PAPER • OPEN ACCESS

Image classification based on RESNET

View the article online for updates and enhancements.

Image classification based on RESNET

Jiazhi Liang

Computer technology, Northwest MuZu University, LanZhou, GanShu, China

*Corresponding author's e-mail: 601186348@qq.com

Abstract. at present, neural networks are becoming more and more complex, from several layers to dozens of layers or even more than 100 layers.The main advantage of deep network is that it can express very complex functions. It can learn features from different levels of abstraction, such as edge features at lower levels and complex features at higher levels.However, the use of deep networks is not always effective, because there is a very big obstacle - the disappearance of gradients: in very deep networks, gradient signals tend to approach zero very quickly, which makes the gradient descent process extremely slow.Specifically, in the process of gradient descent, the weight matrix product operation must be carried out in every step of back propagation from the last layer to the first layer, so that the gradient will drop exponentially to 0.(in rare cases, there is the problem of gradient explosion, that is, the gradient grows exponentially to the overflow in the process of propagation). Therefore, in the process of training, it will be found that with the increase of the number of layers, the rate of gradient decrease increases.Therefore, by deepening the network, although it can express any complex function, but in fact, with the increase of network layers, we are more and more difficult to train the network, until the proposal of residual network, which makes it possible to train deeper network[1].

1.Introduction

Computer vision recognition is a classic field of artificial intelligence, which has been widely concerned by academia and industry[2].The cifar-10 data set used in this example is composed of 60000 32×32 RGB color images, which are classified into 10 categories, including aircraft, car, bird, cat, elk, dog, frog, horse, boat and truck, of which 50000 are training pictures and 10000 are test pictures.Cifar-100 is more detailed than cifar-100.The most important feature of cifar-10 data set is that the recognition is transferred to universal objects and applied to multi classification.Cifar-10 data is stored in a numpy array of $10\,000 \times 3072$, the unit is uint8s, where 3072 means that a 32×32 color image is stored. The first 1024 bits are R values, the middle 1024 bits are g values, and the last 1024 bits are b values.

2.Resnet

Deep convolution neural network has a series of major breakthroughs in image classification[3]. However, in the development of deep learning, when we start to consider the convergence of deeper network, there is a degradation problem, that is, in the deepening neural network, the accuracy will first rise, and then reach saturation. If the depth is increased, the accuracy will decrease.As the error increases in both training set and test set, it is known that the influence is not caused by over fitting.RESNET enhances its network feature extraction ability through cross layer feature fusion, and network performance gradually improves with the deepening of network.The research team tested the deeper RESNET in an acceptable time, and compared several deep learning models, which proved that RESNET has better classification performance than other models, and can improve the accuracy by

increasing the depth.RESNET (residual neural network) is proposed by he Kaiming team of Microsoft Research Institute, which is called residual network in Chinese.The fundamental motivation of RESNET design is to solve the degradation problem of neural network, that is, when the neural network is deeper, the training error rate is higher[4]. To solve this problem, the team proposed a residual structure.The function of network layer is reprogrammed as residual function of input of each layer.In mathematical statistics, the concept of residual is the difference between the actual observation value and the estimated value (fitting value).

There are many kinds of RESNET residual components, which can even be defined according to the project requirements.Figure 2 shows the residual component of resnet-20 used in this paper, which solves the degradation problem well.The residual component is composed of two convolution layers and an identity mapping. The convolution kernel size is 3×3. Therefore, the input and output dimensions of the residual component are the same and can be added directly.When the step size is 1, after batch regularization, relu activation and convolution of the input in RESNET, the padding layer is the original input layer; when the step size is 2, the input of RESNET will do the same operation again, and then average pooling will be carried out to get the filling layer. Finally, the input of output layer is the output of filling layer plus the output of residual component[5].
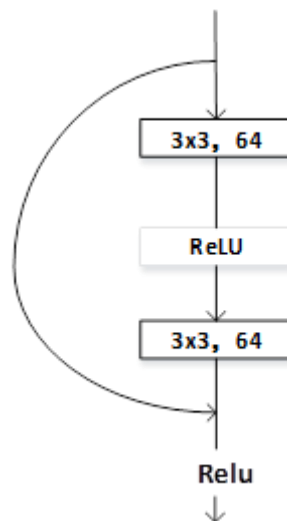

Figure 1. residual component of resnet-20.

3.Inception-Resnet
In the initiation RESNET module, a 1x1 extended conv operation is added at the end of the initiation subnet to make its output width (number of channels) the same as the input width of the subnet, so as to facilitate the addition[6].

4.Inception-resnet v1
The revenue RESNET V1 network is mainly used to compare the performance of the inception V3 model. Therefore, the calculation of the inception subnet used by it is reduced compared with the conventional inception module[7]. This is to ensure that its overall computational / memory overhead is similar to that of inception v3,Only in this way can we ensure the fairness of the comparison (after all, Google's view is: the design of CNN deep network is to pursue the performance optimization under the condition of limited computing and memory)[7]. The pictures below show the inception RESNET modules used in inception RESNET V1 and the connection modules between them.
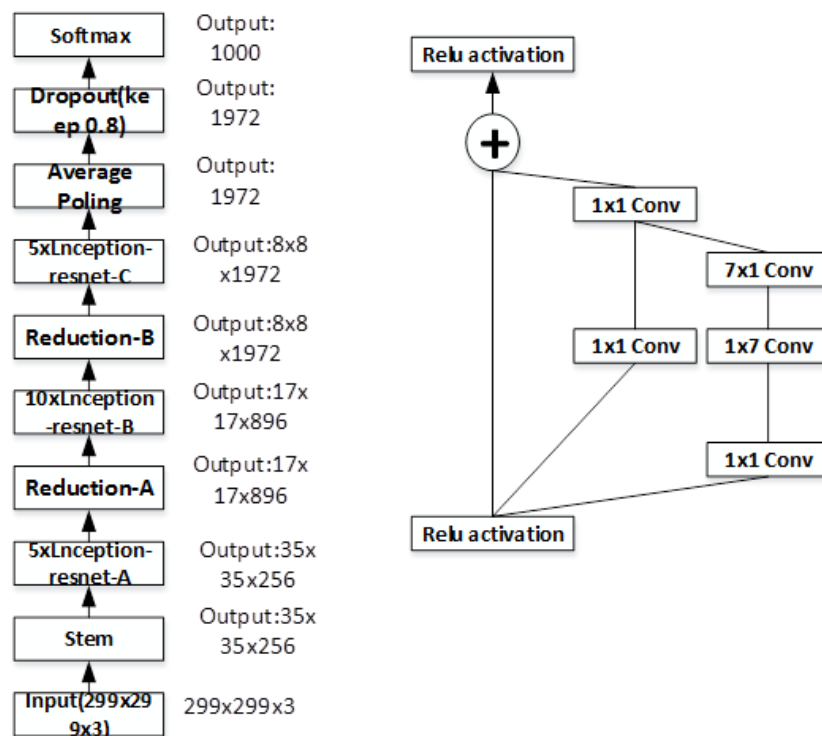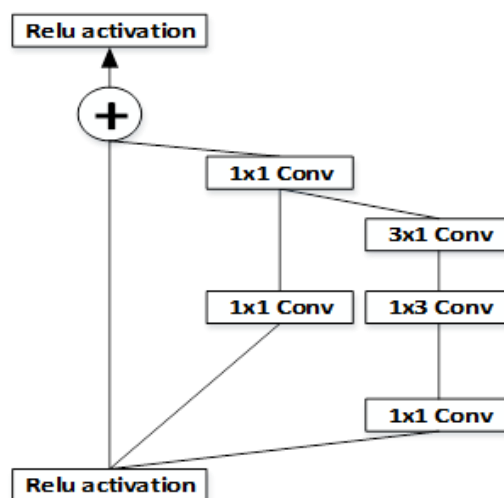
Figure 2. Inception RESNET_Modules used in V1.
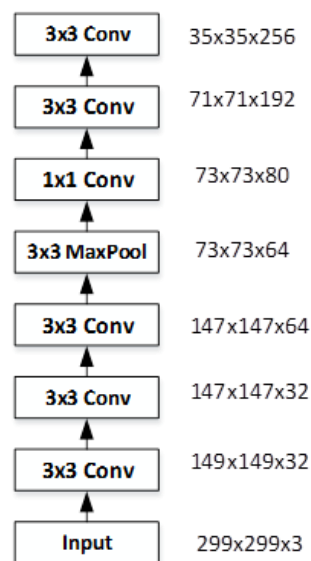
Figure 3. concept RESNET_C module used in V1.

Figure 4. concept RESNET_V1 network input module.

5.Program implementation

5.1. Pre requirements

The pandas library, numpy library, opencv and tensorflow (1.0.0 +) need to be pre installed before running.

5.2. Document organization structure

Among them, cifar10_input.pyIt includes functions of downloading, extracting and preprocessing cifar-10 image.resnet.pyThe RESNET structure is defined.cifar10_train.pyResponsible for training and validation.cifar10_test.pyResponsible for testing images..cifar10_main.pyThe starting file for program execution, including execution training and testing, can start the program by executing this file.

5.3. Parameters

This paper uses imagedatagenerator to enhance data, the importance of data in deep learning is self-evident, and the contradiction between the lack of data and big data is particularly prominent.At this time, data enhancement is particularly necessary, and the advantages are also well reflected in various papers;

Figure 5. data enhancement effect.

5.4. Training

Train () defines all classes about the training phase. The main idea is to run train_OPFLAGS.train_Steps times.If steps%FLAGS.report_If freq = = 0, it will immediately verify, train, and write all summaries on the tensorboard.

5.5. Testing

The test() function in the train() class will help users predict, and it will return a model [num_test_images, num_The softmax probability of labels.The user needs to prepare and preprocess the test data and pass it to the function.

5.6. Experimental results and analysis

5.6.1. Test error

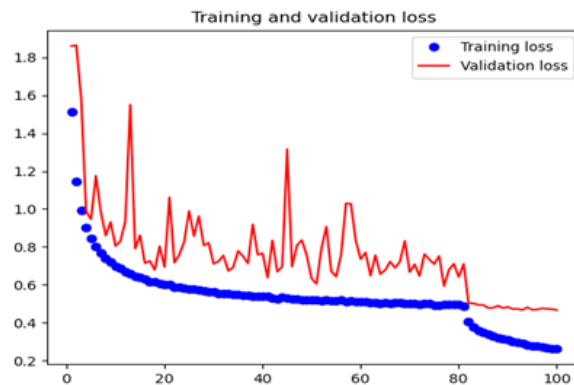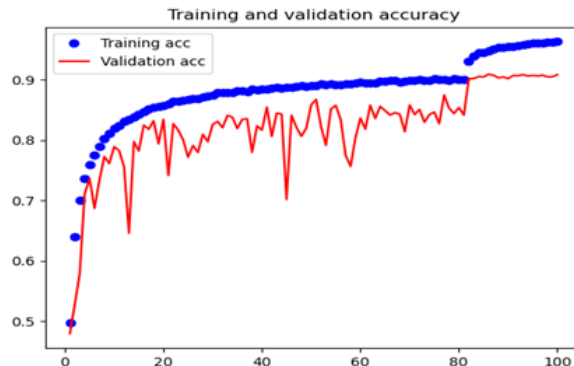Figure 6. error curve.

5.6.2. The test is accurate

Figure 7. accuracy.

6.Analysis

As can be seen from the error curve in Figure 7, the training error and test error will be very large in the early stage of network resnet20-v1.With the increase of iteration times, the training error and verification error decreased significantly, from 0.898438 to 0.117188, 0.1280 and 0.348188 in step 9775, from 0.898438 to 0.912 and 2.3543 to 0.117188, 0.1280 and 0.348188 respectively.With the increase of the number of RESNET network layers, the test error of the network in cifar-10 does decrease. When the number of layers reaches 110, the network performance reaches the best.When the RESNET network exceeds a certain level, it is difficult to optimize the network. It may be that the training network produces over fitting phenomenon.On datasets, using regularization, such as maxout and dropout, will yield better results.However, in this network, we do not use maxout or dropout, but simply use regularization to design network architecture.

7.Conclusion

This paper proposes a neural network model for image recognition based on deep learning.Compared with the traditional vehicle logo recognition method, it integrates multiple features, and can extract features independently for recognition, thus avoiding the tedious and one-sided feature selection manually.The experimental results show that this method has good robustness and accuracy, has strong resistance to noise pollution, and can effectively improve the accuracy of image recognition.

Reasons

The technical terms, paragraph punctuation and quotation format of this paper are accurate and conform to the academic standards. The full text focuses on the theme and has a clear point of view.

At the same time, this paper maintains a high level of writing while putting forward innovative ideas. The full text language is concise and accurate, the argument is clear and rigorous, which can better elaborate and support the views and propositions put forward by him.

Reference

[1]   Huang G., Liu S., Maaten L., et al. CondenseNet: An Efficient DenseNet using LearnedGroup Convolutions. In Conference on Computer Vision and Pattern Recognition, 2752-2761,2018.

[2]   Sandler M, Howard A, Zhu M., et al. MobileNetV2: Inverted Residuals and LinearBottlenecks. In Conference on Computer Vision and Pattern Recognition, 4510-4520, 2018.

[3]   Chollet F.. Xception: Deep learning with depthwise separable convolutions. In Conferenceon Computer Vision and Pattern Recognition, 1800-1807, 2017.

[4]   Zhang X., Zhou X., Lin M., et al. Shufflenet: An extremely efficient convolutional neuralnetwork for mobile devices. In Conference on Computer Vision and Pattern Recognition, 6848-6856, 2018.

[5]   Ma N., Zhang X., Zheng H. T., et al. Shufflenet v2: Practical guidelines for efficient cnnarchitecture design. arXiv preprint arXiv:1807.11164, 2018.

[6]   Hu J., Shen L., Sun G.. Squeeze-and-excitation networks. In Conference on Computer Visionand Pattern Recognition, 7132-7141, 2018.

[7]   Zoph B., Vasudevan V., Shlens J., et al. Learning transferable architectures for scalableimage recognition. In Conference on Computer Vision and Pattern Recognition, 8697-8710,2018.

[8]   Liu C., Zoph B., Shlens J., et al. Progressive neural architecture search. arXiv preprintarXiv:1712.00559, 2017.

[9]   G. Wei, H. Ma, W. Qian, et al. Lung nodule classification using local kernel regressionmodels with out-of-sample extension. Biomedical Signal Processing and Control, 40, 1–9, 2018.