Classification of COVID-19 on Chest CT Scans with Higher Order Residual Network

Hao Huang and Sei-ichiro Kamata Waseda University, IPS, Kitakyushu, Japan Email: koukou44@akane.waseda.jp, kam@waseda.jp

Abstract—The spread of the coronavirus disease 2019 (COVID-19) epidemic has resulted in massive loss of life and economic loss all over the world. Computed Tomography (CT) has been proven to help us screen out patents more effectively and accurately, but there are not enough experts to interpret CT images. Reliable automatic Artificial Intelligence (AI) recognition systems for COVID-19 CT scans are incredibly vital due to this situation. Moreover, the backbone network for feature extraction, e.g. the deep Residual Network (ResNet), plays a fundamental role in these advanced systems. However, the original design of residual connection way can hardly combine the local extracted features, which is important for the interpretation of CT images. In recent years, the relationship between ResNets and the dynamical system has drawn wide attention, and to combine extracted features, our study explores the high order numerical differential formula from this perspective: merging several standard residual blocks into one as advanced high-order scheme in several aspects with no extra parameters. Compared to the original residual connection method, the network with our proposed sixth-order residual block achieves 83.11% accuracy on the CCAP dataset and has better stability and convergence speed performance without adding parameters.

Index Terms—COVID-19 CT diagnosis, deep residual network, high order numerical method

I. INTRODUCTION

A. Background

The Coronavirus disease 2019 (COVID-19) is a contagious disease caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), which is extremely infectious among humans. According to the data from World Health Organization (WHO), as of October 2021, it has led to 244 million infected cases and more than 4.96 million deaths [1] and continues to spread worldwide.

Real-time Reverse Transcription Polymerase Chain Reaction (RT-PCR) is considered as the gold standard in clinical diagnosis of COVID-19; however, RT-PCR has three shortcomings: (1) It takes a long time and requires 4-6 hours to get results, making medical institutions unable to respond quickly to patients; (2) Strict laboratory environment requirements, which means the results are not reliable enough due to sampling contamination or operational errors; (3) Low sensitivity, which will result in the inability to quarantine patients who were actually infected with the COVID-19 and received a negative RT-PCR result, increasing the risk of the epidemic spreading in the community.

With the continuous development of hardware and image acquisition technology, medical images play an increasingly important role in helping diagnosis. Some studies [2], [3] indicated that chest Computed Tomography (CT) has relatively higher sensitivity for the diagnosis of COVID-19, and the positive results of CT scans are highly consistent with positive RT-PCR results. There were few situations in patients who got negative CT images while the results of RT-PCR were positive. Therefore, CT scans can be regarded as a helpful auxiliary tool at an early stage of diagnosis of COVID-19; however, there is a limitation in applying CT scans to clinical detection of COVID-19, that is, images interpretation is relied on skilled radiologists currently, which will cause radiologists to overwork and lead to increasing of the misdiagnosis. Thus, a reliable Computer-Aided Diagnosis (CADx) system is urgently needed.

In recent years, Artificial Intelligence (AI) methods, especially deep learning, have achieved significant results on computer vision tasks, including medical image recognition.

Less than two years after the outbreak of the COVID-19 epidemic, the use of deep learning models to help clinically detect patients infected with COVID-19 has become a very popular topic. Several works [4], [5] summarize current research and challenges in this area. Many researchers developed their models based on the deep residual neural network (ResNet) [6], and some works have shown promising results. However, the residual block, which is the basic component of ResNet, has a limitation when it is used to interpret CT images because of its own structure. That is, when human radiologists interpret CT images, they often need to combine different areas of the image for observation to eliminate interference from similar blood vessels or tissues, but the features extracted from the residual block are independent and have no relationship with each other. This situation leads to the network is easy to be misled by some local features of blood vessels or human tissues similar to the lesions, and make wrong judgments, especially in the deep neural network, the size of the feature map is often compressed to few pixels by pooling operation.

Manuscript received April 18, 2022; revised July 21, 2022.

In this work, we aim to address this problem by redesigning the residual block from a higher order perspective. We first increase weight layers of each residual block by merging several original residual blocks into one, then redesign the residual connection method by introducing the high order numerical differential equations as the guideology. In this way, we redesign the residual block to make extracted features enable to contain information of other regions, and no additional parameters are added compared to the original residual block.

B. Contribution

This paper proposes a new residual connection method based on the high-order numerical differential equations for diagnosing COVID-19, healthy people, and Community-Acquired Pneumonia (CAP). Even though our primary purpose is to screen out the patients suspected of being infected with COVID-19, in practice, COVID-19 and CAP have extremely similar radiological characters (as shown in Fig. 1) while the required treatment strategies are different. Therefore, we need to distinguish COVID-19 from other types of pneumonia. The method we proposed can achieve a higher accuracy (83.11%) on the CCAP dataset, and at the same time, has a faster convergence speed than the original residual block when their parameters are roughly the same.



Figure 1. Some samples of different pneumonias. From left to right are bacteria pneumonia, mycoplasma pneumonia, viral pneumonia and COVID-19.

II. RELATED WORKS

Since the rapid expansion of the epidemic worldwide, many researchers have used deep learning technologies to create AI systems to detect COVID-19 based on Chest X-Rays (CXR) and CT images, which all with promising results.

Khan *et al.* [7] designed CoroNet, a deep Convolutional Neural Network (CNN) model based on Xception architecture, to automatically detect COVID-19 from chest X-rays images. Wang *et al.* [8] proposed COVID-Net, which is customized for diagnosing COVID-19 from Xrays images; a projection-expansion-projection design pattern is widely used in the architecture to improve representation abilities while maintaining computational efficiency. Sethy *et al.* [9] combined deep CNN models and support vector machine classifiers to classify the Xrays images of COVID-19, pneumonia patients, and healthy people. Li *et al.* [10] introduced convolutional Siamese neural networks for COVID-19 severity assessment. This model calculates the Euclidean distance between a CXR image and normal image to indicate the degree of abnormality and use the median distance between a set of normal images and tested image to present the severity.

Some researchers conduct their work based on CT scan images. Singh *et al.* [11] used CNN to distinguish the COVID-19-infected patients as infected or not and used multi-objective differential evolution to tune the hyperparameters of CNN. Shah *et al.* [12] designed a selfdeveloped network CTnet-10 for the diagnosis of COVID-19. Rahimzadeh *et al.* [13] utilized ResNet50V2 as the backbone and applied a feature pyramid network to concatenate the output features of ResNet50V2, which can classify COVID-19 from various scales of features.

Li *et al.* [14] proposed COVNet by using ResNet50 as the backbone and combining a max-pooling operation to the output of ResNet50, aiming to extract two-dimensional local and three-dimensional representative features. In order to address the data deficiency problem, a selfsupervised transfer learning method called Self-Trans was proposed by He *et al.* [15], which exploits contrastive selfsupervised learning to train network weights on target data and applies pre-trained weights to network to reduce the bias of different domains. Wang *et al.* [16] proposed a joint learning method to learn data from two datasets to improve network robustness.

Besides CNN, some researchers try to apply Long Short-Term Memory (LSTM) to diagnosis of COVID-19. Islam *et al.* [17] combined CNN and LSTM, fed output feature maps of CNN to LSTM and get classification result after the fully connected layer. Mohammed *et al.* [18] used bidirectional LSTM to transform the spatial features to forward and backward spatio-axial features vectors.

III. PROPOSED METHOD

The ResNet proposed in 2015 has become one of the most popular models in computer vision. Plenty of stateof-the-art approaches are based on ResNets or variations of ResNet since the elegant skip connection significantly alleviates the degradation problem of plain networks. However, when it comes to medical image recognition field, these variations of ResNets are difficult to achieve better results, thus most researchers still use the original ResNet as the backbone when studying the diagnosis of COVID-19.

In the interpretation of CT images, because the local image morphological features of some tissues and blood vessels in the pulmonary may be similar to the lesions, it is very important to judge the lesions in combination with other local features. Although ResNet itself has the advantages of being fast and concise, each standard residual block of ResNet has only two weight layers, which leads to weak combination between the extracted local features. This inspired us to think about and explore the principle of ResNet, that is, the residual connection method.

In recent years, some researchers [19] have begun to discuss the relationship between ResNet and dynamical

systems. Reference [19] demonstrates that the expression of ResNet:

$$output = F(x) + x, \tag{1}$$

It can be regarded as a first order Euler discretization of a continuous transformation. This view combines residual neural networks with numerical differential equations. Therefore, we attempt to use higher-order numerical differential equations as the guiding ideology to reconstruct the residual connection to enhance the relevance of the local features extracted by ResNet.

A. Proposed Higher Order Residual Block

In this study, we regard the original residual block (1) as the first order block, the high order residual block means the residual block after we redesign the residual connection by introducing high-order numerical differential method. For instance, the residual connection method of a second-order residual block is designed by utilizing Runge-Kutta second-order method, the other high order residual blocks are the same.

In order to enhance the combination of extracted local features, we attempt to increase the number of weight layers in a residual block. We merged several standard residual blocks into one residual block, then used high order numerical differential equations as a guideline to reconstruct the residual connection method. The implementation of our proposed high order residual block is shown as Fig. 2(a), n means n-th order residual block.

As an example, Fig. 2(b) shows the construction of sixth order residual block. We merged six original residual blocks into one block and used sixth-order Runge-Kutta-Nystrom (RKN) formula [20] as the guideline to redesign residual connection of residual block. The symbols are described as

$$k_{1} = F(x)$$

$$k_{2} = F\left(\frac{1}{3}hk_{1} + x\right)$$

$$k_{3} = F\left(\frac{1}{25}h(4k_{1} + 6k_{2}) + x\right)$$

$$k_{4} = F\left(\frac{1}{4}h(k_{1} - 12k_{2} + 15k_{3}) + x\right)$$

$$k_{5} = F\left(\frac{1}{81}h(6k_{1} + 90k_{2} - 50k_{3} + 50k_{4}) + x\right)$$

$$k_{6} = F\left(\frac{1}{75}h(6k_{1} + 36k_{2} + 10k_{3} + 8k_{4}) + x\right)$$

$$output = \frac{1}{192}h(23k_{1} + 125k_{2} - 81k_{5} + 125k_{6}) + x,$$

where h is a trainable parameter used to enable the network to automatically learn the extent of combination. Initializing it to zero also reduces unnecessary calculations caused by random initialization of weights. Each weight layer contains the batch normalization-conv3x3-relu operation [21].



Figure 2. (a) The implementation of high order residual block; (b) The illustration of sixth order RKN residual block.

B. Overall Network Architecture

The overview architecture of our proposed network is illustrated as Fig. 3. The size of input image is 224x224x3, use a convolution kernel of size 3x3 for downsampling. It passes through four high order residual blocks, the feature map size is halved block by block and the number is doubled block by block, then passes through the average pooling and fully connected layer to obtain the classification result.

In this study, we implemented four kinds of high order residual block, they are second-order, third-order, fourthorder and sixth-order residual block, the compare the network composed of them and the network composed of original residual block (the first order block). According to the illustrated implementation method in Fig. 2(a), the second-order block has four weight layers, the third-order block has six weight layers, the fourth-order block has eight weight layers, and the sixth-order block has 12 weight layers. For fair comparison, we need to ensure that the network used for comparison have the same parameter size. The detail parameter size is shown as Table I, we designed the residual block of each network has a total of 48 weight layers.



Figure 3. The overall architecture of proposed network.

TABLE I. PARAMETER SETTING OF EACH NETWORK

Layer	Number of blocks					
name	Original	Second	Third	Fourth	Sixth	
High order block 1	6	3	2	1	1	
High order block 2	6	3	2	2	1	
High order block 3	6	3	2	2	1	
High order block 4	6	3	2	1	1	

IV. EXPERIMENTS AND RESULTS

A. Dataset

We use the CCAP dataset released by Tao Yan [22], which contains five categories of CT images collected from multiple hospitals, including normal cases, mycoplasma pneumonia, bacterial pneumonia, viral pneumonia and COVID-19. In detail, it collected 7,410 CT images from 154 individuals, 2,103 normal images from 50 individuals, 784 mycoplasma pneumonia images from 18 patients, 1,644 bacterial pneumonia images from 32 patients, 844 viral pneumonia images from 12 patients and 2,035 COVID-19 images from 42 patients, respectively.

Given the uneven data distribution, we divided the data into three categories, normal, CAP and COVID-19. Then we split the experimental data into the training set, validation set and test set according to the ratio of 80:10:10 as shown in Table II.

Category	Training (80%)	Validation (10%)	Test (10%)	In total (patients)
Normal	1,683	210	210	2,103(50)
CAP	2,618	327	327	3,272(62)
COVID-19	1,629	203	203	2,035(42)
In total	5,930	740	740	7,410(154)

TABLE II. COMPOSITION OF EXPERIMENTAL DATA

Since the CAP classification contains three subclasses, when designed this classification, we first split each subclass into the training set, validation set and test set according to the ratio of 80:10:10, and then combine them separately.

B. Pre-processing

There is usually some unnecessary background information in the CT images, affecting the network's performance. The most common sources include the bed of the machine, the clothes of the patient, the noise of the sampling process and the imaging process, etc.

We preprocessed the images as shown in Fig. 4 to make sure pulmonary regions are the primary basis for diagnosis.



Figure 4. Process of image preprocessing.

First, we take a horizontal and vertical line in the image, read the pixels on these two lines and draw their histogram. Second, we can easily determine the approximate location of the pulmonary area from the histogram, estimate a rectangular area slightly larger than the lung, and get a masked image. Then apply the active contour method [23] to this masked image to get the contour of the human body and segment it. Finally, to alleviate the obvious area edges in the image, we filled the background with the average value of pixels on the human body.

In this way, we eliminate the information on the CT images that we do not care about and only kept the human body.

C. Implementation Details

The experiments are conducted on a workstation consisting of an NVIDIA GeForce RTX 3090 graphics card, Intel i7-10700K processor, and 16GB memory. The code is implemented in PyTorch. The optimizer used for training was Adam, and the learning rate change strategy is to reduce the learning rate when the validation accuracy has not changed for 10 consecutive epochs with a reduction factor of 0.1. The size of input images is resized to 224x224x3, the batch size is 32, the initial learning rate is 0.0001, the loss function is cross-entropy, and the total epoch is 100.

D. Results

1) Compasion with original residual block

We first compare our proposed high order block with the original standard residual block. We can see from Table III, our designed high order residual connection method gets better accuracy in the classification task. Also, from the training curves in Fig. 5, we see that our residual connection method with a high-order scheme has faster converge speed, usually converges around 25th epoch, and the original residual connection method needs to train more than 40 epochs to converge. Meanwhile, the proposed method can achieve higher accuracy both in training set and validation set.

TABLE III. RESULTS OF DIFFERENT METHODS

Method	Accuracy	Method	Accuracy
Original	0.7797	CTnet-10 [12]	0.7851
Second order	0.8216	ResNet50-FPN	0.8068
		[13]	
Third order	0.8216	COVNet [14]	0.8122
Fourth order	0.8284		
Sixth order	0.8311		



Figure 5. Comparison of the training curves of the proposed network and the original network. (a) Training accuracy curve; (b) Training loss curve.

2) Compasion with some existing works

We then compare our model with some existing works. Among then, COVNet and ResNet50-FPN use ResNet50 as the backbone, and CTnet-10 is based on VGG. We see that our proposed method also performs better. The training curve is shown as Fig. 6. Because the sixth-order residual network is the best performing network among the proposed networks, only the training curve of the sixth-order residual network is drawn for comparison. In terms of convergence speed, COVNet is faster than my proposed network. In other respects, our model has higher accuracy, and the curve is smoother, more stable, and doesn't oscillate.



Figure 6. Comparison of the training curves of the proposed network and some existing works. (a) Training accuracy curve; (b) Training loss curve.

V. CONCLUSION

This study explores the ResNet from a principle perspective to develop a high accuracy diagnosis system for the diagnosis of COVID-19. We redesigned the residual connections with a higher-order scheme based on high order numerical method that obtained several noticeable improvements. Compared with the original residual connection method, the proposed method performs much better on the COVID-19 classification task in terms of accuracy, stability and convergence speed with the same parameter size.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Hao Huang proposed ideas, conducted experiments, analyzed the data and wrote the original draft; Sei-ichiro Kamata supervised and leaded this research, provided experimental resources, reviewed and edited the original draft.

REFERENCES

- [1] WHO (Oct. 2021). Coronavirus (COVID-19) Dashboard. [Online]. Available: https://covid19.who.int/
- [2] T. Ai, et al., "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases,"*Radiology*, vol. 296, no. 2, pp. E32-E40, 2020.
- [3] M. Karam, et al., "Chest CT versus RT-PCR for the detection of COVID-19: Systematic review and meta-analysis of comparative studies," JRSM Open, vol. 12, no. 5, pp. 1-10, 2021.
- [4] A. Mohammadi, et al., "Diagnosis/prognosis of COVID-19 chest images via machine learning and hypersignal processing: Challenges, opportunities, and applications," *IEEE Signal Processing Magazine*, vol. 38, no. 5, pp. 37-66, 2021.
- [5] S. Bharati, P. Podder, M. R. H. Mondal, and V. B. S. Prasath, "Medical imaging with deep learning for COVID-19 diagnosis: A comprehensive review," *International Journal of Computer Information Systems and Industrail Management Applications*, vol. 13, arXiv:2107.09602, 2021.
- [6] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.
- [7] A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest xray images," *Computer Methods Programs Biomed.*, vol. 196, p. 105581, 2020.
- [8] L. D. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, p. 19549, 2020.
- [9] P. K. Sethy and S. K. Behera. (2020). Detection of Coronavirus Disease (COVID-19) based on deep features. *Preprints*. [Online]. Available: doi:10.20944/preprints202003.0300.v1.
- [10] M. D. Li, et al., "Automated assessment and tracking of COVID-19 pulmonary dis ease severity on chest radiographs using convolutional Siamese neural networks," *Radiology: Artificial Intelligence*, vol. 2, article e200079, 2020.
- [11] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 39, no. 7, pp. 1379-1389, 2020.
- [12] V. Shah, et al., "Diagnosis of COVID-19 using CT scan images and deep learning techniques," *Emerg. Radiol.*, vol. 28, pp. 497-505, 2021.
- [13] M. Rahimzadeh, A. Attar, and S. M. Sakhaei, "A fully automated deep learning-based network for detecting COVID-19 from a new and large lung CT scan dataset," *Biomedical Signal Processing* and Control, vol. 68, article 102588, 2021.

- [14] L. Li, *et al.*, "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65-E71, 2020.
- [15] X. H. He, et al. (April 2020). Sample-efficientdeep learning for COVID-19 diagnosis based on CT scans. [Online]. Available: https://doi.org/10.1101/2020.04.13.20063941.
- [16] Z. Wang, Q. D. Liu, and Q. Dou, "Contrastive cross-site learning with redesigned net for COVID-19 CT classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2806-2813, 2020.
- [17] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics in Medicine Unlocked*, vol. 20, article 100412, 2020.
- [18] A. Mohammed, et al., "Weakly-supervised network for detection of COVID-19 in chest CT scans," *IEEE Access*, vol. 8, pp. 155987-156000, 2020.
- [19] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. Dubenaud, "Neural ordinary differential equations," *Advances Neural Information Processing Systems*, vol. 31, pp. 6571-6583, 2018.
- [20] J. R. Dormand, M. E. A. El-mikkawy, and P. J. Prince, "Families of Runge-Kutta-Nystrom formulae," *IMA Journal of Numerical Analysis*, vol. 7, pp. 235-250, 1987.
- [21] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. European Conference on Computer Vision*, vol. 9908, 2016, pp. 630-645.
- [22] T. Yan. (July 2020). CCAP IEEE dataport. [Online]. Available: https://dx.doi.org/10.21227/ccgv-5329.
- [23] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321-331, 1988.

Copyright © 2022 by the authors. This is an open access article distributed under the Creative Commons Attribution License (<u>CC BY-NC-ND 4.0</u>), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.



Hao Huang was born in Haikou, Hainan, China in 1996. From 2014 to 2018. He studied biomedical engineering in Beijing Institute of Technology and obtained a bachelor's degree in engineering. From 2020 to 2022. He majored in information, production and systems in Waseda University and obtained a master's degree in engineering. The main research field is medical images classification based on deep learning method.



Sei-ichiro Kamata was born in Kitakyushu, Japan. He obtained the engineering doctor's from Kyushu University of degree Technology. The main research fields include image processing, pattern recognition, signal processing, etc. Professor Kamata is the member of Institute of Electrical and Electronics Engineers, Image Information Media Society. Information Processing Society, and Electronics and Information Society.