Research on Image Classification Algorithm Based on Convolutional Neural Network

LiHua Luo
College of Software Engineering, Guangdong University of science and technology, Dongguan, Guangdong 52300, China
luolihua@gdust.edu.cn

Abstract. Nowadays, we are in the information age. Pictures carry a lot of information and play an indispensable role. For a large number of images, it is very important to find useful image information within the effective time. Therefore, the excellent performance of the image classification algorithm has certain influence factors on the result of image classification. Image classification is to input an image, and then use a certain classification algorithm to determine the category of the image. The main process of image classification: image preprocessing, image feature extraction and classifier design. Compared with the manual feature extraction of traditional machine learning, the convolutional neural network under the deep learning model can automatically extract local features and share weights. Compared with traditional machine learning algorithms, the image classification effect is better. This paper focuses on the study of image classification algorithms based on convolutional neural networks, and at the same time compares and analyzes deep belief network algorithms, and summarizes the application characteristics of different algorithms.

Keywords: Image classification algorithm; Convolutional neural network; Deep belief network algorithm.

1. Introduction
In today's society, deep learning is progressing rapidly. Computer vision is one of the active research directions in deep learning applications, and image classification as a typical classification task is a basic and important research topic in the field of computer vision. In the process of image classification, feature extraction plays a vital role in improving the accuracy of classification. Therefore, for image classification tasks, deep learning has been widely studied and applied with its powerful feature extraction capabilities, and has attracted attention in the field of machine learning since it was proposed. Deep learning is proposed to solve the problems existing in machine learning feature extraction. Compared with the feature extraction of shallow learning models that rely on manual selection, the deep network structure of deep learning can extract data features layer by layer, making the features more obvious and easier to classify. Among them, convolutional neural networks have developed the fastest. The performance improvement of the convolutional neural network comes at the cost of a deeper network, more training parameters, and huge computer resource consumption. It is to apply the convolutional neural network to image recognition, in the hope that it can be automatically recognized...
through machine learning. Thereby reducing the task of image recognition, while providing more objective analysis results [1].

2. Analysis and development trend of research status at home and abroad
Image classification is a basic task in computer vision, and it is also a task for comparing almost all benchmark models. Image classification is to use a given label set. When an image is input, find a suitable category label in the label set, and assign the category label to the input image. The image needs to be preprocessed before classification. The original image contains a lot of interference, noise, etc., so the image needs to be cropped, denoised, and enhanced. After preprocessing, the interference and noise of the image is reduced, which helps to improve the classification accuracy of the image [2].

In traditional machine learning algorithms, image classification requires the extraction of image features to describe the image. When the entire image is used as the input of the classification algorithm, the amount of data calculated by the algorithm is huge. Secondly, the image contains redundant information such as background, which will reduce the classification efficiency and the accuracy rate. The main purpose of feature extraction is to reduce the dimensionality of the original image, map the original image to a low-dimensional feature space, and obtain low-dimensional sample features that best reflect the essence of the image or distinguish it. Image features are divided into four categories: color features, texture features, shape features, and spatial relationship features. After extracting different features, the features need to be fed into different learning algorithms as input [3].

With the continuous deepening of research on convolutional neural network technology, the application fields of this technology are becoming wider and wider. Convolutional neural network is a pattern classification method developed in recent years, which combines neural network technology and deep learning theory. Convolutional neural network has good learning ability for high-dimensional features and better generalization ability. It has been widely used in image recognition, speech recognition, pedestrian detection and other fields [4].

3. Image feature extraction
Image classification and recognition mainly go through image preprocessing, feature extraction, classifier design and other steps. Among them, feature extraction is an extremely important link. Image feature extraction is the core issue. The method and result of feature extraction will directly affect the accuracy of classification [5].

1. Color feature: Color feature is a global feature that describes the color information corresponding to an image or a local area of an image. There are many ways to describe color characteristics, which are mainly divided into color histograms, color sets, color moments, color aggregation vectors, and color correlation graphs [6].

2. Texture features: Texture features describe the repeated local patterns and arrangement rules in the image, and reflect the surface properties corresponding to the image or image area. The advantage of the texture feature is that the texture feature will not change no matter how the image is rotated, and it has strong anti-interference ability. However, when the resolution of the image changes, the texture characteristics may change significantly. In addition, illumination and reflection may also affect the texture characteristics [7].

3. Shape feature: The shape feature uses a feature description based on the image contour and the area boundary. However, its mathematical model is still not perfect [8].

4. Comparison of classification algorithms

4.1. Deep Belief Network Algorithm
The Deep Belief Network is a multi-layer generative graphics model, which is obtained by training and modifying the restricted Boltzmann machine in a greedy manner. The most notable feature of deep learning is that it can split a complex problem into several simple problems, first solve these simple problems in turn, and then integrate them to solve complex problems. As a widely used network model
in the field of deep learning, deep belief network is a multi-layer network structure. By continuously abstracting the data through multiple hidden layers in the deep belief network, the essential characteristics of the data can be better extracted. The pre-training stage and the fine-tuning stage are two stages that play different roles in the deep belief network training process. The pre-training stage uses an unsupervised learning algorithm, and the fine-tuning part uses a supervised classification algorithm. As an algorithm of deep learning, deep belief network algorithm can also be used in image classification. There are some shortcomings in the fine-tuning of part of the error back propagation algorithm of the deep belief network [9].

RBM is a bidirectional undirected graph containing two changeable layers. The key assumption made by RBM is that the hidden layer is conditionally independent given the explicit layer, and vice versa. Consider an RBM with n explicit layers and m hidden layers. Each explicit layer is only related to m hidden layers and is independent of other explicit layers. That is, the state of this explicit layer is only affected by m hidden layer nodes. For each hidden layer node, it is only affected by n explicit layers. This feature makes RBM training easier. Its energy function is defined as:

$$E(v, h) = - \sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n} \sum_{j=1}^{m} \omega_{ij} v_i h_j$$

Among them, ai and bj are the deviations of the visible layer vi and the hidden layer hj, respectively. wij is the weight between the explicit layer i and the hidden layer j. ai, bj are the binary states of the explicit and hidden layers.

As a multi-layer belief network, each layer of the DBN is a restricted Boltzmann machine, which is stacked on top of each other to form the DBN. The hidden layer of the first layer of RBM is regarded as the visible layer of the second layer of RBM. The second layer of RBM will learn the feature distribution of the hidden layer of the first layer of RBM. The input of the first layer of RBM is the input layer of the entire network. As the layers are stacked, the network will learn increasingly complex combinations of features from the original data [10].

4.2. Convolutional Neural Network

Convolutional Neural Network (Convolutional Neural Network) is a feed-forward neural network designed specifically for image recognition problems, imitating the multi-layer process of human image recognition: the eye observes the image; some cells in the cerebral cortex perform preliminary processing and find the shape Edge, direction; abstract judgment shape; further judgment of image category. The structure diagram of the convolutional neural network constructed is shown in Figure 1.

![Fig. 1 Structure diagram of convolutional neural network](image)

Compared with general neural networks, the most obvious feature of CNN is the addition of convolutional layers and pooling layers. For image translation, deformation and scaling, the convolutional neural network has good fault tolerance. Traditional feature extraction requires preprocessing of the original data, while CNN reduces the preprocessing process and can directly process the multidimensional original data of the image. Convolutional neural networks have three notable features: local receptive fields, weight sharing and pooling layer [11].
The input data is an image, which is represented by a digital matrix. The convolutional layer contains a large number of convolution filters. These filters slide in the original image according to a certain rule, and perform convolution operations on the area corresponding to the original image. Each filter can get a feature map, and each convolution The image features extracted by the product filter are different. The parameters in the filter are initialized by the normal distribution function, and the gradient of the error between the current network output and the real output is reversed, and the parameters in the filter are updated layer by layer, and finally As a result, the error between the output of the network and the real output is as small as possible. After the convolution process, an activation function is used to limit the range of the convolution output, which simplifies the computational complexity of the neural network without causing the problem of gradient disappearance, and enables the neural network to converge quickly [12].

The data sampling in the sampling layer is to reduce the dimensionality of the data, and slide in the feature map in turn, and only retain the largest pixel value in the window or the average pixel value in the window. After this operation, the feature map is reduced to four times the area of the original feature map. One part, greatly reducing the amount of network calculations. The output of the fully connected layer is a column vector, and the convolutional neural network can be understood as an encoder, where the output of the fully connected layer is the encoding vector of the image. The numerical value represents the probability that the input picture belongs to the category, and the sum of all the elements of the vector is one [13]. The optimization process is to continuously update the parameter values of the filters in the convolutional network to maximize the probability of a certain category of the final output vector, which is the true category of the input image [14]. Due to the shared convolution kernel, the convolutional neural network greatly reduces the number of parameters of the network model and improves the computational efficiency. Using multiple convolution kernels can extract a variety of different features of the image, avoiding a single feature from adversely affecting the classification results. The low-level convolution kernel extracts features with low semantics such as edges and lines, and the high-level convolution kernel extracts features such as shape and object contours [15].

Comparing the deep belief network algorithm and the convolutional neural network algorithm, you can see that the classification accuracy of the convolutional neural network algorithm far exceeds the deep belief network algorithm, because the deep belief network algorithm extracts a single feature of the image as input, it can only express part of the image information, and cannot accurately describe the image. Each convolution filter of the convolutional neural network can be regarded as a separate image feature extractor [16]. The low-level filter extracts low-semantic features, and the high-level filter extracts high-level semantic features. The features extracted by a large number of hierarchical filters can express an image more completely. The classification accuracy rate has been significantly improved [17].

5. Conclusion
This paper makes a comparative study of deep belief network algorithm and convolutional neural network algorithm. In the deep belief network algorithm, when the same image feature is used as input and different classification algorithms are used, the classification accuracy rates obtained vary greatly. When different image features are used as input, when the same image classification algorithm is used, the difference in classification accuracy obtained is small. Convolutional neural network algorithms excel in the task of image classification. This research has done a preliminary study on the deep learning image classification method based on convolutional neural network, which can achieve high recognition accuracy.

References
[1] Cheng G E, Sun G Q. Research on Image Classification Based on Convolutional Neural Network. 2018.
[2] Jiang Xuan. Research on scene image classification algorithm based on deep learning [D]. Beijing University of Posts and Telecommunications. 2019.
[3] Li D, Deng L, Cai Z. Research on image classification method based on convolutional neural network [J]. Neural Computing and Applications, 2020(8).

[4] Zhou Y T. Research on Image Classification Method Based on Improved Multi-channel Convolutional Neural Network Model [J]. Journal of Jiamusi University(Natural Science Edition), 2019.

[5] Tang K, Wen Z, Chen Y, et al. Research on Garbage Image Classification Method Based on Convolutional Neural Network [J]. International Journal of Computer Applications Technology and Research, 2020, 9(1):037-040.

[6] Tan Z, Li K. Differential evolution with mixed mutation strategy based on deep reinforcement learning [J]. Applied Soft Computing, 2021: 107678.

[7] Chen Y. Sports Sequence Images Based on Convolutional Neural Network[J]. Mathematical Problems in Engineering, 2021, 2021(4):1-14.

[8] Wang Chao. Research and implementation of image classification technology based on convolutional neural network [J]. Computer Knowledge and Technology, 2016(35).

[9] Jing G U, Zhang K, Zhu Y. Research on weld defect image classification based on convolutional neural network [J]. Journal of Applied Optics, 2020, 41(3):531-537.

[10] Qiang W, Li X, Shi C. Research on image classification based on HP — Net convolutional neural networks[C]// 2017 3rd IEEE International Conference on Computer and Communications (ICCC). IEEE, 2017.

[11] Dong J, Li X. An Image Classification Algorithm of Financial Instruments Based on Convolutional Neural Network [J]. Traitement du Signal, 2020, 37(6):1055-1060.

[12] Daqin W, Haiyan H. The Algorithm Research of Image Classification Based on Deep Convolutional Network [C]/ 2018:231-233.

[13] Wu Z, Jiang S, Zhou X, et al. Application of image retrieval based on convolutional neural networks and Hu invariant moment algorithm in computer telecommunication[J]. Computer Communications, 2020, 150:729-738.

[14] Aizezi Y, Jiamali A, R Abudurexiti, et al. Research on Image Recognition Method Based on Deep Learning Algorithm [C]// 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2018.

[15] D Song, Lu K, Dai X. Research on mail classification algorithm based on improved convolutional neural network [J]. Journal of Physics: Conference Series, 2021, 1871(1):012097 (8pp).

[16] Yunyan, Wang, Chongyang, et al. Image Classification Based on transfer Learning of Convolutional neural network [C]// The 38th China Control Conference. 0.

[17] Wang J, Li H, Yin S, et al. Research on Improved Pedestrian Detection Algorithm Based on Convolutional Neural Network [C]// 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 2019.