Satellite remote sensing image classification based on multiple deep learning algorithms

Shiqi Ren

College of information engineering (College of artificial intelligence), Yangzhou university, Jiangsu, Yangzhou, 225000, China

Corresponding author's e-mail: 937327378@qq.com

Abstract. This paper explores the use of satellite remote sensing technology to observe and measure the earth's surface and classify remote sensing images using various deep learning algorithms. By importing AlexNet, VggNet, GoogleNet and MobileNet models, 90% of the data are randomly selected for training and 10% for testing, and the changes in the loss and accuracy of the validation set during training are recorded, as well as the accuracy of the best round of epochs for training.During the training process, the validation set's loss gradually decreased and converged, and the accuracy of the validation set gradually increased and stabilised. The results show that all four models are able to classify remote sensing images well, among which the highest accuracy is MobileNet model, which reaches 99.8%, and Googlenet model is the second one, which reaches 97.9%.AlexNet and VggNet models also have higher accuracy, which are 93.3% and 95.4%, respectively. Overall, the results of this paper provide an effective solution for satellite remote sensing image classification and provide a reference for the application of deep learning algorithms in remote sensing.

Keywords: Image classification, MobileNet, Googlenet.

1. Introduction

Satellite remote sensing technology refers to the use of remote sensing sensors carried by satellites to observe and measure the Earth's surface and to obtain information about the Earth's surface, including topography, geomorphology, vegetation, climate, hydrology and other kinds of information [1,2]. Satellite remote sensing technology has become an indispensable and important means in the fields of national resource management, environmental protection, and disaster monitoring.

Satellite remote sensing image classification refers to the classification of pixel points in satellite remote sensing images according to the category to which they belong, and common classifications include land, water bodies, vegetation, etc [3,4]. Traditional remote sensing image classification methods are mainly based on feature extraction and machine learning algorithms, such as support vector machine (SVM) and decision tree (DT). These methods require manual feature extraction and are less efficient for complex scenes and large-scale data processing.

In recent years, deep learning algorithms have been widely used in satellite remote sensing image classification. Deep learning algorithm is a machine learning algorithm based on neural network structure, which has the advantages of automatic feature extraction and efficient processing of large-scale data [5]. Deep learning algorithms are mainly classified into two categories in satellite remote

sensing image classification: methods based on convolutional neural networks (CNN) and methods based on recurrent neural networks (RNN) [6].

CNN-based methods are currently the most widely used deep learning algorithms in satellite remote sensing image classification, and their main idea is to automatically extract image features using multi-layer convolution and pooling operations. For example, feature extraction is performed by feeding satellite remote sensing images into a pre-trained deep convolutional neural network and classification is performed by fully connected layers [7,8]. In addition, there are some improved CNN-based methods, such as null convolution and residual networks, which can better handle complex scenes and large-scale data.

RNN-based methods, on the other hand, are mainly applied to time-series remote sensing data classification, such as meteorological prediction, vegetation growth monitoring and other fields [9]. The main idea is to use recurrent neural network structure to model time-series data and classify them through the output layer [10]. For example, meteorological data are modelled using Long Short-Term Memory (LSTM) networks and classified through fully connected layers.

Deep learning algorithms have great potential in satellite remote sensing image classification, this paper selects remote sensing image public dataset to classify remote sensing images based on a variety of deep learning algorithms, and compares the classification effect of various deep learning algorithms classifiers, to provide a certain foundation for subsequent research, and hope that it will be more widely used in future research.

2. Source of data sets

The dataset used in this paper is from the satellite image classification dataset-RSI-CB256, which is a new large-scale benchmark dataset containing multiple instances for RS image scene classification. The dataset consists of four types of images, namely, cloudy, desert, greenland and water for four scenes, which includes a total of 5631 images, and the dataset is divided into a training set and a validation set according to a ratio of 9:1. Examples of the four types of images (cloudy, desert, green field and water respectively) are shown in Figure 1.
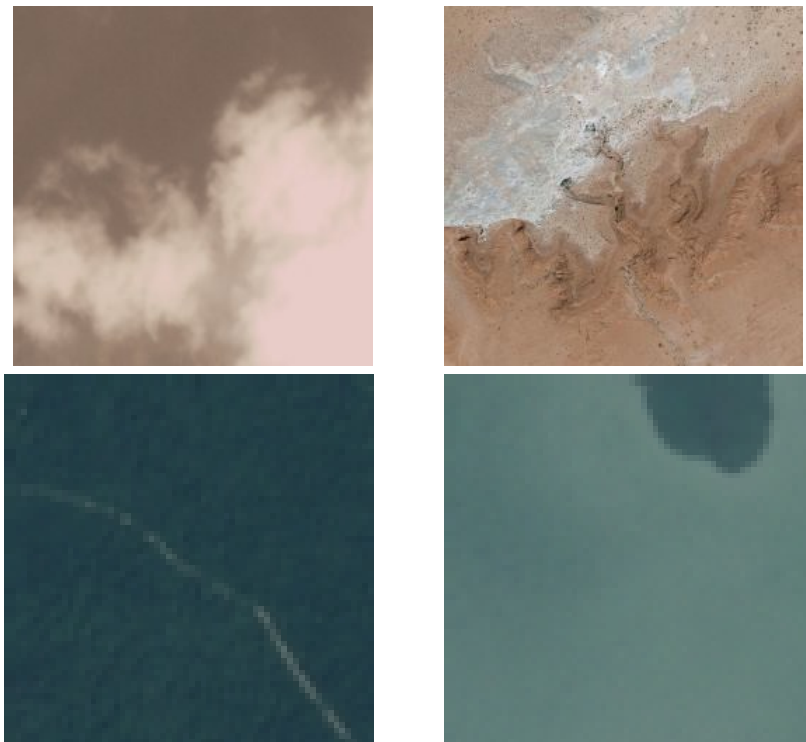
Figure 1. Selected data sets.
(Photo credit: Original)

3. Method

3.1. AlexNet

AlexNet is a deep convolutional neural network that consists of 8 layers of neural networks, 5 of which are convolutional, 2 are fully connected and 1 softmax output layer. The model uses ReLU activation functions and Dropout regularisation to reduce overfitting.AlexNet also uses data enhancement techniques including random cropping, level flipping and colour transformations.

The innovation of AlexNet is the introduction of GPU-accelerated computation and the use of two GPUs for parallel computation. This approach greatly improves training speed and model accuracy.AlexNet also proposes a local response normalisation technique to enhance the contrast of the feature maps.

3.2. VggNet

VggNet is a deep convolutional neural network consisting of 16-19 layers of convolutional neural networks.VggNet is characterised by the use of small size 3x3 convolutional kernels instead of larger size convolutional kernels, which reduces the number of parameters and improves the model's ability to generalise.VggNet also employs multiple pooling layers to reduce the dimensionality of the feature maps, and adds fully-connected layers to achieve classification.

VggNet performed well in the ImageNet competition and has also become one of the classic models in deep learning.VggNet's structure is simple and straightforward, easy to understand and implement, and is therefore widely used in areas such as image classification, target detection, and semantic segmentation.

3.3. GoogleNet

GoogleNet is a deep convolutional neural network consisting of 22 convolutional layers.GoogleNet uses an Inception architecture, whereby convolutional kernels of different sizes are applied in parallel to the input data and stitched together to form a higher dimensional feature map.GoogleNet also uses a 1x1 convolutional kernel to reduce the dimensionality of the feature map, and a global average pooling layer instead of a fully connected layer, thus reducing the number of model parameters.

GoogleNet has achieved excellent results in the ImageNet competition and has proposed several innovative ideas and techniques. The most important of these is the Inception architecture, which drastically reduces the number of model parameters and the amount of computation, and improves model accuracy.

3.4. MobileNet

MobileNet is a lightweight convolutional neural network that drastically reduces model size and computation while maintaining high accuracy.Instead of the traditional convolution operation, MobileNet employs Depthwise Separable Convolution, which decomposes the standard convolution into two steps: depthwise convolution and point-by-point convolution. MobileNet also uses linear bottlenecks and residual concatenation to further improve model accuracy.

The innovation of MobileNet is the introduction of depth-separable convolution, which can significantly reduce the number of model parameters and the amount of computation, so that the model can be run in real time on mobile devices.MobileNet has been widely used in the fields of face recognition, image classification and target detection.

4. Result

The AlexNet, VggNet, GoogleNet and MobileNet models were imported respectively, 90% of the data were randomly selected for training and the remaining 10% for testing, the loss and accuracy changes of the validation set during training were recorded, and the accuracy of the best round of epoch for training was recorded, the batch size is set to 16, the learning rate is set to 0.0002, and the equipment used for the experiment is a 2080 graphics card with 32G RAM.
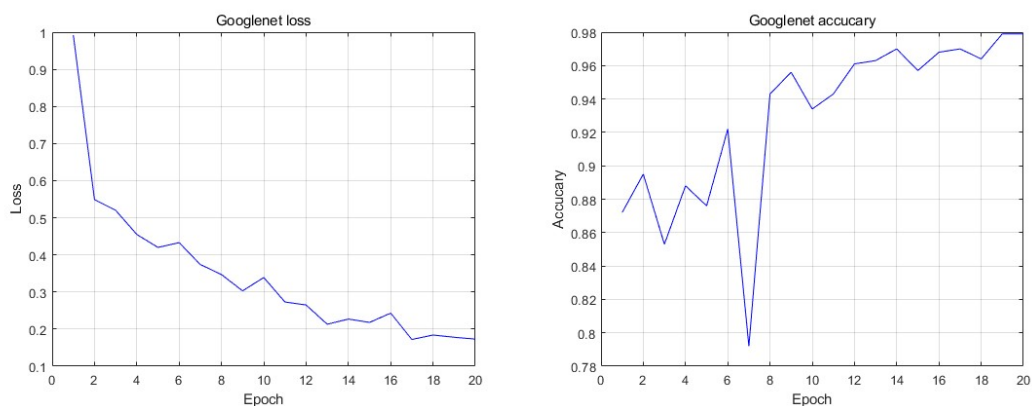
The changes in the loss and accuracy of the validation set during the training of the AlexNet model are shown in Fig. 2, the changes in the loss and accuracy of the validation set during the training of the VggNet model are shown in Fig. 3, the changes in the loss and accuracy of the validation set during the training of the GoogleNet model are shown in Fig. 4, and the changes in the loss and accuracy of the validation set during the training of the MobileNet model are shown in Fig. 4. The variation of loss and accuracy of validation set during training of MobileNet model is shown in Fig. 5.
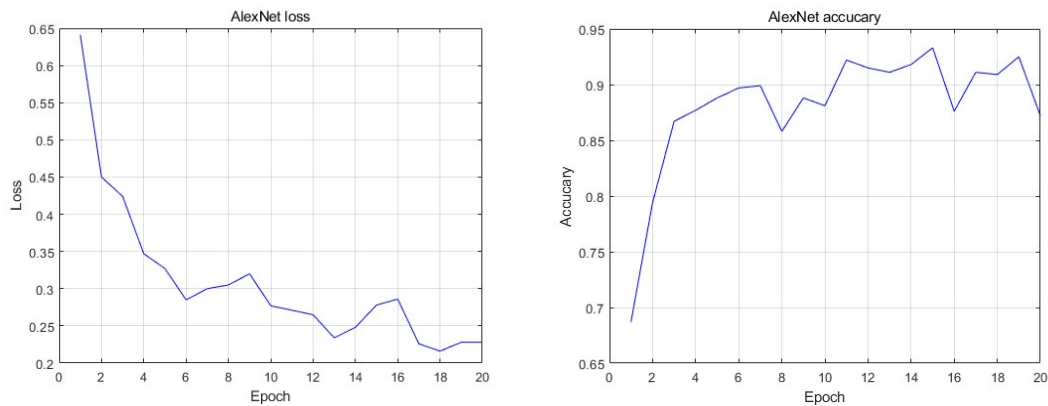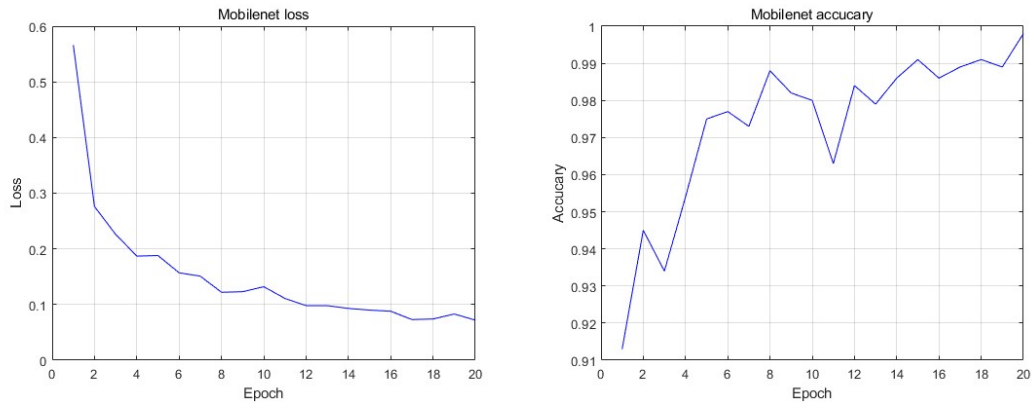

Figure 2. Loss and accuracy of the validation set during AlexNet model training.
(Photo credit: Original)
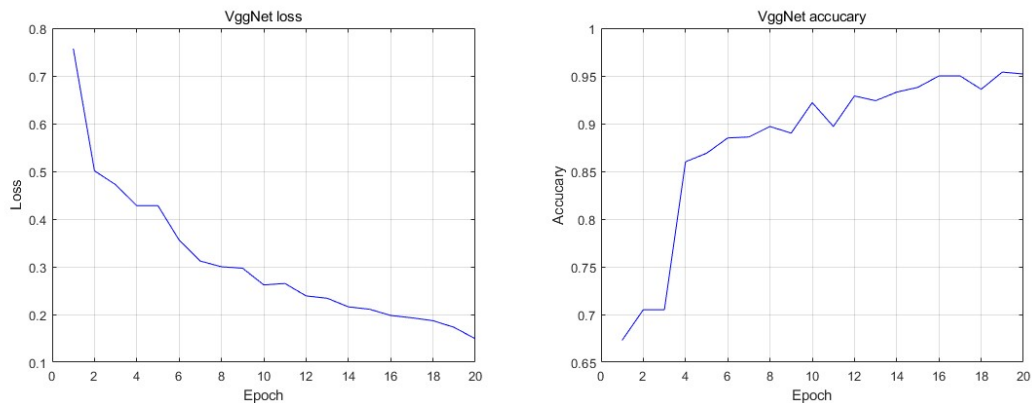
Figure 3. Loss and accuracy of the validation set during training of VggNet models.
(Photo credit: Original)

Figure 4. Loss and accuracy of the validation set during GoogleNet model training.
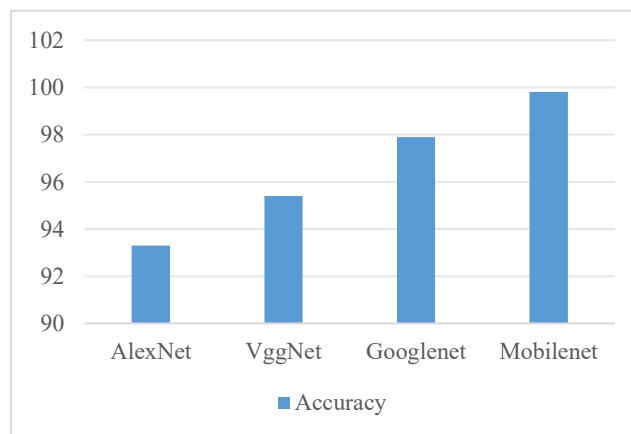(Photo credit: Original)

Figure 5. Loss and accuracy of the validation set during MobileNet model training.
(Photo credit: Original)

As can be seen from the change curves of the loss and accuracy of the validation set during the training process of AlexNet, VggNet, GoogleNet and MobileNet models, the loss of the validation set gradually decreases and tends to converge during the training process, and the accuracy of the validation set gradually increases and gradually stabilises.

The epoch with the best training accuracy for each model and the corresponding accuracy are counted, and the results are shown in Table 1 and Figure 6.

Table 1. Indicators for model evaluation.

Model	Best epoch	Accuracy(%)
AlexNet	15	93.3
VggNet	19	95.4
Googlenet	19	97.9
Mobilenet	20	99.8

Figure 6. Indicators for model evaluation.
(Photo credit: Original)

From the prediction results of each model, it can be seen that the highest accuracy is the Mobilenet model with 99.8%, followed by the Googlenet model with 97.9%. In addition, the AlexNet and VggNet models also achieved an accuracy of 93.3% and 95.4%, and all four models were able to classify remote sensing images well.

5. Conclusion

In this paper, four deep learning algorithms (AlexNet, VggNet, GoogleNet and MobileNet) are used to classify the remote sensing image dataset obtained by using satellite remote sensing technology and their classification effects are compared. During the training process, the LOSS of the validation set gradually decreases and tends to converge, and the ACCURACY of the validation set gradually increases and stabilises. After comparing the prediction results, we found that the MobileNet model has the highest accuracy of 99.8%, followed by the Googlenet model with an accuracy of 97.9%.The AlexNet and VggNet models are also able to classify remote sensing images well, with accuracies of 93.3% and 95.4%, respectively. Therefore, all four deep learning algorithms can be effectively applied in remote sensing image classification tasks.

References

[1] Huaxiang S .MBC-Net: long-range enhanced feature fusion for classifying remote sensing images[J].International Journal of Intelligent Computing and Cybernetics,2024,17(1):181-209.

[2] Long H ,Chen T ,Chen H , et al.Principal space approximation ensemble discriminative marginalized least-squares regression for hyperspectral image classification[J].Engineering Applications of Artificial Intelligence,2024,133(PA):108031-.

[3] Qiu C ,Zhang X ,Tong X , et al.Few-shot remote sensing image scene classification: Recent advances, new baselines, and future trends[J].ISPRS Journal of Photogrammetry and Remote Sensing,2024,209368-382.

[4] Meyer F D M ,Gonçalves A J ,Bio F M A .Using Remote Sensing Multispectral Imagery for Invasive Species Quantification: The Effect of Image Resolution on Area and Biomass Estimation[J].Remote Sensing,2024,16(4):

[5] Xujian Q ,Lei X ,Anxun H , et al.Simplified Multi-head Mechanism for Few-Shot Remote Sensing Image Classification[J].Neural Processing Letters,2024,56(1):

[6] Zhao X ,Zhang M ,Tao R , et al.Fractional Fourier Image Transformer for Multimodal Remote Sensing Data Classification.[J].IEEE transactions on neural networks and learning systems,2024,35(2):2314-2326.

[7] Feng S ,Gao M ,Jin X , et al.Fine-grained damage detection of cement concrete pavement based on UAV remote sensing image segmentation and stitching[J].Measurement,2024,226113844-.

[8] Li H ,Li L ,Zhao L , et al.ResU-Former: Advancing Remote Sensing Image Segmentation with Swin Residual Transformer for Precise Global–Local Feature Recognition and Visual–Semantic Space Learning[J].Electronics,2024,13(2):

[9] Hongning Q ,Zili L .Building classification extraction from remote sensing images combining hyperpixel and maximum interclass variance[J].International Journal of Image and Data Fusion,2024,15(1):86-103.

[10] Wang W ,Yi X.A monitoring method of surface vegetation distribution in the Yellow River Basin based on remote sensing image segmentation[J].International Journal of Environmental Technology and Management,2024,27(1-2):37-48.