

Research Article

Image Recognition Algorithm Based on Artificial Intelligence and Machine Learning

Lipeng Si , Baolong Liu, and Yanfang Fu

School of Computer Science and Engineering, Xi'an Technological University, Xi'an 710021, Shaanxi, China

Correspondence should be addressed to Lipeng Si; silipeng@xatu.edu.cn

Received 23 March 2022; Revised 20 April 2022; Accepted 25 May 2022; Published 5 July 2022

Academic Editor: Liming Chen

Copyright © 2022 Lipeng Si et al. 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.

At present, most of the existing image recognition methods are not only cumbersome in process but also require manual design of functions, resulting in poor recognition results and time-consuming training. This study explores image recognition algorithms based on artificial intelligence and machine learning, which can simulate the hierarchical structure of the human brain and nervous system, realize automatic extraction of complex features, and have powerful data representation capabilities. In this study, the structure of artificial neuron is introduced first, and the training process of the neural network is introduced in detail. At the same time, this study improves the traditional training algorithm and proposes two new machine learning models for different application scenarios, which effectively improves the performance of the optimal model. In relevant tests, the effect is significant. On the basis of these improved algorithms, an online recognition system is designed and implemented. The recognition accuracy rate of the system for the hidden layer reaches more than 90%, which verifies the effectiveness of the technology. At the same time, after repeated experiments, the results show that the multiview algorithm effectively solves the problem that the recognition results of the traditional multiview algorithm are affected by the size of the target contour.

1. Introduction

Image recognition is the input image obtained from vision. Through a series of calculation, analysis, and learning processes, the object recognition in the scene, the description of the relationship between the objects in the scene, and the recognition of the scene are obtained [1]. In short, image recognition. It is to use computers to realize human's visual understanding of images, for example, fingerprint recognition technology, face detection, and recognition technology in various embedded systems such as smart phones and tablet computers [2]; surface defect detection technology and product shape recognition technology in industrial automated inspection systems [3]; robot vision in artificial intelligence, autonomous driving of unmanned cars, and automatic human-machine tracking [4]; and military target tracking, missile guidance, and air remote sensing [5]. Nowadays, the demand for intelligent automation equipment and instruments is becoming more and more urgent [6].

Generally speaking, the existing image recognition methods can be divided into methods based on expert knowledge and methods based on machine automatic learning [7, 8]. In contrast, machine-based automatic learning methods are based on the "machine" self-learning ability [9]. The idea is to design machines with self-learning capabilities by relying as little as possible on "human" experience [10]. For example, given only the structure of the machine and some simple operation measures, the machine can automatically learn the specific rules hidden in the image data [11, 12]. Using "machines" instead of "humans" to learn can reduce the workload of manual design [13, 14]. Therefore, it is more common to rely on the advantages of "machines" in computing speed to solve these problems.

This study proposes a novel research direction of the image recognition algorithm, which can effectively improve the quality and effect of image recognition and can also provide new ideas for the research of artificial intelligence and machine learning technology.

2. Proposed Method

2.1. The Structure of Artificial Neurons. The artificial neural network can be regarded as a directed graph connected by directed weighted arcs with artificial neurons as nodes. The structure of the artificial neuron is shown in Figure 1.

If a local feature appears in one location, it may also appear in any other location. Mathematically, the operation of the feature map based on the convolution kernel corresponds to discrete convolution.

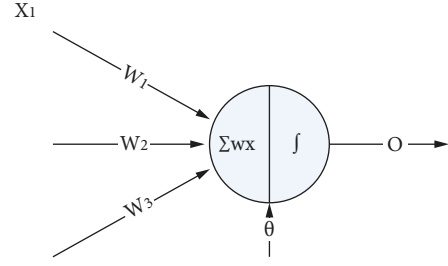


FIGURE 1: Structural diagram of an artificial neuron.

2.2. The Training Process of the Neural Network. By using the min-batch method, back propagation is performed to update the parameters in the neuron.

$$v_{n+1} = 0.85v_n - 0.001\alpha A - \left(\frac{\partial L}{\partial W} | w_n\right) D_n, \quad (1)$$

$$w_{n+1} = w_n + v_{n+1}.$$

The whole training process and training results are shown in Figure 2.

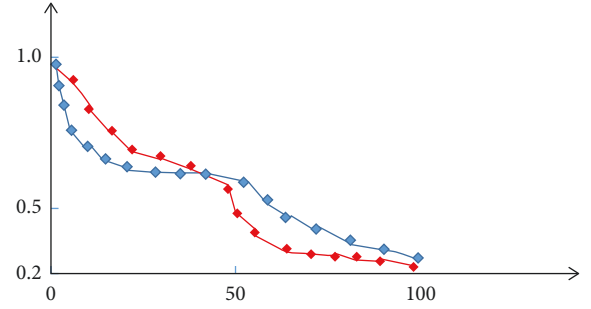


FIGURE 2: Training process of the deep neural network.

2.3. Neural Network Training Result Analysis. Based on the above learning model, the convolution kernel of the first layer of the neural network is obtained, as shown in Figure 3.

The segmentation training of the convolutional layer is not performed. In contrast, the overall performance of the model in this study is better. This article integrates multiple models and merges their output with postresort, as given in Table 1.

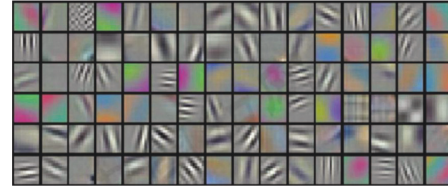


FIGURE 3: The convolution kernel graph obtained by training.

2.4. Improvement of the Model

2.4.1. Optimizes with Optimizer. It is set at the beginning of training to initialize the cumulants of gradients and square cumulants.

$$v_{dw} = 0, v_{db} = 0; \quad s_{dw} = 0, s_{db} = 0. \quad (2)$$

Assuming that in the training phase of the t -round training, two basic momentum parameter updates can be calculated at first.

$$\begin{aligned} v_{dw} &= \beta_1 v_{dw} + (1 - \beta_1) dw, \\ v_{db} &= \beta_1 v_{db} + (1 - \beta_1) db, \\ s_{dw} &= \beta_2 s_{dw} + (1 - \beta_2) dw^2, \\ s_{db} &= \beta_2 s_{db} + (1 - \beta_2) db^2. \end{aligned} \quad (3)$$

Since the average value of the moving index at the beginning of the iteration will cause a large difference between the initial value and the initial value, the deviation of the above value needs to be corrected.

TABLE 1: Comparison of the methods in this study with other methods.

Model	p -1 error rate (%)	p -5 error rate (%)
Sparse coding	46.9	28.1
SIFT + FVs	45.2	25.5
1KCNN	40.5	18.5
1MyCNN	40.1	18.0
5MyCNN	37.5	17.1

$$\begin{aligned} v_{dw}^c &= \frac{v_{dw}}{1 - \beta_1^t}, \\ v_{db}^c &= \frac{v_{db}}{1 - \beta_1^t}, \\ s_{dw}^c &= \frac{v_{dw}}{1 - \beta_1^t}, \\ s_{db}^c &= \frac{v_{db}}{1 - \beta_1^t}. \end{aligned} \quad (4)$$

Through the above formula, the correction value of the cumulative parameter evaluation can be obtained in the t -round iteration process, and the weight and deviation can be updated.

$$w = w - a \frac{v_{dw}^c}{\sqrt{s_{dw}^c + \epsilon}} \quad (5)$$

In summary, the learning rate decay method can make the network’s early training converge quickly, and in the later training, it can converge to the optimal value of the function to the greatest extent.

2.4.2. *Improvement of the Multiview Method.* By comparing the recognition error rates of different methods, Table 2 is obtained.

Observing the data in the table, we find that the error rate and accuracy of the improved algorithm are lower than those of the previous and traditional test methods. Of course, the improved algorithm is not perfect; it also has some problems, the number of improved images is too much, and the problem of frequent testing is unavoidable.

2.5. *Level Classification Algorithm.* The probability of misclassification of m and K are calculated by the following formulas. X is used to indicate the number of objects to be identified, and ϵm and K are used to represent the frequency of misclassification of m to K .

$$\rho_{m,k} = \frac{\epsilon_{m,k}}{\sum_{i=1}^N \epsilon_{m,i}} \quad (6)$$

The principle of undirected graph $G = (V, E)$ is applied. A class is a node, the set of nodes is V , and the set of edges is E . The weight of the edge between two points is used as the probability value of the two classes. Then, use the global minimum cut algorithm to repeatedly find the global minimum cut on this graph, and then, we can get the set of errors that are easiest to classify, as given in Table 3.

3. Experiments

The platform used in the experiment is Python++Theano. SVM is used as the baseline test. Each pixel is used as the input of the SVM algorithm. We can test the effectiveness of the proposed method. Under the Java platform, use LIBSVM for testing.

3.1. *The Number of Samples, the Number of Nodes, and the Number of Layers.* This part of the experiment is based on the MATLAB R2009a computing platform, and the SVM classifier uses the LIBSVM toolbox. The number of training samples used in the experiment is 1000, and the number of test samples is 200. As given in Table 4. The image recognition network model is shown in Figure 4.

Two kinds of training samples and test samples are separately removed, and different hidden nodes and layers are selected to compare the effects of training samples, nodes, and layers. Among them, the training sample of Figure 5 is 1000, and the testing sample is 200.

Figure 6 shows the comparison of the prediction accuracy of this method and the CNN method for different hidden layer nodes when there are two hidden layers. The

TABLE 2: Comparison of recognition error rates for different test methods.

Model	p -1 error rate (%)	p -5 error rate (%)
CNN1	40.01	17.56
CNN5	37.52	17.40
Multiview 1	40.12	17.55
New multiview 1	38.54	18.12
Multiview 5	37.10	16.50
New multiview 5	36.50	16.00

TABLE 3: Basic classification and number of categories.

Basic classes on data sets	Number of corresponding subclasses
Dog	110
Cat	36
Car	49
Leaf	88
Aircraft	61
Handbag	25

TABLE 4: Number of samples.

Category	0	1	2	3	4	5	6	7	8
Training sample size	67	172	87	76	56	98	95	200	123
Number of samples tested	12	23	17	20	14	21	19	10	32
Training sample	479	663	588	593	635	534	501	550	462
Test sample	85	176	166	111	110	87	87	99	89

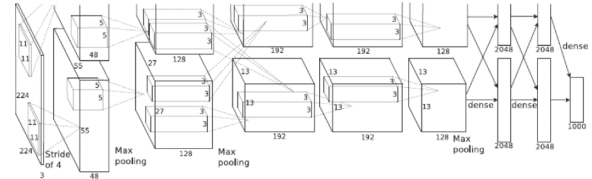


FIGURE 4: Network model of image recognition.

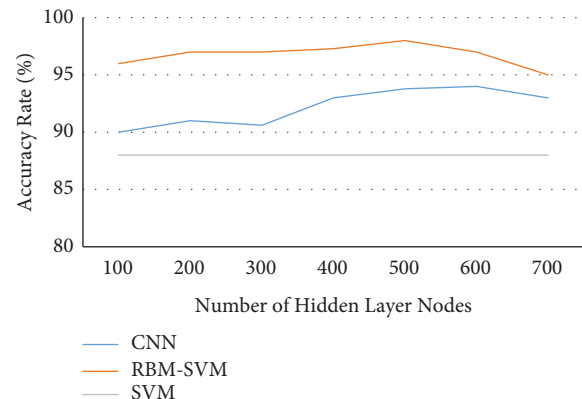


FIGURE 5: Prediction accuracy rate (%) of different methods in single hidden layer 1000/200.

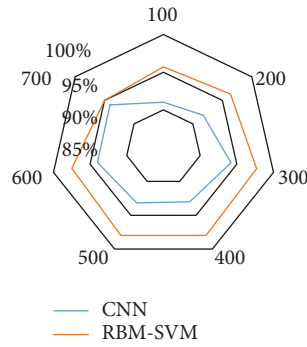


FIGURE 6: Prediction accuracy of two layers of the hidden layer (%).

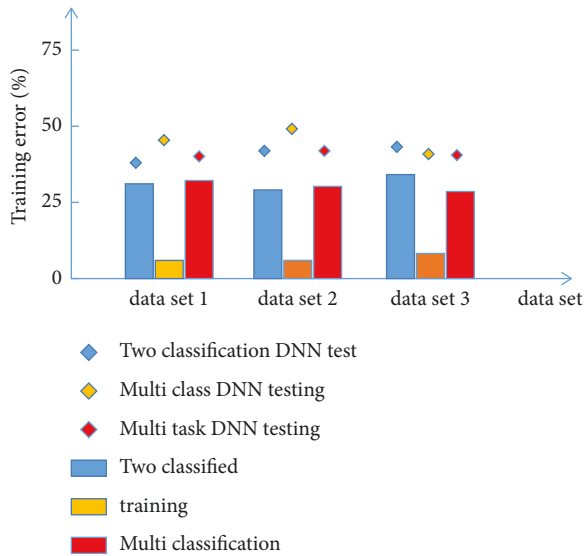


FIGURE 7: Two-class, multiclass, and multitask DNN experimental results.

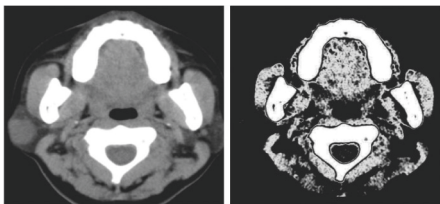


FIGURE 8: Noise reduction processing chart.

training sample is 5000, the test sample is 1000, and the number of the first layer hidden layer is 450.

4. Discussion

4.1. Experimental Comparison with Nondeep Learning Algorithms. As shown in Figure 7, the prediction probability of the class in the multiclass DNN is shown. The multiclass DNN performs the worst when the data are unevenly distributed.

From the processing results of Figures 8 and 9 (picture source network library), the performance of the multitask

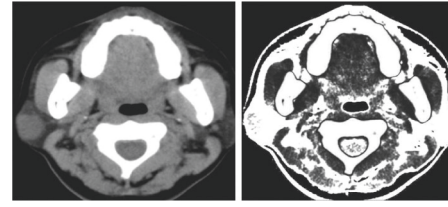


FIGURE 9: Boundary processing diagram.

DNN based on ring training is stable, and the final results on different datasets can basically remain unchanged.

5. Conclusions

In image recognition methods, artificial intelligence machine learning algorithms have the advantages of parallel processing, self-learning capabilities, and flexible nonlinear expression capabilities compared with traditional algorithms. This study proposes a multiview algorithm based on artificial intelligence machine learning. After repeated experiments, the results show that the multiview algorithm effectively solves the problem that the recognition result of the traditional multiview algorithm is affected by the size of the target contour. The structure of the traditional model is optimized, and some problems in the field of traditional image recognition are solved. Through the improvement of related algorithms, the error rate of image recognition is reduced, which has special reference significance for the design and development of related systems in the future.

There are still many deficiencies in the research of this study. The depth and breadth of the research in this study are not enough, and some interference factors involved in the practice of image recognition are not considered. The evaluation of the algorithm is also restricted by many factors. The level of research is also limited, and the research on image recognition algorithms is still in the preliminary stage. In the future work, based on the existing technology and level, we will improve the performance of recognition image recognition accuracy from more angles and continuously optimize the research method.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Science and Technology Program of Shaanxi Province. (2017XT-02) and International S&T Cooperation Program (2021KW-07).

References

- [1] S. Kiakojoori and K. Khorasani, "Dynamic neural networks for jet EngineDegradation prediction and prognosis," in

- Proceedings of the International Joint Conference on Neural Networks*, pp. 2531–2538, Beijing, China, July 2015.
- [2] S. Shatnyi, “Neural network-based information technology for biomedical signal processing,” in *Proceedings of the International Conference on Microwaves, Radar, and Wireless Communication*, pp. 1–3, Gdansk, Poland, June 2016.
- [3] J. Li, K. Ouazzane, H. B. Kazemian, and M. S. Afzal, “Neural network approaches for noisy language modeling,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 11, pp. 1773–1784, 2013.
- [4] D. Reid, A. J. Hussain, and H. Tawfik, “Spiking neural networks for financial DataPrediction,” in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1–10, Dallas, TX, U.S.A, August 2017.
- [5] B. H. Gary, L. Honglak, and L. M. Erik, “Learning hierarchical representations for face verification with convolutional deep belief networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, June 2012.
- [6] M. H. Rabie, I. A. E. Khodary, and A. T. Assem, “A particle swarm optimization algorithm for the continuous absolute p-center location problem with Euclidean distance,” *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 12, 2013.
- [7] B. Yelbay, Ş. İ. Birbil, and K. Bülbül, “The set covering problem revisited: an empirical study of the value of dual information,” *Journal of Industrial and Management Optimization*, vol. 11, no. 2, pp. 575–594, 2015.
- [8] B. Crawford, R. Soto, R. Cuesta, and F. Paredes, “Application of the artificial bee colony algorithm for solving the set covering problem,” *The Scientific World Journal*, vol. 2014, no. 1, pp. 1–8, Article ID 189164, 2014.
- [9] A. E. Zade, A. Sadegheih, and M. M. Lotfi, “A modified NSGA-II solution for a new multi-objective hub maximal covering problem under uncertain shipments,” *Journal of Industrial Engineering International*, vol. 10, no. 4, pp. 185–197, 2014.
- [10] G. E. Hinton, “Learning multiple layers of representation,” *Trends in Cognitive Sciences*, vol. 11, no. 10, 2007.
- [11] Z. Wang, H. Wu, G. Zhao et al., “Corrigendum to “One-dimensional finite difference modeling on temperature history and freezing time of individual food,”” *Journal of Food Engineering*, vol. 79, no. 2, pp. 502–510, 2007.
- [12] S. A. G. Kustrin and R. Beresford, “Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research,” *Journal of pharmaceutical and biomedical analysis*, vol. 22, no. 5, pp. 717–727, 2015.
- [13] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *Bulletin of Mathematical Biology*, vol. 5, no. 1-2, pp. 115–133, 1943.
- [14] C. Helma, T. Cramer, S. Kramer, and L. D. Raedt, “Data mining and machine learning techniques for the identification of mutagenicity inducing substructures and structure activity relationships of noncongeneric compounds,” *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 4, pp. 1402–1411, 2004.