Haze Image Classification Method Based on Alexnet Network Transfer Model

Lu Guo¹, Jing Song²*, Xin-rui Li², He Huang², Jing-jing Du², Yong-chao He², Cheng-zhuang Wang²

¹Xi’an ASN Technology Group Company, Xi’an, China
²School of Electronic and Control Engineering, Chang’an University, Xi’an, China

*Corresponding author e-mail: 1158940754@qq.com

Abstract: In recent years, the serious haze weather frequently appears in China. In order to improve the efficiency of the traffic image dehazing algorithm, it is important to effectively identify whether there is fog and haze in the images. Aiming at the problems of large amount of data sets, low training efficiency, long time and low accuracy in traditional image classification, this paper combines the knowledge of transfer learning and proposes a traffic haze image classification method based on Alexnet network transfer model. First, the image size is normalized and pre-processed. Then, according to the Alexnet network model, the trained multi-classification model is transferred to the haze image binary classification model to form a new Alexnet network transfer model. At last, the network parameters are retrained according to the training parameters in the transfer model, and applied to the haze image classification. Compared with the traditional classification algorithm, the proposed algorithm greatly reduces the amount of data needed for training, and the classification accuracy of the transfer network can reach 97%, which is 9% higher than the traditional SVM algorithm. It can effectively identify haze images.

1. Introduction

In recent years, haze weather has been frequent and has been listed as one of the top ten disasters in the world. The image obtained by the traffic image device contains a lot of noise, and the image is gray and unclear [1], which has a serious impact on traffic monitoring. In view of the problem of clearing traffic foggy images, many scholars have proposed different defogging algorithms, but blindly using the defogging algorithm to deal with degraded images will cause a picture with no fog to add a certain noise, resulting in a certain degree of image distortion. It is not conducive to image recognition, and time-consuming to the original fog-free image, the processing efficiency is low. Therefore, it is very important to effectively identify whether a picture has haze before performing the defogging process, that is to say, it is of great practical significance and application value to classify and recognize the acquired traffic image before processing the image, and the image processing efficiency is improved.

In the classification and recognition of the acquired traffic images, the traditional image classification algorithms are divided into two categories, one is image classification based on image space, the other is image classification based on feature space. Image space classification methods is usually based on some underlying features, such as grayscale, color, shape location, texture, and so on. Literature [2] and [3] classify images using gray histogram features and texture features respectively. Literature [4] uses texture, edge, and color histogram blending features to classify images. Generally,
SVM or BP should be used as classifier when using image space method to classify images. The image space-based classification method requires a large number of data sets, and the system is complex, but the classification effect is better. The commonly used classification methods are: image classification based on support vector machine [5] and artificial neural network [6], etc. At present, these methods are the most used, most commonly used and most effective methods. The classification method of feature space needs to transform the image into the feature space first, then extract the image features in the feature space, and realize the classification of the image by these features [7]. The commonly used transform space methods include K-L transform, wavelet transform and so on. Although the classification method based on feature space has the advantages of reducing the data dimension and computational complexity, the feature correlation to be extracted is generally strong and difficult to obtain, which will affect the results of image classification.

Traditional SVM, BP neural network and other algorithms have very important value for haze image classification. However, in the past, researchers need to extract image features purposefully from images, such as depth of field information, image contrast and other features. Although this method is effective, it also has great limitations. First of all, it may make some human mistakes. Image features are not obvious, and it is likely that there will be omissions in classification, which will reduce the accuracy of classification. Secondly, a large number of data samples are needed for training in image classification to ensure the correctness of the prediction. This work of collecting training samples wastes a lot of time and labor costs. If the machine can automatically recognize image features, humans can be freed from this repetitive work, which not only to ensure the accuracy of classification, but also to improve efficiency.

Aiming at the above problems, this paper proposes a haze image classification algorithm based on Axnet network migration model. The accuracy of classification is improved by using Alexnet network model. The migration learning algorithm can still maintain good classification effect while the number of training samples is small, which greatly improves the work efficiency.

2. Artificial neural network

2.1 Introduction of artificial neural network

The history of the deep learning algorithm can be traced back to the artificial neural network (ANN), which is the further development of ANN. The traditional artificial neural network (ANN) makes the computer work more close to human beings by simulating the simple model of human brain and some learning methods. Deep learning is the deeper development of ANN. It is a multi-layer structure network model. The network has certain complexity and strong reliability. Although the traditional learning algorithm has been well developed in some fields, the network structure of the traditional algorithm is simple, and even some problems can not be expressed by this simple structure. The traditional learning algorithm still has some limitations. Deep learning algorithm makes the network model deeper and more complex on the basis of the original network model, extracting features from each layer of the network, mining useful information of data to achieve the purpose of learning something. And the process of feature extraction is automatic and needs no supervision. Arel et al. introduced the popular depth learning algorithms and the research directions in recent years, and listed the advantages of depth learning methods [8].

The idea of deep learning is to combine the features extracted from the network by constructing a multi-level neural network structure. Because the components of the edge information of objects are similar, the basic structure of the first layer of different objects is similar. In-depth learning allows the machine to automatically extract object features, reducing labor to improve efficiency, while avoiding mistakes in manual selection. However, the number of feature extraction at each level needs to be completed according to experience. Too many features make the amount of calculations larger, and too few features fail to complete the correct representation. The structure of deep learning is shown in Figure 1.
The traditional neural network has only one hidden layer, and the depth learning network structure has many hidden layers. The depth learning can extract the image features by building a multi-layer hidden layer structure. By extracting the image features layer by layer, we can find the most suitable base and select the appropriate number of features to improve the prediction accuracy.

2.2 Convolution Neural Networks

Convolutional neural network is a multi-layer neural network model, which is the most widely used network model in depth learning. It has been applied in many fields, such as image classification [9], target detection [10]. The trained filters and local neighborhood pooling operations alternately act on the input image and then output a series of complex features. Convolutional neural networks is an improvement to the BP, it have a variety of structures, the most commonly used is the 1998 LeCun et al. proposed convolutional neural networks - LeNet5 [11], its network structure as shown in Figure 2.
The C1 and C3 layers are the feature extraction layers, and the input of each layer is the output of the previous layer. After the local features are extracted, the positions of other features are related to them, so their positions are determined accordingly. S2 and S4 layers are feature mapping layers, also known as sub-sampling layers. Each computing layer of the network consists of multiple feature mappings, each feature mapping into a plane, and the weights of neurons on the plane are shared, that is, all neurons have equal weights. In the feature mapping structure, the sigmoid function is used as the activation function of the network, and the convolution result is compressed to a fixed range, so that the output range of each layer is controllable.

3. Transfer learning

Transfer learning is to obtain a compact and effective target task model by learning the data labeled in the source target task and a small amount of annotated data in the target task, and then applies the learning feature representation method to the learning task of the target domain. It is a new machine learning method. The goal of transfer learning is to use the acquired knowledge to help or promote the learning of new knowledge [12]. That is, the existing knowledge learning can transfer some concepts or methods through transfer learning to help or promote future learning. In computer vision, we can solve the problem of insufficient training samples by adjusting the classifiers of other classes. For example, in recent studies, convolutional neural networks are trained using ImageNet image set of large dataset, and then used for scene classification and target location. In addition, Zhu et al. migrated the migration knowledge learned from text files to image classification, and solved the problem of migration learning between different domains very well.

4. Haze image classification method based on alexnet network transfer model

4.1 Convolution neural network ---Alexnet

Alexnet was proposed by Krizhevsky et al. and won the championship in the ImageNet competition in 2012. It has more than doubled the accuracy of ImageNet and has unique advantages in image recognition. It is the most classic model in image processing domain [14] - [15]. Alexnet has a total of eight layers of network architecture, including five convolution layers and three fully connected layers. By training the network with large-scale data, up to 1000 kinds of image classification can be achieved. The specific structure is shown in Figure 3.

Fig 3. Alexnet network structure diagram

The network input image size is 227 * 227 * 3, 227 is the width and height of the image, 3 represents the RGB image of the three channels. Because the training images are of different sizes, the image size should be normalized to 227 * 227 pixels before training. In the traditional Alexnet network, ReLu activation function, maximum pooling layer and norm normalization are added after the first layer and the second layer respectively. The third layer only uses convolution layer and ReLu activation function, and the fifth layer is similar to the first layer, but not normalized by norm. The sixth to eighth layers are the fully connected layers, and the last layer uses the soft max classifier to classify the images...
by 1000. In addition, Dropout is used in this paper to properly discard parameters to prevent overfitting of the model.

4.2 Haze image classification based on transfer model

It is well known that traditional artificial neural network models often require a large amount of data to achieve better modeling accuracy. Similarly, using Alexnet network to classify images requires a large number of data sets to train the network to achieve a higher accuracy rate [16]. However, in practice, it is often difficult to obtain a large amount of training data in a general modeling task, and the excessive number of training sets can improve the accuracy rate, which is not only increases the complexity of the network training speed, making network training more difficult, but also brings the problem of over-fitting. In order to improve the accuracy and reduce the difficulty of network training, this paper proposes a haze image classification algorithm based on Alexnet network transfer.

As is show in Fig4, the AlexNet network is trained on the ImageNet classification task (Source task) and test the training accuracy. Then transfer the last three parameters to the haze image recognition task (Target task) to realize the transfer of network parameters. Because of the difference between the original network and the data (Data set 1 & 2) set used by the transferred network, when transferring the original network parameters to the task of haze image recognition, the new network should be trained, and the classification accuracy can be optimized by adjusting the parameters repeatedly.

5. Experiment

The software test platform used in this experiment is Matlab2017a, the hardware resources are: Intel(R)i5-3230, 2.4Ghz, graphics card is Nvidia GTX960TI, GPU memory is 4GB. In the experiment, a data set containing 800 images was used as the experimental data, including 400 foggy images and 400 clear images. In order to avoid mixing and overlapping of training set and test set, 800 images were randomly scrambled and 350 images were randomly selected from the two kinds of images as training set, and the remaining 100 images were used as test set. The 1000 classifications of Alexnet network are adjusted to 2 classifications. The network iterations are set for 7 epochs, 14 training iterations per round, 98 times in total, and the initial learning rate is set to 0.002. The learning rate of each iteration is automatically reduced to half of the original. Under the same experimental conditions, the SVM algorithm [17] is compared with the classification results of the proposed algorithm. The experimental results are shown in Figures 5 and table 1,2.
The Experiment results show that the training accuracy can reach 1, and the test set accuracy can reach 0.98. It can be seen from the results in Figure 4 and Table 1, the accuracy curves of test set and training set are very close, that is, there is no over-fitting phenomenon, and the accuracy approaching 1 and 0.97 respectively. Comparing with Table 1, the classification accuracy of the propose algorithm is 0.09 higher than that of SVM algorithm, and the classification accuracy is significantly improved.

**Acknowledgments**

This work was financially supported by Scientific Innovation Practice Project of Postgraduates of Chang’an University(2018076)

**References**

[1] WANG Gui-ping, SONG Jing, DU Jing-jing, et al. “Haze Defogging Algorithm for Traffic Images Based on Improved Gradient Similarity Kernel”. China Journal of Highway and Transport, 2018, 31(6), pp. 264-271, 280.
[2] Swain, Michael J., and D. H. Ballard. "Color indexing." International Journal of Computer Vision 7.1(1991), pp. 11-32.

[3] Haralick R, Shanmugam K, Dinstein I. Textural features for image classification[J]. IEEE Transactions on Systems, Man and Cybernetics, 1973, 23, pp. 610-621.

[4] Wan Hua-lin, M. U. Chowdhury. “Image Semantic Classification by Using SVM”. Journal of Software, 2003, 14(11), pp. 1892-1899.

[5] Li Jing, Yao Ming-hai. “Research of Semantic Image Classification Based on Support Vector Machine”. Computer Technology and Development, 2010, 20(2), pp. 75-78.

[6] Zhang Chong, Shi Qing-xuan, Miao Xiu-fen, et al. “Handwritten Numeral Recognition Based on BP Neural Network”. Computer Vision and Pattern Recognition IEEE, 2008, 18(6), pp. 126-128, 163.

[7] Hu Jian-ce, Wu Guo-ping. “BP Neural Network Based on principle Component Analysis in Multi-spectral Remote Sensing Images Classification”. Science of Surveying and Mapping, 2009, 34(3), pp. 137-139.

[8] Arel I, Rose D C, Karnowski T P. “Deep Machine Learning--a New Frontier in Artificial Intelligence Research”. IEEE Computational Intelligence Magazine, 2010, 5(4), pp. 13-18.

[9] Dan, Ciregan, U. Meier, and J. Schmidhuber. “Multi-column deep neural networks for image classification.” Computer Vision and Pattern Recognition IEEE, 2012, pp. 3642-3649.

[10] Sermanet, Pierre, et al. "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks." Eprint Arxiv (2013).

[11] LeCun Y, Bottou L, Bengio Y, et al. “Gradient-based Learning Applied to Document Recognition”. Proceedings of the IEEE, 1998, 86(11), pp. 2278-2324.

[12] Pan S J, Yang Q. “A Survey on Transfer Learning”. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10), pp. 1345-1359.

[13] Zhu, Yin, et al. "Heterogeneous transfer learning for image classification." AAAI Conference on Artificial Intelligence AAAI Press, 2011, pp. 1304-1309.

[14] Krizhevsky, Alex, I. Sutskever, and G. E. Hinton. "ImageNet classification with deep convolutional neural networks." International Conference on Neural Information Processing Systems Curran Associates Inc. 2012, pp. 1097-1105.

[15] Ding Peng-li, Li Qing-yong, Zhang Zhen, et al. “Diabetic Retinal Image Classification Method Based on Deep Neural Network”. Journal of Computer Applications, 2017, 37(3), pp. 699-704.

[16] Lv Hong-meng, Zhao Di, Chi Xue-bin. “Deep Learning for Early Diagnosis of Alzheimer’s Disease Based on Intensive AlexNet”. Computer Science, 2017, 44(6), pp. 50-59.

[17] Hu Zhong-yi, Liu Qing, Guo Jian-ming, et al. “An Automatically Detection Method of Hazing Images Based on SVM Classification”. Computer Simulation, 2015, 32(2), pp. 342-346.