

Research Article

Improved Convolutional Neural Image Recognition Algorithm based on LeNet-5

Lijie Zhou¹  and Weihai Yu²

¹Yantai Vocational College, Yantai 264670, Shandong, China

²Yantai Research Institute of Education Science, Yantai 264003, Shandong, China

Correspondence should be addressed to Lijie Zhou; ytzhoulijie@163.com

Received 23 July 2022; Accepted 21 September 2022; Published 8 October 2022

Academic Editor: Arvind Kumar

Copyright © 2022 Lijie Zhou and Weihai Yu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Convolutional neural network (CNN) is a very important method in deep learning, which solves many complex pattern recognition problems. Fruitful results have been achieved in image recognition, speech recognition, and natural language processing. Compared with traditional neural network, convolutional weight sharing, sparse connection, and pooling operations in convolutional neural network greatly reduce the number of training parameters, reduce size of feature map, simplify network model, and improve training efficiency. Based on convolution operation, pooling operation, softmax classifier, and network optimization algorithm in improved convolutional neural network of LeNet-5, this paper conducts image recognition experiments on handwritten digits and face datasets, respectively. A method combining local binary pattern and convolutional neural network is proposed for face recognition research. Through experiments, it is found that adding LBP image information to improved convolutional neural network of LeNet-5 can improve accuracy of face recognition to 99.8%, which has important theoretical and practical significance.

1. Introduction

An image is a visual form that describes state of objective world or transforms energy into a two-dimensional plane. Image recognition is based on image recognition algorithm, which is a technology for computer to analyze specific types of objects contained in images [1]. Image recognition is used in many fields, such as biomedical pathological diagnosis, military dangerous environment reconnaissance, and intelligent portrait recognition in field of life services. Images contain a lot of useful information for human beings [2]. Effective image recognition is of great significance to progress of whole society.

Image data not only contains information about human beings but also brings huge profits to companies that master image data. The technology giant Google Inc. used data to bring about \$54 billion in revenue in the United States in 2009. This makes people see importance of mining information from images. With rapid development of modern

industries, image data has exploded and accumulated at an unprecedented speed. How to extract information from images has become a problem worthy of study. With deepening of information mining, more and more algorithms are used to mine information in images [3]. Machine learning is known for its efficient mining ability, adaptive ability, and learning ability. It is an algorithm with excellent performance suitable for image processing and big data. The convolutional neural network algorithm is an excellent deep learning algorithm specially designed to process image data.

2. Related Work

The idea of neural network is to let machine learn way of human thinking to simulate way that humans obtain information from nature to deal with problems. Based on this, scientists put forward artificial neural network theory (artificial neural network for short ANN) [4, 5] proposed back propagation (BP) algorithm of artificial neural network. The

BP algorithm was applied to shallow forward neural network model, and a hidden layer was added to solve XOR gate problem that perceptron could not solve [6]. The computational complexity of optimization problem is reduced, and the algorithm becomes most basic algorithm of neural network. From then on, research and application of neural network began to recover. In the 1860s, concept of convolutional neural networks was proposed. Reference [7] found that visual image of a small area in cat's visual cortex is only processed by one neuron, which is receptive field. Reference [8] proposed that this is first prototype of a convolutional neural network, in which visual system is modeled and can be recognized when objects are displaced or even slightly deformed. The neurocognitive machine proposed by scientists contains two types of neurons, one neuron S-cells extract features to filter convolution kernel of convolutional neural network, and other neuron C-cell is antideformation for convolutional neural network. Activation function and pooling: convolutional neural networks are developed on basis of multilayer perceptrons. Reference [9] proposed to be derived based on an efficient BP training method and first achieved success in field of English handwriting recognition. Convolutional neural network is the first artificial neural network that has been successfully trained in deep learning [10]. Subsequently, various aspects of convolutional neural network have been continuously improved, making it a most potential and successful model. Convolutional neural network is widely used. There are applications in face recognition, pedestrian detection, natural language processing for robot navigation, medicine discovery, disaster climate discovery, and even Go AI programs [11]. At present, there are many research studies on convolutional neural network in world, mainly aiming at improving the structure of convolutional neural network. Domestic research is still in its infancy.

Image recognition is the basis of image classification and detection, and correct image classification results are of great significance to development of computer vision. Image recognition is a research that analyzes image categories by extracting feature information of images. Because image information is very complex, it is particularly important to select effective features [12]. Although algorithms for feature extraction are continuously proposed, traditional algorithms are difficult to meet performance and efficiency requirements of massive images. After more than 50 years of development in image recognition, a large number of excellent machine learning algorithms have emerged [13]. Shallow neural network algorithms such as artificial neural network (ANN), support vector machine model (SVM), Gaussian mixture model (GMM), and conditional random field (CRF) are correct in classifying a series of objects such as fruits, rock formations, and peanuts rate of more than 90%. The shallow neural network also has problems in obtaining good classification [14]. For example, if you need to manually select features, training effect after deepening number of layers is not ideal [15]. Deep learning is a deep network algorithm that automatically learns image features [16]. Deep learning is divided into supervised learning (Supervised Learning) and

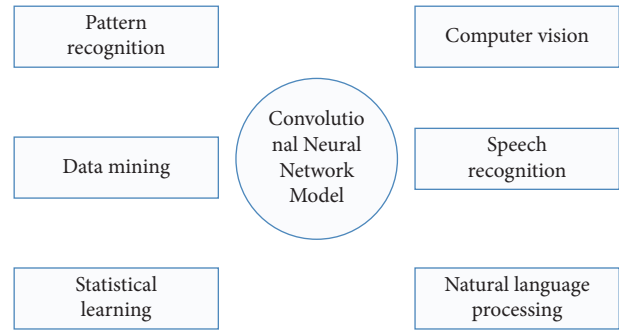


FIGURE 1: Deep learning application areas.

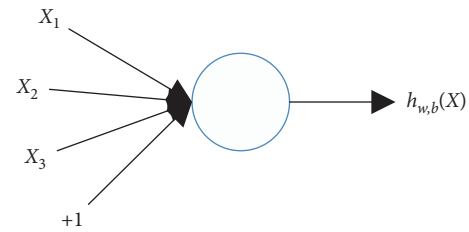


FIGURE 2: Neuron model.

unsupervised learning (Unsupervised Learning) according to whether there are labels [17].

2.1. Convolutional Neural Network Model. More and more scholars in various countries have joined in the research of machine learning and various high-tech companies. Machine learning is becoming more and more intelligent and has made outstanding achievements in various fields, and its results are continuously applied in practice. Figure 1 shows the application of deep learning in various fields.

Neuron is a basic processing unit of neural network, generally has multiple inputs and one output. The basic structure model is shown in Figure 2.

The convolution operation has two excellent properties: sparse connection, that is, in convolution operation, one input neuron is only connected to corresponding part of output neurons, and one output neuron is only connected to corresponding part of input neurons. Figure 3(a) shows the sparse connection method of convolution, and Figure 3(b) shows the full-connection method. It can be clearly seen that each output neuron is connected to all input neurons in the full-connection method, while in convolution sparse connection method, an output neuron is only connected to corresponding part of input neurons, for example, output neuron s_3 is only connected to input neurons x_2 , t_3 , and x_4 .

Parameter sharing, that is, parameters of each part of convolution operation are shared. Figure 4(a) shows the convolution operation, and Figure 4(b) shows full-connection operation. It can be seen that weight parameters of each connection in full-connection operation are different. Therefore, when the number of input and output neurons is 5, the required number of parameters is 25. In convolution operation, because of the nature of parameter sharing, such as

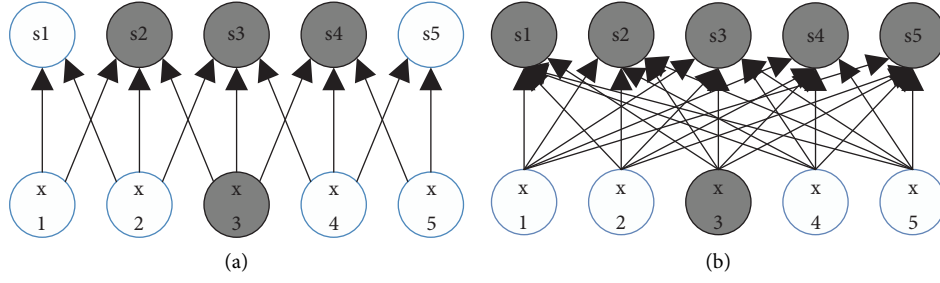


FIGURE 3: Sparse connection and full connection ((a) convolution and (b) full connection).

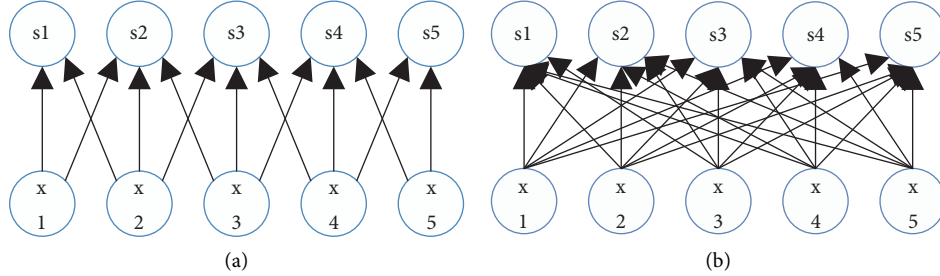


FIGURE 4: Parameter sharing ((a) convolution and (b) full connection).

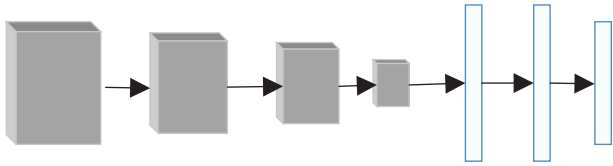


FIGURE 5: Schematic diagram of convolutional neural network node connection.

weight parameters of x and s , connection is same, so when the number of input and output neurons is 5 and size of convolution kernel is 3, the number of parameters required is 3.

The convolutional neural network inputs an image, and when image is only black and white, the image can be abstracted into a two-dimensional matrix in space. When image is color, there are three color channels, and nodes of each layer of neural network are composed of three RGB vectors, which can abstract image into a three-dimensional matrix in space. Figure 5 is a schematic diagram of connection of convolutional neural network nodes. It can be seen from figure that multiple nodes in the upper layer of convolutional neural network are connected to a single node in the lower layer of neural network.

3. Softmax Algorithm

In softmax, what needs to be solved is a multiclass problem (assuming k classification problem), so value range of class label y is k different integer values, and different values represent different classes. For training set $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$ with m samples, there are $y^{(i)} \in \{1, 2, \dots, k\}$, of which $i \in \{1, 2, \dots, m\}$. Given a sample $(x^{(i)}, y^{(i)})$, we can use hypothesis function to calculate probability value $p(y^{(i)} = j|x^{(i)})$ that sample belongs to each category j , that is, estimate sample $x^{(i)}$ for each category probability of

outcome. So, suppose output of function is a k -dimensional numeric vector (the vector element values sum to 1), where i component of vector represents estimated probability value of sample belonging to i class. Suppose function is of the following form:

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)}; \theta) \\ p(y^{(i)} = 2|x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k|x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}, \quad (1)$$

where $\theta_1, \theta_2, \dots, \theta_k \in R^{n+1}$ is a model parameter. If a $k \times (n+1)$ matrix is used to represent parameter θ , there are

$$\text{following: } \theta = \begin{bmatrix} \theta_1^T \\ \theta_2^T \\ \vdots \\ \theta_k^T \end{bmatrix}.$$

The cost function of softmax algorithm is as follows:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right], \quad (2)$$

where $1\{\cdot\}$ is an indicative function. For minimization problem of $J(\theta)$, there is no closed-form solution, and generally an iterative optimization algorithm (such as gradient descent method) is used. After derivation, gradient formula is as follows:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} (1\{y^{(i)} = j\} - p(y^{(i)} = j|x^{(i)}; \theta))]. \quad (3)$$

In the gradient descent method, each iteration needs to be updated as follows: $\theta_j = \theta_j - \alpha \nabla_{\theta} J(\theta)$, ($j = 1, \dots, k$).

Usually, softmax cost function will add weight attenuation term [18], which can solve numerical problem caused by parameter redundancy of softmax. The cost function with weighted attenuation term is as follows:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^k e^{\theta_i^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^k \sum_{j=1}^n \theta_{ij}^2, \quad (4)$$

where $\lambda > 0$ is the weight attenuation coefficient. Its gradient is calculated as follows:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} (1\{y^{(i)} = j\} - p(y^{(i)} = j | x^{(i)}; \theta))] + \lambda \theta_j. \quad (5)$$

Suppose a labeled sample set $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$ contains M samples. The following equation describes the use of batch gradient descent method to solve neural network model. For a single sample (x, y) , cost function is given as follows:

$$J(W, b; x, y) = \frac{1}{2} \|h_{W,b}(x) - y\|^2. \quad (6)$$

Then, cost function of whole m samples is defined as follows:

$$\begin{aligned} J(W, b) &= \left[\frac{1}{m} \sum_{i=1}^m J(W, b; x^{(i)}, y^{(i)}) \right] + \frac{\lambda}{2} \sum_l \sum_i \sum_j (W_{ji}^{(l)})^2 \\ &= \left[\frac{1}{m} \sum_{i=1}^m \frac{1}{2} \|h_{W,b}(x^{(i)}) - y^{(i)}\|^2 \right] + \frac{\lambda}{2} \sum_l \sum_i \sum_j (W_{ji}^{(l)})^2. \end{aligned} \quad (7)$$

The gradient descent method updates parameters w and B according to the following formula:

$$\begin{aligned} W_{ij}^{(l)} &= W_{ij}^{(l)} - \alpha \frac{\partial}{\partial W_{ij}^{(l)}} J(W, b), \\ b_i^{(l)} &= b_i^{(l)} - \alpha \frac{\partial}{\partial b_i^{(l)}} J(W, b), \end{aligned} \quad (8)$$

where α is the learning rate, and the most critical step is to calculate partial derivative. Let's introduce effective calculation method of partial derivative: back-propagation algorithm.

4. Results

We trained five different convolutional neural networks, and relationship between loss value and the number of iterations during training process is shown in Figure 6. In order to clearly see difference of each curve, we discard selection process of first 1000 times because their loss value is relatively large. The figure depicts training process from 1000 to

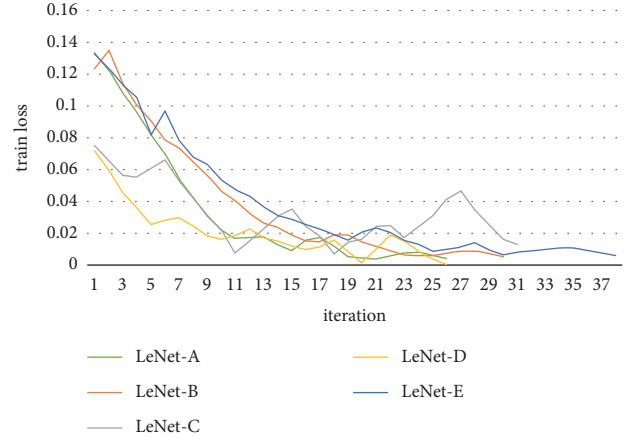


FIGURE 6: Training loss value vs. number of iterations.

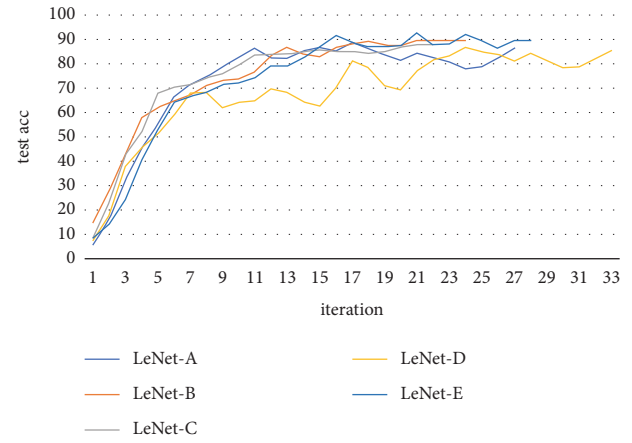


FIGURE 7: Test accuracy vs. iterations.

10000 epochs. It can be seen from the figure that when the number of generations reaches 3000, the loss value of network training has achieved a better effect. Among all networks, LeNet-E network can achieve lowest loss value, indicating that in case of fitting, parameters learned by its network structure can better express image features, that is, deeper network structure can better fit data features. It can be found from the figure that network loss value of LeNet-D and LeNet-E networks is very unstable during training. It shows that although dropout strategy helps to improve recognition rate of network, it affects stability of network training.

We also plot the relationship between test accuracy and number of iterations, as shown in Figure 7. It can be found that LeNet-D and LeNet-E networks have higher recognition rates than other three networks, indicating that the depth of network and dropout strategy have a greater impact on recognition rate. From the curve of LeNet-A, it can be found that with the increase of number of generations, recognition accuracy has a downward trend, which may be overfitting. When the width of network is not enough to represent characteristics of input data, the recognition rate of test set will decrease with the increase of number of network selections.

5. Conclusion

Convolutional neural network is a very important feature learning model in machine learning. It has strong representation ability in image recognition and has been widely used in the field of pattern recognition. This paper conducts in-depth research on convolutional neural networks, introduces in detail network structure of convolutional neural networks and specific operations of convolution and pooling, and introduces back-propagation algorithm used in network optimization. Based on LetNet-5 network model, five different network structures are constructed for handwritten digit recognition. The effects of network depth, number of hidden layer convolution kernels, and feature dimension on recognition accuracy are tested. The experimental results show that deeper and wider network structure is more conducive to improvement of recognition rate, larger feature dimension, and better recognition rate. Finally, a face recognition method based on local binary pattern and convolutional neural network is proposed, which uses stitching of original input image and local binary pattern image as the input of convolutional neural network, so that convolutional network not only learns the global features of face but also local structural features of face. Through experiments, it is found that added LBP information helps to further improve face recognition rate.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

- [1] G. Lou and H. Shi, "Face image recognition based on convolutional neural network," *China Communications*, vol. 17, no. 2, pp. 117–124, 2020.
- [2] Y. Xiao and X. Xie, "Application of novel gabor-DCNN into RGB-D face recognition," *International Journal on Network Security*, vol. 22, no. 3, pp. 532–539, 2020.
- [3] W. Yang, Q. Liu, S. Wang et al., "Down image recognition based on deep convolutional neural network," *Information processing in agriculture*, vol. 5, no. 2, pp. 246–252, 2018.
- [4] T. Wang, C. Lu, G. Shen, and F. Hong, "Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network," *Peer Journal*, vol. 7, Article ID e7731, 2019.
- [5] A. S. Shamsaldin, P. Fattah, T. A. Rashid, and N. K. Al-Salihi, "A study of the convolutional neural networks applications," *UKH Journal of Science and Engineering*, vol. 3, no. 2, pp. 31–40, 2019.
- [6] X. Wang and W. Zhang, "Anti-occlusion face recognition algorithm based on a deep convolutional neural network," *Computers & Electrical Engineering*, vol. 96, Article ID 107461, 2021.
- [7] G. Wei, G. Li, J. Zhao, and A. He, "Development of a LeNet-5 gas identification CNN structure for electronic noses," *Sensors*, vol. 19, no. 1, p. 217, 2019.
- [8] T. Li, D. Jin, C. Du et al., "A better way to monitor haze through image based upon the adjusted LeNet-5 CNN model," *Signal, Image and Video Processing*, vol. 14, no. 3, pp. 455–463, 2020.
- [9] L. Wan, Y. Chen, H. Li, and C. Li, "Rolling-element bearing fault diagnosis using improved LeNet-5 network," *Sensors*, vol. 20, no. 6, p. 1693, 2020.
- [10] C. Zhang, X. Yue, R. Wang, N. Li, and Y. Ding, "Study on traffic sign recognition by optimized Lenet-5 algorithm," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 1, Article ID 2055003, 2020.
- [11] W. Cui, Q. Lu, A. M. Qureshi, W. Li, and K. Wu, "An adaptive LeNet-5 model for anomaly detection," *Information Security Journal: A Global Perspective*, vol. 30, no. 1, pp. 19–29, 2021.
- [12] Y. Hou and Z. B. Chen, "LeNet-5 improvement based on FPGA acceleration," *Journal of Engineering*, vol. 2020, no. 13, pp. 526–528, 2020.
- [13] Y. Zhu, G. Li, R. Wang, S. Tang, H. Su, and K. Cao, "Intelligent fault diagnosis of hydraulic piston pump combining improved LeNet-5 and PSO hyperparameter optimization," *Applied Acoustics*, vol. 183, Article ID 108336, 2021.
- [14] P. Wang, W. Q. Hong, D. L. Tsai, and M. S. Jhou, "Ransomware classification using LeNet-5 convolutional neural networks," *Communications of the CCISA*, vol. 26, no. 2, pp. 21–48, 2020.
- [15] Y. Sun, S. Liu, T. Zhao et al., "A new hydrogen sensor fault diagnosis method based on transfer learning with LeNet-5," *Frontiers in Neurobotics*, vol. 15, Article ID 664135, 2021.
- [16] Q. Gan, F. Su, K. Lian, D. Xu, and Y. Dong, "Rice classification algorithm research based on modified LeNet-5 model," *Journal of Anhui Agricultural University*, vol. 46, no. 3, pp. 549–553, 2019.
- [17] M. R. Alwanda, R. P. K. Ramadhan, and D. Alamsyah, "Implementasi metode convolutional neural network menggunakan arsitektur LeNet-5 untuk pengenalan doodle," *Journal Algoritme*, vol. 1, no. 1, pp. 45–56, 2020.
- [18] E. Xi, "Image classification and recognition based on deep learning and random forest algorithm," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 2013181, 9 pages, 2022.