Research on the Application of Deep Learning Algorithms on Image Classification

Zijian Wang

Information Engineering University, Zhengzhou, Henan, 450000, China

Keywords: Deep learning algorithms, Image classification, Application study

Abstract: In recent years, the research of deep learning (DL) has received widespread attention. Compared with the traditional shallow network model, the multi-layer network structure of the deep learning model can express complex functions more effectively, thereby learning features with stronger characterization ability, and applying deep learning to image classification can effectively improve Classification accuracy. However, deep learning still faces severe challenges in practical applications: the important basic unit model of deep learning-Restricted Boltzmann Machine (RBM) has high computational complexity and model in the training process. The likelihood of training data is not high enough, which leads to a long training time for deep learning models.

1. Introduction

Today's Internet and Internet technologies are developing rapidly. With the rise of new entertainment tools such as WeChat and Momo, and the widespread popularization of handheld terminal devices equipped with digital cameras such as tablets and smartphones, the image data on the Internet has increased dramatically. These images cover all aspects of human life and spread a lot of useful information. In recent years, how to quickly and effectively extract and analyze the semantic information contained in these images and apply them to practical problems has become the focus of research in the fields of image classification and recognition, image search, image understanding and analysis. With the deepening of image research, intelligent image classification algorithms such as fuzzy set methods, decision tree classification methods, knowledge-based classification methods, and machine learning methods continue to emerge. The fuzzy set method is obtained through experience. Although it can handle some relatively fuzzy problems well, it has certain uncertainty and subjectivity. Decision tree classification is a relatively good classification method, which is derived by imitating human thoughts, but it also has shortcomings such as high dependence, difficult combination of classification decision rules and expert systems, and inability to make full use of the spatial characteristics of the classification objects. Knowledge-based classification methods do not have adaptive capabilities. When experience and knowledge are disturbed by external factors, the classification effect of this method is poor. Therefore, compared with the above three methods, the method of machine learning has received increasing attention from scholars because of its profound theory and remarkable effect. The deep learning algorithm is an important branch of machine learning, and its excellent analysis and modeling capabilities also provide a new direction for the research of image classification technology. The deep learning algorithm simulates the multi-layer neural network to build a model based on the hierarchical characteristics of the human visual system's processing of information. From low-level edge features to targets or shapes, to high-level overall targets, target behaviors, etc., the higher-level feature representations can express the semantics of the image. The deep learning model is composed of a large number of simple neurons. The neurons of each layer input to the higher layer. Through the nonlinear relationship between the input and output, the low-level features are combined into a higher-level abstract representation. Discover the distributed characteristics of the observed data, the top-down learning method, forming a multi-level abstract representation, and the multi-level feature learning is automatic and without manual intervention. Then, through the learned network structure, the input sample data is mapped to various levels of features, and the output of the uppermost unit is classified and identified using existing matching algorithms and models. The

  DOI: 10.25236/mmmce.2021.036

original intention of deep learning research is mainly to apply to image recognition. So far, although deep learning has been applied to many fields such as speech, image, text, emotion, etc., about 70% of the papers published in the field of deep learning are about image recognition. Deep learning has huge advantages in image recognition and classification.

2. The Emergence of Deep Learning

Marvin Minsky and Seymour Papert in 1969 discussed the shortcomings of the first generation of neural networks incapable of handling the easy XOR problem, which made researchers less and less confident in neural networks. Therefore, since 1970, neural networks the research and development of the network has entered its first low ebb. The physicist John Hopfield proposed a new neural network model in 1982, called the Hopfield model, and proposed the definition of energy for the first time. The Hopfield model has strong fault tolerance and can reconstruct a complete data image from incomplete or distorted data images. Following the first-generation neural network perceptron, Hopfield's model is called the second-generation neural network. Unlike the perceptron, the Hopfield model no longer uses pre-designed features, but is automatically adapted by the algorithm during the training process. Therefore, the problem of narrow application range is solved. During training, this type of neural network uses a back propagation algorithm (Back Propagation, BP) to backpropagate sample errors from the output layer to the input layer. The weights of the network are in the process of propagation. Continuously adjust to improve the likelihood of the neural network for the sample. At the same time, although the hidden layer can strengthen the expressive ability of the network, its introduction also increases the complexity of the network. Compared with the first-generation neural network, although the second-generation neural network can handle more problems, it still has obvious shortcomings: Supervised learning is its only way, and labels should be used when training data. Most of the data in the real environment is unlabeled. The training time is long, especially when there are too many hidden layers, but reducing the number of hidden layers will reduce the accuracy; The backpropagation algorithm is easy to fall into a local optimal solution. Therefore, the development of neural networks has entered the second trough. Finally, this stalemate ceased to exist in 2006. Hinton first proposed the theory of deep networks and deep learning, and published it in the journal Science, which aroused people's interest in deep learning. Attention. The deep learning architecture is composed of multiple layers of non-linear computing units, which can learn complex functions that represent high-level abstract concepts. In the process from low to high, low-level output is used as high-level input. Through multiple input data learning, high-level feature representation of its structural information can be obtained, which can be used in specific problems such as classification, regression and information retrieval. The birth of deep learning has given new hope for the study of artificial intelligence. However, with the current level of research, people are still unable to control what the characteristics learned at each layer of nodes are, and it is far from enough to realize the idea of replacing the human brain with computers. Therefore, deep learning has become a hot research direction in the current machine learning field. Looking at the current main deep learning algorithms, compared with the traditional second-generation neural network, the main method for deep learning algorithms to overcome the learning efficiency bottleneck is to treat the multi-layer model as a stack of multiple RBMs or unit models similar to this, and then The purpose of training a multi-layer model is achieved through layer-by-layer learning of these RBMs. Therefore, further research on the RBM model is one of the core contents of deep learning and has very important significance.

3. Application of Deep Learning in Image Classification

The image classification method based on convolutional neural network is to simulate the human brain visual system to classify images through the model structure of alternating convolutional layer and down-sampling layer. The convolutional layer strengthens the original signal and improves the signal-to-noise ratio to a certain extent. At the same time, according to the principle of image local

correlation, the down-sampling layer samples the image between domains to extract useful information while reducing the amount of data; at the same time; The characteristics of system parameter reduction and weight sharing solve the problem of long training time to a certain extent.

The first layer is 28×28 for the input layer, because the pixels of the MNIST handwritten font image are 28×28. The second layer is the convolutional layer C1, which performs convolution operations on the input layer to extract features. In the second layer, 4 convolution filters are used, that is, C1 contains 4 feature maps; since each neuron in C1 is connected to the neighborhood of the input layer, and the neighborhood size is 5×5, you feel The size of the wild window is also 5×5, and the window moving step size is 1. Therefore, the size of the feature map of this layer is 24×24. The third layer is the down-sampling layer S2. The S2 layer mainly performs down-sampling processing on C1 to ensure that the amount of data is reduced as much as possible while retaining useful information. Since the image is locally related, the information of the entire area of the image in a certain neighborhood can be expressed with only one pixel, and down-sampling uses this principle. If the neighborhood of downsampling is too large, information will be lost, and if it is too small, the downsampling effect cannot be achieved. The fourth layer is a convolutional layer C3. C3 convolves the feature map of the previous layer S2. It should be noted that the relationship from S2 to C3 is not one-to-one. S2 contains 4 12×12 feature maps. But here the C3 layer contains 12 feature maps. The receptive field window of layer C3 is 5×5, and the moving step is 1, so the size of all feature maps in this layer is 8×8. All the feature maps in C3 are mapped using the feature maps in the S2 layer during the convolution process, and then superimposed. The fifth layer is the down-sampling layer S4, which is the same as S2, and the neighborhood size is 2×2. Therefore, the S4 layer has 12 feature maps, of which the size of the feature map is 4×4. The sixth layer is the output layer. Since there are 10 sample labels, the number of nodes in the output layer is 10. At the same time, all the feature maps of the S4 layer are modified into one-dimensional vectors by the algorithm, and they are directly connected to the output layer. In the convolutional neural network, the convolutional layer of the upper layer is mapped to obtain the down-sampling layer, so the up-sampling layer and the down-sampling layer are equal in the number of feature maps.

Based on the image classification method of DBNs, the layered training method can greatly improve the training speed. At the same time, after using the layered idea, the model's ability to express the semantics of complex images will also be greatly improved. According to the pictures in the MNIST library the number of input layer units is set to 784; the two hidden layers are defined as 100 units, because there are 10 numbers from 0 to 9 in the database, and the output layer of 10 units is selected; the final system model The structure is 28×28-100-100-10. The meaning of 28×28-100-100 is: the input layer and two hidden layers of DBNs, the number of sample types is 10, and the neural network structure has received data at this time. When training DBNs, the learning rate of both layers of RBM is 0.5, and the number of RBM training iterations is set to 10. After the training is completed, the weights learned by DBNs are transmitted to the network, and the weights are fine-tuned; the number of iterations of NN is set to 100, the learning rate is initially set to 0.01, and the Sigmoid function is used as the activation function of the neural network part. DBNs are fully connected, so there are 10,000 weights between the two hidden layers, and the number of weights between the input layer and the first hidden layer is 78,400, and the number of weights between the last two layers is 1000 A. After the training is completed, the system performs feature extraction on the test samples, and the output layer is a 1×10 vector. Only one bit of the vector is 1 and the remaining nine are 0; then it is combined with the label (a 1×10 vector) For comparison, if the two vectors are the same, the test result obtained by this test sample is correct.

4. Conclusion

At present, the use of deep learning algorithms for image recognition, on the one hand, the deep network used for recognition still has problems such as complex model structure and long training time. On the other hand, as an important basic model in deep learning, restricted Boltzmann machines have been widely studied and applied in many fields in recent years. The restricted Boltzmann machine provides a powerful tool for people to solve intelligence problems, and

provides new technologies and new ideas for research in other fields. However, there are still some problems on the restricted Boltzmann machine. On the one hand, due to the high computational complexity of the normalized parameters of the restricted Boltzmann machine, it is impossible to directly obtain the joint distribution between the hidden unit and the visible unit, and only some sampling methods can be used to obtain its approximate value. The process greatly increases the complexity of RBM training. On the other hand, the learning rate of the restricted Boltzmann machine has a great impact on the learning performance, so the choice of learning rate is very important.

References

[1] Xu Peng, Lin Sen. Traffic classification method based on C4.5 decision tree. Journal of Software, 2009, 20(10): 2692-2704.

[2] Wang Shuangcheng, Du Ruijie, Liu Ying. Learning and Optimization of Continuous Attribute Complete Bayesian Classifier. Chinese Journal of Computers, 2012, 35(10): 2129-2138.

[3] Ge Junwei, Sha Jing, Fang Yiqiu. BP algorithm with chaotic learning rate. Computer Engineering, 2010, 36(23):168-170.

[4] Lin Xumei, Mei Tao. A CMAC learning rate optimization method based on adaptive genetic algorithm. Journal of System Simulation, 2013, 17(12): 3081-3084.

[5] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems, 25(2):2012.