasons, as shown in Figure 4.

2.4. Proportion of Smoke Pixel

Smoke concentration and the proportion of smoke pixels in one image are different with forest fire stage. At the beginning of fire, thin scattered smoke pixels account for a small amount in the image; however, in the middle stage of fire, the entire image is nearly occupied by densely spread smoke. The proportion distribution of smoke pixels is shown in Figure 5.

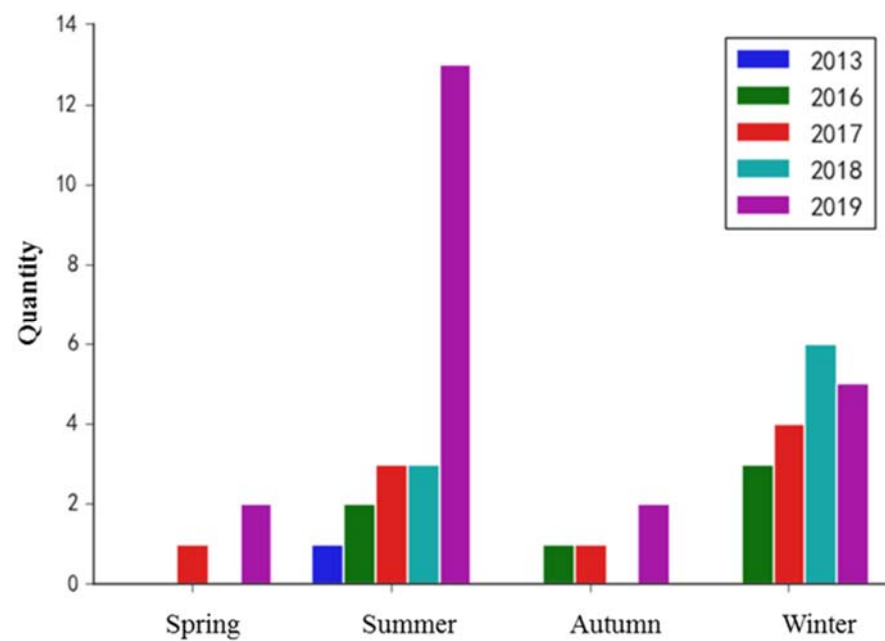


Figure 4. Period of fire occurrence.

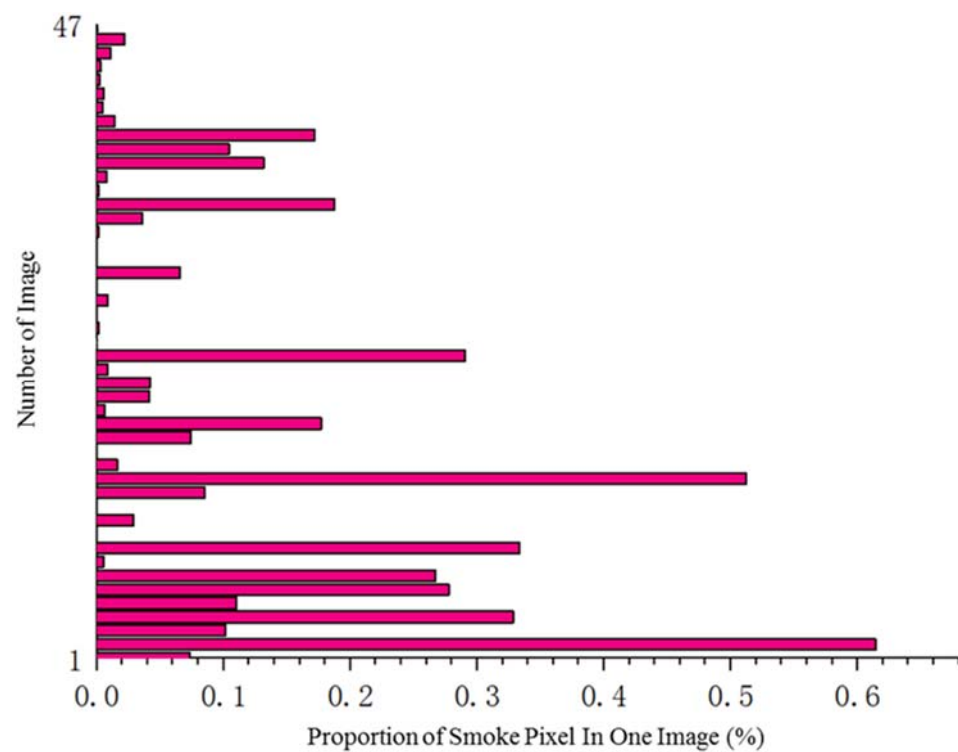


Figure 5. The proportion of smoke pixels of different images.

2.5. Training and Validation Dataset

To reduce overfitting, data augmentation was performed, including random cropping, vertical and horizontal mirroring operations on the images. As a result, the dataset in this study contains a total of 47 multispectral forest fire smoke images, composed of RGB, NIR, SWIR and mid-infrared bands. Thirty-four images are randomly selected as training data, 5 images are used as verification data, and 8 images are used as test data.

3. Methods

As a dense prediction problem, the task of smoke classification in satellite image is to make a prediction at each pixel. Based on the Unet network structure, Smoke-Unet, fused into residual blocks and attention model, was put forward to segment smoke in satellite images in this paper.

As seen in Figure 6, Smoke-Unet consists of a contraction path on the left side and an expansive path on the right side. The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3×3 convolutions (padded convolutions), each followed by a linear unit (ELU) and a 2×2 max pooling operation with stride 1 for downsampling. At each downsampling step, we double the number of feature channels. Every step in the expansive path consists of an upsampling of the feature map followed by a 2×2 convolution (“up-convolution”) that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3×3 convolutions, each followed by a ELU. The cropping is necessary due to the loss of border pixels in every convolution. Because the resolution of the remote sensing image is smaller (one pixel for Landsat with a resolution of 30 m), downsampling will have a catastrophic effect on these local small target features, resulting in the problem of vanishing gradients for many network layers. Therefore, Smoke-Unet is designed to only downsample three times. The steps of convolution and downsampling are alternately performed three times to obtain a high-dimensional feature map and then the spatial resolution is restored through the three-time symmetrical convolution and upsampling operations. The feature map with the same resolution was fused through a skip connection to compensate for the loss of detail caused by downsampling.

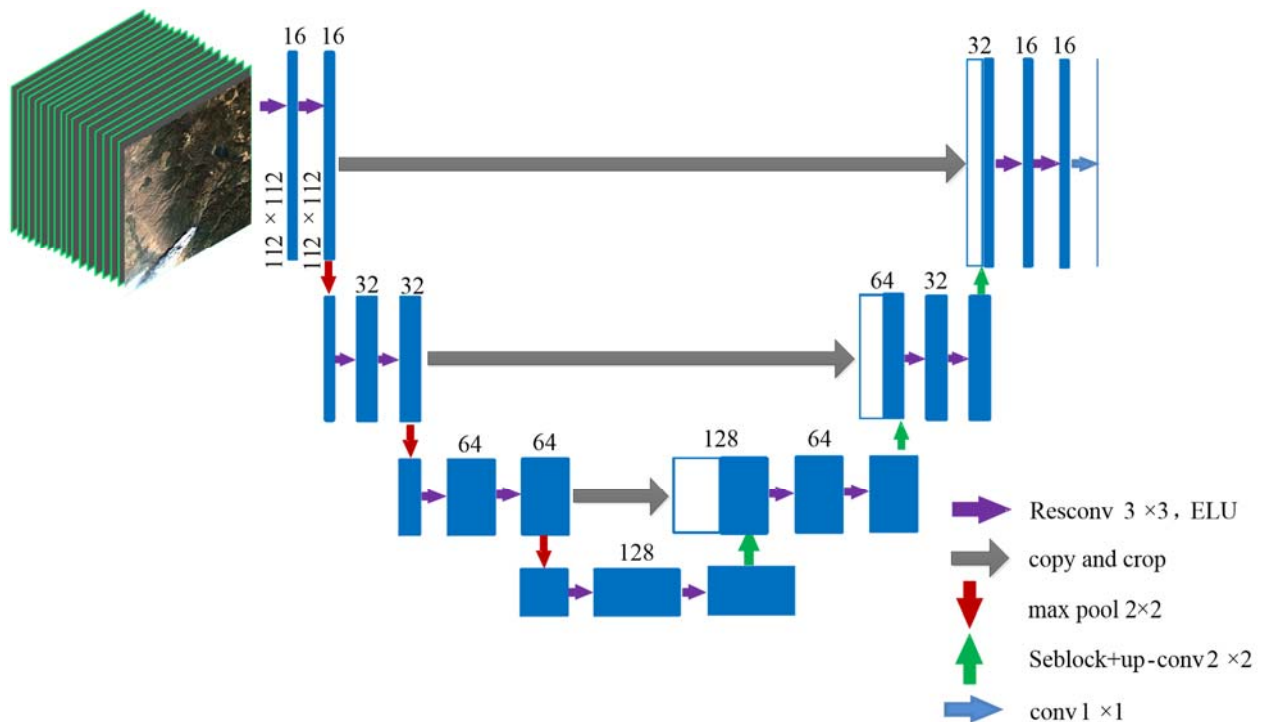


Figure 6. Smoke-Unet.

In order to improve the feature learning ability of the network, ResBlock, a residual block is added to the convolution block to enhance the feature extraction ability. The residual block with skip connection structure can enhance the robustness of the network and improve the performance of the network. The skips structure between layers can fuse coarse semantic and local appearance information. This skip feature is learned end-to-end to improve the semantics and spatial precision for the output. Remote sensors onboard

satellite have so many spectral channels that too much irrelevant information leads to difficulty in extracting feature. In order to emphasize effective information and reduce the interference of invalid band information, the SEBlock module based on the attention mechanism is added to the Smoke-Unet network structure. In the attention model, the focus process can be imitated by setting the weight coefficient. The key attention areas can be set with larger weight coefficients, which represent the importance of the information in these areas, while other areas can be set with smaller coefficients to filter invalid information. Through considering different degree of importance for information, the efficiency and accuracy of information processing can be greatly improved. At the final layer, a 1×1 convolution is used to map each 16-component feature vector to final smoke class. In total, the network has 15 convolutional layers.

4. Results and Discussion

In this section, three kinds of semantic segmentation experiments were made on our dataset. By comparing the experimental results, the performance of Smoke-Unet was evaluated and the sensitivity of band and remote sensing parameters was analyzed.

4.1. Experimental Environment

The network structure uses the Keras architecture and several related image processing libraries, the programming language uses Python 3.5. The specific configuration is shown in Table 2.

Table 2. Deep learning environment configuration.

Programming Environment	Auxiliary Library	Hardware Configuration	Other Software
Python3.5	Shapely	CPU:InterE5-2620v3@2.4 GHz	
Tensorflow1.9	Opencv2.2	GPU:NVIDIA TITAN X	ENVI5.3
CUDA8.0	Tifffile0.12	RAM:16 GB	ArcGIS10.3
cuDNN10.0	Rasterio1.1.2	Numba0.26.0	Scikit_image0.12.3
Keras2.2.0	h5py2.6.0		

4.2. Implementation Details

The input of the Smoke-Unet network is the multichannel remote sensing image and the index of the multi-remote sensing feature. The data have 13 channels, as shown in Table 3. The schematic diagram of the network is shown in Figure 6.

Table 3. Bands and remote index.

Number	Data Type	Item	Band
1	Band Data	Multispectral Band	1–7, 10
2	Band Data	Panchromatic Band	8
3	Remote Sensing Index	EVI	/
4	Remote Sensing Index	NBR	/
5	Remote Sensing Index	AOD	/
6	Remote Sensing Index	BT	/

During the model training, the back-propagation optimization algorithm uses the stochastic gradient descent (SGD) algorithm, the learning rate is 1×10^{-3} , the momentum is 0.9, the learning rate attenuation is 0.1, the loss function is the joint loss function, and the evaluation function is Jaccard similarity function. The batch size is 128. Considering the computing resources, there are 25 iterations in total, and shuffle is used to disrupt the order of training samples in each epoch. After each round of iteration is completed, the Jaccard coefficient, Accuracy, F1 and other indicators of the training set and the validation set are calculated.

4.3. Implementation Details

In the field of deep learning image segmentation, the similarity coefficient is an important indicator to measure the accuracy of image segmentation. Jaccard similarity coefficient is used in this paper to evaluate the similarity and difference between image targets. The larger the value of Jaccard, the more similar the two targets. For two sets A and B, the Jaccard coefficient is the ratio of the intersection and the union of the two, defined as:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \quad (1)$$

$$0 \leq J(A, B) \leq 1, \quad (2)$$

4.4. Ablation and Comparative Analysis

In order to verify the role of residual block and attention mechanism of Smoke-Unet, the ablation experiments were made in wildfire smoke segmentation based on remote sensing satellite images. As shown in Table 4, Res-Unet means the network combined Unet with the residual module. Atten-Res-Unet means the network integrated the attention mechanism module with Res-Unet. The results of semantic segmentation were evaluated by metrics such as Jaccard, Accuracy, Recall and F1. In order to validate the effectivity more extensively, other common semantic segmentation networks such as FCN [46], Segnet [47] and PSPnet [48] have been compared. The results are compared in Table 4 and Figure 7.

Table 4. Ablation and comparative analysis of different models.

Network	Dataset	Loss	Jaccard	Accuracy	Recall	F1
Unet	Train	0.844	0.657	0.801	0.753	0.773
	Validation	1.889	0.699	0.694	0.781	0.735
Res-Unet	Train	0.851	0.690	0.805	0.829	0.813
	Validation	1.636	0.59	0.701	0.944	0.805
Atten-Res-Unet	Train	1.514	0.703	0.835	0.816	0.823
	Validation	1.926	0.654	0.696	0.894	0.782
FCN	Train	1.479	0.735	0.845	0.852	0.844
	Validation	1.974	0.58	0.711	0.811	0.758
Segnet	Train	1.532	0.712	0.831	0.835	0.828
	Validation	1.708	0.665	0.761	0.841	0.799
PSPnet	Train	1.406	0.748	0.845	0.871	0.851
	Validation	1.901	0.581	0.751	0.812	0.765
Smoke-Unet	Train	0.759	0.752	0.923	0.917	0.918
	Validation	1.134	0.644	0.725	0.838	0.775

It can be seen from Table 4 that Jaccard coefficient, accuracy, recall rate, F1 and other indicators of Smoke-Unet have been improved to varying degrees. Compared with the original Unet network architecture, the Jaccard coefficient on the training set is increased by 14.46% and the Jaccard coefficient on the verification set is reduced to a certain extent. The accuracy on the training set is increased by 15.23% and the accuracy on the validation set is increased by 4.47%. The recall rate on the training set was increased by 21.78% and the recall rate on the verification set was increased by 7.30%. F1 on the training set is increased by 18.76% and F1 on the validation set is increased by 5.44%. It can be concluded that the proposed network performs better than the original Unet network, and it can be seen from Table 4 that Smoke-Unet is better than other common semantic segmentation networks. The specific segmentation image is shown in Figure 7.

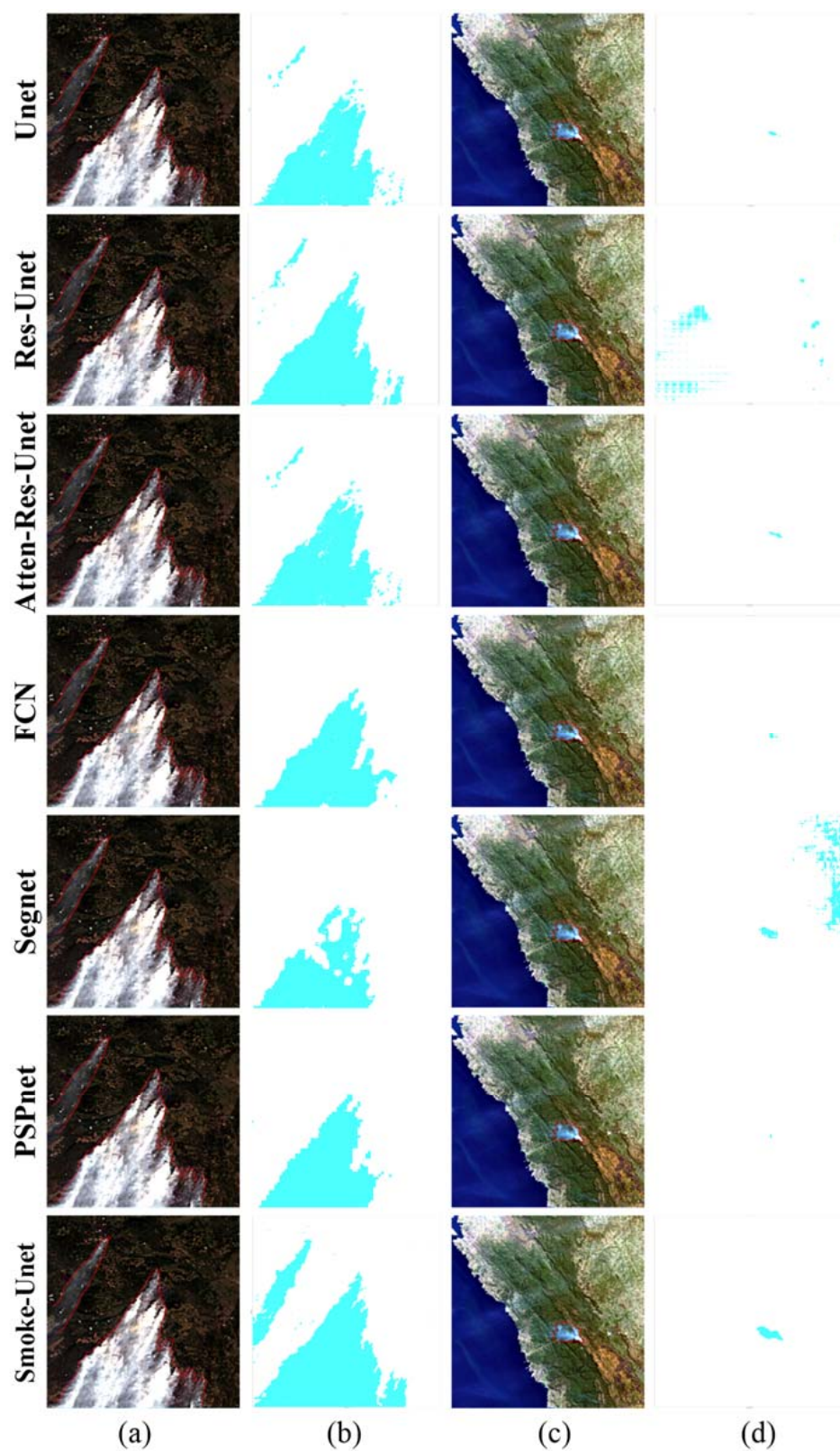


Figure 7. The results of segmentation of different networks. (a) Image acquired over British Columbia, Canada, on 4 August 2017, the smoke is depicted in red line area; (b) The segmentation results of smoke over British Columbia, the smoke pixels are depicted in aqua color; (c) Image acquired over New Zealand area, on 7 Feb 2019, the smoke is depicted in red line area; (d) The segmentation results of smoke over New Zealand area, the smoke pixels are depicted in aqua color.

In Figure 7a, the smoke contains a wide range of dense smoke and scattered diffuse thin smoke, and the land cover includes vegetation, bare soil, and some cirrus clouds. In Figure 7c, the smoke, located near the fire point, is thin and has a relatively small range, and the land cover includes sea water, seashore, bare land, vegetation and so on.

It can be seen from Figure 7b,d that the Unet network can roughly segment the smoke pixels in different images. In Figure 7b, Res-Unet can effectively segment the smoke pixels, because the number of smoke pixels in the diffusion area at the upper left of Figure 7b has increased, while in Figure 7d there is an over-segmentation by Res-Unet, and some pixels are incorrectly segmented as the smoke pixel. In Figure 7b, Atten-Res-Unet can effectively segment the smoke pixels, as the number of smoke pixels in the diffusion area at the upper left of Figure 7b has increased, while the under-segmentation exists in Figure 7d, resulting that some pixels are not identified. The segmentation effects using FCN, SegNet and PSPnet are worse than Unet-based methods. It can be seen from Figure 7b,d that the Smoke-Unet network has a better recognition performance than the other networks when segmenting a wide range of dense smoke and a small area of thin smoke.

4.5. Sensitivity Analysis

With the increasing number of high-resolution images and dimensional channels of data, the information redundancy generated by high-dimensionality makes it difficult to effectively utilize the rich information of remote sensing images. Based on the above-mentioned forest fire smoke detection algorithm, this section will analyze and discuss the influence of different band combinations of multispectral data and remote sensing parameters on the accuracy of the algorithm.

4.5.1. Sensitivity of Bands

In order to evaluate the band sensitivity, the segmentation experiments based on different band combination were made on our dataset. The data source distribution is shown in Table 5. The test images contain a large proportion of smoke, small proportion of smoke, the land cover includes bare land, vegetation, seashores and highly reflective ground.

Table 5. Details of different bands combination.

Number	Data Type	Data Dimension	Band
1	RGB	3	Band 2~4
2	RGB + NIR	4	Band 2~5
3	RGB + TIRS1	4	Band 2~4,10
4	RGB + SWIR2	4	Band 2~4,7
5	RGB + SWIR1 + SWIR2	5	Band 2~4,6,7
6	RGB + SWIR1 + NIR	5	Band 2~6
7	RGB + TIRS1 + SWIR2	5	Band 2~4,7,10
8	TIRS1	1	Band 10
9	NIR + SWIR1/2 + TIRS1	4	Band 5~7,10
10	SWIR1 + NIR + Blue	3	Band 2,5,6
11	Multiple	8	Band 1~7, Band 10
12	Multiple + Pan	9	Band 1~7, Band 10~11
13	All data	11	Band 1~11

From Table 6, Figures 8 and 9, it can be found that the segmentation result of smoke is the best when the input band is RGB and SWIR2. Compared to all the data bands as the input, Jaccard with the input of RGB and SWIR2 increases by 6.5%. When the input is all data source, it can effectively segment a wide range of smoke. However, compared with the segmentation result of the RGB data source, the smoke pixel with the input of all band data has the problem of under-segmentation for a small area of smoke, especially in the downwind diffusion area. It shows that too much data will interfere with the network parameter learning and degrade the performance of the network.

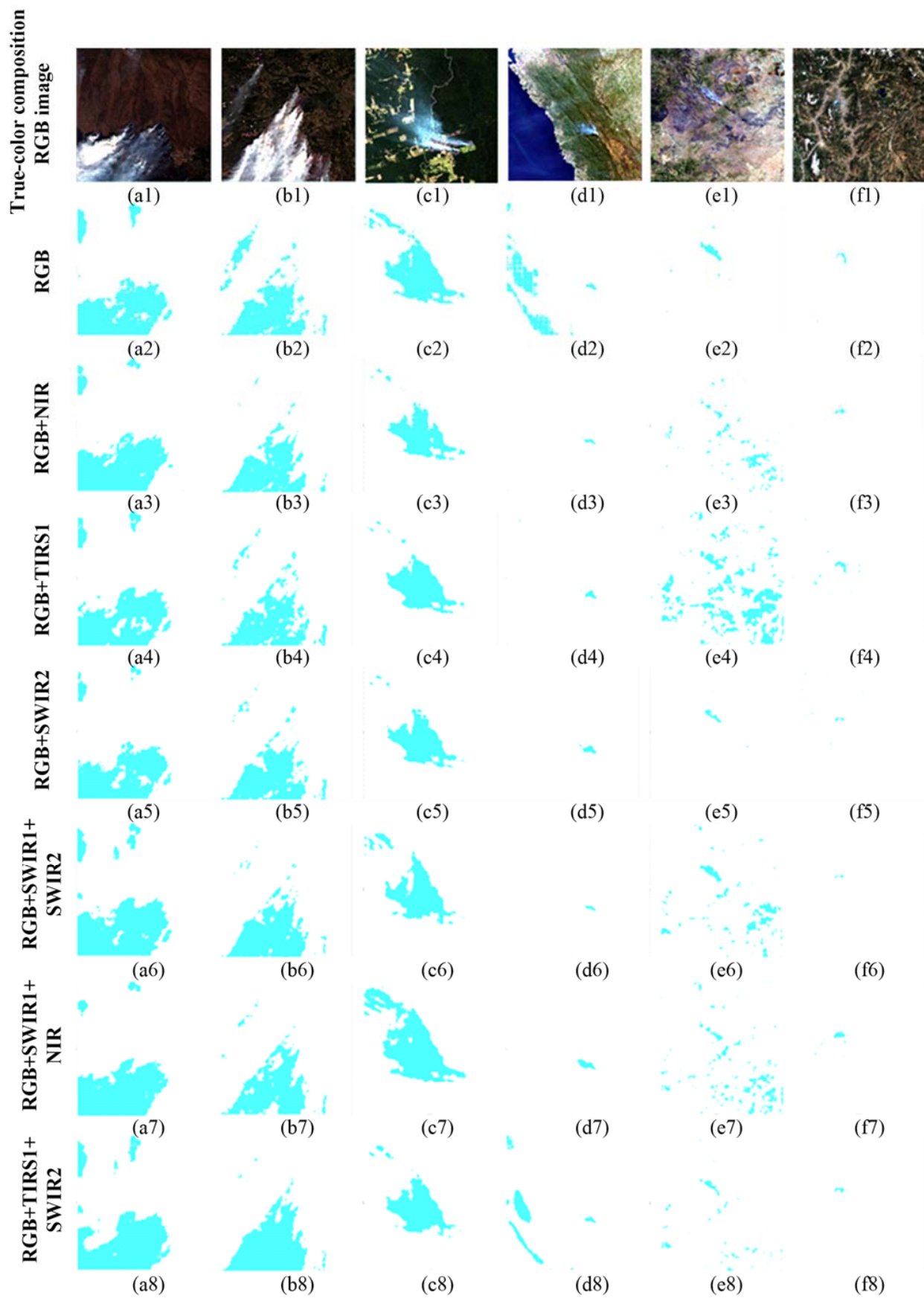


Figure 8. Cont.

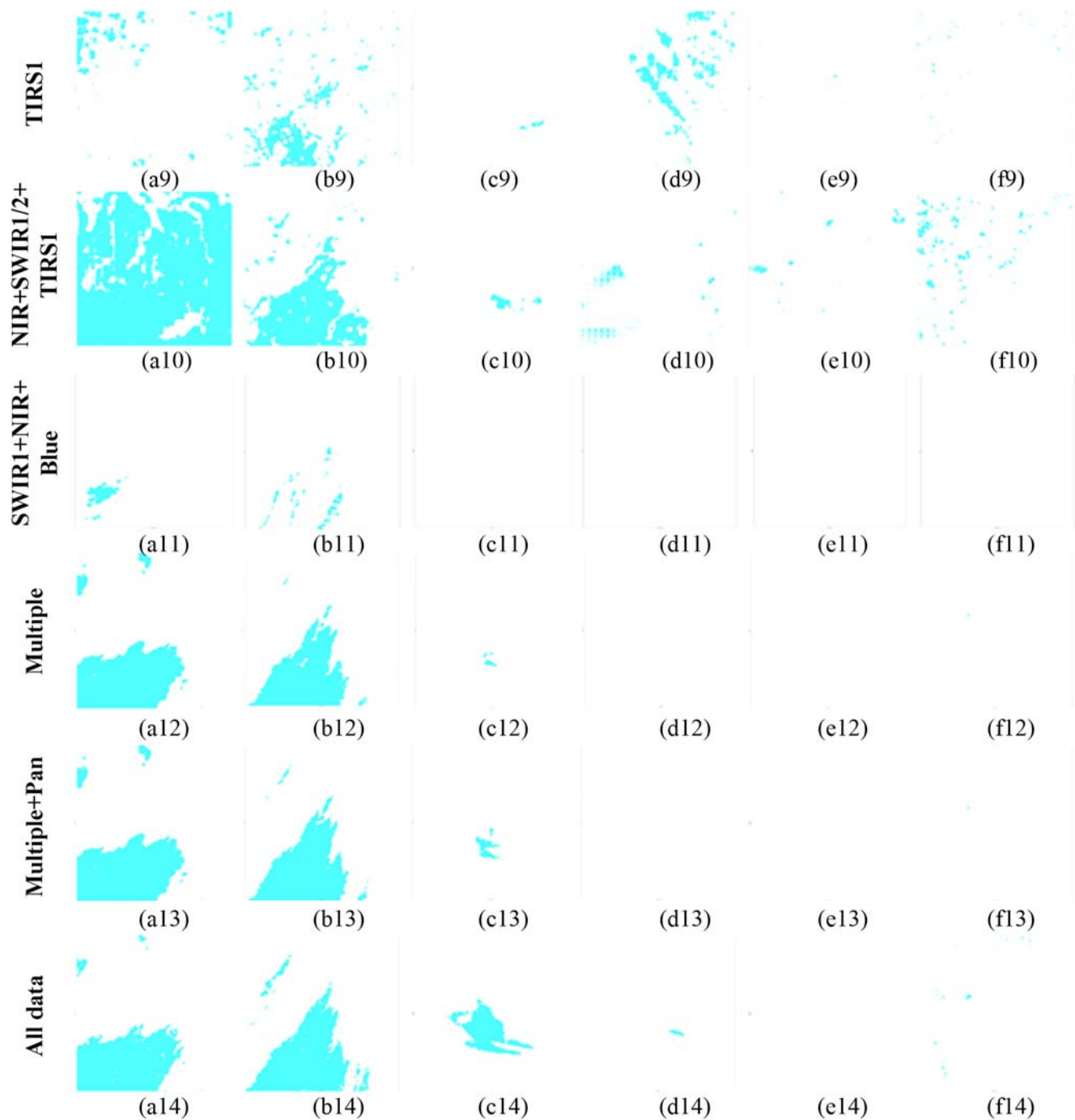


Figure 8. The first line shows true-color composition RGB images of smoke plumes. **(a1–a14)** Siberia area, Russia, on 17 March 2018; **(b1–b14)** British Columbia, Canada, on 4 August 2017; **(c1–c14)** Amazon region, Brazil, on 9 August 2019; **(d1–d14)** New Zealand area, on 7 Feb 2019; **(e1–e14)** Zambia, on 26 June 2017; **(f1–f14)** Liangshan region, China, on 21 May 2019. All rows except the first are segmentation results of smoke with different input data, the smoke pixels are depicted in aqua color.

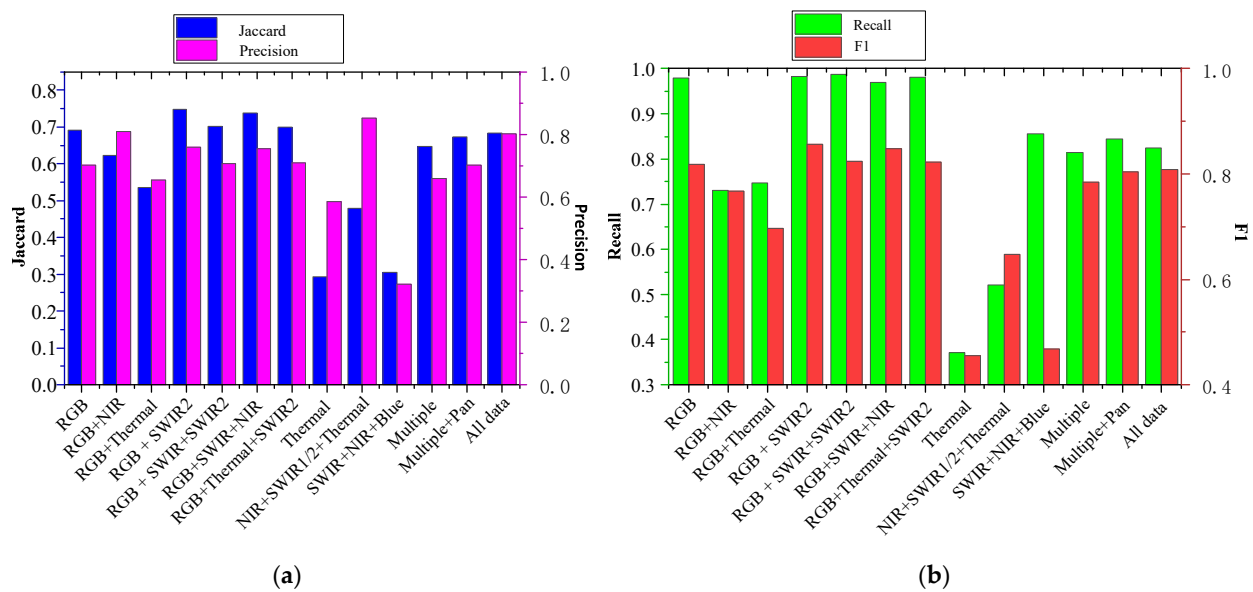


Figure 9. The segmentation results of smoke with variety bands combination. (a) The result of Jaccard and Accuracy; (b) The result of recall and F1.

Table 6. The segmentation results of different bands combination.

Number	Data Type	Jaccard	Accuracy	Recall	F1
1	RGB	0.692	0.701	0.980	0.818
2	RGB + NIR	0.623	0.809	0.730	0.767
3	RGB + TIRS1	0.535	0.653	0.747	0.697
4	RGB + SWIR2	0.748	0.759	0.982	0.856
5	RGB + SWIR1 + SWIR2	0.701	0.707	0.988	0.824
6	RGB + SWIR1 + NIR	0.737	0.753	0.970	0.848
7	RGB + TIRS1 + SWIR2	0.700	0.709	0.981	0.823
8	TIRS1	0.294	0.585	0.371	0.455
9	NIR + SWIR1/2 + TIRS1	0.479	0.852	0.522	0.648
10	SWIR1 + NIR + Blue	0.305	0.322	0.855	0.468
11	Multiple	0.646	0.658	0.814	0.784
12	Multiple + Pan	0.673	0.701	0.844	0.804
13	All data	0.683	0.801	0.825	0.809

In order to better distinguish smoke from clouds, the spectral characteristics of smoke and cloud in different bands were compared. As shown in Figure 10, the image contains smoke (heavy smoke numbered 2; smoke near the fire point numbered 5; thin smoke in the diffusion area numbered 3 and 4) and clouds (numbered 1). To highlight the features, the logarithmic transformation was made to the image. The spectral characteristics of different objects in each band of the multispectrum are shown in Figure 11.

It can be seen from Figure 11a,b that clouds and dense smoke have very similar spectral characteristics in the RGB band (Band 3~5); therefore, it is difficult to distinguish dense smoke with clouds by the naked eye. However, the pixel values of the two are quite different in the SWIR2 band (Band 8), which may be the reason why the smoke pixels can be better distinguished by using RGB and SWIR2. From Figure 11b,c, it shows that the spectral characteristics of heavy smoke and thin smoke are greatly different, which makes the task of smoke recognition challenging.

4.5.2. Sensitivity of Remote Sensing Parameters

In order to evaluate the sensitivity of different remote sensing feature indexes to forest fire smoke, EVI, NBR, BT and AOD were respectively combined with RGB and SWIR2 as shown in Table 7 to evaluate the impact on the smoke segmentation.

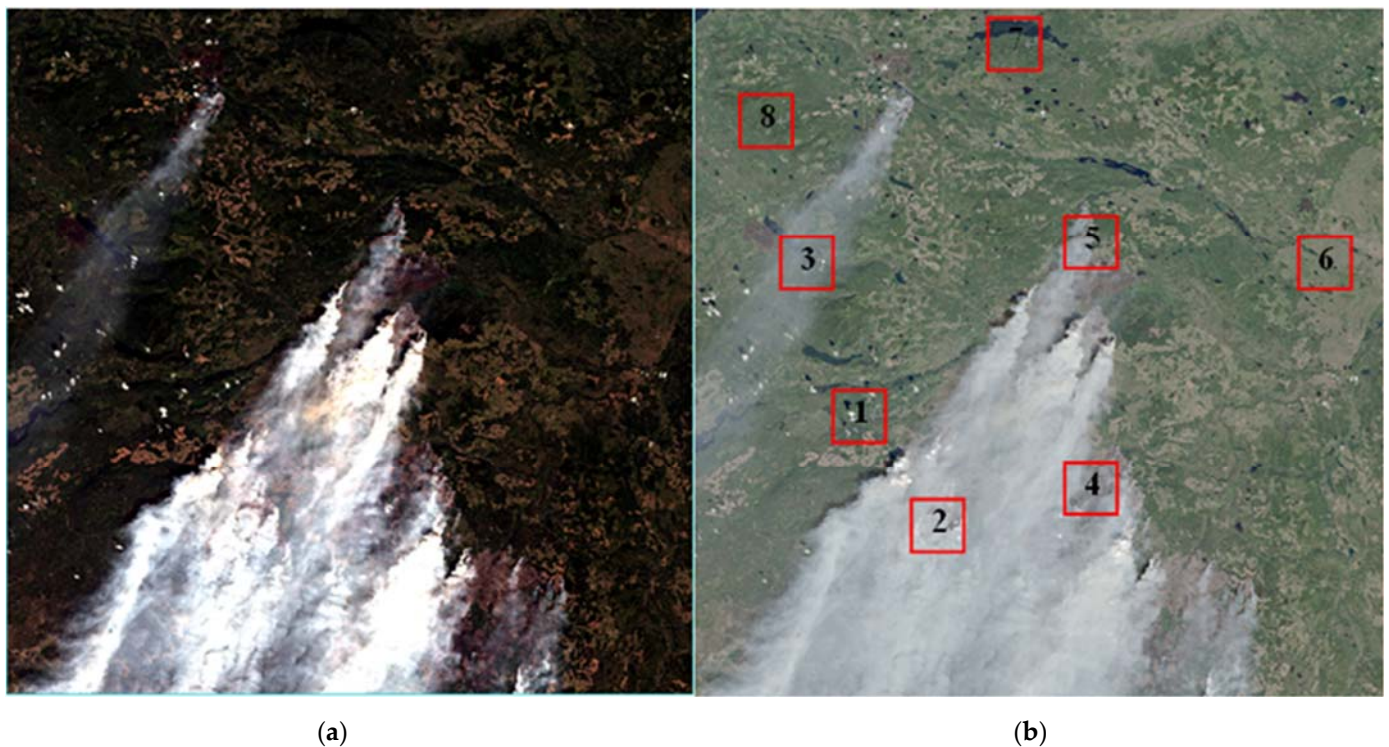


Figure 10. The image of smoke acquired over British Columbia, Canada, on 4 August 2017. (a) The true-color composition image. (b) The image of smoke after logarithmic transformed. Different targets are marked with numbers 1 through 8. (1) The cloud; (2) The heavy smoke; (3) The thin smoke over area 3; (4) The thin smoke over area 4; (5) The smoke over the hot spot; (6) The soil; (7) The water; (8) The vegetation.

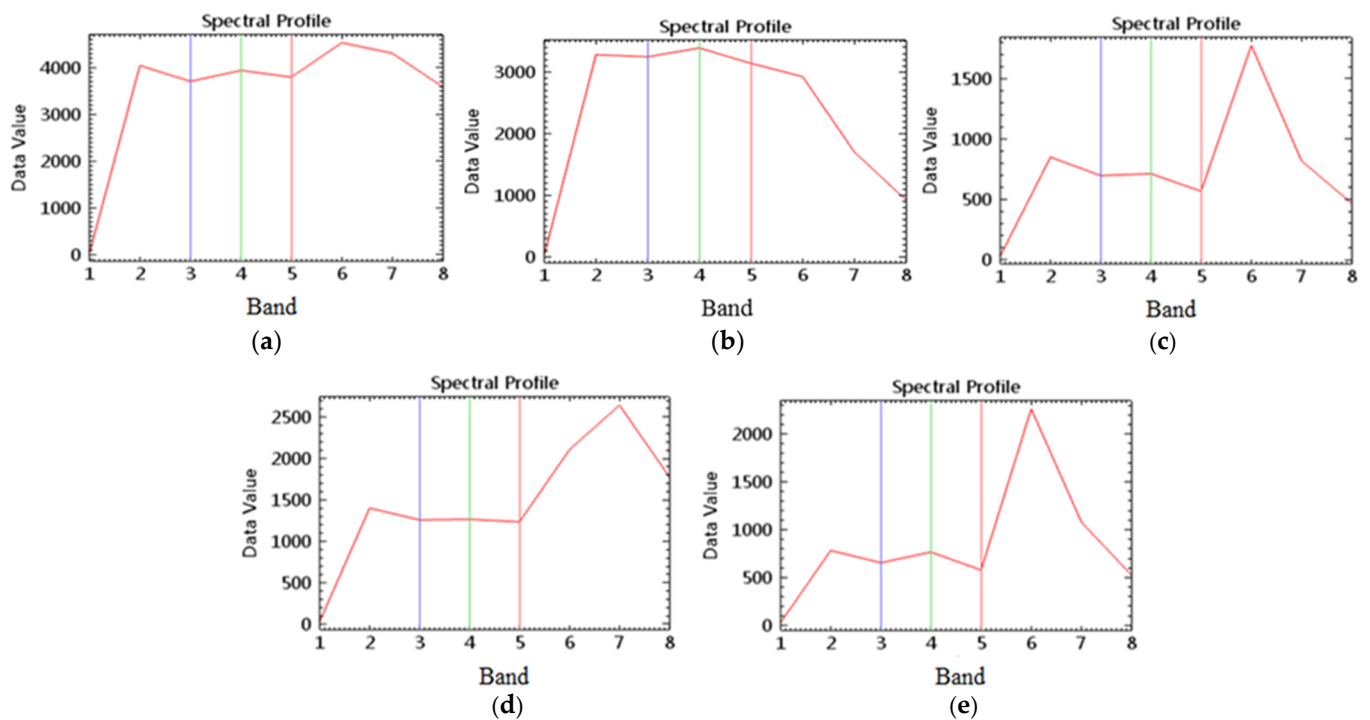


Figure 11. The spectral profile of different objects. (a) The profile of cloud on area 1; (b) The profile of heavy smoke on area 2; (c) The profile of thin smoke over the area 3; (d) The profile of thin smoke over the area 4; (e) The profile of smoke over the hot spot (the fire point) on area 5.

Table 7. Fusion of different remote sensing features.

Number	Data Type	Data Dimension
1	RGB + SWIR2 + EVI	5
2	RGB + SWIR2 + NBR	5
3	RGB + SWIR2 + BT	5
4	RGB + SWIR2 + AOD	5

As shown in Figure 12, both EVI and NBR do not contribute to forest fire smoke segmentation and BT help to identify high temperature abnormal points, resulting in under-segmentation of smoke pixels.

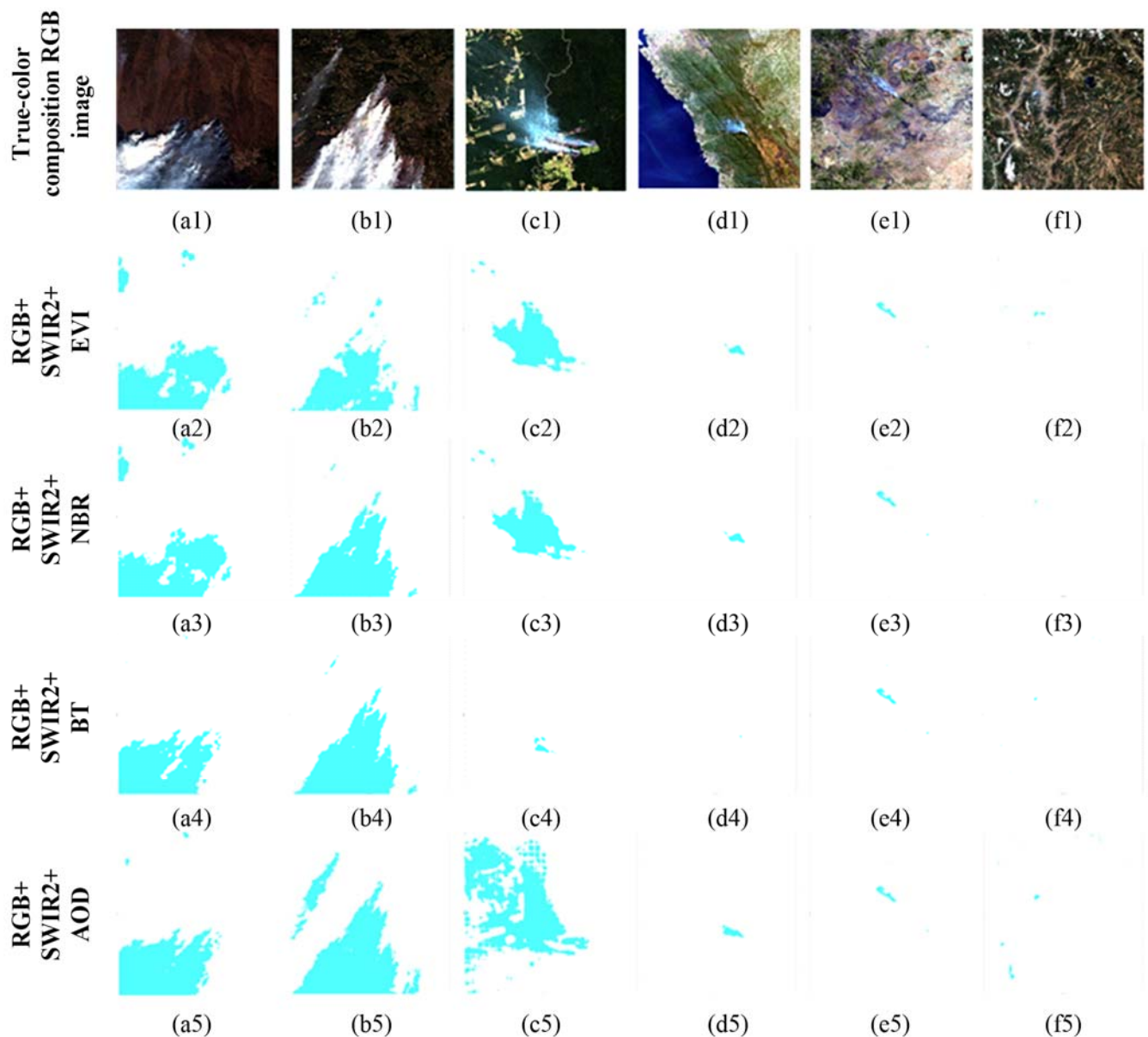


Figure 12. The first line is true-color composition RGB images of smoke plumes. (a1–a5) Siberia area, Russia on 17 Mar 2018; (b1–b5) British Columbia, Canada, on 4 August 2017; (c1–c5) Amazon region, Brazil, on 9 August 2019; (d1–d5) New Zealand area, on 7 February 2019; (e1–e5) Zambia, on 26 June 2017; (f1–f5) Liangshan region, China, on 21 May 2019. All rows except the first are segmentation results of smoke with multiple bands and remote sensing indexes, the smoke pixels are depicted in aqua color.

In Figure 12(c5), the upper left area is the smoke plume diffusion area, and a large number of smoke pixels that could not be identified by visual interpretation were segmented. This may be a result from the increasing aerosol concentration in this area due to the large amount of carbon oxides and nitrogen oxides contained in forest fire smoke. In Figure 12(f5), some mis-segmentation was made because much smaller smoke area and fewer smoke pixels are prone to be mis-recognized by image noise. Therefore, it can be concluded that the segmented smoke pixels significantly increase, especially for the thin smoke in the downwind diffusion zone, when AOD is added as the input of RGB and SWIR2.

5. Conclusions

In order to solve the difficulty of detecting forest fire smoke in remote sensing images, this study proposed the Smoke-Unet network to segment forest fire smoke and analyzed the sensitivity of remote sensing satellite data and remote sensing index used for wildfire detection. This paper first constructed a multispectral remote sensing smoke dataset containing different years, seasons, regions and land cover. Second, Smoke-Unet, which combined an improved Unet network with attention mechanism and residual block, was put forward in this paper and verified by comparing with other methods on the experiments. Third, the sensitivity of different spectral band combinations of multispectral data and the remote sensing index to the wildfire smoke segmentation were analyzed by the experiments. The results show that the smoke pixel accuracy rate using the proposed Smoke-Unet is 3.1% higher than that of Unet and RGB, SWIR2 and AOD bands are verified as the sensitive band combination and the remote sensing index for wildfire smoke segmentation, which could effectively segment the smoke pixels in remote sensing images. This proposed method under the RGB, SWIR2 and AOD bands can help to segment smoke by using high-sensitivity band and remote sensing index and makes an early alarm of forest fire smoke. However, some problems need to be further solved in subsequent studies. A large amount of mixed spectrum phenomenon in the diffusion area makes it much difficult to label thin smoke plume in the downwind direction by visual interpretation. How to exploit the feature-extraction advantages of deep learning methods to better interpret remote sensing images requires a lot of exploration.

Author Contributions: Conceptualization, Z.W. and P.Y.; data curation, P.Y.; formal analysis, P.Y.; funding acquisition, C.Z.; methodology, P.Y.; project administration, C.Z.; software, P.Y.; supervision, H.L., C.Z., J.Y., Y.T. and W.C.; validation, Z.W., P.Y., C.Z., J.Y., Y.T. and W.C.; visualization, Z.W. and P.Y.; writing—original draft, Z.W. and P.Y.; writing—review and editing, Z.W., H.L., C.Z., J.Y., Y.T. and W.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 31971668.

Informed Consent Statement: Not applicable.

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

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

References

1. Sun, G.; Hallema, D.; Asbjornsen, H. Ecohydrological processes and ecosystem services in the Anthropocene: A review. *Ecol. Process.* **2017**, *6*, 35. [\[CrossRef\]](#)
2. Hansen, M.C.; Loveland, T.R. A review of large area monitoring of land cover change using Landsat data. *Remote Sens. Environ.* **2012**, *122*, 66–74. [\[CrossRef\]](#)
3. Houghton, R. Historic role of forests in the global carbon cycle. In *Carbon Dioxide Mitigation in Forestry and Wood Industry*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 1–24.
4. Canadell, J.G.; Raupach, M.R. Managing forests for climate change mitigation. *Science* **2008**, *320*, 1456–1457. [\[CrossRef\]](#)
5. Bowman, D.M.J.S.; Balch, J.K.; Artaxo, P.; Bond, W.J.; Carlson, J.M.; Cochrane, M.A.; D’Antonio, C.M.; DeFries, R.S.; Doyle, J.C.; Harrison, S.P.; et al. Fire in the earth system. *Science* **2009**, *324*, 481–484. [\[CrossRef\]](#)

6. Allen, C.D.; Macalady, A.K.; Chenchouni, H.; Bachelet, D.; McDowell, N.; Vennetier, M.; Kitzberger, T.; Rigling, A.; Breshears, D.D.; Hogg, E.H.T. A global overview of drought and heat-induced tree mortality reveals emerging climate change risk for forests. *For. Ecol. Manag.* **2010**, *259*, 660–684. [\[CrossRef\]](#)
7. Hirsch, K.G.; Corey, P.N.; Martell, D.L. Using expert judgment to model initial attack fire crew effectiveness. *For. Sci.* **1998**, *44*, 539–549.
8. Korontzi, S.; McCarty, J.; Loboda, T.; Kumar, S.; Justice, C. Global distribution of agricultural fires in croplands from 3 years of Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Glob. Biogeochem. Cycles* **2006**, *20*, GB2021. [\[CrossRef\]](#)
9. Dwyer, E.; Pinnock, S.; Gregoire, J.M.; Pereira, J.M.C. Global spatial and temporal distribution of vegetation fire as determined from satellite observations. *Int. J. Remote Sens.* **2000**, *21*, 1289–1302. [\[CrossRef\]](#)
10. Csiszar, I.; Denis, L.; Giglio, L.; Justice, C.O.; Hewson, J. Global fire activity from two years of MODIS data. *Int. J. Wildland Fire* **2005**, *14*, 117–130. [\[CrossRef\]](#)
11. Dozier, J. A method for satellite identification of surface temperature fields of subpixel resolution. *Remote Sens. Environ.* **1981**, *11*, 221–229. [\[CrossRef\]](#)
12. Matson, M.; Stephens, G.; Robinson, J. Fire detection using data from the NOAA-N satellites. *Int. J. Remote Sens.* **1987**, *8*, 961. [\[CrossRef\]](#)
13. Robinson, J.M. Fire from space: Global evaluation using infrared remote sensing. *Int. J. Remote Sens.* **1991**, *12*, 3–24. [\[CrossRef\]](#)
14. Kaufman, Y.J.; Tanré, D. Algorithm for Remote Sensing of Tropospheric Aerosol From MODIS. In *NASA MODIS Algorithm Theoretical Basis Document*; Goddard Space Flight Center: Greenbelt, MD, USA, 1998; Volume 85, pp. 3–68.
15. Giglio, L.; Descloitres, J.; Justice, C.O.; Kaufman, Y.J. An enhanced contextual fire detection algorithm for MODIS. *Remote Sens. Environ.* **2003**, *87*, 273–282. [\[CrossRef\]](#)
16. Giglio, L.; Schroeder, W.; Justice, C.O. The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sens. Environ.* **2016**, *178*, 31–41. [\[CrossRef\]](#)
17. Justice, C.O.; Román, M.O.; Csiszar, I.; Vermote, E.F.; Wolfe, R.E.; Hook, S.J. Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *J. Geophys. Res.* **2013**, *118*, 9753–9765. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Wolfe, R.E.; Lin, G.; Nishihama, M.; Tewari, K.P.; Tilton, J.C.; Isaacman, A.R. Suomi NPP VIIRS prelaunch and on-orbit geometric calibration and characterization. *J. Geophys. Res.* **2013**, *118*, 508–511, 521. [\[CrossRef\]](#)
19. Csiszar, I.; Schroeder, W.; Giglio, L.; Ellicott, E.; Vadrevu, K.P.; Justice, C.O. Active fires from the Suomi NPP Visible Infrared Imaging Radiometer Suite: Product status and first evaluation results. *J. Geophys. Res.* **2014**, *119*, 803–816. [\[CrossRef\]](#)
20. Schroeder, W.; Oliva, P.; Giglio, L.; Quayle, B.; Lorenz, E.; Morelli, F. Active fire detection using Landsat-8/OLI data. *Remote Sens. Environ.* **2015**, *185*, 210–220. [\[CrossRef\]](#)
21. Li, Z.; Nadon, S.; Cihlar, J. Satellite-based detection of Canadian boreal forest fires: development and application of the algorithm. *Int. J. Remote Sens.* **2000**, *21*, 3057–3069. [\[CrossRef\]](#)
22. Li, Z.; Kaufman, Y.J.; Ichoku, C.; Fraser, R.; Trishchenko, A.; Giglio, L.; Yu, X. A Review of AVHRR-based Active Fire Detection Algorithms: Principles, Limitations, and Recommendations. 2000. Available online: http://www.fao.org/GTOS/gofc-gold/docs/fire_ov.pdf (accessed on 29 September 2021).
23. Csiszar, I.A.; Morissette, J.T.; Giglio, L. Validation of active fire detection from moderate resolution satellite sensors: The MODIS example in northern Eurasia. *Remote Sens.* **2006**, *44*, 1757–1764. [\[CrossRef\]](#)
24. Genet, H.; McGuire, A.D.; Barrett, K.; Breen, A.; Euskirchen, E.S.; Johnstone, J.F.; Yuan, F. Modeling the effects of fire severity and climate warming on active layer thickness and soil carbon storage of black spruce forests across the landscape in interior Alaska. *Environ. Res. Lett.* **2013**, *8*, 045016. [\[CrossRef\]](#)
25. Giglio, L.; Kendall, J.D.; Justice, C.O. Evaluation of global fire detection algorithms using simulated AVHRR infrared data. *Int. J. Remote Sens.* **1999**, *20*, 1947–1985. [\[CrossRef\]](#)
26. Schroeder, W.; Oliva, P.; Giglio, L.; Ivan, A.C. The New VIIRS375m active fire detection data product: Algorithm description and initial assessment. *Remote Sens. Environ.* **2014**, *143*, 85–96. [\[CrossRef\]](#)
27. Waigl, C.F.; Stuefer, M.; Prakash, A.; Ichoku, C. Detecting high and low-intensity fires in Alaska using VIIRS I-band data: An improved operational approach for high latitudes. *Remote Sens. Environ.* **2017**, *199*, 389–400. [\[CrossRef\]](#)
28. Elvidge, C.D.; Zhizhin, M.; Hsu, F.C.; Baugh, K.E. VIIRS Nightfire: Satellite Pyrometry at Night. *Remote Sens.* **2013**, *5*, 4423–4449. [\[CrossRef\]](#)
29. Cho, K.; Kim, Y.; Kim, Y. Disaggregation of Landsat-8 Thermal Data Using Guided SWIR Imagery on the Scene of a Wildfire. *Remote Sens.* **2018**, *10*, 105. [\[CrossRef\]](#)
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. Garay, M.J.; Mazzoni, D.M.; Davies, R.; Diner, D. The application of support vector machines to the analysis of global datasets from MISR. In *Proceedings of the Fourth Conference on Artificial Intelligence Applications to Environmental Science*, San Diego, CA, USA, 9–13 January 2005.
32. Mazzoni, D.; Garay, M.J.; Davies, R.; Nelson, D. An operational MISR pixel classifier using support vector machines. *Remote Sens. Environ.* **2006**, *107*, 149–158. [\[CrossRef\]](#)
33. Mazzoni, D.; Logan, J.A.; Diner, D.; Kahn, R.; Tong, L.; Li, Q. A data-mining approach to associating MISR smoke plume heights with MODIS fire measurements. *Remote Sens. Environ.* **2007**, *107*, 138–148. [\[CrossRef\]](#)

34. Li, X.L.; Wang, J.; Song, W.G.; Ma, J.; Telesca, L.; Zhang, Y.M. Automatic Smoke Detection in MODIS Satellites Data based on K-means Clustering and Fisher Linear Discrimination. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 971–982. [\[CrossRef\]](#)
35. Li, X.L.; Song, W.G.; 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. Zhang, Z.; Liu, Q.; Wang, Y. Road extraction by deep residual u-net. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 749–753. [\[CrossRef\]](#)
37. Jeppesena, J.H.; Jacobsena, R.H. A cloud detection method for landsat 8 images based on pcanet. *Remote Sens.* **2018**, *10*, 877.
38. Ba, R.; Chen, C.; Yuan, J.; Song, W.; Lo, S. SmokeNet:Satellites Smoke Scene Detection Using Convolutional Neural Network with Spatial and Channel-Wise Attetion. *Remote Sens.* **2019**, *11*, 1702. [\[CrossRef\]](#)
39. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015. Available online: <https://arxiv.org/pdf/1505.04597.pdf> (accessed on 29 September 2021).
40. Bao, Y.; Liu, W.; Gao, O.; Lin, Z.; Hu, Q. E-Unet++: A Semantic Segmentation Method for Remote Sensing Images. In Proceedings of the 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 18–20 June 2021; pp. 1858–1862. [\[CrossRef\]](#)
41. Li, X.; Du, Z.; Huang, Y.; Tan, Z. A deep translation (GAN) based change detection network for optical and SAR remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *179*, 14–34. [\[CrossRef\]](#)
42. Maratkhan, A.; Ilyassov, I.; Aitzhanov, M.; Demirci, M.F.; Ozbayoglu, A.M. Deep learning-based investment strategy: Technical indicator clustering and residual blocks. *Soft Comput.* **2021**, *25*, 5151–5161. [\[CrossRef\]](#)
43. Kastner, S.; Ungerleider, L.G. Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* **2000**, *23*, 315–341. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Balshi, M.S.; McGuire, A.D.; Zhuang, Q.; Melillo, J.; Kicklighter, D.W.; Kasischke, E. The role of historical fire disturbance in the carbon dynamics of the pan-boreal region: A process-based analysis. *J. Geophys. Res.* **2007**, *112*, 1–18. [\[CrossRef\]](#)
45. Dennison, P.E.; Brewer, S.C.; Arnold, J.D.; Moritz, M.A. Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.* **2014**, *41*, 2928–2933. [\[CrossRef\]](#)
46. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [\[CrossRef\]](#)
47. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6230–6239.