

Design of an Online Detection System for COVID-19 based on Attention Mechanism

Fan Liu, Wei Xiang*, Li Wang, Jintao Zhang, Jiacheng Li

College of Electrical & Information Engineering, Southwest Minzu University, Chendu, 610041, China
Email: 1025605858@qq.com

Abstract. Corona Virus Disease 2019 (COVID-19) is an acute respiratory infection caused by the 2019 novel coronavirus infection that was discovered in some hospitals in Wuhan, China in December 2019. From a global perspective, COVID-19 epidemic is still in a pandemic period. Every country is taking timely preventive measures. Therefore, it is necessary to use Artificial Intelligence (AI) identification to assist radiologists in diagnosing COVID-19. This paper proposes a network model of dynamic self-attention machine for pneumonia images. The model adopts DPN92 and GCNet. It implements CR images and CT image classification of COVID-19. A COVID-19 detection system based on dynamic auto-attention machine is designed. The lung CT or CR images that meet the design requirements will be uploaded through the browser, and the background server will analyze and process the lung images, and finally output the detection category of the lung images.

Keywords: B/S model, SRGAN, GCNet, DPN92, COVID-19, CR, CT, image classification.

1 Introduction

Novel Coronavirus is highly contagious, spreads quickly, has a long incubation period, asymptomatic carrier infection, has no specific drug so far, and returns to positive after recovery [1-2]. Although nucleic acid testing is the most important basis for a patient's diagnosis, it is time-consuming. Chest imaging can help doctors quickly identify lesions and play an important role in the diagnosis of COVID-19. CT images cost more than traditional CR images, which are available in every hospital. However, all hospitals need doctors with experience in CR and CT imaging diagnosis of COVID-19, and the manual diagnosis workload is heavy, time-consuming and energy consuming, which is likely to cause misdiagnosis. Therefore, it is necessary to assist doctors with artificial intelligence (AI) to reduce the workload and improve the detection efficiency. In order to solve this problem, we improved the DPN-92 network based on the attention mechanism, and designed an online pneumonia detection system on this basis, which enables users to realize the pneumonia detection function by browsing the web and uploading pictures.

2 System Design

This system is an online pneumonia detection system, is an open platform, the system uses The B/S mode as the overall structure framework. The whole system can be divided into two parts, foreground browser and background server. The foreground is open to users, who can enter through browsing web pages, and the background is open to managers. The overall structural framework of the system is shown in Figure 1.

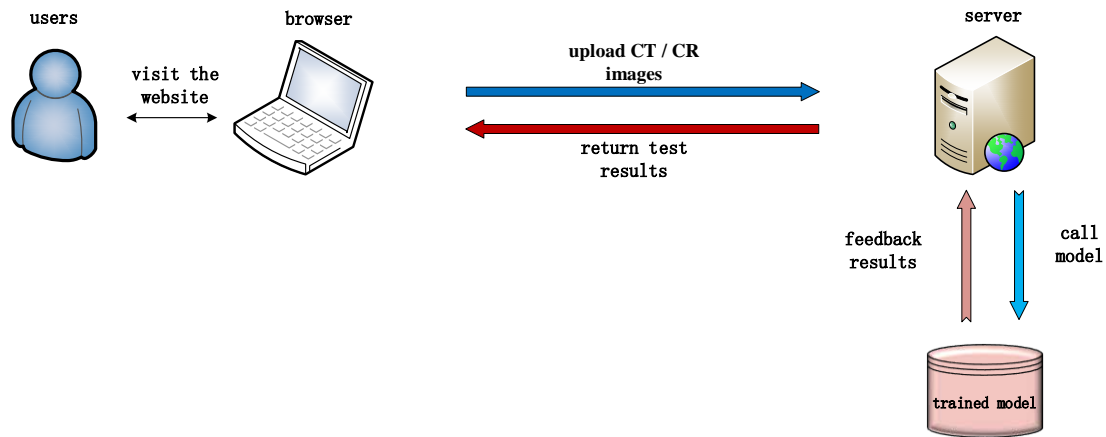


Figure 1. Frame diagram of the overall structure of the system

3 Model Design and Improvement

The COVID-19 detection system we designed focuses on the design of the classification model based on the attention mechanism in the background. Traditional image recognition technology is mainly based on the shallow hierarchical network structure, and the image needs to be pre-processed artificially, thus reducing the accuracy of image recognition. The purpose of deep learning is to build a multi-layer neural network, automatically train, learn and obtain the implicit relationship of data, extract higher-dimensional and more abstract data, and make the obtained features more expressive. Deep learning is widely used because of its strong ability to extract features, high recognition accuracy and fast training speed. In particular, good results have been achieved in face recognition [7], medical image recognition [8-9], remote sensing image recognition [10] and other aspects. Compared with the traditional feature extraction method, the feature extraction method of deep learning can obtain the image features at a deeper level, which makes the image expression ability richer. The research contents of this project mainly include data enhancement, the combination of DPN92[11] model and GCNet[12] model, model deployment and the diagnosis of CR and CT images of pneumonia on the front page [22], etc.

3.1 Data to Enhance

As the acquisition of COVID-19 images has certain privacy and a large number of data sets cannot be obtained, the model is prone to overfitting during training, resulting in insufficient generalization ability. Therefore, it is necessary to enhance the data set. Horizontal and vertical flipping, affine transformation, and cropping are used in this project. In addition, since the CR image data acquired by us are very few and the resolution is low, we also improve the resolution of CR image through enhanced super resolution generation countermeasures network (ESRGAN) to make the data balanced. The central problem with super resolution is the lack of high frequency information and the loss of fidelity at higher resolution. The proposed SRGAN[4] uses the hope-connection Depth residual network (ResNet) to restore the lifelike texture of the image through depth drop sampling, as shown in Figure 2. Its perceptive loss function includes counter loss and content loss. The discriminator network is used to counter the loss, the super resolution image and the original image are different, and the result is closer to the original image by continuous training. Content loss focuses on visual similarity and alleviates the training difficulties of deep networks by introducing residual blocks and hop-joins. ESRGAN (Enhanced super-Resolution Generative Adversarial Networks)[5], published in ECCV2018, is based on the improvement of SRGAN, which is mainly improved in the following three aspects: 1) Improved network structure, anti-loss, and perception loss, using Relativistic Average GAN (RaGAN)[6]; 2) Introduce a residual density unit (RRDB), as shown in Figure 3 and 4; 3) Use VGG features before activation to improve perceived loss.

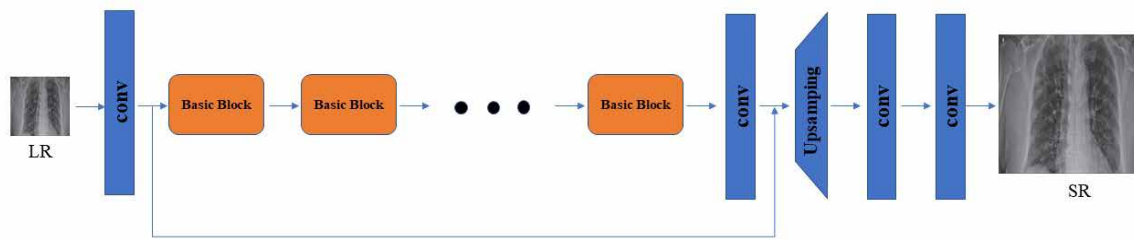


Figure 2. We employ the basic architecture of SRResNet, where most computation is done in the LR feature space. We select “basic blocks” (e.g., residual block, dense block, RRDB) for better performance

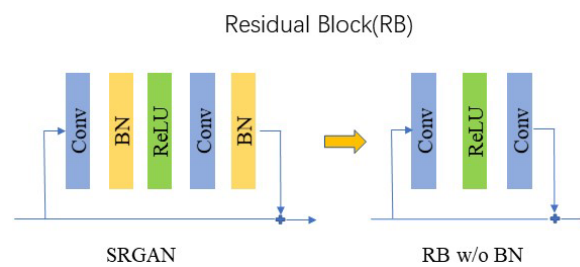


Figure 3. We remove the BN layers in residual block in SRGAN

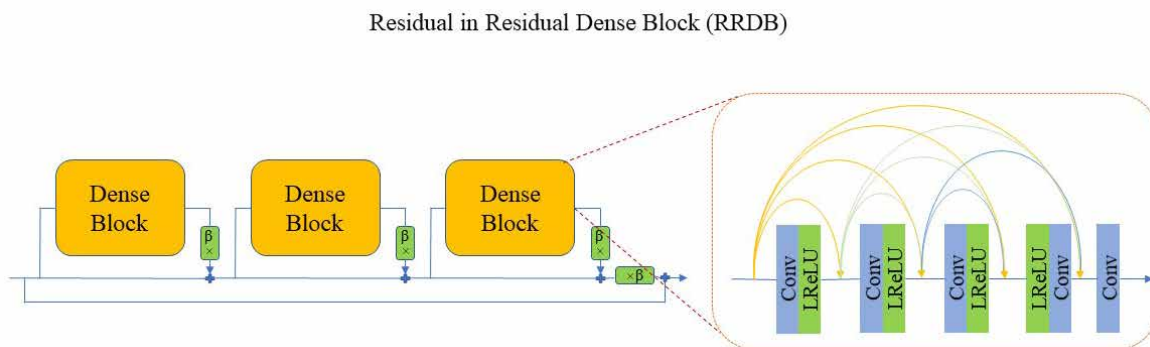


Figure 4. RRDB block is used in our deeper model and β is the residual scaling parameter

3.2 DPN-92

In this project, we used dual Path Networks (DPN) and won the first prize in the Imagenet[27] competition in 2017. DPN mainly fuses ResNet[13] and DenseNet[14]. Essentially, it is the addition of Element-wise that fuses ResNet and the Concatenate of Densenet, but since group convolution is used in the 3×3 convolution of the Prevent layer, it is considered to be ResNet.

Figure 5 (b) for DPN - 92, first of all, after 7×7 convolution, followed by 3×3 largest pool, and then we moved to the left as shown in the operation, [] is a block, cycle 3×3 said the shape of the block, G is a block of ResNet points how much path (i.e., group number), + 16 said DenseNet a block of time increase the number of channels (namely the bar on the left in figure 5 (a) walking down, every time the extra width). Of conv3, of conv4, of conv5, and then softmax.

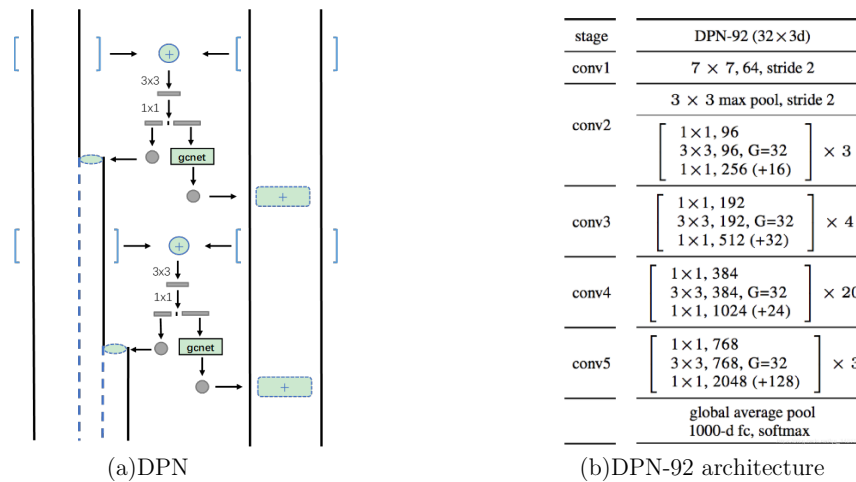


Figure 5. Network architecture

3.3 GCNet

GCNet is similar to SENet and NLNet. In order to overcome the problem of large computation in NLNet, GCNet proposes a Simplified NL block, and due to its similarity with THE STRUCTURE of SENet, GCNet is obtained by combining it with the improvement of SENet.

SENet[15] uses the global context to re-calibrate the weights of different channels and adjust channel dependencies. However, this approach does not take full advantage of the global context information. The goal of capturing long distance dependencies is to gain a global understanding of the visual scene, which is useful for many computer vision tasks, such as image classification, video classification, target detection, semantic segmentation, and so on. NLNet[16] models long distance dependencies through self-attention mechanisms.

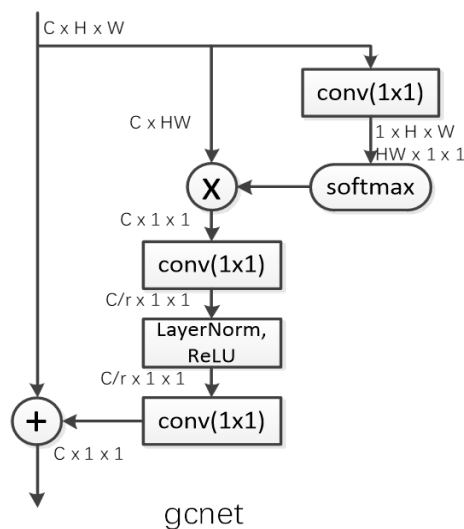


Figure 6. GCNet architecture

As shown in Figure 6, GCNet is able to establish effective long-distance dependencies like NLNet while reducing computation like SENet. The steps are as follows: (1) global attention pooling[17] is used for context modeling. (2) Prevent inter-channel dependencies. (3) Broadcast element-wise addition[19] is used for feature fusion. In the simplified version of NLNet, the Transform module has a number of parameters. To gain the benefit of the SENet lightness, the 1×1 convolution Transform module is used

instead, where R is the reduction rate. Because the two layer Normalization modules increase the difficulty of optimization, the addition of a layer normalization layer in front of ReLU also acts as a regularization to improve generalization ability.

3.4 Design Process of On-line Inspection System

The online detection system model deployment step is divided into four steps. Step 1: Create a Python file in the corresponding directory on the server for routing binding; Step 2: Obtain the uploaded pictures through Post request, verify the legitimacy of the uploaded pictures, and then call different deep learning modules to detect the pictures; Step 3: The returned results are programmed in Python and the related variables are defined. Then, Flask is used to insert the defined variables into the HTML and the test results are displayed. Step 4: Run the Nginx middleware on the server. After setting the domain name and port number, the user can access the website through the browser and test it.

4 Test and Display the Results

4.1 Model Training and Test Results

Optimization is a difficult problem in deep learning. The first is to extend the data set to prevent overfitting on the training set. Then the stochastic gradient descent method is used to minimize the loss function by training the neural network. Its value is the standard that measures network good or bad. The learning rate was optimized by the AdaGrad[20] algorithm, which divided the learning rate by a cumulative squared gradient. During the iteration, the learning rate changed with the cumulative squared gradient. Because the historical accumulation gradient will be larger and larger, the change of learning rate will be smaller in the later stage of training, resulting in unsatisfactory effect. AdaDelta optimizes the learning rate of AdaGrad calculation, calculating the cumulative gradient not for all items, but for a few items of fixed size, avoiding the problem of too large a cumulative gradient.

In the final testing stage, we divided a part of the data set given by the official as the test set, and the test set was not repeated with the training set to avoid overfitting [23].

We used ResNet50[24] to compare with our model. The training and test results of the comparison are shown in Table 1 and Table 2 below. The results show that the training and test results of DPN-92+GCNet model are far better than those of ResNet50.

Table 1. CT training and testing results

category	graphics	ResNet50	DPN-92+GCNet	ResNet50	DPN-92+GCNet
covid-19	3821	81.56%	89.87%	82.22%	92.88%
pneumonia	3277	82.22%	89.14%	83.65%	92.36%
normal	3096	81.32%	90.01%	81.84%	93.02%

Table 2. CR training and testing results

category	graphics	ResNet50	DPN-92+GCNe	ResNet50	DPN-92+GCNet
covid-19	216	80.45%	90.61%	82.54%	93.03%
pneumonia	1321	87.36%	89.54%	89.79%	91.75%
normal	1254	88.24%	89.74%	89.98%	92.69%

4.2 Display of System Results

The page presents the pneumonia detection section, which contains static images and a brief introduction to the related functions of each subsystem. Pneumonia CT and CR detection can be carried out on the home page. Users can enter the relevant domain name address to access the server through the browser and choose to use the target subsystem. Users can upload pictures through an upload

button, and the feedback results can return relevant data in the form of a table for reference and analysis by the detector [25]. In the final testing stage, the performance and accuracy of the network model we proposed were evaluated. The corresponding pictures were extracted from the test set as the experimental results, and the displayed results were compared with the labels corresponding to the pictures to verify the accuracy of the model. Figure 7 is the page tested with the online CT detection system. Figure 8 is the page for testing with the online CR detection system.

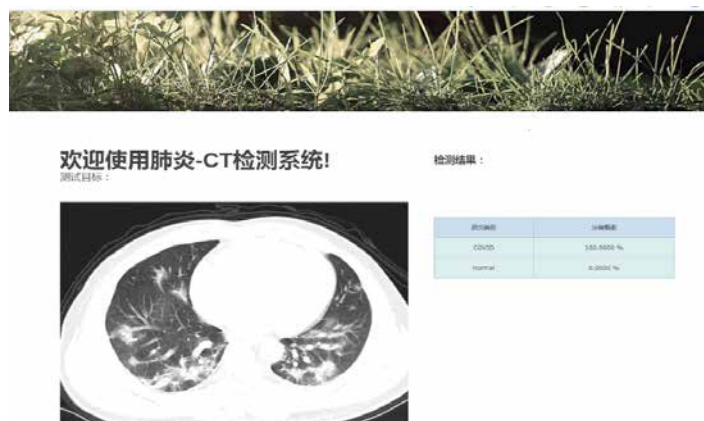


Figure 7. CT images of lung were detected online



Figure 8. CR images of lung were detected online

5 Discussion

Our detection system uses the DPN92 model combined with GCNet, then fuses the features, processes the fused features, obtains new features, and finally outputs the final results through the convolutional layer and the full connection layer. In the face of many problems such as uneven data set distribution, different picture quality, small sample size and difficult model optimization, the accuracy rate is improved.

The final test results showed that the accuracy rate of lung CT detection with the disease could reach more than 90%. Some CT images of the lungs without the disease showed certain deviation, but the final judgment result was more accurate. By deploying to the web page, the system is given flexibility, which is more convenient for users to use and can also be used as a reference for medical judgment. There is still a lot of room for improvement in the system itself. Through testing, the optimal number of iterations in the self-created model can be mined. To solve the problem of data set quality and number,

it is also possible to generate new data set by deforming the existing data set after strict screening of the data set, which need further study.

Acknowledgments. This work was supported by the Southwest Minzu University Graduate Innovative Research Project (Master Program CX2019SP24). A special acknowledgement should give to Southwest Minzu University for its experimental conditions and technical support.

References

1. F. Wu, S. Zhao, B. Yu, et al. A new coronavirus associated with human respiratory disease in China[J], *Nature*,2020,579 (7798):265–269.
2. C. Huang, Y. Wang, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan[J], *Lancet* ,2020,395 (10223) :497–506.
3. I. J. Goodfellow et al. Generative adversarial nets[C]//*Neural Information Processing Systems*.2014: 2672–2680.
4. Ledig C, Theis L, Huszar F, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[J]. *IEEE International Conference on Computer Vision*, 2016.
5. Wang X, Yu K, Wu S, et al. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks[C]// *European Conference on Computer Vision*. Springer, Cham, 2018.
6. Jolicoeur-Martineau, Alexia. The relativistic discriminator: a key element missing from standard gan[C/OL]. (2018-7-5)[2020-10-20].arXiv:1807.00734.
7. Shirahama K, Grzegorzec M. Towards large-scale multimedia retrieval enriched by knowledge about human interpretation[J]. *Multimedia Tools and Applications*, 2016, 75(1):297-331.
8. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. *Nature*, 2017, 542:115–8.
9. Geert L, Thijs K, Babak EB, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017, 42: 60–88.
10. Xia G S, Bai X, Ding J, et al. DOTA: A Large-scale Dataset for Object Detection in Aerial Images[C]// *Conference on Computer Vision and Pattern Recognition*. 2018.
11. Yunpeng Chen, Jianan Li, Huaxin Xiao, et al. Dual Path Networks[C]//. *Conference and Workshop on Neural Information Processing Systems*.2017.
12. Juneja M, Thakur N, Thakur S, et al. GC-NET for classification of glaucoma in the retinal fundus image[J]. *Machine Vision and Applications*, 2020, 31(5).
13. Targ S, Almeida D, Lyman K. Resnet in Resnet: Generalizing Residual Architectures[J]. *ICLRW*, 2016.
14. Muhammad A, Rizwan N, Dong K, et al. IrisDenseNet: Robust Iris Segmentation Using Densely Connected Fully Convolutional Networks in the Images by Visible Light and Near-Infrared Light Camera Sensors[J]. *Sensors*, 2018, 18(5):1501-.
15. Hu J, Shen L, Albanie S, et al. Squeeze-and-Excitation Networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017(99).
16. Surhone L M, Tennoe M T, Henssonow S F. NLnet[M]. *Betascript*, 2010.
17. Huang X, Zheng S. Self-learning Attention Global Pooling Based Image Representation for Visual Place Recognition[C]// *Chinese Automation Congress*.2019.
18. Zbyněk Zajíc, Zelinka J, Jan Vaněk, et al. Convolutional Neural Network for Refinement of Speaker Adaptation Transformation[C]// *International Conference on Speech and Computer*. Springer, 2014.
19. Bernengo M G. Vitiligo is an independent favourable prognostic factor in stage III and IV metastatic melanoma patients: results from a single-institution hospital-based observational cohort study[J]. *Annals of Oncology*, 2010, 21(2):409-414.
20. Hadgu A T, Nigam A, Diaz-Aviles E. Large-scale learning with AdaGrad on Spark[C]// *IEEE International Conference on Big Data*. IEEE, 2015.
21. Qu Z, Yuan S, Chi R, et al. Genetic Optimization Method of Pantograph and Catenary Comprehensive Monitor Status Prediction Model Based on Adadelta Deep Neural Network[J]. *IEEE Access*, 2019, 7:23210-23221.
22. Kaggle. Available online: <https://www.kaggle.com/andrewmvd/convid19-xrays>
23. Xu Z, Shi L, Wang Y, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome[J]. *The Lancet Respiratory Medicine*, 2020.

24. Greenspan H, Van Ginneken B, Summers RM. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique[C]// IEEE Transactions on Medical Imaging.2016, 35:1153–1159.
25. Deng L, Yu D, et al. Deep learning: methods and applications[J]. Foundations and Trends®in Signal Processing. 2017, 7:197–387.