Wildfire Detection Based on Multi-source Spatial Data

Yigao Wang(23020221154121, AI)[1], Lijuan Weng(23020221154124, AI)[1], Yongyi Wu(23020221154126, Information)[1], Pengyu Xu(23020221154132, AI)[1], Yuchen Tian(23020221154114, Information)[1]

[1] College of Information, Xiamen University, Xiamen, Fujian, China

Abstract

Forest is an important part of the global ecosystem and is of great significance to the sustainable development of the ecological environment. As the rising temperatures of the earth, forests become increasingly dry and wildfires occur more frequently, which brings great losses to global forest resources and the safety of individuals' lives. Therefore, wildfire detection is of great significance for environment protection. Recently, thermal infrared remote sensing shows the advantages of wide coverage and long observation time, which provides an important technology for the identification of large-scale and all-weather wildfires. In this paper, we use remote sensing data to detect wildfires by leveraging deep learning algorithm. Specifically, we first investigate the characteristics of wildfires in thermal infrared remote sensing imagery and make a dataset, and then train a Faster R-CNN model to detect hotspots. Next, we filter out nonforest areas to discover wildfire by fusing the global forest cover map. Results show that our method can accurately detect wildfires with F1-score of 0.881.

Introduction

Forests are important ecosystems in the world, providing human beings with rich and diverse natural resources. In addition, forests are of great value in water and soil conservation, climate regulation, and disaster prevention. However, Under the general trend of global warming, forests have become drier and wildfires have greatly increased in frequency and severity (Farooq et al. 2022). In November 2018, wildfires erupted in northern California (Sills et al. 2019), which destroyed 15,336 acres of land and caused a large number of casualties and property losses. On March 30, 2019, the wildfire in Liangshan Prefecture killed 31 people and burned about 19 hectares of forest. Wildfires not only seriously damage the balance of forest ecosystem, but also pose a great threat to the safety of individuals' lives and property (Graham, Mccaffrey, and Jain 2004). China has rich forest resources and a wide variety of plants. Once a wildfire occurs, it will destroy a large number of precious trees and plant resources, causing the soil to lose its ability to infiltrate and retain water. In severe cases, it will cause other natural disasters such as flash floods, mudslides and so on. On the other hand, the spread of wildfires poses a serious threat to all kinds of infrastructure in forest areas. Therefore, it is of great significance to discover an all-weather and large-scale method of wildfire detection.

Traditional wildfire detection methods rely on manual patrols, but it is inefficient and high-cost. To detect wildfires more accurately, researchers have proposed a lot of methods, such as sensors, visible light detection systems, and so on. However, all of them have certain shortcomings. Visible light cannot work around the clock, and sensors are prone to false alarms caused by environmental interference. Given the above problems, the gradually mature satellite remote sensing technology solves these shortcomings with its advantages of wide coverage and independence of visible light (Xue-Qiong et al. 2010). Remote sensing realizes the observation of the earth through the sensors and radar on satellites, and thermal infrared remote sensing can record the thermal radiation of ground objects. The stronger the thermal infrared radiation of the object, the higher the gray value of the object in the thermal infrared imagery. Therefore, wildfires generate high-intensity thermal infrared radiation at high temperatures and appear as bright white areas.

Detecting wildfires in satellite remote sensing imagery usually relies on visual interpretation, which is time-consuming and labor-intensive. In recent years, the emergence of deep learning has successfully solved the problem of difficult feature extraction of image data. Deep learning imitates the principle of the human brain's visual system to process images hierarchically, and develops into an unsupervised feature learning model (Fergus et al. 2012). In this paper, we use deep learning technology to solve the problem of wildfire detection on the basis of thermal infrared remote sensing image data. However, to build such a wildfire detection model, the following issues must be addressed:

- **Large scale imagery.** With the development of remote sensing technology, the resolution of satellite remote sensing images has developed to a new level. The training process of deep learning requires repeated iterations to achieve better results. Faced with the large amount of data brought by high resolution, it is difficult for deep learning methods to quickly learn important information while ensuring the quality of detection. Therefore, we need to preprocess the remote-sensing image data.
- **Small object detection.** Compared with the background area of remote sensing imagery, the propor-

tion of wildfires object is too small. Most of the object detection algorithms based on deep learning use the convolutional neural network as the backbone network. Due to the downsampling of the convolutional neural network, small objects tend to disappear in deep networks passing through multiple downsampling layers. Therefore, it is necessary to select a suitable feature extraction network for object detection.

- **Nonwildfire Area Filter.** Some hotspots that are not caused by wildfires also show bright white areas in thermal infrared remote sensing imagery, which are similar to wildfires. Therefore, it is also important to exclude these hotspots and extract the real wildfire.

To address the above issues, we propose a two-stage wildfire detection framework. In the first stage, we first cut the original imagery into small patches. Next, we performed a series of data enhancement on them. Then, we utilize Faster R-CNN, which uses ResNet-50+FPN as the backbone, for hotspot detection. In the second stage, with the aid of the GIS platform and the global forest cover map, we use multi-source data fusion to filter non-wildfire areas. Finally, we used the JS API of the Gaode map and flask to complete the visualization of the wildfire.

In summary, the main contributions of this paper include:

- In order to reduce the manpower and material resources consumed by traditional wildfire detection methods, we proposed a wildfire detection framework based on multi-source spatial data to obtain detect wildfires from remote sensing imagery.
- We successfully utilize Faster R-CNN, which uses ResNet-50+FPN as the backbone, to effectively detect the hotspots in the thermal infrared remote sensing imagery. Moreover, the nonwildfire areas in the hotspot detection results are filtered out by multi-source spatial data.

Related Work

The concept of deep learning (Hinton and Salakhutdinov 2006) was proposed by Hinton et al. in 2006. The convolutional neural network in the field of deep learning is designed to solve the problem of image recognition. Its design is inspired by the perception of the outside world by the biological visual cortex cells. Nowadays, the research upsurge of convolutional neural network algorithms continues to rise, and many scholars at home and abroad are trying to apply this deep learning technology to fire detection.

HICINTUKA Jean Philippe et al. proposed a convolutional neural network to identify fire in video data. This fire detection method based on depth domain has a powerful function of extracting fire features. In order to balance efficiency and accuracy, the model structure is adjusted to make it better applicable (Philippe and Zhou 2019; Ren et al. 2016). Due to the irregular shape, color change and indescribable texture of fire smoke, Yingshu Peng et al. proposed a smoke detection algorithm that combines extraction of smoke suspicious areas and deep learning. First,

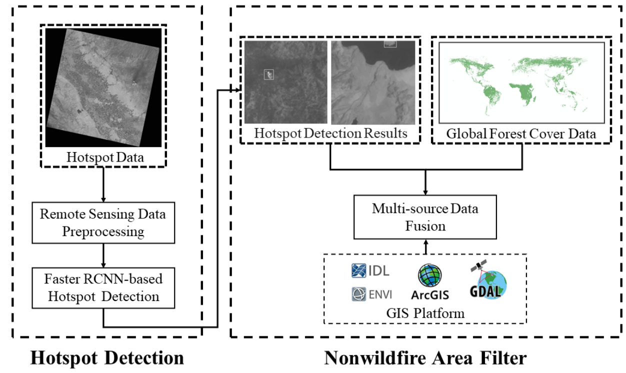


Figure 1: Framework Overview.

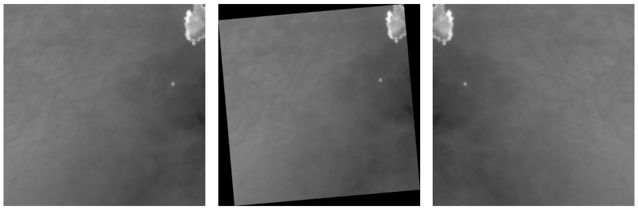
the algorithm is designed to extract smoke suspicious areas of the image, and then it serves as the input of the convolutional neural network (Peng and Wang 2019). Faisal Saeed and others initially proposed three kinds of deep neural networks. After training and verification, two algorithms, Adaboost-LBP model and convolutional neural network, were finally used for forest fire detection of video images. The accuracy of model prediction was very high (Saeed et al. 2020). Jivitesh Sharma et al. developed a fire detection system using two pretrained convolutional neural networks, namely VGG-16 and ResNet-50. The data set contains more non-fire images than fire images. By creating an unbalanced dataset to simulate the real scene, the accuracy of ResNet-50 is slightly better than VGG-16 in the experimental results, indicating that the deeper convolutional neural network structure is more challenging. It has good performance on the dataset (Sharma et al. 2017).

Proposed Solution

As shown in Figure 1, our framework consists of two parts, which specifically includes hotspot detection and nonwildfire area filter. In the first stage, we collected a decade of thermal infrared remote sensing images containing hotspots in California and its vicinity and made them into a dataset of VOC 2012 standard format after data preprocessing. After that, we build a hotspot detection model based on Faster R-CNN to obtain hotspot detection results. In the second stage, we filter out the nonwildfire areas in the hotspot detection results with the global forest cover map through multi-source data fusion. Next, we would elaborate on these three parts respectively.

Hotspot Detection

Different from natural color imagery, remote sensing imagery is mainly shot at high altitudes, which covers a wide range of objects and complex backgrounds. At the same time, wildfire usually shows irregular shape in remote sensing imagery, and there are some problems such as scattered object distribution and small scale. Therefore, we need to consider how to preprocess remote sensing data, and then select a suitable object detection model to better detect



(a) Data cutting (b) Data rotation (c) Data flip

Figure 2: Data preprocessing operation.

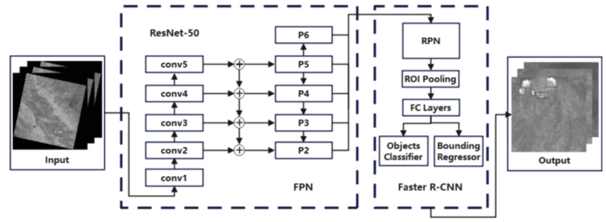


Figure 3: Faster R-CNN network structure with Resnet50+FPN as the backbone.

hotspots.

Image Preprocessing The training process of deep learning requires repeated iterations to achieve better results. Faced with the large amount of data brought by high resolution, it is difficult for deep learning methods to quickly learn important information while ensuring the quality of detection. Therefore, for remote sensing imagery, we first cut the large image of the original size into a series of small images, which not only makes the network have the ability to process these data but also enlarges the image and improves the detection result. Then, we performed a series of data enhancement operations on the data, such as flipping, rotating, cropping, etc., to enrich the distribution of training data and improve the generalization and robustness of the model.

Faster R-CNN With the development of deep learning, object detection has developed rapidly in recent years. The current object detection algorithms can be roughly divided into two categories: one-stage and two-stage. Compared with the one-stage algorithm, the ROI pooling in the two-stage algorithm will resize the object. The feature of the small object will be amplified, and its feature outline will be clearer, so the detection results are more accurate and the missed detection rate is lower. Among two-stage detection models, Faster R-CNN, as the representative of the two-stage network, has an excellent performance in problems such as multi-scale and small objects by virtue of its superior performance (Ren et al. 2017). Therefore, we use Faster R-CNN model for hotspot detection in this paper.

Faster R-CNN is a region-based convolutional neural network framework, which is mainly composed of the following four parts:

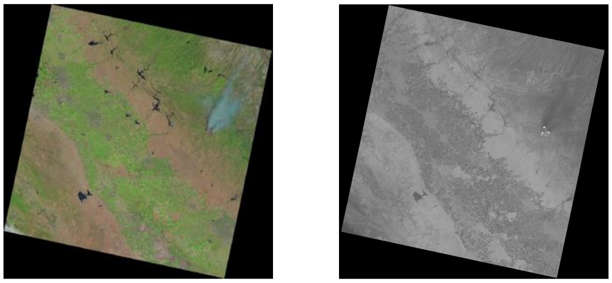
Feature extraction network. Feature extraction networks are used for feature extraction. The image is input into the network, and then the corresponding feature map is generated through the convolutional layer. The commonly used feature extraction networks of Faster R-CNN are ZFNet (Zeiler and Fergus 2014) and VGG-16 (Simonyan and Zisserman 2015). However, because of their relatively simple network structures and the limited extraction of deep-level features, it is difficult to learn deep semantic information of images. Usually, the deep network structure can improve the performance of the model, and the model will get a better training effect. However, as the depth of the network becomes deeper, the performance of the network becomes worse, and the degradation problem occurs. In order to solve this problem, Kaiming He et al. proposed a residual network structure (He et al. 2016) to solve the problem of stochastic gradient disappearance, and a deeper network can also extract richer semantic information. Based on the excellent performance of ResNet, we use ResNet-50 as the backbone network of Faster R-CNN for feature extraction in this paper, so as to obtain better detection results. Since we use ResNet-50 as the backbone, the deep network layer leads to excessive feature downsampling multiples, which may easily cause small object missed detection and classification errors. Therefore, we utilize FPN to realize multi-scale information fusion. It integrates low-level detailed information with high-level semantic information so that the low-level can obtain more context information when performing small object detection. FPN's utilization effectively solves the problem that it is difficult to detect hotspots, which account for a small proportion and are dense. Figure Figure 3 shows the Faster R-CNN structure with ResNet-50+FPN as the backbone.

RPN. The full name of RPN is Region Proposal Network. In Faster R-CNN, RPN is used to generate candidate frames. RPN has two tasks. One is object classification, which is to judge whether the anchors are the foreground area or the background area through softmax. The other is candidate frame regression, which is to obtain more accurate candidate areas by adjusting the anchors.

RoI pooling. RoI Pooling is a region of interest pooling layer. The candidate frame generated by RPN is projected onto the feature map obtained by the feature extraction network to obtain the corresponding feature matrix, and then the feature maps of different sizes are uniformly scaled to a 7×7 size through the RoI Pooling layer matrix.

Classification and regression. In this part, we pass the corresponding feature matrix generated by the RoI pooling layer into softmax and bounding box regression for further classification and regression.

In this work, we first input the dataset into ResNet-50 through the bottom-up path to produce four feature maps of different scales. Next, four feature maps are used as the input of FPN. After up-sampling and lateral connection, new feature maps are obtained respectively. Then, we input the new feature maps to RPN, which is used to extract the region proposal and output it to the RoI pooling layer. At the same time, we also input the new feature maps to the RoI pooling layer to extract the corresponding feature maps for each re-



(a) Natural Color Imagery (b) Thermal Infrared Imagery

Figure 4: Remote sensing imagery of hotspot.

gion proposal. Finally, the feature map is passed through a series of full connection layers to output the results and the coordinates of the hotspot prediction boxes.

Nonwildfire Area Filter

After hotspot detection, we found that some high-temperature areas such as human activities, desert high-temperature areas, also have high gray values, which means they show similar features to the wildfire in the remote sensing imagery and are easily misidentified as wildfires. Therefore, we take advantage of the multi-source data fusion to filter out nonwildfire areas. Since the images in the hotspot detection results have been cut, we first map the prediction box coordinates back to the original remote sensing images through GDAL(Ya-dong et al. 2010). The coordinate transformation formula is as follows:

$$XG = GT(0) + X * GT(1) + Y * GT(2) \tag{1}$$

$$YG = GT(3) + X * GT(4) + Y * GT(5) \tag{2}$$

where, *XG*, *YG* are the geographic coordinates; *X*, *Y* are the row and column coordinates; *GT(0)*, *GT(3)* are the longitude and latitude of the upper left corner; *GT(1)*, *GT(5)* are the horizontal and vertical resolution; *GT(2)*, *GT(4)* are the rotation coefficients.

Then, we leverage the global forest coverage map to filter out nonwildfire areas. Specially, we transform the prediction box coordinates in the hotspot detection results into the coordinates in the forest coverage map. Then, we traverse the pixels of the forest cover map, counting the number of pixel points in the prediction box with a value of 1 (forest) and dividing by the total number of pixels in the prediction box to obtain the probability that each prediction box contains forest. We set the threshold at 0.1. As long as the probability is greater than 0.1, we think that there are forests in the prediction box, otherwise there is no forest. By traversing all prediction boxes, nonwildfire areas can be filtered out.

Experiments

Dataset Description

Remote sensing imagery is obtained from the USGS Landsat-8 satellite, which is equipped with operational land imager(OLI) and thermal infrared sensor (TIRS). We mainly use the 10-TIR band, which has a wavelength range of 10.60–11.19 μm , and a spatial resolution of 30 meters(Sibo et al. 2021). As shown in Figure 4.

We download 103 satellite remote sensing imageries of hotspots in California and its surrounding areas over the past decade. The percentage of cloud cover in these imageries is small, which makes it easy to observe hotspots. After a series of data preprocessing, we get 3,871 training images, 1,658 validation images and 72 testing images.

Table 1: Hotspot Detection Results

Method	Precision	Recall	F1-score
MobileNet-V3	0.665	0.926	0.774
EfficientNet-B0	0.693	0.939	0.797
ResNet-50	0.741	0.916	0.819
MobileNet-V3+FPN	0.737	0.943	0.827
EfficientNet-B0+FPN	0.745	**0.949**	0.835
ResNet-50+FPN	**0.843**	0.923	**0.881**

Evaluation Metrics

In order to evaluate the performance of the proposed framework, we use the Precision, Recall, and F1-score as the evaluation metrics. If a real hotspot is correctly identified as a hotspot, we call it a *true positive (TP)*. If a false hotspot is identified as a hotspot, we call it a *false positive(FP)*. If a real hotspot is identified as a false hotspot, we call it a *false negative (FN)*. Based on the definition, these metrics can be calculated as follows:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1 - socre = \frac{2 * Precision * Recall}{Precision + Recall} \tag{5}$$

Baseline Methods

In order to present the effectiveness of our framework, we compare our method with some representative baseline methods in hotspot detection.

MobileNet-V3 (Howard et al. 2019): This method adopts depthwise separable convolutional filters to build lightweight deep neural networks. The network introduces two simple global hyperparameters: width multiplier and resolution multiplier, which can effectively balance latency and accuracy.

EfficientNet-B0 (Tan and Le 2019): This method uses Neural Architecture Search to balance input resolution, depth and width of the network at the same time, which achieves a better performance in object detection.

ResNet-50: This method adopts Residual Blocks to realize more deep neural network, which improves the efficiency and accuracy of network depth deepening and avoids the problem of network degradation.

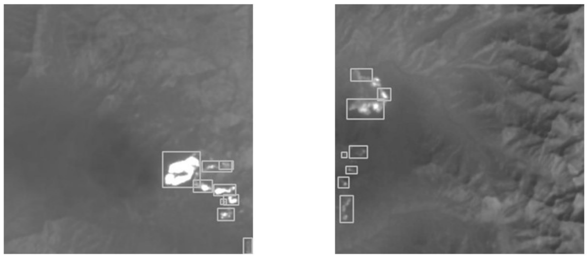


Figure 5: Part of the test dataset imagery prediction results.

MobileNet-V3+FPN: This method combines MobileNet-V3 and FPN to realize hotspot detection. FPN achieves multi-scale information fusion, which can help to better detect small objects.

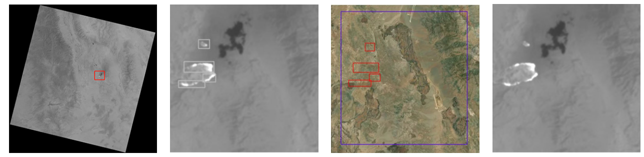
EfficientNet-B0+FPN: This method combines EfficientNet-B0 and FPN to realize hotspot detection. FPN achieves multi-scale information fusion, which can help to better detect small objects.

Results

Hotspot Detection Results In this part, we use different methods to compare the performance on the hotspot dataset to show the effectiveness of our model on the hotspot detection. We compare the overall accuracy of different methods in TABLE.1. We can see that ResNet-50+FPN method achieves an F1-score of 0.881, which is better than other baseline methods. First, we used three different backbones without FPN for experiments. We can see that MobileNet-V3 and EfficientNet-B0 perform well, but ResNet-50 outperforms them, which shows that the residual network structure of ResNet can well solve the degradation problem caused by the deepening of the network and extract deeper semantic information. Next, we introduce FPN based on these three methods respectively. It can be seen that their performance has been improved to a certain extent, which shows the effectiveness of FPN on small and dense hotspots in remote sensing images, and ResNet-50+FPN still performs the best among them. Overall, the ResNet-50+FPN method can successfully detect hotspots with high detection accuracy.

Finally, the prediction of some samples is shown in Figure 5, which proves that the prediction effect of the hotspot is in a good performance.

Nonwildfire Area Filter Results Figure 6(a) shows the detection result of a hotspot imagery with nonwildfire areas, which is enlarged as shown in Figure 6(b). It can be seen that the bright hotspots are boxed up. However, according to the corresponding natural color imagery Figure 6(c), the prediction box areas are nonforest, indicating that there are no wildfire in this area. Therefore, we implement the nonwildfire area filtering. The final result is shown in Figure 6(d), which indicates that the nonwildfire areas is successfully filtered.



(a) Detection results (b) Zoom in (c) Natural color imagery (d) After filtering out

Figure 6: Nonwildfire area filtering for remote sensing imagery containing nonwildfire areas.

Figure 7: Wildfire visualization.

WILDFIRE VISUALIZATION

In the process of wildfire detection, we obtain the latitude and longitude coordinates of prediction boxes, from which we can obtain the geographic information of the wildfire areas. Moreover, we also incorporate the wildfire detection results into visualization. Next, we collect the information after the wildfire on Twitter and make it into a word cloud. Through the Twitter data, we further show the reliability of the results of wildfire detection and burned area estimation. What's more, we can preliminarily analyze the changes in public opinion on the wildfire on Twitter after the wildfire through the word cloud, which provides a certain reference for the public to obtain key information about wildfires.And it can display data more intuitively. Finally, the visualization results are shown in Figure 7.

CONCLUSION

In this paper, we propose a two-stage framework for wildfire detection based on multi-source spatial data. First, we use the Faster R-CNN model to realize hotspot detection. Second, with the multi-source data fusion, we filter out the nonwildfire areas to get a reliable wildfire area. Moreover, we realize the visualization of wildfire detection. Results show that our framework can accurately detect the wildfire.

In the future, we plan to deepen this work from two following directions. On the one hand, we plan to adopt other algorithms for wildfire detection to further improve the performance of our framework. On the other hand, we consider to integrate more remote sensing data and public opinion data to further conduct sentiment analysis, risk prediction and so on.

References

Farooq, I.; Kachroo, M. M.; Din, A. E.; Fawzy, M.; and Ahmad, P. 2022. Elevation in wildfire frequencies with respect to the climate change. *Journal of Environmental Management*, 301.

Fergus, R.; Ranzato, M.; Salakhutdinov, R.; Taylor, G.; and Yu, K. 2012. Deep learning methods for vision. In *Proceedings of the CVPR*. Citeseer.

Graham, R. T.; Mccaffrey, S.; and Jain, T. B. 2004. Science basis for changing forest structure to modify wildfire behavior and severity. *usda forest service rocky mountain research station general*.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Hinton, G. E.; and Salakhutdinov, R. R. 2006. Reducing the Dimensionality of.

Howard, A.; Sandler, M.; Chen, B.; Wang, W.; Chen, L.-C.; and Tan, M. e. a. 2019. Searching for MobileNetV3. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*.

Peng, Y.; and Wang, Y. 2019. Real-time forest smoke detection using hand-designed features and deep learning. *Computers and Electronics in Agriculture*, 167: 105029.

Philippe, H. J.; and Zhou, W. 2019. Video Based Fire Detection Systems on Forest and Wildland Using Convolutional Neural Network. 36(2): 9.

Ren, S.; He, K.; Girshick, R.; and Sun, J. 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *NIPS*.

Ren, S.; He, K.; Girshick, R.; and Sun, J. 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 39(6): 1137–1149.

Saeed, F.; Paul, A.; Karthigaikumar, P.; and Nayyar, A. 2020. Convolutional neural network based early fire detection. *Multimedia Tools and Applications*, 79(3).

Sharma, J.; Granmo, O. C.; Goodwin, M.; and Fidje, J. T. 2017. Deep Convolutional Neural Networks for Fire Detection in Images.

Sibo, D.; Chen, R.; Zhaoliang, L.; Mengmeng, W.; Hanqiu, X.; and et al., L. H. 2021. Reviews of methods for land surface temperature retrieval from Landsat thermal infrared data. *National Remote Sensing Bulletin*.

Sills, J.; Leverkus, A. B.; Murillo, P. G.; Doña, V.; and Pausas, J. G. 2019. Wildfires: Opportunity for restoration? *Science*, 363(6423): 134.2–135.

Simonyan, K.; and Zisserman, A. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations*.

Tan, M.; and Le, Q. V. 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.

Xue-Qiong, W. U.; Qin, X. L.; Cheng, L. I.; Zu-Wei, T.; Xiong, Y. Q.; Yang, D. F.; and Zhang, R. 2010. Analysis of Current Forest Fire Monitoring System in China. *Inner Mongolia Forestry Investigation and Design*.

Ya-dong, L.; Qing-yuan, L.; Hai, T.; and Ke-xue, Q. 2010. Open source library GDAL and its application in image stitching. *Digital Technology Application*.

Zeiler, M.; and Fergus, R. 2014. Visualizing and Understanding Convolutional Networks. In *ECCV 2014*.