Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/22/20231216

Research on COVID-19 X-ray image recognition algorithm

Haolei Chen

School of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom, SO17 1BJ

Haolei.chen.work@outlook.com

Abstract. With the Covid-19 pandemic, early identification is considered an important measure to fight against this epidemic. Currently, the risk of secondary infection is high, so it is needed a method that can rapidly identify infected individuals. In this research, two algorithms are used, VGG16 and ResNet50, to classify the lung X-ray images from Kaggle and compare the performance of the two algorithms that ResNet50 has better accuracy than VGG16 on the Chest Radiography Database. Next, the CLAHE image enhancement algorithm is introduced in this research to pre-process the images in the dataset. Finally, the two algorithms, VGG16 and Resnet50, are retrained on the processed dataset and compared with their results. And the results show that the two algorithms are improved but ResNet50 still has better accuracy than VGG16. Thus, the CLAHE algorithm is good in this task, which not only significantly enhances the image quality, but also improves the performance on classification of the neural network model.

Keywords: VGG16, ResNet50, CLAHE.

1. Introduction

In December 2019, Covid-19 was first discovered, and it rapidly broke out all over the world. Millions of patients being diagnosed brought new challenges to the medical systems in various countries and regions. At present, the concern about Covid-19 has been reduced in various countries around the world, but the Coronavirus is different from other influenza and second infections may occur. Therefore, it is very important to accelerate the progress of Covid-19 diagnosis, and CT has become the first choice for the detection because physicians can use a patient's chest X-ray image to determine whether the patient has a coronavirus infection [1] or another lung disease. However, the chest X-ray image of coronavirus has very similar symptoms to viral pneumonia [2], and mistakes may be made by manual diagnosis. So, image recognition algorithms provide us with an extra efficient way to diagnose novel coronaviruses, and Visual Geometry Group Network with 16 layers(VGG16) and Residual Network with 50 layers(ResNet50) are widely used in image classification problems. For this consideration, aim of this research is to compare the classification performance of the two algorithms on chest X-ray images. In addition, the quality of images varies due to different standards in the image collection process. And CLAHE algorithm can improve the image quality by enhancing the contrast of the image. So, Contrast Limited Adaptive Histogram Equalized(CLAHE), a medical image enhancement algorithm, is used to process the images and obtain the classification results of two machine learning algorithms after image enhancement algorithm processing.

2. Methodology

In this section, the dataset used in this research is first introduced, and then the two machine learning algorithms are presented to use in this research as well as the image enhancement algorithm.

2.1. Database

Both datasets used in this research are from Kaggle. COVID-19 Radiography Database [3, 4]: The dataset was created by researchers and medical doctors from several countries and regions. It contains 33,920 images, including normal images, COVID-19 pneumonia images, and viral pneumonia images. In this research, 80% of this dataset is randomly selected as the training set and the remaining 20% as the validation set. Chest X-ray (Covid-19 & Pneumonia): The three projects in this dataset are provided by the University of Montreal, the University of California and the University of Waterloo, respectively. And the images in this dataset have three labels, Covid, normal and viral pneumonia, with a total of 6432 images. In this research, this dataset is only used as a test set to verify the generalization ability of the model.

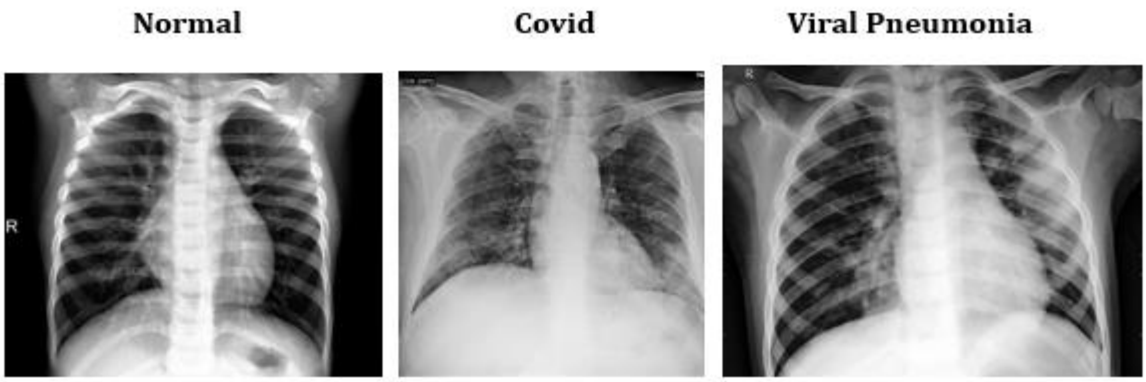

Figure 1. Example of an image in a dataset.

As Figure 1 shows, the image resolutions from the two datasets are not the same. For the machine learning algorithm to perform feature extraction more easily, the resolution of the images is adjusted to 256*256 pixels.

2.2. Data Enhancement（CLAHE）

Due to different acquisition standards as well as sources, the images of X-rays in this research have different qualities, such as different sizes, resolutions, grey and so on. In particular, the different brightness and contrast of images are not conducive to the extraction of image features by the classification algorithm. Therefore, this research adopts the most commonly used enhancement algorithm in medical image processing, the contrast-limited adaptive histogram equalization (CLAHE) algorithm. The CLAHE algorithm is a very classical algorithm that enhances the contrast of an image by changing the histogram of the image into an approximately uniform distribution. It first divides the graph into many non-overlapping blocks, then calculates the grey histogram of each segmented block and the corresponding transformation function, and finally averages the pixels of the rectangular block into individual grey levels in which the average pixel of the grey level can be found by equation (1). Therefore, CLAHE not only enhances the contrast of the image but also represents the key features in the image well without causing a loss of information.

$$\overline{N_a} = \frac{N_{xp} - N_{yp}}{N_g} \tag{1}$$

DOI: 10.54254/2755-2721/22/20231216

N_g is the number of grey levels, N_{xp} represents the number of pixels in the X-axis direction on the segmented block, and N_{yp} represents the number of pixels in the Y-axis direction on the segmented block.

Next, a threshold value N_L needs to be set according to Equation (2), and the grayscale greater than N_L is cropped and assigned to each grey level. where S is the intercept coefficient.

$$N_L = S\overline{N_a} \tag{2}$$

Finally, the histogram of each segmentation block is equalized, and the transformed grey value is obtained using the transform function. figure 2 shows the comparison between the result of image enhancement and the original image.

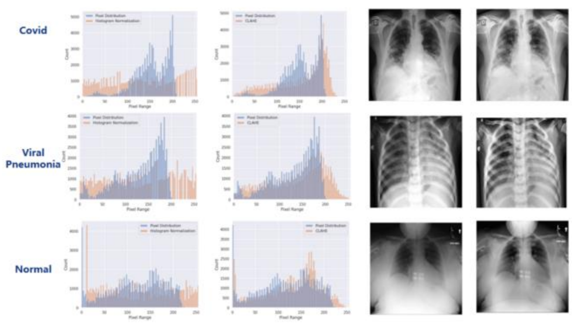

Figure 2. Results of CLAHE image enhancement algorithm.

As figure 2 shows, after the CLAHE algorithm enhancement, the contrast of the image is significantly enhanced, and the shadows in the lungs are more clearly defined.

2.3. VGG16

VGGNET is a convolutional neural network model proposed by Simonyan and Zisserman, which achieved second place in the ImageNet image classification competition in 2014 [5]. VGG16, as one of the best classification algorithms in VGGNET, is often used to classify images, and it has been applied to medical image classification and recognition and has shown good results [6]. Its structure is shown in figure 3.

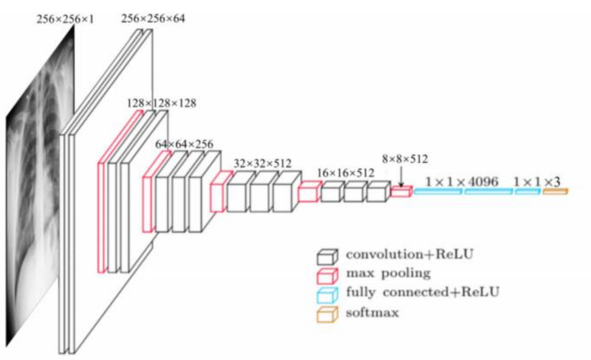

Figure 3. The structure of VGG16.

VGG16 is widely used in the medical field. Albashish et al. used VGG16 to classify breast cancer and achieved high accuracy [7]. And Abdar et al. used VGG16-based migration learning to diagnose neo-coronavirus and achieved 90% accuracy [8].

The VGG16 network consists of five convolutional segments plus one fully connected segment, in which the five convolutional segments contain 13 convolutional layers and the one fully connected segment contains three fully connected layers, so the VGG16 network has a total of 13+3=16 layers.

In this research, VGG16 is used to recognize 3 different types of X-ray images, normal, Covid, and viral pneumonia. First, the image size is set to the same size of 256*256*1 as input, and then the channel becomes 64 after convolution through the first layer. Next, the channel increases twice with each convolution layer, and the image size becomes one-half with each pooling layer. After 13 convolutional layers, 5 pooling layers and 3 fully connected layers, the classification result of the image is output.

2.4. ResNet50

In general, increasing the depth of machine learning will improve the learning ability of the network and increase the accuracy of network classification, because the more complex structure of the network is, the more features can be learned. But it is proved that as the network becomes deeper, the network will degenerate. In other words, the loss value of the model decreases in the early stage but when the number of network layers reaches a bound, the loss value does not decrease but increases. To address this phenomenon, KaiMing et al. proposed the ResNet network, which won the 2015 ImageNet image classification competition [9]. The main idea of the ResNet network is to allow the transfer of the original input information directly to the later layers in the network. This structure allows the network not to be adversely affected by gradient disappearance. Thus, ResNet50 is widely used in various fields for image recognition classification research.

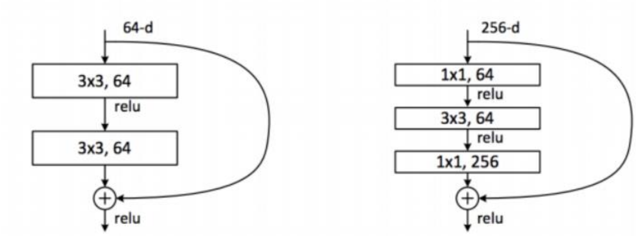

Figure 4. The structure of" skip connection" in ResNet [9].

In Figure 4, the basic residual block is on the left and the bottleneck residual block is on the right. It is worth mentioning that the bottleneck residual block speeds up the training time of the network by changing the dimensionality of the feature map, without affecting the training accuracy.

Walvekar et al. implemented the Resnet network for the classification of CT images of coronavirus pneumonia and obtained an accuracy of 96.23% [10], which means Resnet50 is a good approach to classifying coronavirus pneumonia images. Its structure is shown in figure 5.

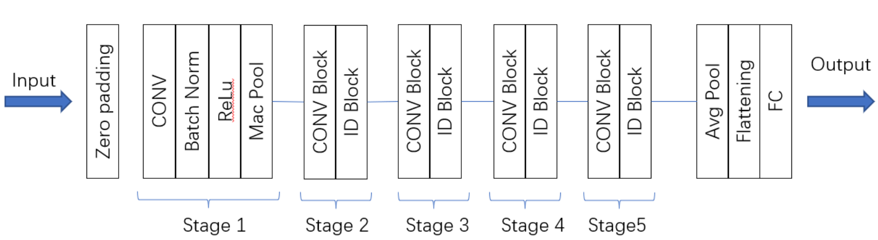

Figure 5. The structure of ResNet50.

The structure of the Resnet network consists of five stages and a fully connected layer. The first stage consists of convolution, normalization, ReLU [11] and pooling operations. Through the first stage, the network depth becomes 64, and after 5 stages, the final output is 3.

DOI: 10.54254/2755-2721/22/20231216

3. Evaluation Indicators

Accuracy represents the number of samples that the classifier can correctly classify as a percentage of the total number of samples.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{3}$$

Precision indicates the probability of a true positive sample being correctly predicted among those whose prediction is positive.

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

The recall represents the probability that the samples with positive predictions account for all positive samples.

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

F1-score is the evaluation index of precision and recall, in general, they are negatively correlated, the higher the F1-score the better

$$F1 - score = 2 \times \frac{Recall \times Accuracy}{Recall + Accuracy} \tag{6}$$

TP is the number of positive class samples predicted to be positive, FN is the number of positive class samples predicted to be negative, TN indicates the number of negative class samples predicted to be negative, and FP is the number of negative class samples predicted to be positive.

In medical classification, accuracy is an important method to test model performance. Therefore, accuracy is used to verify the ability to the classification of the algorithms in this research.

4. Algorithm Framework

In order to identify patients with COVID-19 infection quickly and accurately, this research adds the CLAHE data enhancement algorithm to the commonly used algorithms, VGG16 and ResNet50.

These algorithms firstly preprocess the input image, segment the image and enhance the contrast of the image; then divide the data into the training set, test set and validation set, where the training set and validation set are from the same dataset and the test set is from another dataset that it aims to test the generalization of the model; next, the models are trained to extract the features of the X-ray images. Finally, the quantified features are classified, and the triple classification results are output. The structure of this process is shown in figure 6.

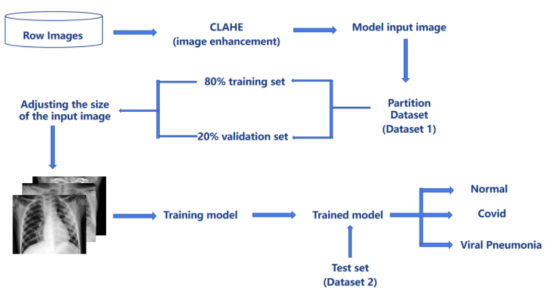

Figure 6. The structure of the algorithm.

5. Performance

In this section, the first part is the performance of the classical algorithm, VGG16 and ResNet50. And the second part is the result of merging the CLAHE algorithm into the classical algorithm.

DOI: 10.54254/2755-2721/22/20231216

5.1. Performance without CLAHE

5.1.1. *VGG16.* The accuracy curve of VGG16 is shown in figure 7. At an epoch equal to 24, the accuracy of the model starts to decrease, and the loss starts to increase, which indicates that there may be overfitting of the model and the accuracy of the model reaches 96%. Therefore, training should be stopped at this point. VGG16 performs best at an epoch equal to 24 where the accuracy of test set is 92.7%. Because the first dataset is used for training and validation and the second dataset only as a test set and they are two independent datasets, the accuracy of 92.7% represents that the VGG16 model has good classification and strong generalization on the lung X-ray images.

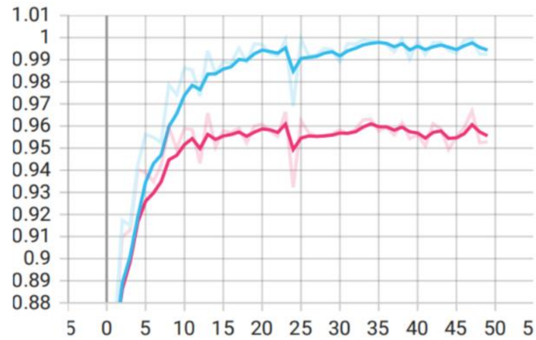

Figure 7. Accuracy of the VGG16.

5.1.2. ResNet50. In order to find the suitable stopping point, an early stop algorithm is used, and the patience and tolerance of the algorithm are set to ask 4 and 0 respectively. The accuracy of ResNet50 on the training set and validation sets was obtained as shown in Figure 8. ResNet50 was able to achieve an accuracy of 96% in the validation set and 95% in the test set. As with VGG16, the dataset used by Resnet50 for training and validation and the dataset used for testing are two completely different datasets, so the 95% accuracy in the test set can also indicate that the model has good generalization.

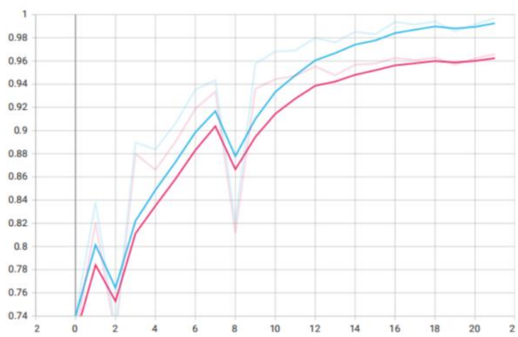

Figure 8. Accuracy of the ResNet50.

5.2. Performance with CLAHE

5.2.1. VGG16. After adding the CLAHE data enhancement algorithm to the 4,1,1 algorithm, it can be seen that the accuracy of the algorithm increases from 96.5% to 97.1% on the validation set and from 92% to 94% on the test set, which indicates that the CLAHE image enhancement algorithm can enhance the classification performance of VGG16.

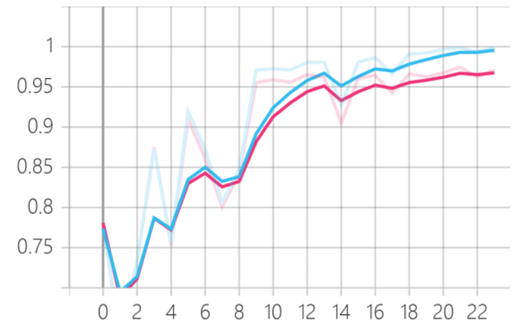

Figure 9. Accuracy of the VGG16 with CLAHE.

5.2.2. ResNet50. After adding the CLAHE data enhancement algorithm to the ResNet50 algorithm of 4.1.2, it can be seen that the accuracy of the algorithm increases from 96% to 97% on the validation set and from 95.2% to 95.8% on the test set, indicating that the CLAHE image enhancement algorithm can enhance the classification performance of ResNet50

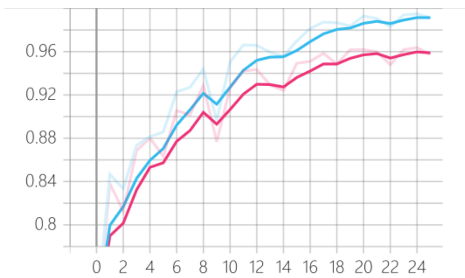

Figure 10. Accuracy of the ResNet50 with CLAHE.

5.3. Comparison

By comparing the above four algorithms, both two algorithms show strong generalization and good accuracy. Compared with VGG16, ResNet50 has better classification results on the Covid dataset. In addition, the image enhancement algorithm, CLAHE, can significantly enhance the contrast of the image and improve the accuracy of the classification algorithm to a great extent.

The improvement of the accuracy of the VGG16 algorithm by the CLAHE is significant, but the ResNet50 algorithm is not satisfactory. This may be because the classical Resnet algorithm itself achieves a high accuracy rate. or because the effect of the CLAHE algorithm structure, which can pass image features from the upper layer directly into the lower layer is the same as that of the image contrast enhancement algorithm, and both algorithms extract more image features

6. Conclusion

In this research, the aim is to classify COVID-19 X-ray images by neural networks and implement two basic classification algorithms, the VGG16 and ResNet50. The difference in accuracy between the VGG16 network on the first data set and the second data set is 3.3%, which means that the model generalizes well. On the other hand, ResNet50 exhibits better performance than VGG16. ResNet50 possesses 95.2% accuracy on the second dataset, indicating that the ResNet50 model has a stronger generalization capability. In addition, applying the medical image enhancement algorithm (CLAHE) to the VGG16 and ResNet50 models, it can be seen that the accuracy of both algorithms increases on the second dataset, 94% and 96%, respectively.

This research compares four Covid image recognition algorithms with high accuracy, but the accuracy of the algorithm is not yet optimal due to hardware limitations. The free GPU computing resources of COLAB used in the experiments of this research do not invoke the GPU performance well, and the accuracy can be improved by increasing the number of iterations and increasing the Batch-size if a better hardware environment is available.

In future improvements, the accuracy rate can be improved in three aspects. Firstly, in terms of data enhancement, the U-Net-based rib subtraction algorithm is considered to improve data quality and consider using Generative Adversarial Network to increase the amount of data. Secondly, in terms of network parameters, increasing the number of network layers, such as comparing the classification accuracy of Resnet50, Resnet101, and Resnet152 in the COVID-19 dataset. Finally, in terms of model structure, it may be possible to consider using VGG16 and Resnet50 dual-channel algorithms.

In conclusion, ResNet50 shows better classification ability than VGG16 in classifying new coronary pneumonia images, and the CLAHE algorithm can largely improve the accuracy of the model, which is a good data enhancement algorithm.

References

[1] Wenling Wang, Yanli Xu, Ruqin Gao, Roujian Lu, Kai Han, Guizhen Wu, and Wenjie Tan. Detection of sars-cov-2 in different types of clinical specimens. Jama, 323(18):1843– 1844, 2020.

[2] Adam Jacobi, Michael Chung, Adam Bernheim, and Corey Eber. Portable chest x-ray in coronavirus disease-19 (covid- 19): A pictorial review. Clinical imaging, 64:35–42, 2020.

[3] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, “Can AI help in screening Viral and COVID-19 pneumonia?” IEEE Access, Vol. 8, 2020, pp. 132665 - 132676. Research link

[4] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S. and Chowdhury, M.E., 2020. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images. Research Link

[5] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[6] Javier Civit-Masot, Francisco Luna-Perej´on, Manuel Dom´ınguez Morales, and Anton Civit. Deep learning system for covid-19 diagnosis aid using x-ray pulmonary images. Applied Sciences, 10(13):4640, 2020.

[7] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat and N. Ahmad Almansour, "Deep CNN Model based on VGG16 for Breast Cancer Classification," 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 2021, pp. 805-810, doi: 10.1109/ICIT52682.2021.9491631.

[8] Abolfazl Karimiyan Abdar, Seyyed Mostafa Sadjadi, Hamid Soltanian-Zadeh, Ali Bashirgonbadi, and Mehran Naghibi. Automatic detection of coronavirus (covid-19) from chest ct images using vgg16-based deep-learning. In 2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME), pages 212–216. IEEE, 2020.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[10] Walvekar, S., & Shinde, D. S. (2020, May). Detection of COVID-19 from CT images using resnet50. In 2nd International Conference on Communication & Information Processing (ICCIP).

[11] Liang XL, Xu J. Biased ReLU neural networks. Neurocomputing, 2021, 423: 71–79. [doi: 10.1016/j.neucom. 2020.09.050]