Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2022.030106
ISSN 2790-2692 Vol. 3, Issue 1: 69-75

Deep Learning Algorithm Based on Multi-layer Convolutional Neural Network

Malik Teekaraman*

GLA University, India

**corresponding author*

Keywords: Multi-layer Channel, Convolutional Neural Network, Deep Learning, Algorithm Application

Abstract: Convolutional neural network (CNN) is a kind of deep neural network(NN) composed of convolutional computation, which has excellent image feature extraction ability. This paper mainly studies the application of deep learning algorithm based on multi-layer CNN. This paper first introduces the model structure of CNN in detail, and lists the advantages of CNN and the model′ s parameter update method. Then, aiming at the problem of insufficient feature extraction in image recognition, this paper designs a multi-layer CNN structure (Mc-cnn). Multiple single view CNNs are used to train the image input of a certain view respectively, and the learned features are fused. The experimental results can verify the effectiveness and feasibility of the model network designed in this paper for image recognition.

1. Introduction

Deep learning can directly take the original data form as the input of the network model, process the original data into the feature expression required by a specific task through the model, extract features layer by layer, and finally achieve the task goal. The whole process does not need artificial operation, it is said that deep learning reduces the labor costs, and among them, the most basic work is to study network model, and the core content is the characteristics of learning, will also contain feature abstraction, etc., often when finish the work division of labor cooperation, also contains the meaning of the word " depth " . CNN algorithm is relatively representative. Currently, it is widely used in computer vision and image processing. In a sense, CNN is also a part of NN, which is operated by convolution operation. The sample data is deeply mined and learned, and then the data features are extracted. CNN can have many convolutional layers, the number of convolutional layers varies from a small number to a large number, the network structure Settings are also varied, and the performance of CNN is also improved year by year. At the same time, the CNN in the field of image classification application has an irreplaceable position, image classification, as the name implies, is an image input to the system, then output the image content belong to which kinds of problem, often study this problem in the field of computer vision, its application in the actual scene is very much also.

Image recognition technology has a history of many years, early there are many image recognition methods, among which the classic is two edge detection algorithm, differential operator and Canny, and corner detection algorithm. Differential operator edge detection algorithm is based on the first mathematical derivative, through the gradient matrix, binarization and other operations, the image edge will be finally obtained. Canny edge detection algorithm needs to implement filtering processing on the image, the amplitude of the gradient also needs to be calculated, and the direction of the gradient, non-maximum suppression and detection of image edge is also necessary. Generally speaking, traditional image recognition algorithms can only extract low-level image features and contribute to obtaining high-level feature information. In view of this, NN algorithms applied to image recognition appear. In recent years, image recognition technology tends to adopt NN algorithms, such as CNN algorithm, attention NN, autocoding NN, generative adversarial network and pyramid NN. In the late 20th century, CNN algorithm was applied to handwritten digit character recognition task, and the model adopted was LeNet. In recent years, GoogLeNet series of CNN models and ResNet have been successively applied to image recognition tasks.

In this paper, multi-layer convolution NN model, make full use of image information, multiple points of view and based on the depth of the clustering algorithm, on the premise of no class label information automatically learning image class standard, achieve tag data, fast and efficient automatic tagging these data sets can be used as a subsequent model training data, improve the model classification accuracy.

2. Multi-layer CNN Image Recognition Algorithm

2.1. Basic Structure of CNN

The structure of CNN follows a fixed pattern, that is, the convolutional layer is connected to the pooling layer, and then to the fully connected layer. Finally, the network is trained according to the loss function and the back propagation algorithm.

(1) Convolutional layer

As the core of the CNN, the convolutional layer is mainly composed of convolution calculation and activation function, and its main function is to extract image features. For detailed calculation, see Equation (1).

$$y_i^l = f(w_i^l * x_i^l + b_i^l) \tag{1}$$

Where, y, W, x and b are the output value, convolution kernel parameter, input value and bias of the ith neuron in layer L, respectively.

* is the convolution operation, and f() is the activation function.

Activation function is another important part of CNN, which is mainly used to increase the nonlinear expression of image features by CNN. Common activation functions such as Softplus, Softsign, Mish, etc. Activation functions have various properties: differentiability, monotonicity and non-saturation to ensure that CNN can carry out stable back propagation in the deep convolutional layer: The finite range ensures that CNN can update parameters stably. Fewer parameters will reduce the operation cost of CNN.

(2) Pooling layer

The pooling layer only contains pooling operation, which is a special convolution operation, and it has the same " kernel " as the convolution calculation to slide the eigenvalues of each position. Different from convolution operation, the " kernel " of pooling operation has no parameters, and the step size of the kernel is usually greater than or equal to the width of the kernel. The function is: since the features extracted by the convolution layer are affected by the sample conditions, they contain a lot of noise. Therefore, CNN removes the features containing noise through pooling layer to increase the generalization ability of the network. As the number of network layers increases, the features become more abstract. The pooling layer can combine multiple convolution kernels to further improve the feature expression ability.

Different from the convolution layer, the pooling layer in state Q has various forms, among which the two most widely used pooling forms are as follows:

Maximum pooling selects the largest value in the feature region as the output, which has the advantage of good sparsity. This is the same activation mechanism of human brain vision system for features, so it can achieve good experimental results in various computer vision tasks.

The average pooling takes the average value as the average value of the output feature region of the " convolution kernel " . Its advantage is that it can retain relatively complete features and is suitable for image recognition tasks with less sample noise.

(3) Fully connected layer

The fully connected layer consists of neurons connected in a fully connected way and classifiers at the end. Its main function is to complete the image recognition task of the network. It usually has two layers:

The number of neurons in the first layer is usually the same as the number of image pixels, and each pixel corresponds to the neuron one by one.

The number of neurons in the second fully connected layer is usually the same as the number of recognized categories, and the neurons in this layer represent the feature mapping of pixels onto the image category.

The classifier is mainly reflected in the output mode of neurons in the second fully connected layer. Different classifiers lead to different output modes of neurons. The most common classifier is Softmax classifier, which is derived from Logistic regression and can be regarded as a combination of multiple binary classifiers. The specific calculation of Softmax classifier in CNN is shown in Equation (2).

$$\bar{y}_{ij} = \frac{e^{o_j}}{\sum_{j=1}^{N} e^{o_j}}; j = 1,2,...,N; i = 1,2,...,M \tag{2}$$

Where, Oj represents the output value of the JTH neuron in the output layer, y represents the possibility that the ith sample is the JTH category, N is the number of neurons in the last layer (the number of predicted categories), and M is the number of samples.

2.2. Multi-layer CNN Recognition Model

The convolution NN model (Mc-cnn) constructed in this paper is mainly composed of three parts: the traditional CNN, the improved cross-layer connection based on the residual module and the Wider module optimized on the inception-Resnet-V2 module. Cross-layer connection can make the training of the deep network easier, and the Wider module can obtain multi-scale image feature information. The residual module introduced in Mc-cnn network does not use ReLu function. It does not use the complete residual model in ResNet, but it draws lessons from this idea.

The Wider module plays an extremely important role in the whole Mc-cnn network and contributes greatly to the final good image classification performance of the network. There are two branches in the width module, and the convolution kernel size is 1*1 and 3*3 respectively. One branch refers to the cross-layer connection, and the output filter banks of the last two layers of the two branches form an output vector through the sum operation.

There are four cross-layer connections in each block of the Mc-cnn model, of which one is outside the Wider module (except the last block of the model) and the other three are inside the Wider module. Each Wider module of the Mc-cnn model contains two parallel networks. In this way, the width of the network can be increased to learn data features from a new perspective, and each network adopts a different size of convolution kernel. The information confluence of each network uses a sum operation, which is also used in ResNet. ReLu was added after each Wider module. In addition, using dropout after each block can avoid overfitting. The global average pooling layer does not bring the number of parameters to the model; the Softmax layer implements probabilistic classification, which is contained in a block package at the end. Mc-cnn network can not only deal with overfitting, alleviate gradient disappearance, prevent network saturation, etc., but also accelerate the training speed of the model, so as to achieve better image classification effect.

Inception-resnet-v2 module contains many branches and uses 4 convolution kernels of 1*1 size and 3 convolution kernels of 3*3 size. Besides, the identity mapping in the residual module is also used separately, so the structure is relatively complex. The Mc-cnn model optimizes its structure and parameters, which greatly reduces the number of branches and the number of convolution operations. Only 3 convolution kernels of 3*3 size and 2 convolution kernels of 1*1 size are used, and no single identity mapping operation is used. Although the module of inception-Resnet-V2 has been greatly simplified, the Mc-cnn module continues to maintain the richness of receptive field. The concept of receptive field can be understood as follows: convolutional layers are not independent of each other, and there is a mapping relationship between them. There is a certain relationship between the pixels of different layers of the picture, and this relationship from the latter layer to the previous layer is the receptive field. The richness of receptive field is helpful to improve the performance of image classification.

3. Image Recognition Simulation Experiment

3.1. Experimental Environment

The hardware environment and software environment of this experiment are shown in Table 1.

Table 1. Configuration table of experimental software and hardware environment

<table>
<tr><th>Experimental environment</th><th colspan="2">Configuration instructions</th></tr>
<tr><td rowspan="3">Hardware environment</td><td>CPU</td><td>Intel Core i7-10700</td></tr>
<tr><td>GPU</td><td>NVIDIA RTX 2080</td></tr>
<tr><td>Memory</td><td>16GB</td></tr>
<tr><td rowspan="3">Software environment</td><td>System</td><td>Windows 10</td></tr>
<tr><td>Algorithmic programming environment</td><td>TensorFlow 1.7</td></tr>
<tr><td>Programming language</td><td>Python 3.6</td></tr>
</table>

3.2. Experimental Data Set

Fish4Knowledge dataset is a commonly used data set for fish image classification. Fish images in it are all taken from real underwater environment. This dataset contains a total of 23 fish species and a total of 27,370 images, among which fish images are affected by the shooting Angle and illumination. There is a big difference in background and posture.

4. Analysis of Experimental Results

4.1. Performance Comparison Before and after Model Improvement

Table 2. Performance comparison before and after improvement

	CNN	MC-CNN
Accuracy(%)	97.15	99.37
Classification of time(s)	87.56	61.02

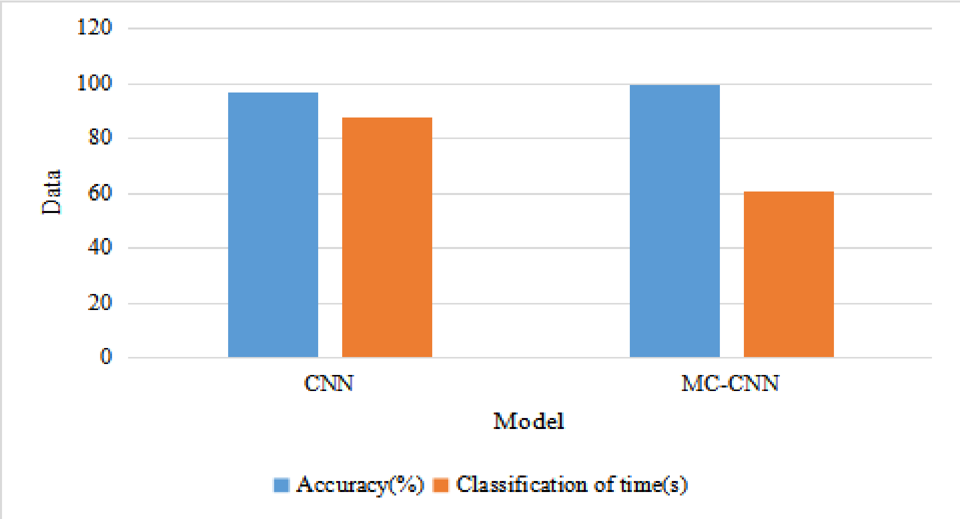

Figure 1. Performance comparison of CNN before and after improvement

As shown in Figure 1 and Table 2, the classification accuracy of the Mc-cnn model is higher than that of the CNN model, indicating that the Mc-cnn model extracts more meaningful image features for classification and verifies the effectiveness of the multi-layer model. Moreover, the classification time of the model is greatly reduced, and the classification efficiency is improved, which verifies the effectiveness of multi-layer on the improved model.

4.2. Different Model Performance

Table 3. Performance comparison results of different models

	ResNet50	MC-CNN	GoogleNet	VGG16
Accuracy(%)	97.92	99.37	99.06	94.87
Classification of time(s)	201.58	61.02	107.59	412.78

As shown in Table 3 and Figure 2, this paper also compares the performance of the algorithm with other commonly used convolutional networks for image classification, such as ResNet50, GoogLeNet and VGG16. The proposed multi-layer CNN model has certain advantages over other models in test set classification accuracy and test time, and can meet the needs of real-time image recognition.

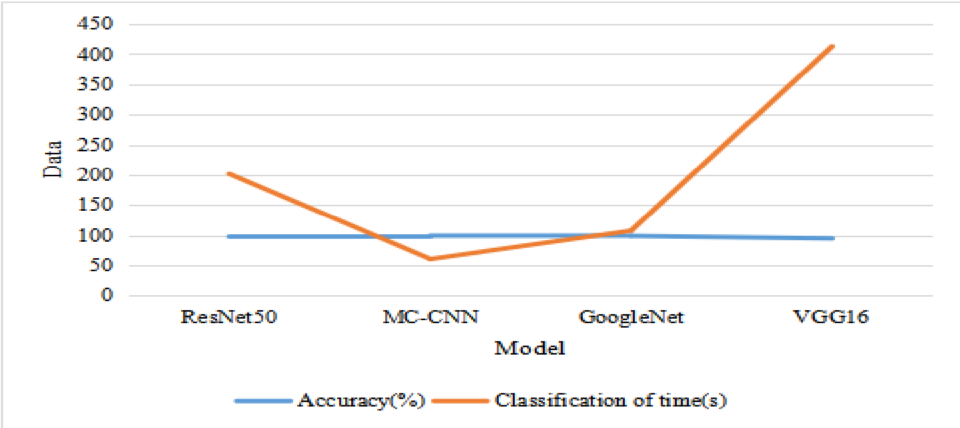

Figure 2. The test results of the four models were compared on the dataset

5. Conclusion

Convolution NN by the animal′ s brain signal transmission principle of visual cortex neurons inspired a deep learning NN model is designed, through its own convolution and pooling layer network structure, special sharing, sparse connection with local awareness, parameter characteristics of applications in the field of all kinds of recognition shows powerful feature extraction and accurate classification ability. In this paper, features are not fully utilized in image classification, which affects the classification results of the model. This paper proposes an image recognition algorithm based on multi-layer CNN (Mc-cnn), which can fully learn the features of the image and improve the accuracy of image recognition by fusing the low-level features and high-level features. In this paper, multi-layer CNN is applied to the problem of image recognition. Although part of the problem is solved, there are still some problems to be solved. How the image classification algorithm based on multi-layer CNN determines the weight relationship between each channel needs to be further solved. The next step is to apply the optimization algorithm to CNN to automatically learn and determine the hyperparameters of the model.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Kaur P, Kaur H. Kernel CNN Deep Learning Algorithm To Classify LIVER Disease. International Journal of Engineering and Advanced Technology, 2019, 8(6):2002-2007. https://doi.org/10.35940/ijeat.E7826.088619

[2] Haenssle H A, Fink C, Schneiderbauer R, et al. Man against machine: diagnostic performance

of a deep learning CNN for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology, 2018, 29(8):1836-1842.
[3] Luo Y, Fan Y, Chen X. Research on optimization of deep learning algorithm based on CNN. Journal of Physics: Conference Series, 2021, 1848(1):012038 (5pp). https://doi.org/10.1088/1742-6596/1848/1/012038
[4] Shahril R, Saito A, Shimizu A, et al. Bleeding Classification of Enhanced Wireless Capsule Endoscopy Images using Deep CNN. Journal of Information Science and Engineering, 2020, 36(1):91-108.
[5] Reddy A R, Rao A N. An active model for ranging by deep CNN and elephant herding optimization algorithm (DCNN-EHOA) in WSNs. International Journal of Pervasive Computing and Communications, 2021, ahead-of-print (ahead-of-print).
[6] Anjani I A, Pratiwi Y R, Nurhuda S. Implementation of Deep Learning Using CNN Algorithm for Classification Rose Flower. Journal of Physics: Conference Series, 2021, 1842(1):012002 (11pp). https://doi.org/10.1088/1742-6596/1842/1/012002
[7] Putra B, Amirudin R, Marhaenanto B. The Evaluation of Deep Learning Using CNN (CNN) Approach for Identifying Arabica and Robusta Coffee Plants. Journal of Biosystems Engineering, 2022, 47(2):118-129. https://doi.org/10.1007/s42853-022-00136-y
[8] Giovanni, L, F, et al. Lung nodules diagnosis based on evolutionary CNN. Multimedia Tools and Applications, 2017, 76(18):19039–19055. https://doi.org/10.1007/s11042-017-4480-9
[9] Mukilan P, Semunigus W. Human and object detection using Hybrid Deep CNN. Signal, Image and Video Processing, 2022, 16(7):1913-1923. https://doi.org/10.1007/s11760-022-02151-0
[10] Alotaibi G, Awawdeh M, Farook F F, et al. Artificial intelligence (AI) diagnostic tools: utilizing a CNN (CNN) to assess periodontal bone level radiographically—a retrospective study. BMC Oral Health, 2022, 22(1):1-7. https://doi.org/10.1186/s12903-022-02436-3
[11] Gautam A, Singh V. CLR-based deep convolutional spiking NN with validation based stopping for time series classification. Applied Intelligence, 2020, 50(3):830-848. https://doi.org/10.1007/s10489-019-01552-y
[12] Rohan A, Rabah M, Kim S H. CNN-based Real-Time Object Detection and Tracking for Parrot AR Drone 2. IEEE Access, 2019, PP(99):1-1. https://doi.org/10.1109/ACCESS.2019.2919332
[13] Ruchai A N, Kober V I, Dorofeev K A, et al. Classification of Breast Abnormalities Using a Deep CNN and Transfer Learning. Journal of Communications Technology and Electronics, 2021, 66(6):778-783. https://doi.org/10.1134/S1064226921060206
[14] Deepa B G, Senthil S. Predicting invasive ductal carcinoma tissues in whole slide images of breast Cancer by using CNN model and multiple classifiers. Multimedia Tools and Applications, 2022, 81(6):8575-8596. https://doi.org/10.1007/s11042-022-12114-9
[15] Hur T, Kim L, Park D K. Quantum CNN for classical data classification. Quantum Machine Intelligence, 2022, 4(1):1-18. https://doi.org/10.1007/s42484-021-00061-x
[16] Wang Y, Zhang L, Shu X , et al. Feature-sensitive Deep CNN for Multi-instance Breast Cancer Detection. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021, PP(99):1-1.
[17] Cui R, Azuara G, Scalea F, et al. Damage imaging in skin-stringer composite aircraft panel by ultrasonic-guided waves using deep learning with CNN:. Structural Health Monitoring, 2022, 21(3):1123-1138. https://doi.org/10.1177/14759217211023934
[18] Kumari N, Bhatia R. Efficient facial emotion recognition model using deep CNN and modified joint trilateral filter. Soft Computing, 2022, 26(16):7817-7830. https://doi.org/10.1007/s00500-022-06804-7