Type of the Paper (Article, Review, Communication, etc.)

Forest fire recognition based on dynamic feature similarity of multi-view images

Zhang Li [1], DiDi Wu [2]

[1] Affiliation 1; 3052206871@qq.com Shanghai Jiao Tong University 1954 Huashan Road, Xuhui District, Shanghai, China
[2] Affiliation 2; 34555676@qq.com Shanghai Jiao Tong University 1954 Huashan Road, Xuhui District, Shanghai, China
* Correspondence: 3052206871@qq.com; (optional; include country code; if there are multiple corresponding authors, add author initials)

Abstract: Forest fire identification is important for forest resource protection. Effective monitoring of forest fires requires the deployment of multiple monitors with different viewpoints, while most traditional recognition models can only recognize images from a single source. By ignoring the information from images with different viewpoints, these models produce high rates of missed and false alarms. In this paper, we propose a graph neural network model based on the similarity of dynamic features of multi-view images to improve the accuracy of forest fire recognition. The input features of the nodes on the graph are converted into relational features of different gallery pairs by establishing pairs (nodes) representing different viewpoint images and gallery images. The new feature library relationship is used to update the image gallery with dynamic features in order to achieve the estimation of similarity between images and improve the image recognition rate of the model. In addition, to reduce the complexity of image pre-processing process and extract key features in images effectively, this paper also proposes a dynamic feature extraction method for fire regions based on image segment ability. By setting the threshold value of HSV color space, the fire region is segmented from the image, and the dynamic features of successive frames of the fire region are extracted. The experimental results show that, compared with the baseline method Resnet, this paper's method is more effective in identifying forest fires, and its recognition accuracy is improved by 2%. And the scheme of this paper can adapt to different forest fire scenes, with better generalization ability and anti-interference ability.

Keywords: forest fire; image recognition; graph neural network;

1. Introduction

Forest fires are a serious threat to forest ecosystems, and early detection of the fire source before it turns into a catastrophic event is the key to prevent the fire from spreading and creating significant hazards1. With the rapid development of computer vision technology, forest fire monitoring based on computer vision has become a current research hotspot in the field of forest fire prevention. Flame detection is the most common method for identifying forest fires, and most fire identification systems are based on flame detection 2.

In the process of fire identification, color identification is one of the earliest methods used. Through the recognition of motion, space and time characteristics of flame color pattern, the effect of flame recognition is achieved 3.However, the color recognition method requires a smaller recognition distance and a larger flame size. In recent years, researchers have proposed using the CNN method to identify forest fires. Khan Muhammad uses the deep CNN model to identify the fire area, avoiding the cumbersome and time-consuming process of feature extraction, and automatically learning the rich features in the original fire data4.Although this method has high recognition accuracy, it can only identify static fire images5.C. Emmy Pema proposed to use the background subtraction method to find moving pixels, and then use the color model to find the color areas of the flame, and analyze these areas in time and space to identify irregular and flickering fire features6.Liu and ZC proposed an Adaboost classifier based on the characteristics of HOG (Histogram of Oriented Gridients) has been used to make primary prediction of forest fire area image, and then convolutional neural networks (CNN) and support vector machine (SVM) have been used to carry out the secondary recognition of the fire area7.Wang YuanBin combined the traditional image processing technology with the convolutional neural network and introduced the adaptive collection method. This algorithm can segment the flame region and learn the features in advance 8.

These schemes only analyze images under a single viewpoint, but in practice, effective monitoring of forest fires often requires the deployment of multiple monitors with different viewpoints, and images from a single viewpoint do not fully characterize the full picture of the monitoring point, which makes fire identification based on images from a single viewpoint not effective8. More often, the presence of fire-like interference in the recognition process they cannot be handled effectively, leading to false alarms. Thus, to reduce the complex pre-processing process of forest fire images and effectively extract key features from the images, this paper proposes a dynamic feature extraction method for fire regions based on image segment ability, which uses dynamic features as model inputs to improve the robustness of the network in recognizing forest fires from different perspectives.

But to look at information from images from different perspectives, you first have to identify which images are for the same fire surveillance scenario. In the same forest fire monitoring scene, images taken from different monitoring perspectives often have the same fire characteristics, such as fire area, background color, heat radiation, etc. Based on these features, researchers designed various metrics to investigate the similarity between images from different perspectives 8.However, these studies pay more attention to the similarity between the two images as a whole, but ignore the inherent similarity between the whole. For example, when estimating the similarity between the detection image and the target image, most feature learning and measurement learning only train and test the pair relationship between a single image pair, while ignoring other relationships between images from different sources. These relationships mean that some hard positive or negative pairs are difficult to obtain an appropriate similarity score. Because only limited relationship information between samples can be used for similarity estimation. To overcome this problem, you need to find differences in the valuable images. C. C. Loy used the method of manifold learning to map images into manifolds to make the local geometry of images more smooth, so as to investigate the similarity between images 10.M. Ye used the reordering method, and the similarities between the sorted images were also used to estimate the local similarity between the images 11.However, the methods of manifold learning and reordering are mostly unsupervised, which makes it difficult to evaluate the effectiveness of the methods.

In recent years graph neural networks (GNNs) have received attention from researchers for their strong ability to generalize graph data 12. GNNs deliver messages in the structure of a graph, and by decomposing the graph, the final representation of the nodes can be obtained and the classification of the nodes can be performed. GNNs use graph node representation, which makes training end-to-end and facilitates the learning of feature representations compared to stream learning and rearrangement order. The network combines graph computation and deep learning to obtain a deep learning framework with robust similarity estimation and recognition.

Thus, this paper proposes the use of GNN model to achieve forest fire recognition under multi-view images. For a small batch of images consisting of multiple images, the initial visual features and images of the learned images are first learned in pairs supervised; then, each pair of images is processed as a node on the graph, which is responsible for generating the similarity score of the graph. In addition, the pairwise relational features associated with each node are updated and optimized by propagating deeply learned messages among the nodes. Based on this, image recognition is performed using feature fusion weights to obtain robust similarity estimates for images from different viewpoints.

The main outcomes of this thesis are as follows.

1) A graph neural network model based on multi-view image similarity (MVSGNN) is proposed, which generates a graph to represent the pairwise relationships (nodes) between images based on images from different viewpoints, and uses the updated feature relationships of the nodes to estimate the similarity between images. Since this model can synthetically examine the information from images from different viewpoints, it can improve the recognition rate of forest fires.

2) In order to reduce the complex pre-processing process of images and extract key features from images effectively, a fire region feature extraction method based on image segmentation is proposed. By using the extracted features as the input of MVSGNN model, it helps to improve the robustness of identifying forest fires.

3) The experimental results on different fire datasets show that the method proposed in this paper can better identify forest fires in different scenes with strong generalization ability and anti-interference capability.

2.Related Work

Most forest fire recognition algorithms use computer vision techniques for forest fire recognition in terms of color features, texture features and motion features of flames. 13studied the dynamic behavior and irregularity recognition of fires in RGB and HSI color spaces now. 14used the property of separation of color components from luminance in YCbCr space to design classification rules. 15investigated the shape of flames and the motion of rigid objects, and proposed to intelligently extract features using optical flow information and flame behavior to distinguish flames accordingly. 16combined shape, color and motion attributes to form a multi-expert system framework for real-time flame recognition. 17found experimentally that flames in HSV color space show lower chromaticity. In18, based on the RGB color model, the flame pixel points are first extracted, and then the flames are recognized based on their growth and disorder features.19 calculated the motion direction of the fire by fast estimation method and accumulated the motion direction to time to identify the fire according to the fire spreading characteristics, and proposed a fire identification algorithm based on spectral, spatial and temporal features and fuzzy logic features, and designed a real-time forest fire alarm system based on this.

With the continuous development of deep learning technology, the combination of computer vision technology and deep learning for fire identification has become a new idea. For example, 20designed a convolutional neural network for forest fire recognition, and used the alternative random initialization parameter method for the problem of small training sample size in the network training process, and achieved a better fire classification effect.21 combined traditional recognition methods with neural networks, and firstly used AdaBoost and LBP (Local Binary Pattern) algorithms for initial recognition of images to extract flame candidate areas, and then used convolutional neural networks for feature extraction and classification of candidate areas.22 applied deep confidence network for flame recognition. 23 trained ResNet network 24 using deviated data by exploiting the quantitative difference between fire images and normal images, and then used the network to recognize flames. 25proposed a multilayer noise reduction automatic coding network algorithm and applied it to more than a dozen different scenarios including forest fires. 26proposed a cascaded convolutional neural network algorithm, which uses two independent convolutional networks to identify static and dynamic features of flames separately and combines the results of the two networks to determine whether they are flames or not. 27designed DnCNN networks to recognize flame images and compared them with networks such as VGG and ZF-Net. These models are able to accurately recognize the same viewpoint image, while the recognition of multiple source samples is not accurate. For this reason, 28 introduced that GNNs can effectively use intergraph relational information to improve image recognition. Bruna et al. 29 proposed two constructions of deep convolutional networks on graphs (GCN), one is based on the spectrum of graph Laplacian, which is called spectral construction. Another is spatial construction, which extends properties of convolutional filters to general graphs and uses the spatial structure GCN for disaster behavior recognition.

Most of the current work is based on images from the same perspective to identify forest fires. Although it is more accurate in extracting features, the images from the same perspective may not be able to identify forest fires sometimes. Therefore, it is more accurate to identify forest fires through images from multiple perspectives. The method using the training data labels to supervise, can more accurately in the messaging in figure characteristics of fusion weights, thereby effectively fire image recognition under different Angle of view, but to be more accurate feature fusion, as far as possible to extract more accurate and original features of the integration of figure messaging similarity score, so the following describes how to extract dynamic characteristics of the forest.

3 Dynamic characteristics

Since forest fire is continuously developing, the dynamic characteristics of fire features from continuous video frames are very important for fire detection, which makes it possible to identify fire from the similar objects. The dynamic characteristics of the fire area segmented from video images are described in the following three aspects.

3.1 Image segmentation

The HSV (hue, saturation, value) model provides a more humane way of describing color than the RGB color model. The way of neural network perceives color is closely related to the HSV component29. The HSV color space can be defined as

$$V = \{x|\ x(H) \in [0,429], x(S) \in [0,265], x(V) \in [0,265]\} \qquad (1)$$

The $x(h), x(s), x(v)$ are the H, S, and V component values of x, the pixel in the HSV color space is x, and, respectively. Thus, we can obtain the fire color distribution from the sample image containing the forest fire region, whose sample color values form the pixel component values as shown in Figure 1. Gaussian mixture model was used to represent the fire shape, and the pixel points whose colors are within the range of the distribution model are used as fire pixel points.

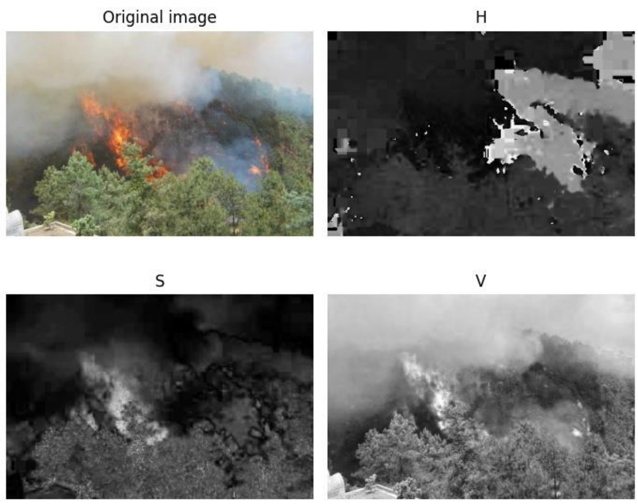

Figure 1. H、S and V component display

To further reduce the computational effort, three 2D projection planes are used instead of the 3D distribution model, the color of the flame on the fireplace sample is projected on the HS, HV and SV planes. In each plane, the extent of the color distribution can be easily represented by one or two rectangles, so that a relatively simple 2D color distribution can be defined.

$$Vx = \{x|\ x(H) \in [0,37] \cup [335,429], x(S) \in [69,265]\} \qquad (2)$$

Based on the color range, the image is segmented and candidate fire areas are obtained, as shown in Figure 2, which can clearly segment the forest fire scene.

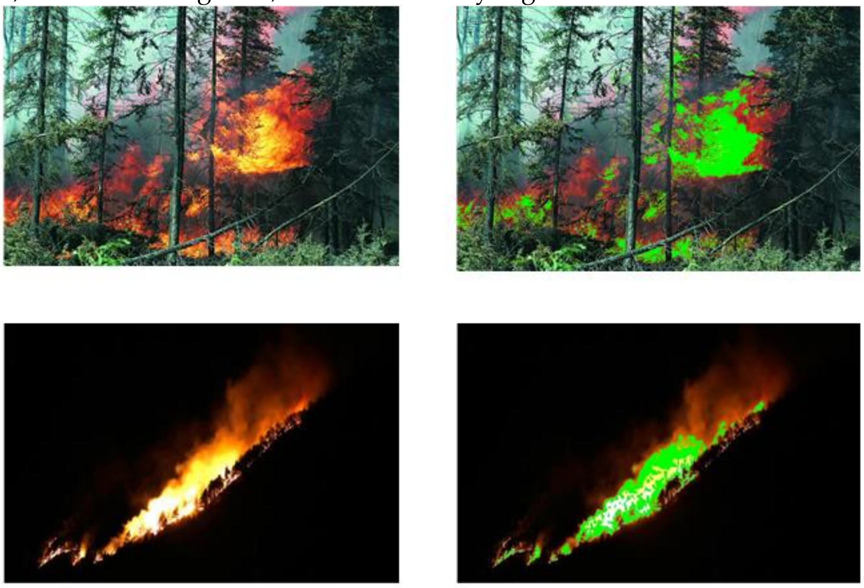

Figure 2. Color-based fire area segmentation

3.2 Extraction of fire features

In this subsection, features such as area, roundness and contour of the fire area are acquired for the fire area segmented in the image. Forest fires are initially unstable flames and the number of fire pixels increases with the fire area, so area is an important feature of fires 30. To identify the degree of area variation of a fire, the change in size of the fire area can be calculated from two consecutive images. If the result exceeds a predefined threshold, the fire growth is judged to occur. However, due to the complexity of the fire shape, it is difficult to achieve complete feature extraction by area alone, so boundary chain codes and roundness are added to refine the fire features.

Boundary chain code and roundness: given a segmented fire area, using Laplace operator to retrieve its boundary, then its connected boundary chain code can be easily retrieved, whereby the perimeter of the boundary L can be easily calculated. the roundness of the fire area is calculated from its perimeter and area, which can indicate the quality of the fireplace space form, i.e., the additional complicated the form, the larger its price. And the roundness helps to get rid of the recognition interference of irregular bright objects in the early fire recognition.

Contour: Since the shape of the fire area varies due to air flow, the degree of its fire can be measured by calculating the contour undulation , assuming that there are N points on the boundary and they are in the plural form $\{z_i | z_i = x_i + jy_i\}$, where, (x_i, y_i) is the coordinate of the ith point of the fire zone boundary crossed clockwise. The discrete Fourier transform of z_i is obtained as:

$$F_w = \frac{1}{M} \sum_{i=1}^{M} z_i \exp\left(-j \frac{2\pi}{M} iw \right) \tag{3}$$

where Φ represents the center of gravity of the one-dimensional boundary. According to reference 417 only a few dozen Fourier coefficients are really needed to describe the profile, and based on experience the first 32 are chosen $D=(|\Phi_1|_2,|\Phi_2|_2,\cdots,|\Phi_{32}|_2)$ The difference of two consecutive Fourier is :

$$D_i = \sum_{w=1}^{32} \left\| \Phi_w^{i'} \right\|_2 - \left\| \Phi_w^{i-1} \right\|_2 \tag{4}$$

If Di it is greater than Td and lasts longer than Tm, where Td and Tm are statistical thresholds from the experiment, it means that a drastic change in shape has occurred and a fire may have occurred.

3.3 Dynamic characteristics of fires

With the spread of forest fires, the dynamic characteristics of continuous fire image features are important for fire detection24. We define a dynamic characteristic containing n continuous images. In order to ensure real-time fire detection, n should be a relatively small number. In general, the characteristic frequency of flame flicker is about 10 Hz and the recorded video has 30 frames per second. Based on the real scene requirements, the value of n is set to 5 to define the dynamic characteristics for every 5 consecutive images of the fire features. Therefore, construct associate $n \times m$ matrix for the flame options within the image, set range the quantity of consecutive pictures $n=5$ and therefore the number of flame options (area, conformation and contour) $m=3$, roundness and contours. Assume that $\zeta(i,j)$ is an element of the matrix corresponding to the i[th] image and the j[th] flame feature, based on which the dynamic properties are represented by the mean and mean squared deviation.

$$\psi(j)=\frac{1}{n}\sum_{i=1}^{n}\zeta(i,j) \tag{5}$$

$$\xi(j)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\zeta(i,j)-\psi(j))^2} \tag{6}$$

Thus, for any forest fire image, there are associated dynamic features, i.e., the mean and mean squared deviation of the image matrix. In the machine learning model of this paper the above image segmentation information, fire features and fire dynamic features are fed into the model as auxiliary information along with the forest fire images for training.

4 The proposed graph neural network

To evaluate the forest fire recognition algorithm, the test dataset is divided into a detection set and a probe library image set. Given a pair of detection images and image pairs with different viewpoints, the goal of the forest fire recognition model is to robustly determine the visual similarity between the detection images. In this paper, the model is trained in small batches, and the detection images and probe library images of different image pairs are evaluated separately in the setup, i.e., the estimated similarity between a pair of images, so that it will be independent of the influence of other image pairs. The graph neural network model based on multi-view image features proposed in this paper is given in Figure 3. It takes a detector and multiple library images as inputs to create a graph, and each node models one detector image-probe library image. It outputs the similarity score of each probe library image. During end-to-end training, the deep learning information is propagated among the nodes to update the relational features associated with each node to obtain more accurate similarity score estimates.

4.1 Graphical representation and node characteristics

In this paper MVSGNN framework, Each node represents a pair of probe library images, a probe image library and M images are given to construct an undirected complete graph $G(V, E)$, assuming that $V=\{v_1, v_2, \ldots, v_n\}$ is the set of nodes consisting of probe library image pairs. First, we need to estimate the similarity score of each probe library image response. Generally speaking, the input of any node is encoded between its corresponding probe library images. The scheme in this paper acquires the input relational features as shown in Figure 3a. Each input image will be fed into the CNN for pairwise relational feature encoding when we give a probe image and M images. In this paper we use ResNet-50 27 as CNN structure. The last global average ensemble feature of the two images in ResNet-50 is element-wise subtracted to obtain two-two correlation features. The two-two features are processed into differential features d_i $(i=1,2,...,N)$, it's not only the deep visual relationship between the encoding probe and the i-th image but also input features for the i-th node on the graph. Since the task on the graph is node classification. In the linear classifier, since node classification is a complex task that requires inputting the input features of each node and obtaining the output similarity score, the pairwise relationship between nodes is not considered. The loss function of the model in this paper is in the form of cross-entropy as follows.

$$L=-\sum_{i=1}^{N} y_i \log\left(f\left(d_i\right)\right)+\left(1-y_i\right)\log\left(1-f\left(d_i\right)\right)$$

(7)

where denotes the classifier as a sigmoid function 28. denotes the label of the i^th^ detector library image pair, and 1 indicates that the detector and the i^th^ library image belong to the same identity.

Figure 3 depicts the basic model and the deep messaging implementation of the graph architecture in this paper. The basic model in Figure 3a can be used not only to obtain the similarity of detector image pairs for deep message passing and to update the relational features of detector image pairs. The similarity of detector image pairs can also be calculated. Figure 3b in order to deliver more effective messages, the image relationship features di is first fed into a two-layer messaging network for feature encoding. Use the similarity score of the detection gallery, the detection gallery relational features are fused to derive the message passing and feature fusion scheme, the objective function is shown in equation 10.

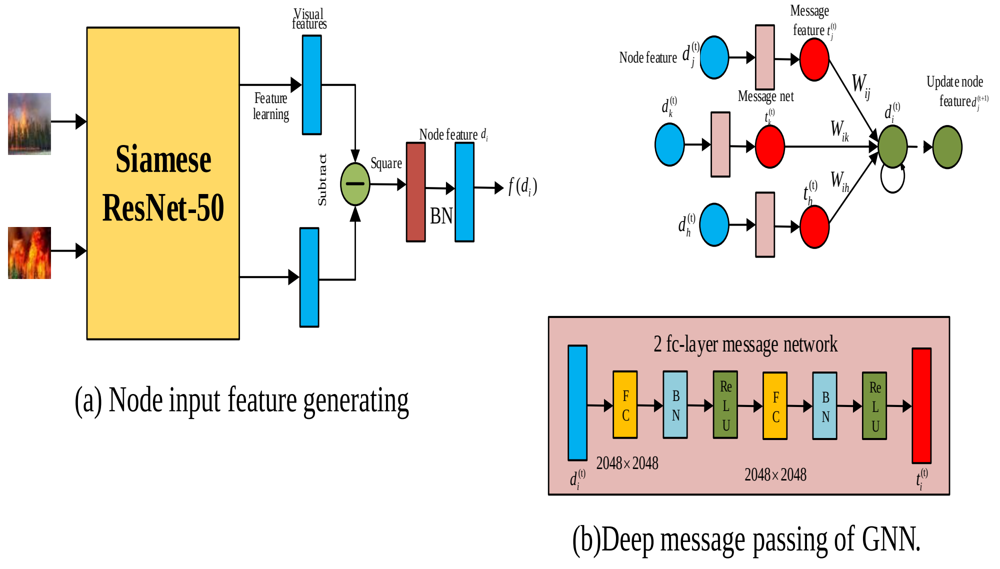

Figure 3. Basic model and proposed graph architecture

4.2 Similar guidance

Clearly, the simple node classification model (Equation (7)) ignores the valuable information between different probe library pairs. In order to utilize this important information, it is necessary to create edges E on the graph G. G is fully connected and E denotes the set of relationships between different probe library pairs, where Wij is a scalar edge weight. It denotes the importance of the relationship between node i and node j, which can be calculated as

$$W_{ij} = \begin{cases} \dfrac{\exp\left(S\left(g_i, g_j\right)\right)}{\sum_j \exp\left(S\left(g_i, g_j\right)\right)}, & i \neq j \\ 0, \quad i = j \end{cases} \tag{8}$$

Where g_i and g_j are the i^{th} and j^{th} images. $S\left(g_i, g_j\right)$ is a pairwise similarity estimation function that estimates the similarity scores between g_i and g_j is modeled in the same way as the node classification model.

To enhance the pairwise relationship features of nodes with other nodes' information, deeply learned messages are propagated among all connected nodes. The node features are then updated to a weighted sum fusion of all input messages and the original features of the nodes. Before message delivery starts, each node first encodes the deep message to be sent to other nodes connected to it. The input relational features of the nodes are fed into a message network that has two fully connected layers containing BN and ReLU 31to generate deep message, as shown in Figure 3(b). This process learns messages that are more suitable for node-relationship feature updates,

$$t_i = F\left(d_i\right) \text{ for } i = 1, 2, \ldots, N \tag{9}$$

where F denotes the 2-level FC subnet used to learn the propagated deep messages. After obtaining the edge weights W_{ij} and deep messages t_i of each node, the update scheme of the node relationship features d_i can be expressed as:

$$d_i^{(1)} = (1-\alpha) d_i^{(0)} + \alpha \sum_{j=1}^{N} W_{ij} t_j^{(0)} \quad \text{for } i = 1, 2, \ldots, N \tag{10}$$

where $d_i^{(1)}$ denotes the ith image relational feature, $d_i^{(0)}$ denotes the ith input relational feature, and $t_j^{(0)}$ denotes the deep message from node j. α denotes the weighting parameter that balances the fused features and the original features. The weighted fusion of relational features can be performed iteratively as follows

$$d_i^{(t)} = (1-\alpha) d_i^{(t-1)} + \alpha \sum_{j=1}^{N} W_{ij} t_j^{(t-1)} \quad \text{for } i = 1, 2, \ldots, N \tag{11}$$

where t is the number of iterations. The refined relational feature $d_i^{(t)}$ can replace the relational feature d_i in equation (7) for loss function calculation and GNN training. The training equation (11) can be back-propagated for framework structure and model update.

5 Forest fire data production

There are few open-source image datasets related to forest fire recognition and the clarity of these images is low. In this paper, a large number of videos of forest fires are collected and produced into a supervised forest fire recognition dataset by manual annotation, and the dataset is open-sourced for use by subsequent researchers. The production of this dataset requires the implementation of crawling the forest fire image website, and then cutting out and saving the fire part to achieve the fire dataset.

5.1 Related Technologies

Since there are a large number of forest fire images on the Internet, we can obtain image resources on the Internet through crawling techniques.

Crawling is a technique that automatically obtains corresponding information or resources on the Internet according to a certain purpose. This paper uses a crawling technique based on the Python language, where the toolkits used for crawling in this paper are Requests, Beautiful Soup18.

OpenCV32is a library of open source (API) functions for computer vision and this paper uses its Python interface, where Cascade and Classifier are cascaded classifiers for target recognition in OpenCV, which is used by using Local Binary Pattern (LBP) to import specific classifier file, such as an image classifier recognition of targets 24.

5.2 Concrete implementation

The most critical part of the process of producing forest fire image dataset is the need to reduce the difficulty of manual rechecking while maintaining quality and speed, i.e., the results of processing using the program can only contain a very small number of non-fire images, so the filtering module is designed in this paper. The process of making forest fire image dataset in this paper consists of four steps, the first is image collection using crawlers, the second is fire recognition module, which uses OpenCV module to realize cropping of fires, and the third is screening module, which uses DLib module to filter the cropped images to remove unqualified images.

The specific method is shown in Figure 4, also the content of the finished dataset is shown in Figure 5.

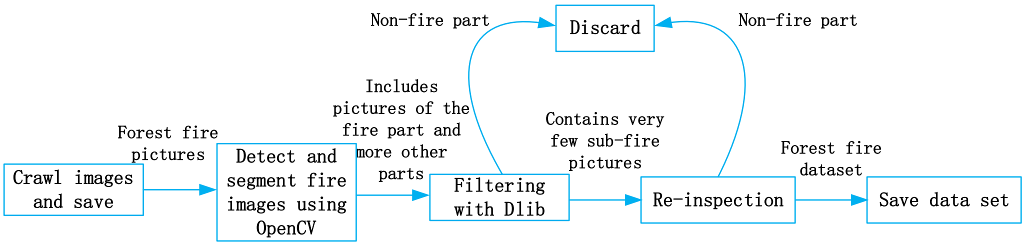

Figure 4. Forest fire dataset production process

Figure 5. Content of forest fire dataset

6 Experiments

6.1 Data set

In this paper, we use the technique proposed in Section 5 to collect a total of 2826 forest fire images on the Web, including images at the beginning of the fire and images while it is burning, and an additional the forest fire dataset 932 non-forest fire images.

The xBD dataset 32 is one of the public datasets of high-resolution satellite images 34with annotations. The natural disaster image dataset updated by MIT, encompasses 19 disaster events and contains the number: 22068 images with an image resolution of 1024 × 1024, where each building has an identifier. We just use this dataset to contain data on forest fires.

6.2 Setup

The models in this paper are implemented using Keras and TensorFlow frameworks, and its implementation platform running environment configuration operating system is

Ubuntu 19.04, GPU is Geforce GTX 1080Ti, Intel i710500U processor, 16 GB memory RAM,1 TB hard disk.

In this paper, we use deep learning models for comparison, ResNet and DenseNet 29, and set the learning rate to 0.01 and the batch size to 64.

Our proposed MVSGNN is based on ResNet-50 for forest fire recognition. All input images are adjusted to 256×128. the base CNN model is first pre-trained with an initial learning rate of 0.01 set on all datasets, and the learning rate is reduced by an element of 10 once 50 epochs, then coaching rate is mounted for 50 training cycles. The weights of the linear classifier used to obtain the image similarity were initialized using the weights of the linear classifier trained in the base model training phase, and the model was optimized using Adam 24 with the weighting parameter α set to 0.9.

6.3 Results

The number of parameters, training time and test results of each model are shown in Table 1.

Table 1. Test results of different models

	xBD Dataset			Forest fire dataset		
Model	ResNet	DenseNet	MVSGNN (our)	ResNet	DenseNet	MVSGNN (our)
Number of layers	50	50	50	50	50	50
Total parameters	23602904	1477058	768626	21702355	1142587	548633
Training time/s	724.94	1153.9	867.8	701.35	1023.44	804.83
Accuracy/%	94.33	96.15	98.33	95.67	97.85	99.02

As can be seen from Table 1, the training time of our model on the two data sets is much less than that of DENSENET. In this paper, images from different perspectives or multiple sources can be effectively used, and the model can be rapidly converged through data accumulation. However, with the same training time as RESNET, the number of parameters in this paper is not as large as that in RESNET, and the accuracy of MVSGNN is 4% higher than that of RESNET. This is because we took the lead in extracting dynamic features from forest fire pictures, so that the model can quickly learn deep information of the pictures. MVSGNN model is also used to guide the similarity of forest fire images from different sources or perspectives. At the same time, the total parameters of our architecture are relatively minimal, reducing the memory overhead. Our model has achieved good accuracy on different training sets, so the model has not produced any fitting. The proposed scheme has strong generalization ability and robustness, and can effectively identify forest fire images from different perspectives, making the proposed scheme able to meet the requirements of forest fire monitoring in different scenes.

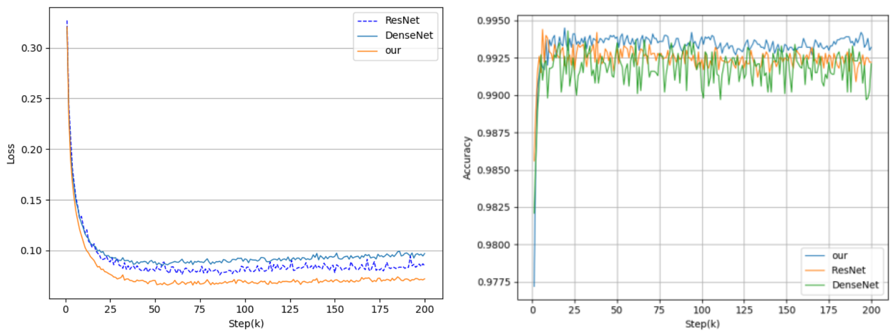

Figure 6. Loss and accuracy during model training

7 Conclusion

In this paper, a graph neural network based on the similarity of forest fire images from different perspectives is proposed, which integrates rich similarity information of gallery into the training process of forest fire image recognition to realize the estimation of similarity between images, so as to ensure that the weighted depth information fusion feature is more effective. In addition, a method of dynamic feature of fire region that can be separated from the image is proposed, which reduces the complex preprocessing process and effectively extracts the key features from the image, thus improving the robustness of forest fire discrimination by the network. We have also contributed a dataset of forest fire imagery that we have produced. Experimental results of the proposed method and several deep learning methods in fire datasets show that our method is adaptable to different fire scenarios and has strong generalization ability and anti-interference ability.

In the future, we plan to design fire identification and monitoring systems with perspective dynamics, for example, deploying drone patrols to regularly patrol and monitor open fires in forest areas via drones.

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

Reference:

1. Zhang J, Zhu H, Wang P, et al. ATT Squeeze U-Net: A Lightweight Network for Forest Fire Detection and Recognition[J]. IEEE Access, 2021, PP(99):1-1.
2. Wu X , Cao Y , Lu X , et al. Patch wise dictionary learning for video forest fire smoke detection in wavelet domain[J]. Neural Computing and Applications, 2021:1-13.
3. Andries, Heyns, Warren, et al. Optimisation of tower site locations for camera-based wildfire detection systems[J]. International Journal of Wildland Fire, 2019, 28(9):651-665.
4. Khan Muhammad . Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications [J]. IEEE Transactions on Systems Man Cybernetics-Systems, 2019, 7(1):1419-1434.
5. Barmpoutis P , Papaioannou P , Dimitropoulos K , et al. A Review on Early Forest Fire Detection Systems Using Optical Remote Sensing[J]. Sensors, 2020, 2020(20):6442.
6. C. Emmy Prema . Efficient Flame Detection Based on Static and Dynamic Texture Analysis in Forest Fire Detection [J]. Fire Technology , 2018(54):255-288.
7. Liu, ZC, Wang, CY,Zhang, K. Research on the identification method for the forest fire based on deep learning[J]. OPTIK,2020,223:1-8.
8. Wang YB , Dang LF ,Ren JY. Forest fire image recognition based on convolutional neural network[J]. Journal of Algorithms & Computational Technology,2019(13):1-11
9. HICINTUKA, Jean, Philippe, et al. Video Based Fire Detection Systems on Forest and Wildland Using Convolutional Neural Network[J]. Journal of Donghua University(English Edition), 2019, v.36(02):40-48.
10. C. C. Loy, C. Liu, and S. Gong. Person re-identification by manifold ranking. IEEE International Conference , 2013:3567–3571.
11. M. Ye, C. Liang, Y. Yu, Z. Wang, Q. Leng, C. Xiao, J. Chen, and R. Hu. Person reidentification via ranking aggregation of similarity pulling and dissimilaritypushing. IEEE Transactions on Multimedia, 2016,18(12):2553–2566.
12. Dasari P , Reddy G K J , Gudipalli A . Forest fire detection using wireless sensor networks[J]. International Journal on Smart Sensing and Intelligent Systems, 2020, 13(1):1-8.
13. Srinivasa P , Sri R T , Aravind D S , et al. Forest Fire Detection Using Convolution Neural Networks[J]. Advanced Science Letters, 2020, 26(6):691-696.
14. Wu Z, Shen C, Van Den Hengel A. Wider or deeper: Revisiting the resnet model for visual recognition[J]. Pattern Recognition, 2019, 90: 119-133.
15. Kanniga E , Jadhav P S . A Study Paper on Forest Fire Detection Using Wireless Sensor Network[J]. International Journal of Psychosocial Rehabilitation, 2019, 23(4):397-407.
16. Abbassi M A E , Jilbab A , Bourouhou A . Efficient Forest Fire Detection System Based on Data Fusion Applied in Wireless

Sensor Networks[J]. International Journal on Electrical Engineering and Informatics, 2020, 12(1):1-18.
17. Qin L , Wu X , Cao Y , et al. An Effective Method for Forest Fire Smoke Detection[J]. Journal of Physics: Conference Series, 2019, 1187:052045-.
18. Kingma D P, Ba J. Adam: A Method for Stochastic Optimization[C]//ICLR (Poster). 2015.
19. Jeong S W , Yoo J . i-FireNet: A Lightweight CNN to Increase Generalization Performance for Real-Time Detection of Forest Fire in Edge AI Environments[J]. Journal of Institute of Control, 2020, 26(9):802-810.
20. Sun X , Sun L , Huang Y . Forest fire smoke recognition based on convolutional neural network[J]. Journal of Forestry Research, 2020:1-7.
21. Heyns A , Plessis W P D , Kosch M , et al. Optimisation of tower site locations for camera-based wildfire detection systems[J]. International Journal of Wildland Fire, 2019, 28(9):651-665.
22. Attri V , Dhiman R , Sarvade S . A review on status, implications and recent trends of forest fire management[J]. Archives of Agriculture and Environmental Science, 2020, 5(4):592-602.
23. Cuomo V , Lasaponara R , Tramutoli V . Evaluation of a new satellite-based method for forest fire detection[J]. International Journal of Remote Sensing, 2001, 22(9):1799-1826.
24. Alkhatib, Ahmad A A . A Review on Forest Fire Detection Techniques[J]. International Journal of Distributed Sensor Networks, 2014, 2014:103-110.
25. Begoña C. Arrue, Aníbal Ollero, Dios JRMD . An intelligent system for false alarm reduction in infrared forest-fire detection[J]. Intelligent Systems & Their Applications IEEE, 2000, 15(3):64-73.
26. Wang L , Qu J J , Hao X . Forest fire detection using the normalized multi-band drought index (NMDI) with satellite measurements[J]. Agricultural & Forest Meteorology, 2008, 148(11):1767-1776.
27. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770–778.
28. Wang, SD, Miao, et al. An Improved Algorithm for Forest Fire Detection Using HJ Data[J]. PROCEDIA ENVIRON, 2012, 2012,13(-):140-150.
29. J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun. Spectral networks and locally connected networks on graphs. arXiv preprint arXiv:1312.6203, 2013.
30. TianY Zhang,HM Hu,B Li. A Naturalness Preserved Fast Dehazing Algorithm Using HSV Color Space[J] .IEEE Acess,2018(6):10644-10649.
31. Giuseppe Marbach, An image processing technique for fire detection in video images[J]. Fire Safety Journal , 2006:285-289.
32. Yuan C , Liu Z , Zhang Y . Vision-based forest fire detection in aerial images for firefihting using UAVs[C]. International Conference on Unmanned Aircraft Systems. IEEE, 2016.
33. Culjak I, Abram D, Pribanic T, et al. A brief introduction to OpenCV[C]//2012 proceedings of the 35th international convention MIPRO. IEEE, 2012: 1725-1730.
34. Saul'Skii V K . Choosing the Structure of Satellite Systems for Meteorology and Forest Fire Detection on the Basis of a Vector Model of Surveying the Earth[J]. Cosmic Research, 2020, 58(4):295-306.
35. Weber E, Kané H. Building disaster damage assessment in satellite imagery with multi-temporal fusion[J]. arXiv preprint arXiv:2004.05525, 2020.