

Fire image detection and classification analysis based on VGG16 image processing model

Fengjun Hou

College Of Engineering, China Agricultural University, Beijing 100083, China

hfj@cau.edu.cn

Abstract. Fire image classification technology refers to the classification and recognition of fire images through computer vision technology in order to take timely countermeasures. With the development of computer vision technology, fire image classification technology has been widely studied and applied. Deep learning techniques have achieved great success in the field of image classification, with researchers classifying and identifying fire images by using deep learning models such as convolutional neural networks. This paper introduces the research background and application of fire image classification technology, and the experimental results of detection and classification analysis of fire image based on vgg16 image processing model. The model can classify and identify fire images well and achieve good prediction effect. The accuracy of training set and test set is stable at 99%, the accuracy and accuracy of the model reach 98%, and the recall rate and F1 score reach 99%. The application of fire image classification technology is of great significance. By using computer vision technology to classify and identify fire images, it can improve the accuracy and efficiency of fire monitoring and early warning system, timely detect and control fire, and reduce the loss caused by fire. At the same time, the continuous development of deep learning technology also provides a broader space for the research and application of fire image classification technology.

Keywords: Vgg16, Fire images, Accuracy.

1. Introduction

Fire image classification technology is a crucial aspect of fire detection and control, which is necessary to minimize the damage caused by this natural disaster to people and property. With the advancement of computer vision technology, researchers have extensively studied and applied various techniques for fire image classification [2]. One of the most commonly used methods is image classification models, such as deep learning models like convolutional neural networks (CNN) [3]. These models have the capability to automatically extract features and classify images, thereby enhancing the accuracy and efficiency of fire image classification [4]. Feature extraction is another widely used image classification method that involves the extraction of local features to classify images. Researchers have employed methods such as SIFT, HOG, and others to classify and recognize fire images, which have significantly improved the accuracy and robustness of fire image classification [5,6,7]. Machine learning technology has also been widely used in fire image classification, where researchers have utilized machine learning algorithms such as support vector machines (SVM) and random forests to classify and identify fire

images [8]. These algorithms can automatically learn the features of flame images and classify and recognize them with high accuracy [9,10].

Building on the above research, this paper proposes an image processing model based on vgg16 to detect and classify fire images, which provides a strong foundation for future research in this field. By adopting this model, the accuracy and efficiency of fire image classification can be further enhanced, leading to more effective fire detection and control measures.

2. Data set visualization

This article chooses is open source data set fire (<https://www.kaggle.com/phylake1337>), used in computer vision tasks of outdoor fire images and the images. The data is divided into 2 folders, the fire_images folder containing 755 outdoor fire images, some of which contain heavy smoke, and the other non-fire_images folder containing 244 natural images.

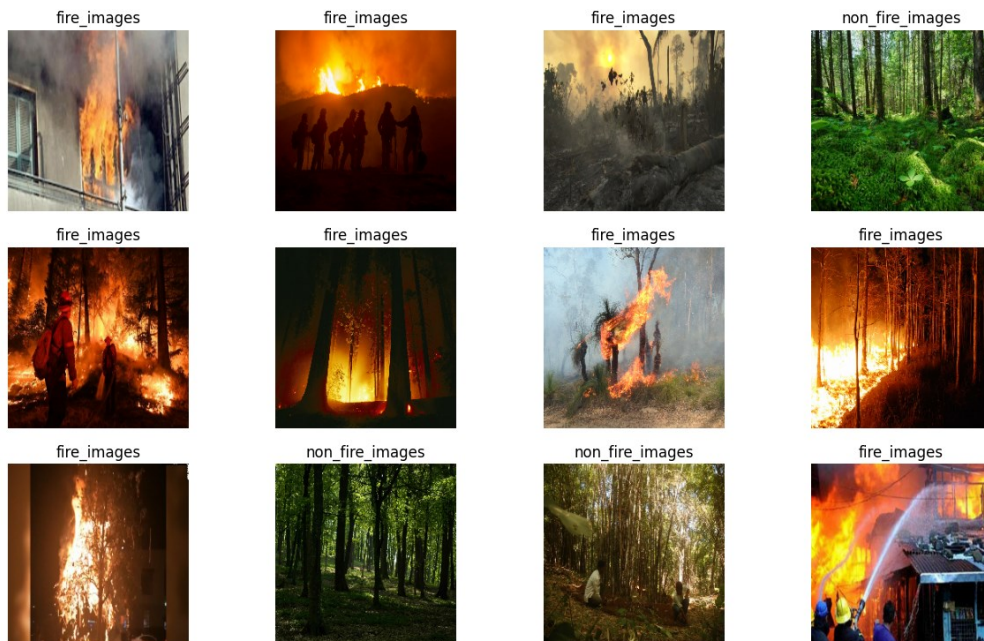


Figure 1. Data set visualization.
(Photo credit: Original)

3. Model principle and model structure of VGG16

VGG16 is a deep convolutional neural network model proposed by a research team at the University of Oxford, which achieved first place in the ImageNet image recognition challenge in 2014. The VGG16 model consists of 16 layers, including 13 convolutional layers and 3 fully connected layers. The unique feature of the VGG16 model is its use of very small convolutional kernels (3x3) and a deeper network structure, which enables it to achieve outstanding performance in image classification and feature extraction tasks.

The input of the VGG16 model is a 224x224 pixel color image, which is processed through a series of convolutional layers and pooling layers. Finally, the model outputs a 1000-dimensional vector, indicating which of the 1000 categories the image belongs to. During this process, the VGG16 model employs several special techniques to improve its accuracy and generalization ability.

The first technique used in the VGG16 model is the use of small convolutional kernels. By using small kernels, the model can reduce the number of parameters and increase the non-linearity of the network, which in turn improves the model's ability to learn complex features from images.

The second technique used in the VGG16 model is the use of deeper network structures. By increasing the depth of the network, the model can learn more complex features and achieve higher accuracy in image recognition tasks.

The third technique used in the VGG16 model is the use of pooling layers. Pooling layers are used to downsample the feature maps generated by the convolutional layers, which reduces the spatial dimensionality of the feature maps and makes them more computationally efficient.

The fourth technique used in the VGG16 model is the use of dropout regularization. Dropout is a technique used to prevent overfitting by randomly dropping out some of the neurons during training, which forces the model to learn more robust features and reduces the risk of overfitting.

Overall, the VGG16 model is a highly effective deep learning model for image classification and feature extraction tasks. Its use of small convolutional kernels, deeper network structures, pooling layers, and dropout regularization make it a powerful tool for a wide range of image recognition applications and we will introduce the model principle of the VGG16 model in detail below.

3.1. Convolution layer

The VGG16 model's convolutional layers consist of 13 convolutional layers and 5 max pooling layers. These layers work together to transform the input image into a set of feature vectors that contain various features such as edges, textures, and colors.

During the convolutional layers, the VGG16 model uses very small convolutional kernels (3x3) to perform convolutions. These small kernels are effective in extracting local features from the image, and by combining multiple kernels, the model can extract more complex features. In the pooling layers, the VGG16 model uses 2x2 max pooling layers to reduce the dimensionality of the feature maps, which reduces the computational cost and the number of parameters. The combination of these convolutional and pooling layers enables the VGG16 model to extract rich image features, which improves the accuracy of the model.

The VGG16 model's convolutional and pooling layers work together to extract a wide range of features from the input image. By using small convolutional kernels and max pooling layers, the model can extract both local and global features, which makes it highly effective in image classification and feature extraction tasks.

3.2. Fully connected layer

After the convolutional and pooling layers, the VGG16 model uses three fully connected layers for classification. These fully connected layers transform the feature vectors into the final classification results. During this process, the VGG16 model employs several special techniques to improve its accuracy and generalization ability.

The first technique used in the VGG16 model is Dropout, which is used to prevent overfitting. Dropout is a technique that randomly sets some of the neuron outputs to zero during training, which prevents the model from overfitting to the training data. In the VGG16 model, Dropout is applied to the fully connected layers, which effectively improves the model's generalization ability.

The second technique used in the VGG16 model is Batch Normalization, which speeds up the training process. Batch normalization is a normalization technique that normalizes the inputs of each layer in the neural network, which speeds up the training process. In the VGG16 model, Batch Normalization is applied to both the fully connected layers and convolutional layers, which effectively improves the model's training speed and accuracy.

The last technique used in the VGG16 model is the Softmax function, which is used for classification. The Softmax function is a commonly used classification function that converts the feature vectors into a probability distribution, indicating the probability of the image belonging to each category. In the VGG16 model, the Softmax function is applied to the last fully connected layer, which generates the probability distribution for each category.

Overall, the VGG16 model's fully connected layers are responsible for the final classification of the input image. By employing Dropout, Batch Normalization, and Softmax function, the VGG16 model

can effectively prevent overfitting, speed up the training process, and generate accurate classification results. The combination of these techniques makes the VGG16 model one of the most powerful deep learning models for image classification and feature extraction tasks.

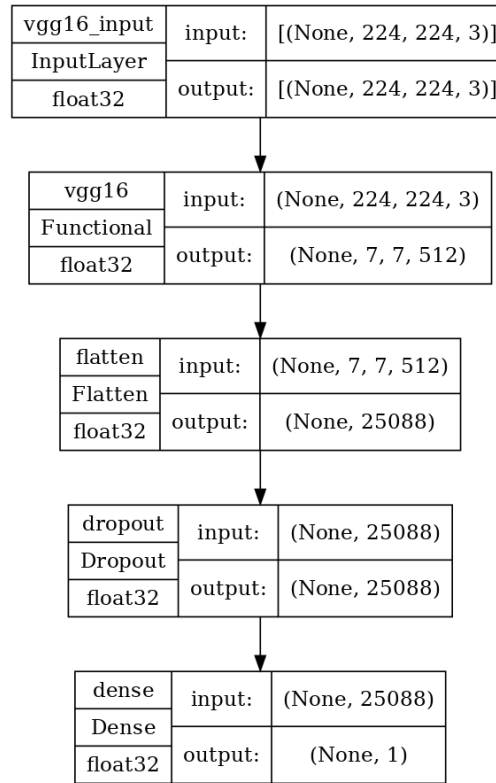


Figure 2. Model structure of VGG16.
(Photo credit: Original)

4. Model principle and model structure of VGG16

The VGG16 model is a deep convolutional neural network that is commonly used in image classification tasks. It consists of 13 convolutional layers and 3 fully connected layers, which are used to extract features from the input image and classify it into one of the predefined categories.

During the training process, the VGG16 model uses the cross-entropy loss function to calculate the error between the predicted output and the actual output. The cross-entropy loss function is commonly used in classification tasks because it can measure the gap between the predicted results of a model and the real results. The output of the Softmax function is used as the input to the cross-entropy loss function, which calculates the gap between the predicted probabilities and the actual probabilities.

To update the model parameters, the VGG16 model uses the backpropagation algorithm. Backpropagation is a common algorithm used in neural networks to calculate the gradients of the loss function with respect to the model parameters. These gradients are then used to update the model parameters using an optimization algorithm.

In the case of the VGG16 model, the stochastic gradient descent algorithm is used to update the model parameters. Stochastic gradient descent is a popular optimization algorithm that updates the model parameters by taking small steps in the direction of the negative gradient of the loss function. This algorithm is applied to both the fully connected and convolutional layers of the model.

By constantly adjusting the model parameters using the stochastic gradient descent algorithm, the accuracy of the VGG16 model can be improved over time. This allows the model to better classify images into the correct categories, making it a powerful tool for image classification tasks. The result of training 50 epochs is shown below:

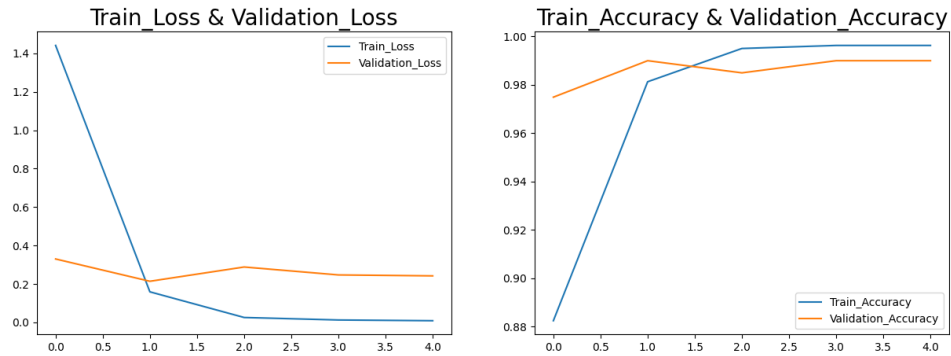


Figure 3. Loss and accuracy.
(Photo credit: Original)

Table 1. Model evaluation parameter.

Evaluation parameter	Accuracy	precision	recall	f1-score
Training set	0.99	0.99	0.99	0.99
Test set	0.98	0.98	0.99	0.99

As can be seen from the loss curve, the loss of the training set and the test set gradually decreased until convergence, and the accuracy of the training set and the test set gradually increased and stabilized at 99%. Finally, the test set was used to test, and the accuracy and accuracy of the model reached 98%, the recall rate and F1 score reached 99%, and the fire image was very accurately classified. The confusion matrix of classification effect is shown in the figure below:

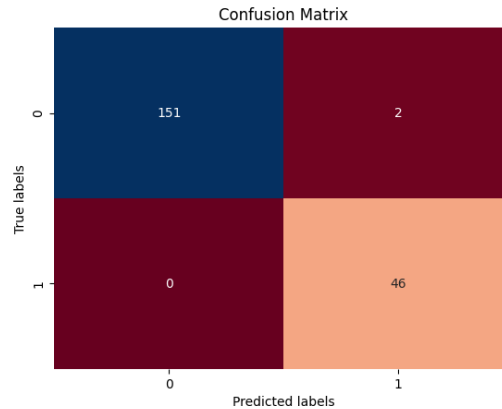


Figure 4. Confusion matrix.
(Photo credit: Original)

According to the confusion matrix, among the 199 fire images in the test set, the classification prediction of 197 images is accurate, and 2 fire images are predicted as non-fire images. The specific prediction results of each image in the test set are shown as follows:

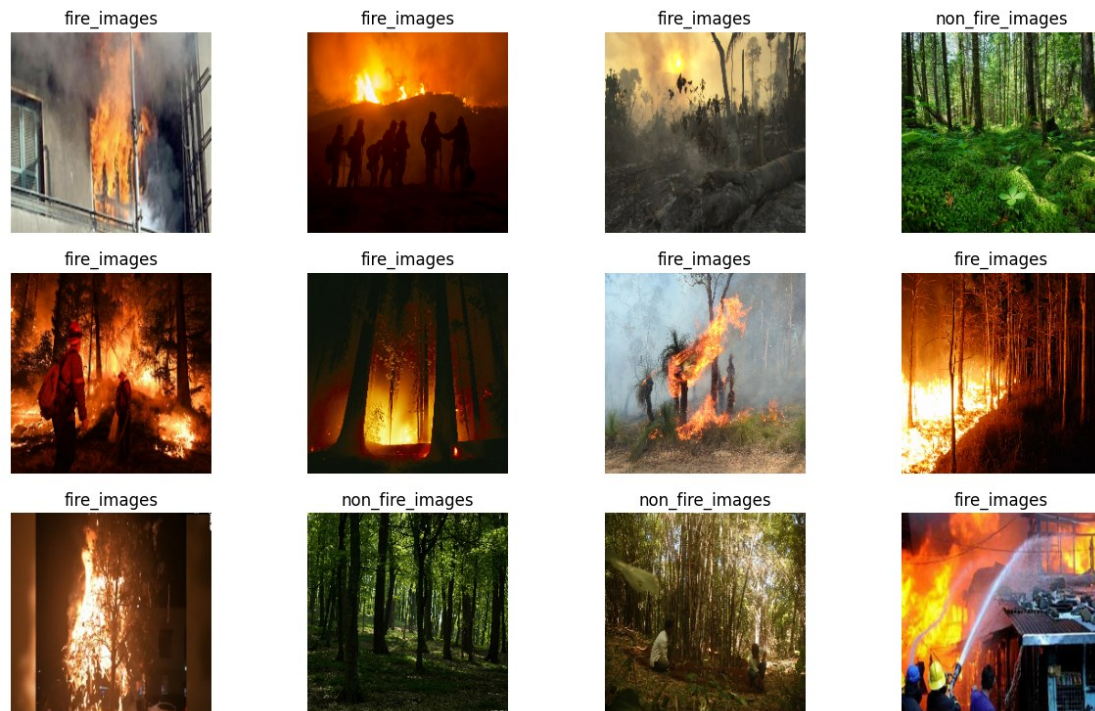


Figure 5. Test set.
(Photo credit: Original)

5. Conclusion

Fire is a kind of natural disaster, and its occurrence often brings great loss to human life and property. Early detection and control of fires is therefore essential to minimize the potential losses caused by these disasters. However, due to the suddenness and complexity of fire, manual detection and control have certain limitations. With the development of computer vision technology, fire image classification technology has become an important tool for fire detection and control.

Fire image classification technology involves the use of computer vision algorithm to classify and identify fire images. The technology can quickly and accurately identify whether there is a fire in the image, so that countermeasures can be taken in time. In the fire image classification technology, the steps of image preprocessing, feature extraction, classifier training and classifier testing are needed. Among them, pre-processing includes image denoising, enhancement and segmentation, and feature extraction is to extract the feature information in the image for classifier training and classifier testing.

The application range of fire image classification technology is very wide, not only can be applied to buildings, factories, warehouses and other indoor places, but also can be applied to forests, grasslands and other outdoor places. In addition, the fire image classification technology can also be combined with other technologies, such as sensor technology, network technology, etc., to achieve more intelligent fire detection and control.

The process of fire image classification typically involves the use of deep learning algorithms, such as convolutional neural networks (CNNs), to analyze and classify images. These algorithms are trained on a large dataset of images, including both fire and non-fire images, to learn the features that distinguish fire images from non-fire images.

During the training process, the CNN learns to identify the unique patterns and characteristics of fire images, such as the color, shape, and texture of flames. Once the CNN has been trained, it can be used to classify new images as either fire or non-fire with a high degree of accuracy.

Fire image classification technology has been widely studied and applied in a variety of settings, including in wildfire detection systems, fire alarm systems, and surveillance systems. By enabling early

detection and response to fires, this technology can help to minimize the damage and loss caused by these disasters, making it an important tool for protecting human life and property.

In recent years, deep learning techniques have achieved great success in the field of image classification. The researchers used deep learning models such as convolutional neural networks (CNN) to classify and identify fire images. Among them, the image processing model based on vgg16 is a common deep learning model, which can automatically extract and classify the features in the image.

In this paper, the image processing model based on vgg16 is used to detect and classify fire images. The model is trained and tested by the data of training set and test set. The results show that the loss of training set and test set gradually decreases until convergence, and the accuracy of training set and test set gradually increases and stabilizes at 99%. When finally tested with the test set, the model achieved 98% accuracy and accuracy, and the recall rate and F1 score reached 99%. This shows that the model can classify and recognize fire images well and achieve good prediction effect.

The application of fire image classification technology is of great significance. By using computer vision technology to classify and identify fire images, it can improve the accuracy and efficiency of fire monitoring and early warning system, timely detect and control fire, and reduce the loss caused by fire. At the same time, the continuous development of deep learning technology also provides a broader space for the research and application of fire image classification technology.

References

- [1] Jun S ,Fengyi Y ,Jiehong C , et al. Nondestructive identification of soybean protein in minced chicken meat based on hyperspectral imaging and VGG16-SVM[J]. Journal of Food Composition and Analysis,2024,125.
- [2] Yumeng Q . Performance comparison among VGG16, InceptionV3, and resnet on galaxy morphology classification[J]. Journal of Physics: Conference Series,2023,2580(1).
- [3] Cheng C . Galaxy morphology classification using VGG16[J]. Journal of Physics: Conference Series,2023,2580(1).
- [4] Li Y ,Li X ,Zhao Q . Multimodal Deep Learning Framework for Book Recommendations: Harnessing Image Processing with VGG16 and Textual Analysis via LSTM-Enhanced Word2Vec[J]. Traitement du Signal,2023,40(4).
- [5] Ying C ,Yinyin C ,Shuangshuang F , et al.VGG16-based intelligent image analysis in the pathological diagnosis of IgA nephropathy[J].Journal of Radiation Research and Applied Sciences,2023,16(3):
- [6] Wilson B ,Serestina V . VGG16 Feature Extractor with Extreme Gradient Boost Classifier for Pancreas Cancer Prediction.[J]. Journal of imaging,2023,9(7).
- [7] Fuzhen Z ,Jiacheng L ,Bing Z , et al.UAV remote sensing image stitching via improved VGG16 Siamese feature extraction network[J].Expert Systems With Applications,2023,229(PA):
- [8] K S ,R S G ,K P , et al.COVID-19 prediction based on hybrid Inception V3 with VGG16 using chest X-ray images.[J].Multimedia tools and applications,2023,
- [9] Soma S ,Biswash N S T ,Heping C .High accuracy keyway angle identification using VGG16-based learning method[J].Journal of Manufacturing Processes,2023,98
- [10] Wen S ,Yang L ,Duan H , et al.Pavement Recognition Based on Improving VGG16 Network Model[J].International Journal of New Developments in Engineering and Society,2023,7(3):