Scholar Publishing Group
International Journal of Neural Network
https://doi.org/10.38007/NN.2021.020207
ISSN 2790-2692 Vol. 2, Issue 2: 47-54

Convolutional Neural Network in Face Recognition in Online Classroom

Amar Singh*

Dhurakij Pundit University, Thailand

**corresponding author*

Keyword: Neural Network, Face Recognition, Online Class, Internet Plus

Abstract: With the arrival of the era of artificial intelligence and data, the "Internet +" education model has achieved a blowout development. At the beginning of 2020, the outbreak of COVID-19 further promoted the online teaching model. However, due to the separation of time and space and the way of single line teaching, online classes lack human communication; when students' learning mood and behavior turn away from the learning classroom, teachers can not correct and give reminders in time, resulting in poor learning effect of students. Therefore, how to use artificial intelligence methods to obtain students' learning behavior from online classes, analyze and judge their classroom status, and form an effective classroom status evaluation mechanism is an important foundation for achieving scientific online education and improving learning effects. In this paper, the following research has been carried out on the online classroom face recognition system based on convolutional neural network. The use of ghost module can effectively reduce the amount of parameters and computation, and still ensure the good face recognition effect of the network.

1. Introduction

With the continuous progress of modern science and technology, the "Internet +" education model has been born. From 2013 to now, it has shown a blowout development. In 2020, COVID-19 ravaged the world, making it impossible for most students to study in the classroom on campus. This also made e-learning a new favorite of current teaching methods, and online classroom is the most important way of learning. Since the large-scale development of MOOC in 2012, online classroom has gradually become a "storm" that changes the way of education. A large number of online education platforms have emerged, such as "Netease cloud classroom", "51CTO", "school online", and so on[1].

Compared with other biological features, face also has more additional information that can be used by us. We can not only directly use faces to complete relevant tasks as above, but also conduct deeper mining and analysis of face images to get more meaningful information for us and use this information for other alications[2]. In China, driver monitoring systems are installed in all major passenger cars, which can greatly reduce the incidence of dangerous driving accidents every year. In hospitals, we can install a face recognition system to judge whether patients are hay or angry by analyzing their facial features. Nowadays, this facial expression recognition has also developed rapidly due to its huge demand. Through the analysis of face images, we can judge a person's gender, age, race and whether he wears masks and glasses[3]. For example, during the epidemic, we can use face recognition technology to judge whether people wear masks to prevent the rapid spread of the virus. In addition to the examples listed above, face and face recognition technology are also alied in many other aspects[4]. Therefore, it is of great significance to study face recognition technology. At present, face recognition technology has made a lot of achievements with great practical significance, but the effect of face recognition under unconstrained conditions is still far from meeting people's requirements. This paper hopes to improve the effect of face recognition under unconstrained conditions by improving the existing neural network model[5]. At the same time, it studies face related fields such as face mask detection and living body detection to meet the requirements in special scenes. Based on this, it is of great value to study face recognition technology under unconstrained conditions from both theoretical and practical perspectives. In terms of its development stage, face detection technology can be divided into traditional face detection stage and face detection era based on deep learning. In terms of methods, it can be divided into four mainstream face detection methods based on skin color model, template matching, feature statistics and deep learning. Many foreign scholars have proposed AdaBoost algorithm based on statistical theory. The core idea is that the weak classifier can form a final classifier with stronger classification effect in some way. The detection speed of this algorithm is fast, but it is easy to be interfered by external environment factors, resulting in poor detection accuracy and low robustness. Since then, deep learning has been widely used in the field of face detection. Compared with the traditional face detection methods, the deep learning method does not need to manually extract face features, and automatically completes the training of the network model with a large amount of data. While effectively improving the accuracy of face detection, it can greatly reduce the amount of computation in the detection process[6].

In this paper, mtcnn is selected as the face detection algorithm. Research on face related fields has greatly facilitated our life. For example, in some scenes with high security requirements, we can add face detection technology to make the system more secure and reliable. During the epidemic, deep learning can also be used to detect the wearing of face masks. This paper mainly focuses on face recognition based on convolutional neural networks, exploring how to improve the recognition rate of faces under unconstrained conditions including face pose changes, and at the same time, lightweight design of neural networks, so that they can be deployed in small embedded devices, and finally complete the accurate recognition of faces in videos.

2. Overview of Related Concepts

2.1. Forward Propagation Algorithm

Face images are input to the network in the form of feature vectors, and the neural network continuously deduces and outputs the final results to solve the classification or regression problems we face. The whole process from input to output depends on the forward propagation algorithm of the network. In the neural network, neurons are the smallest unit. Each neuron has at least one input and only one output. The input of the whole network or the weighted sum of neurons can be taken as its input. It is assumed that layer L-1 and layer l have m and N nerves, respectively

The output of layer l can be summarized as[7, 8]:

$$a^L = f(z^L) = f(W^L d^{L-1} + b^L) \tag{1}$$

In the above formula, a^L and a^{L-1} are n-dimensional and m-dimensional vectors respectively ,Indicates the output of layer L and layer L-1. W^L Layer n×M-dimensional weight matrix, b^L indicates that the layer N is an offset vector. The above is the forward propagation process of the neural network. In the forward propagation process, we hope that the output of the network can aroach the real label of the sample as much as possible[9, 10].

2.2. Back Propagation Algorithm

Back propagation is a method to calculate gradients, rather than the whole learning algorithm for complex neural networks. The main steps of implementing BP algorithm are as follows[11]:

The forward propagation algorithm in the previous section is used to deduce the training sample data of the input neural network layer by layer, and the results are obtained at the final output layer[12].

Calculate the error between the forward propagation output result and the real label of the training sample, and then optimize the network by repeatedly adjusting the parameter weight to make the final output as close as possible to the actual result.

2.3. Convolutional Neural Network

With the continuous improvement of deep learning theory and hardware equipment, and relying on its strong learning and feature extraction capabilities, it has achieved great success in various fields of computer vision[13].

Convolutional neural networks have different learning tasks in different network layers, which can be roughly divided into shallow layer, intermediate layer and deep layer in terms of their functions[14]. Taking face recognition as an example, CNN extracts the edge features of face images by using the shallow layer, then abstracts and combines the obtained edge features into local features by using the middle layer, and finally abstracts and combines each local feature into the global feature or the overall contour of the recognition target in the deep layer[15].

2.4. Activation Function

Most of the problems to be solved in the face field are complex and nonlinear. However, the characteristics of neural network forward propagation determine that it is a linear model. In order to effectively solve this contradiction, researchers proposed adding nonlinear elements (activation functions) to the neural network[16]. The specific method is: in CNN, by adding non-linear elements to its hidden layer, the whole linear network model is changed into a non-linear one, and the non-linear network has stronger expression ability. Common activation functions are as follows[17]:

$$Sigmoid(x) = \frac{1}{1+e^{-x}} \tag{2}$$

$$Tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3}$$

The value range of sigmoid activation function is (0,1) which is smooth, continuous and monotonous, and easy to optimize. However, the problems of large computation and slow convergence also affect the training of the network. As another logical activation function, the tanh activation function has an S-line curve with a value range of (- 1,1) which strengthens the negative input maing of the function, and only the zero input is maed to the nearest non-zero input, eliminating the possibility of the gradient disaearing. The output mean value is 0 (close to the distribution mean value of many samples) which determines its fast convergence speed and fewer iterations compared with sigmoid[18].

3. Online Classroom Face Recognition Based on Convolutional Neural Network

3.1. Channel Attention Mechanism

This paper adds the attention mechanism senet algorithm to the improved residual network resnet50. The senet attention mechanism module mainly acts on the neural network model channel. It is easy to add this module to the existing network model, and the amount of parameters and calculations at the same time is very small, which can improve the effect.

3.2. Depth Residual Identity Maing Module

The aearance of deep learning makes face recognition a great success. However, compared with the front face, many existing face recognition models still perform relatively poorly when dealing with the side face. A key reason is that the number of front faces and the number of side faces are highly unbalanced, so there are more face training samples. In addition, fundamentally speaking, it is very difficult for the neural network to learn the depth representation that keeps the geometry unchanged with large attitude changes through training.

4. Numerical Analysis Results

One of the problems of online classroom is that teachers can't grasp students' classroom status in time, resulting in the decline of teaching quality. Therefore, this paper analyzes students' facial features and judges students' attendance status through facial recognition.

4.1. Analysis of Identification Experiment Results

Table 1. Network hyperparameters

Parameter	Numerical value
Batch-size	16
Initial learning rate	0.1
Momentum	0.9
Weight_decay	0.0001

As shown in Table 1, the experimental batch size is set to 16, and the initial learning rate is 0.1. At the same time, the learning rate of every 30 epochs is attenuated to 0.1 times of the last learning rate. The momentum parameter is set to 0.9. To prevent overfitting of the model, the weight decay is set to 0.0001.

Table 2. Face recognition accuracy of different network models on each dataset

Model	LFW	CFP	IJB-A
ResNet50	98.14%	92.16%	95.62%
SENet+ResNet50	98.37%	92.84%	95.91%
DREAM+ResNet50	98.65%	93.76%	96.07%
SENet+DREAM+ResNet50	99.46%	94.15%	96.23%

As shown in Table 2, the accuracy of LFW, CFP and ijb-a data sets is the highest, reaching 99.46%, 94.15% and 96.23% respectively. In order to further prove the effectiveness of the improved residual neural network, experiments were also conducted based on the four network model structures, and the accuracy of face recognition was compared on the YTF and sllfw datasets. The experimental results are shown in Figure 1 below.

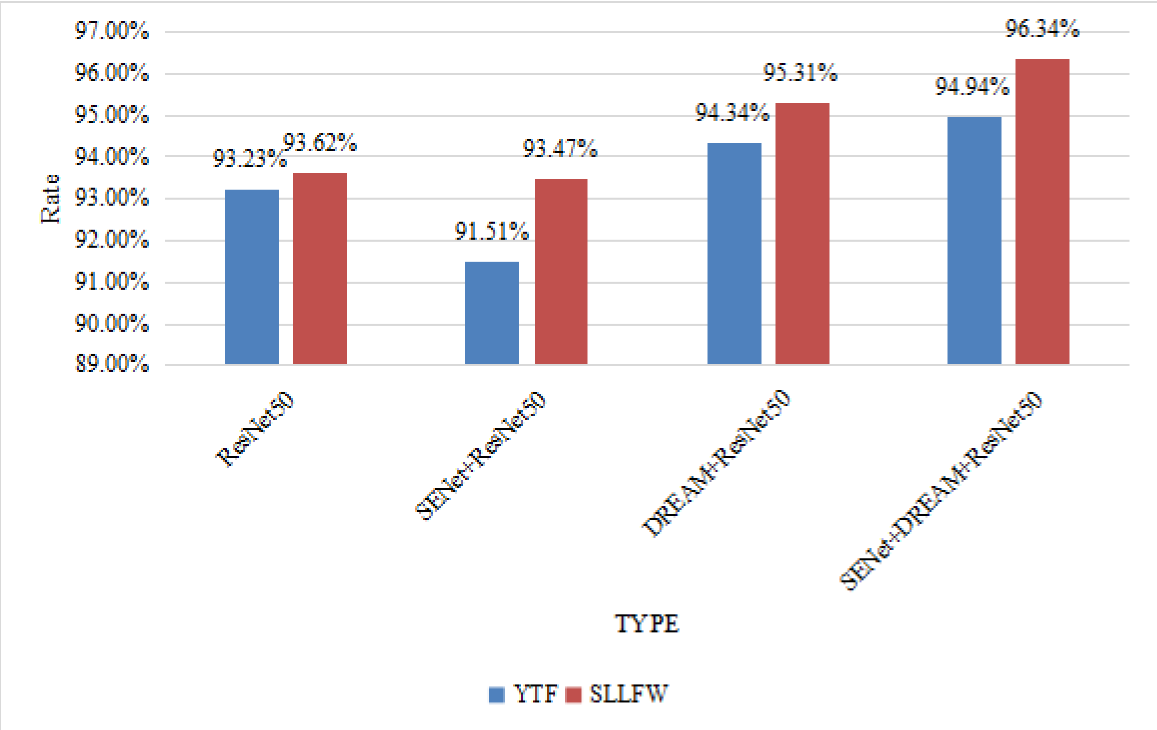

Figure 1. Face recognition accuracy of different network models on YTF and sllfw datasets

It can be seen from the above figure that the network with the attention mechanism and the depth residual identity maing module added at the same time has the highest accuracy, reaching 94.94% and 96.34% on the YTF and sllfw datasets respectively. Compared with the original network resnet50, the accuracy is improved.

4.2. Experimental Analysis of Residual Network Lightweight

Ghost module has two super parameters, D and S. in the experiment, the control variable method is used to determine the impact of this parameter on the network performance. For example, first set s to a fixed value, and then let D adjust in {1, 3, 5, 7}. The experimental results are shown in Table 3 below.

Table 3. Experimental performance effects of different D on self-made face dataset

d	Weights	FLOPs	Acc. (%)
ResNet-50	25.6	213	98.14
1	13	107	98.12
3	12.9	109	98.15
5	12.9	112	97.89
7	12.9	119	97.63

It can be seen from the table that when the value of D is 1, the size of the convolution kernel is too small, which will lead to the loss of spatial information for the feature map. When the value of D is greater than 3, the size of the convolution kernel is too large, which will increase the calculation amount of the network model. Therefore, in subsequent experiments, D is set to 3.

After improving the network model, the senet + dream + RESNET network model has the best effect. Therefore, it is compared with the lightweight neural network model ghost resnet50. Figure 2 shows the face accuracy comparison data of the lightweight improved network model and the optimal network model in Chapter 4 on five different data sets, as shown below.

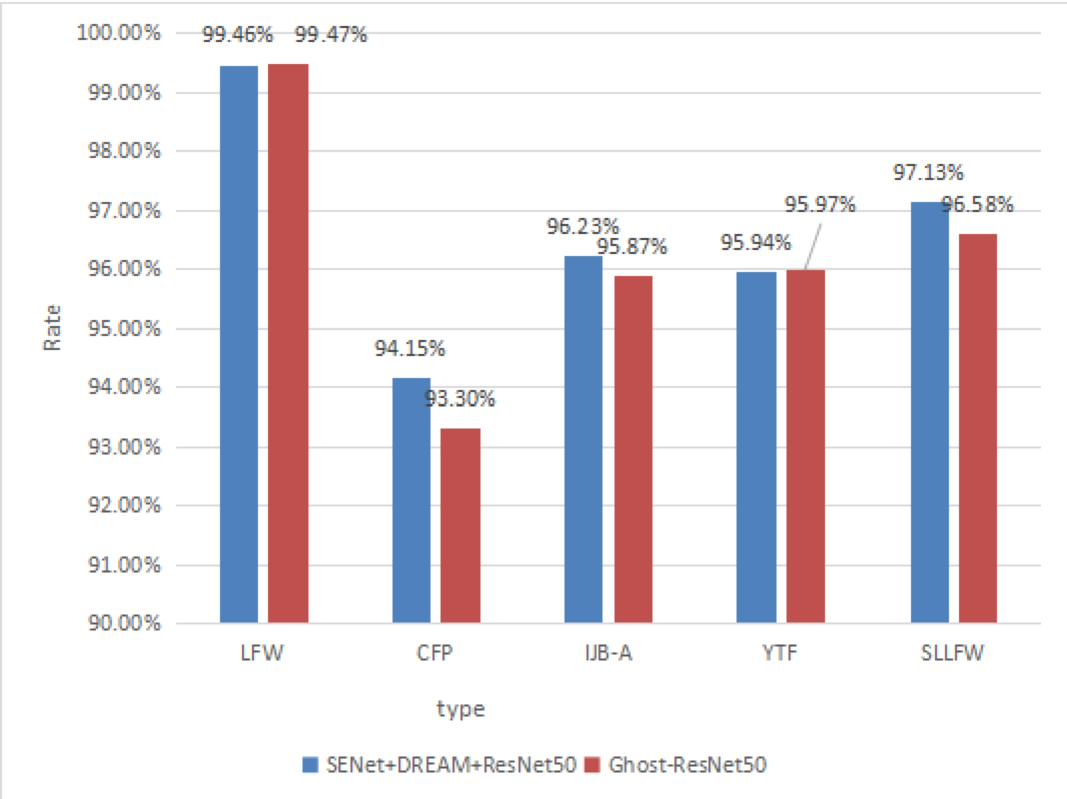

Figure 2. Comparison of accuracy of two different network models on five face data sets

5. Conclusion

As a "storm" that changes the way of education, online classes have more and more obvious advantages. Online classes are not limited by any time and place. Students can learn independently, choose more and more effective learning resources, better improve learning efficiency and stimulate learning interest. In view of the above problems, this paper improves the robustness of neural network to face recognition by improving the existing deep learning algorithm. The arrival of the era of deep learning has greatly promoted the research process and alication of face recognition, but not all problems have been well solved. In addition, there are also many specific aspects of the face that are worth exploring, such as the detection and even recognition of face masks in the epidemic scene. The research on face recognition is still a long and arduous process. We need to calm down and study hard so that face recognition technology can better serve people's lives.

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] Acharya U R, Fujita H, Lih O S, et al. Automated Detection of Coronary Artery Disease Using Different Durations of ECG Segments with Convolutional Neural Network. Knowledge-Based Systems. (2017) 132(sep.15):62-71. https://doi.org/10.1016/j.knosys.2017.06.003

[2] Salehi S, Erdogmus D, Gholipour A. Auto-Context Convolutional Neural Network (Auto-Net) for Brain Extraction in Magnetic Resonance Imaging. IEEE Transactions on Medical Imaging. (2017) (11):1-1.

[3] Evo I, Avramovi A. Convolutional Neural Network Based Automatic Object Detection on Aerial Images. IEEE Geoscience & Remote Sensing Letters. (2017) 13(5):740-744. https://doi.org/10.1109/LGRS.2016.2542358

[4] Cheng G, Wang Y, Xu S, et al. Automatic Road Detection and Centerline Extraction via Cascaded End-to-End Convolutional Neural Network. IEEE Transactions on Geoscience and Remote Sensing. (2017) 55(6):3322-3337. https://doi.org/10.1109/TGRS.2017.2669341

[5] Alotaibi M, Mahmood A. Improved Gait Recognition Based on Specialized Deep Convolutional Neural Network. Computer Vision and Image Understanding. (2017) 164(NOV.):103-110. https://doi.org/10.1016/j.cviu.2017.10.004

[6] Chen F C, Jahanshahi R. NB-CNN: Deep Learning-based Crack Detection Using Convolutional Neural Network and Na ïve Bayes Data Fusion. IEEE Transactions on Industrial Electronics. (2018) 65(99):4392-4400. https://doi.org/10.1109/TIE.2017.2764844

[7] Mao L. Remote Classroom Action Recognition Based on Improved Neural Network and Face Recognition. Journal of Intelligent and Fuzzy Systems. (2021) (1):1-11. https://doi.org/10.3233/JIFS-189803

[8] Mansoora S, Sadineni G, Kauser S H. Attendance Management System Using Face Recognition Method. Journal of Physics: Conference Series. (2021) 2089(1):012078-. https://doi.org/10.1088/1742-6596/2089/1/012078

[9] Thampi S M, Krishnan S, Rodriguez J C, et al. [Advances in Intelligent Systems and Computing] Advances in Signal Processing and Intelligent Recognition Systems Volume 678 || A Combined

Feature Extraction Method for Automated Face Recognition in Classroom Environment. (2018) 10.1007/978-3-319-67934-1:417-426. https://doi.org/10.1007/978-3-319-67934-1_38

[10] Granda L, LRB Guamán, Torres-Carrion P V, et al. Face Recognition Systems in Math Classroom through Computer Vision Traditional Techniques. Communications in Computer and Information Science. (2020) 1194:652-664. https://doi.org/10.1007/978-3-030-42520-3_51

[11] Mustafa H T, Yang J, Zareapoor M. Multi-Scale Convolutional Neural Network for Multi-Focus Image Fusion. Image and Vision Computing. (2019) 85(MAY):26-35. https://doi.org/10.1016/j.imavis.2019.03.001

[12] Griffiths D, Boehm J. Improving Public Data for Building Segmentation from Convolutional Neural Networks (Cnns) For Fused Airborne Lidar and Image Data Using Active Contours. ISPRS Journal of Photogrammetry and Remote Sensing. (2019) 154(AUG.):70-83. https://doi.org/10.1016/j.isprsjprs.2019.05.013

[13] Rashid M, Khan M A, Sharif M, et al. Object Detection and Classification: A Joint Selection and Fusion Strategy of Deep Convolutional Neural Network and SIFT Point Features. Multimedia Tools & Alications. (2019) 78(12):15 https://doi.org/10.1007/s11042-018-7031-0751-15777.

[14] Rai A K, Mandal N, Singh A, et al. Landsat 8 OLI Satellite Image Classification using Convolutional Neural Network. Procedia Computer Science, 2020, 167:987-993. https://doi.org/10.1016/j.procs.2020.03.398

[15] Akhtar Z, Rattani A. A Face in any Form: New Challenges and Oortunities for Face Recognition Technology. Computer. (2017) 50(4):80-90. https://doi.org/10.1109/MC.2017.119

[16] Sakulchit T, Kuzeljevic B, Goldman R D. Evaluation of Digital Face Recognition Technology for Pain Assessment in Young Children. The Clinical journal of pain. (2019) 35(1):18-22. https://doi.org/10.1097/AJP.0000000000000659

[17] Rui C, Lima S, Á Rocha. Alication of Face Recognition Technology Based on CA Algorithm in Intelligent Residential Property Management. Journal of Intelligent & Fuzzy Systems. (2018) 35(3):2909-2919. https://doi.org/10.3233/JIFS-169646

[18] Lonarevi Z, Saraevi M, Hadi M. Analysis of Some Techniques of Biometric Authentication Systems with Emphasis on 2D And 3D Face Recognition Technology. Ekonomski Izazovi. (2018) 7(13):119-139. https://doi.org/10.5937/EkoIzazov1813119L