Research on CT Image Classification Algorithm of COVID-19 Based on Improved ResNet

Xipei Chen, Yicong Zhao, Yanqiu Wang, Bin Yang, Xiaofei Yan

College of Information Science and Engineering, Zaozhuang University, Zaozhuang 277160, China

Abstract

The classification of COVID-19 and other viral pneumonias will help doctors to diagnose new coronary patients more accurately and quickly. Aiming at the classification problem of CT in patients with COVID-19, this paper proposes a CT image classification method based on an improved ResNet50 network based on the traditional convolutional neural network classification model. This paper uses the multiscale feature fusion strategy, combined with the improved attention mechanism to obtain the correlation coefficient between the internal feature points of the feature map, and finally achieves the effect of enhancing the representation ability of the feature map. Through the analysis and comparison of the technical principle, classification accuracy, and other parameters, it shows that the improved algorithm has better adaptive ability and classification ability. Through experiments, the improved ResNet50 classification model has a certain improvement in accuracy, time complexity, and spatial complexity compared with the traditional classification model, and the accuracy rate can reach 90.1 %.

Keywords

ResNet50 model; COVID-19; CT image classification

1 Introduction

At present, nucleic acid detection is the most common method to diagnose COVID-19 [1]. Reverse transcription polymerase reaction (RT-PCR) [2] has become the most mainstream COVID-19 detection technology. RT-PCR can detect RNA viruses in samples obtained from pharyngeal swabs, nasal swabs, sputum, bronchial lavage fluid, alveolar lavage fluid, etc. However, various studies have shown that the accuracy of RT-PCR detection is relatively low, and it usually requires multiple tests to be more accurate. Due to the low sample quality and pharyngeal virus load, nucleic acid detection by pharyngeal swabs is prone to false negatives, with a high retest rate and a long time to wait for nucleic acid detection results [3].

CT imaging technology plays an important role in the detection of COVID-19. Chest CT image is an effective tool to help doctors quickly diagnose COVID-19. However, because the lung characteristics of patients in the early stage of infection are not obvious in CT images, inexperienced doctors will not be able to accurately identify the CT image characteristics of COVID-19, which may lead to misdiagnosis. Using deep learning to analyze lung images (CT images) can reveal many insignificant features in the images and then give clear detection results. Therefore, integrating deep learning into medical images, image processing, target analysis, and other work on CT images, accurately extracting key focus areas and texture features, and screening for the performance characteristics of COVID-19, such as ground glass shadow, paving stone sign, lung consolidation and so on[4].

Aiming at the classification problem of CT images of COVID-19, this paper takes resnet50 based on the improved attention mechanism as the training model, and uses Softmax classifier to build a classification model to assist clinicians in diagnosis and analysis, to reduce clinicians' work intensity and pressure and improve work efficiency.

2 Method of this paper

2.1 Data enhancement

Image data augmentation is the technical processing of existing data to make the data realize greater value without adding data[5]. For the CT image axial plane that has been preprocessed, the data enhancement methods used include flipping and rotation (such as horizontal or vertical flipping, random angle rotation), image transformation (such as color transformation or affine transformation). Random rotation is to randomly select different angles to rotate right or left within the specified angle range; color transformation is to randomly select different brightness, contrast, saturation, or hue within a certain range to adjust the image; affine transformation is to randomly select different rotation angles, stagger angles, translation distances, scaling factors, etc. within a certain range to adjust the image. Some data enhancement results. Some data enhanced images are shown in Figure 1.
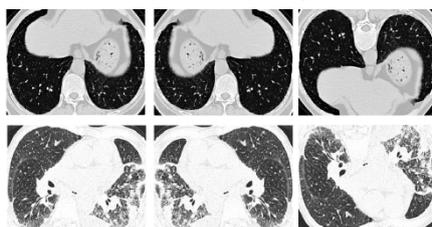

Figure 1. Partial data enhanced image

2.2 Characteristic pyramid structure

In computer vision, the detection of multiscale objects usually takes the image as the input after scaling at different scales, which is used to generate the feature combination of different scale information[6]. This method can effectively express various scale features of pictures, but it requires high computing power and memory of computers. Feature pyramid network (FPN) network is to establish a feature expression structure of different dimensions of pictures of the same size at all levels from bottom to top in the convolution neural network. It can effectively act on the typical convolution neural network model, to generate a feature map with more effective representation ability. In essence, it is a method to strengthen the feature expression of the backbone network.

In the convolutional neural network of ResNet50, conv2_3, conv3_4, conv4_6, conv5_3 in ResNet 50 are used to rebuild FPN. The FPN network structure is shown in Figure 2.
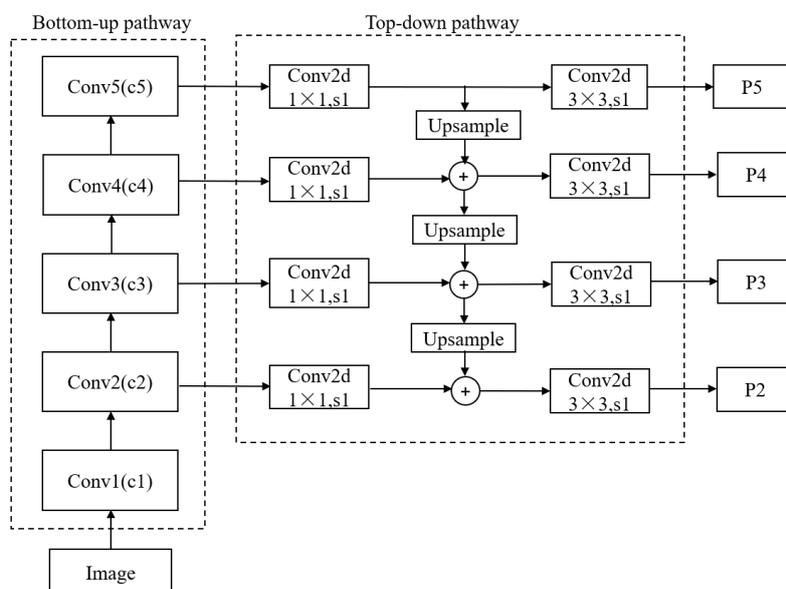

Figure 2. FPN structure diagram of ResNet50

2.3 Improved attention mechanism

In computer vision, the attention mechanism calculates the correlation of different pixels or pixel blocks in a picture to obtain the salient feature information in the picture. Its essence is to obtain the weight distribution of image features, and the core purpose is to obtain key information[7]. Add the attention module before the residual block output. It can effectively calculate the correlation characteristics between feature maps, so that the feature map output to the next residual block contains the correlation characteristics between long-distance features. Because the attention module does not change the characteristics of the size between the input and output. Therefore, the parameter setting of the original network structure is not changed. As shown in Figure 3, the improved attention module is mainly added to the last residual block of conv3, conv4, and conv5 to realize the effective combination of attention module and residual network module.
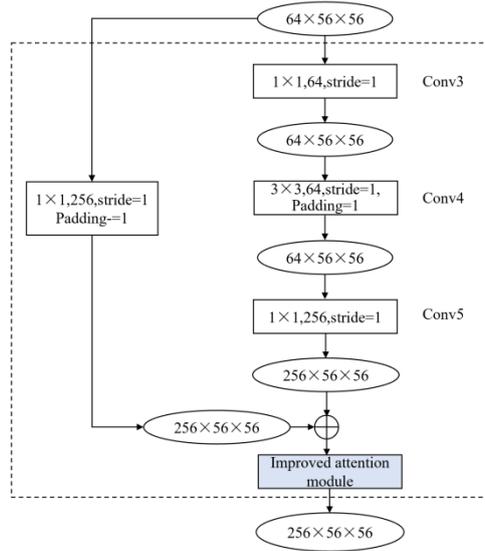

Figure 3. Residual structure of improved attention mechanism

2.4 Loss function and classification function of the model

In the multi-classification medical image classification task using neural network, Softmax is generally used as the activation function of the output layer[8], and category is used as a loss function, cross entropy (multi-category cross entropy loss function) is defined in formula (1).

$$Loss = -\sum y_k \log\left(S(l_k)\right) \qquad (1)$$

Where $S(l_k)$ represents the neuron corresponding to the output layer activated by the softmax function, y_k represents the label corresponding to one hot coding, and the output layer contains k neurons corresponding to K categories[9].

Softmax is defined as formula(2), where Z is a vector and z_i and z_j is one of the elements.

$$Softmax(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \qquad (2)$$

3 Experiments and results

3.1 Experimental environment

The experimental environment of this paper is set up in the Windows Server 2019 operating system.

The deep learning framework is TensorFlow (GPU version) 2.4.0, and the development language is Python 3.7. The computer is HP Z8G4 workstation, and the main experimental environment configuration is shown in Table 1.

Table 1. experimental environment configuration

Items	details
Processor	Intel Xeon (R) Gold 6226, 2.7 GHz
Hardware environment	NVIDIA RTX 2080Ti GPU，Video memory 11GB
Operating System	Windows Server 2019
Deep learning framework	Tensorflow (GPU version) 2.4.0 Keras
General parallel computing architecture	CUDA 11.0

3.2 Experimental process

In the experiment, when the number of training iterations exceeds 35, the loss value of the verification set tends to be stable, and the number of training iterations in this paper is set to 35. As shown in Figure 4. Accuracy indicates the proportion of the correct quantity predicted by the model in the total quantity. The "acc" in the figure refers to the model training accuracy, val_acc refers to the accuracy of the model on the validation set.

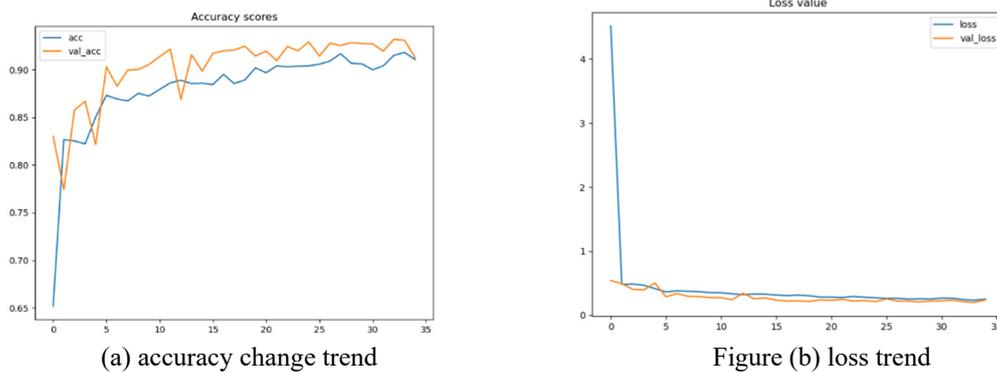(a) accuracy change trend   Figure (b) loss trend

Figure 4. Model parameter training diagram

Loss represents the loss value of the training set, val_loss represents the loss value of the validation set.

In this experiment, we collected about 3000 CT images from the Internet. We showed 1350 CT images of COVID-19 infection, 300 CT images of other viruses, and 1350 CT images of healthy people to the model to test the robustness of the model. We make model predictions on these random images and record the prediction performance of the proposed model. Figure 5 shows the classification of CT images.
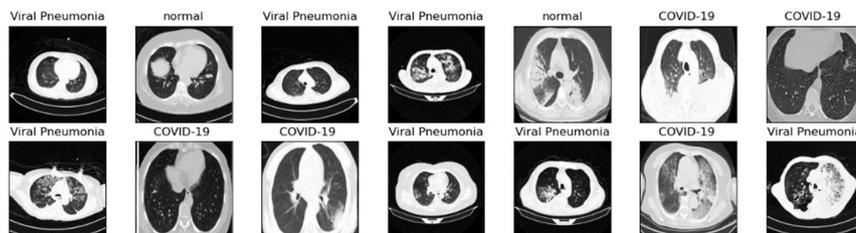

Figure 5. CT image classification

The model successfully detected 2903 CT images, including 1298 images of COVID-19 infection, 1221 images of viral pneumonia, and 184 normal images. The accuracy of the model is 90.1%. As the confusion matrix is shown in Figure 6.
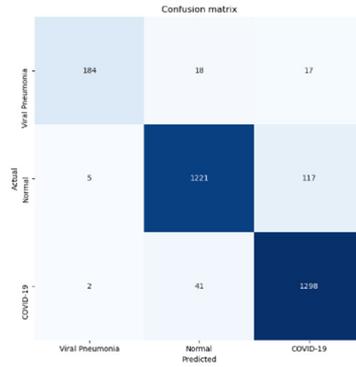
Figure 6. Confusion matrix of the model

3.3　Comparison with the original RESNET classification network

We also tested other neural network methods using the same data set. We established the same experimental setup for the baseline and the proposed method and trained and tested the method on similar data sets. After that, we compared these methods and recorded the speed, accuracy, and other performance indicators.

When testing the speed of various methods, we find that the processing speed of the proposed method is 12.75 FPS on CPU and 39.56 FPS on GPU. However, the original ResNet50 model was 7.89 FPS on the CPU and 23.52 FPS on the GPU. This method is superior to the initial model in speed.

The accuracy of the model proposed in this paper reaches 90.1% after epoch=35, just as shown in Table 2, while ResNet50 model shows 87% accuracy after epoch=35. The proposed model is 3.1% higher than the classical ResNet classification model.

Table 2. Improved ResNet and ResNet comparison

Model	CPU speed（FPS）	GPU speed（FPS）	Accuracy (%)
ResNet50	7.89	23.52	87
Improved ResNet50	12.75	39.56	90.1

4　Conclusions

Accurate and rapid detection of COVID-19 is a challenging diagnostic task. This paper first uses ResNet50 as a pre-training model to study the classification of CT images of COVID-19. Then, based on ResNet50 model, feature fusion and improved attention mechanism are added. The improved ResNet50 is a lightweight and fast feature extraction model. The experimental results show that, compared with ResNet50 model, the improved network structure has fast training speed and 90.1% accuracy, which can meet the detection needs and reduce the pressure of medical workers to a certain extent.

5　Acknowledgments

6　References

[1] WHO, Clinical management of severe acute respiratory infection when novel coronavirus （2019-nCoV） infection is suspected: interim guidance, https://apps. who. int/ir-is/handle/10665/330893, 2021.

[2] LI Shixue, SHAN Ying, Review of research progress in COVID-19. Journal of Shandong University(Medical Edition),vol. 58 no. 3, pp. 19-25,2020.

[3] WANG W, XU Y, GAO R, et al., Detection of SARS-CoV-2 in different types of clinical specimens, JAMA The Journal of the American Medical Association, vol.323 no.18, pp.1843-1844.2020.

[4] Tavare A N, Braddy A, Brill S, et al., Managing high clinical suspicion COVID-19 inpatients with negative RT-PCR: a pragmatic and limited role for thoracic CT, Thorax, vol. 75 no. 7, pp. 537-514, 2020.

[5] Lei J,Li J,Li X,et al.CT Imaging of the 2019 Novel Coronavirus(2019-nCoV)Pneumonia[J].Radiology,2020, 295(1):18.

[6] Pan Y, Guan H, Zhou S, et al., Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV):a study of 63 patients in Wuhan, China. Eur Radiol, vol.30 no.6 pp. 3306-3309 ,2020.

[7] He Kaiming, Zhang Xiangyu, Ren Shaoqiang, et al., Spatial pyramid pooling in deep convolutional networks for visual recognition, IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 37 no. 9, pp. 1904-1919, 2014.

[8] https://zhuanlan.zhihu.com/p/353235794.

[9] WANG S H, FERNANDES S, ZHU Z, et al., AVNC: attention-based VGG-style network for COVID-19 diagnosis by CBAM, IEEE Sensors Journal, vol. 99 no.1.2020