When Convolutional Neural Networks Meet Remote Sensing Data for Fire Detection

Ziwen Li\textsuperscript{1,2}, Yuehuan Wang\textsuperscript{*}\textsuperscript{1} and Shuo Liang\textsuperscript{2}

\textsuperscript{1} National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China

\textsuperscript{2} CETC Key Laboratory of Aerospace Information Application, China

Corresponding author’s Email: yuehwang@hust.edu.cn

Abstract: In this paper, we present a novel end-to-end Dual Fire Detection Network (DFD-Net) for the remote sensing data fire detection task. The proposed network architecture consists of two streams in a parallel fashion, a fire estimate stream is used to detect fire pixels, and a cloud-water stream is built to exclude cloud and water regions. Moreover, the pixel and band attention modules adapted to characteristics of the remote sensing data are proposed. Experimental results on our prepared Himawari-8 data fire detection dataset with ground truth labels demonstrate that the proposed algorithm outperforms existed fire detection methods in various metrics.

Index Terms: Fire detection, remote sensing, dual network, self-attention.

1. Introduction

Fire is one of the serious disasters in the world, which destroys ecosystems and threaten the lives of animals and humans. Remote sensing data fire detection plays a vital role in forest fire prevention and control, especially in the early stage of fire detection.

Himawari-8\textsuperscript{[1]} is a geostationary synchronous satellite system from the Japan Meteorological Agency (JMA), which has a temporal resolution of 10 minutes for global observations, including 16 bands from visible to infrared light, with a spatial resolution of 500-2000 m.

Currently, due to the good resolution of satellite data, many countries have adopted fire detection techniques using remote sensing data. In the current algorithms, the most predominant algorithms are threshold algorithms, which include fixed thresholds algorithms\textsuperscript{[2]} and context-based thresholding algorithms \textsuperscript{[3]}. The fixed threshold algorithm sets thresholds for reflectance and brightness temperature to determine whether a pixel is a fire point or not. However, the fixed-threshold algorithm has significant limitations for regional and seasonal variations, such as alternating dry and wet seasons. Context-based thresholding algorithms overcome the shortcomings to some extent with the surrounding neighborhood information, but they also suffer from the problem that the selection of thresholds is more difficult to determine. In addition, there are algorithms using machine learning methods including Multilayer Perceptron\textsuperscript{[5]}, random forest \textsuperscript{[4]}, decision tree\textsuperscript{[6]}, and logistic regression\textsuperscript{[7]} to detect the fire point.

Deep convolutional neural networks are common in computer vision. In this paper, we adapted the convolutional network to be applied to remote sensing data. We first created a fire detection dataset with ground-truth labels. Further, based on the exploration of the fire detection task and the characteristics of the remotely sensed data, we propose the DFD-Net with two branches: the fire estimate stream is designed to identify fire point pixels, while the cloud-water stream is used to exclude non-fire point
pixel. We also propose two self-attentive modules to obtain significance between bands and pixels. Our algorithms can have the real-time and comparable performance. Overall, our contributions are as follows:

1. We propose a novel end-to-end deep convolutional network. To the best of our knowledge, it is the first deep learning algorithm to tackle the fire detection issue. We separate DFD-Net into two parallel streams, a fire estimate stream for fire assessment and a cloud-water stream for non-fire exclude.

2. The self-attention module is adopted by analyzing the characteristics between bands and pixels. The proposed module can largely improve the performance of remote sensing fire detection.

3. Based on the Himawari-8 satellite remote sensing data, we prepare a remote sensing dataset for the fire detection task with corresponding ground-truth labels.

4. Experimental results demonstrate that our algorithm outperforms previous algorithms both quantitatively and qualitatively by a large margin in all metrics.

2. Proposed method

2.1. Dual Network Architecture

We design the novel dual network architecture for fire detection. To the best of our knowledge, DFD-Net is the first time that the deep learning is used for fire detection tasks.

The network structure is shown in figure 1. Accordingly, we divide the Dual Fire Detection network into two parts, respectively fire estimate stream and cloud-water stream. We integrate these two components into the unified network architecture. the last part is the merge block, which fuses the fire feature and non-fire feature and gets the final result.

![Figure 1. The structure of our Dual Fire Detection Network.](image)

2.1.1. Fire Estimate Stream. The fire estimate stream is used to detect the fire point and the input data bands include tbb_07, tbb_14, tbb_15, SOZ band. The fire estimate stream consists of pixel attention block, band attention block, and convolutional block. More specially, after feeding the data into the fire estimate stream, we use two pixel and band attention block to get the self-attention weights. Then, we use a convolutional and full connection layer to get a feature vector. It’s worth noting that we did not use the pooling layer, because the pooling layer inevitably loses some feature information. Instead, we use a learnable convolutional layer for the downsampling operation.

2.1.2. Cloud-water Stream. The cloud-water stream is used to exclude non-fire point pixels such as clouds and water, and the input bands are albedo_03, albedo_04, albedo_06 bands. Unlike the fire estimate stream, fewer number of attention block and convolutional blocks are used than the fire estimate stream. This is mainly due to the cloud-water correlation band is only weakly correlated, while the spatial distribution is more concentrated (e.g., clouds and water are always blocked). Finally, we get the final feature map and convert it into a feature vector in the same way.
2.1.3. **Merge Block.** The feature vectors will be merged here with the element-wise product. Finally, we use the SoftMax function to get the confidence level. When testing, we choose the value with a confidence level greater than 0.5 as the decision result of the fire detection.

2.2. **Self-Attention Module**
Different bands and pixels have different importance, and this inhomogeneity cannot be effectively exploited if all pixels and bands are treated equally. We address this problem by proposing two self-attentive mechanisms for the fire detection task. The details are elaborated below.

2.2.1. **Band Attention Block.** Each band has a different importance for fire detection. For example, in the Himawari-8 satellite data, the tbb_07 band is used to measure the absolute temperature, while the tbb_14 band and the tbb_15 band are used to measure the relative temperature and to detect contextual fires. Therefore, we introduce a band attention strategy to models of correlations between bands.

![Figure 2. The proposed (a) band attention block and (b) pixel attention block.](image)

As shown in figure 2, we first use the GAP and GMP to a C*1*1 attention vector respectively. Then element-wise addition is adopted to fuse the two attentional feature vectors. Next, we use two 1*1 Conv to do the nonlinear transformation. The ReLU and sigmoid function constraint the value to [0,1]. In the end, we could obtain the final attention result by using an element-wise product.

2.2.2. **Pixel Attention Block.** Because the dispersion of fires is uneven, it is necessary to give different attention to different pixels. To this end, we introduce pixel attention for fire detection that adapts to give different degrees of interest to different pixels.

As illustrated in figure 2, we first extract initial spatial weights by using the 1*1 convolutional layer to redistribute the pixel weights. With those two convolutional layers, we can obtain the pixel attention map with the shape 1*H*W. Then we duplicate the pixel attention map to the same size as the input feature map. Ending with the element-wise product, we acquire a feature map with rational weights.

2.3. **Dataset**
Based on the characteristics of the fire detection mission, The data selection principle is to select only those data that are highly relevant to fire detection and to eliminate irrelevant attributes.

The bands that are relevant to fire are tbb_07, tbb_14, tbb_15. Considering the influence of clouds and water, albedo_03, albedo_04, and albedo_06 are selected as features. Finally, SOZ was chosen as a judgment factor considering the difference in fire temperature during the day and night. In total, the selected bands for fire detection are albedo_03, albedo_04, albedo_06, tbb_07, tbb_14, tbb_15, and SOZ. We chose a neighborhood size of 21*21 in our experiments, and it was verified that the neighborhood of 21*21 is the most appropriate range that contributes to the fire point determination of the center pixel.

Based on the reference information from the Himawari-8 satellite, data blocks with labels can be obtained. During sampling, care is taken that the number of fire and non-fire data blocks should be kept...
equal to avoid unbalanced data sets. The data blocks are divided into the train and test set by random sampling in a ratio of 7:3.

3. Experiments

3.1. Implementation Details
The training set contains a total of 800 data and the test set containing 154 point data. It is important to note that the test set is completely independent of the training set and is not used for the training process.

The training process of the model is performed on a GPU with a machine model NVIDIA TITAN V. The cross-entropy loss function is adopted. The initial learning rate is set to 0.001 and the MultiStepLR method is used to adjust the learning rate. The specified training Epoch is 150 times, and the optimizer uses the Adam optimization method.

3.2. The Results
The proposed algorithm is compared with the effects of other machine learning algorithms, including support vector machine, multilayer perceptron, decision tree, logistic regression, and random forest. These models were also compared quantitatively using various metrics, as can be seen in Table 1, DFD-Net achieved the highest scores on all metrics, because CNN has stronger classification ability.

Table 1. The evaluation results of all algorithms.

| Metric/Method       | SVM  | LR  | MLP  | RF  | DT  | DFD-Net |
|---------------------|------|-----|------|-----|-----|---------|
| Overall accuracy    | 0.6  | 0.89| 0.73 | 0.88| 0.97| 0.99    |
| Precision           | 0.21 | 0.84| 0.69 | 0.90| 0.99| 0.99    |
| Recall              | 0.94 | 0.93| 0.75 | 0.86| 0.96| 1.00    |
| F1-score            | 0.34 | 0.88| 0.71 | 0.88| 0.97| 0.99    |

Figure 3. ROC curve for each algorithm on fire detection.

The performance of these algorithms can be reflected by the ROC curve and AUC value. As you can see from the Figure 3, the ROC curve for the DFD-Net model is at the top with an AUC of 0.99, which is greater than the second-ranked decision tree. This means that the performance of DFD-Net is the best among all algorithms.

4. Conclusion
In this paper, we address the fire detection task with a novel dual convolutional neural network and self-attention module. Moreover, a fire detection dataset is created for training model parameters and testing algorithm performance. We also compare the performance of the algorithm with other machine learning
algorithms, such as support vector machines, logistic regression, decision tree, random forest, multilayer perceptron, etc. The experimental results verify that the proposed DFD-Net has higher accuracy and can be applied to solve the fire detection problem in a real-time manner.

5. References

[1] Basho K, Date K, Hayashi M, et al. An introduction to Himawari-8/9—Japan’s new-generation geostationary meteorological satellites[J]. Journal of the Meteorological Society of Japan. Ser. II, 2016, 94(2): 151-183.

[2] Kaufman Y J, Justice C O, Flynn L P, et al. Potential global fire monitoring from EOS-MODIS[J]. Journal of Geophysical Research: Atmospheres, 1998, 103(D24): 32215-32238.

[3] Giglio L, Schroeder W, Justice C O. The collection 6 MODIS active fire detection algorithm and fire products[J]. Remote Sensing of Environment, 2016, 178: 31-41.

[4] Remo R, Chavicol E. Developing a random forest algorithm for MODIS global burned area classification[J]. Remote Sensing, 2017, 9(11): 1193.

[5] Li X, Song W, Lian L, et al. Forest fire smoke detection using back-propagation neural network based on MODIS data[J]. Remote Sensing, 2015, 7(4): 4473-4498.

[6] Pashynska N, Snytyuk V, Putrenko V, et al. A decision tree in a classification of fire hazard factors[J]. Восточно-Европейский журнал передовых технологий, 2016 (5 (10)): 32-37.

[7] Bisquert M, Caselles E, Sánchez J M, et al. Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data[J]. International Journal of Wildland Fire, 2012, 21(8): 1025-1029.