Pisco Med Publishing

Covid-19 Diagnosis Based on CT Images Through Deep Learning and Data Augmentation

Ruiwen Hu[1]*, Tianrun Wang[2], Yaxing Jing[1], Tangyu Xu[3], Feiyu Chen[2]
1. Huazhong University of Science and Technology, Wuhan 430074, China.
2. Xiamen University, Xiamen 361005, China.
3. Soochow University, Suzhou 215026, China.

Abstract: Coronavirus disease 2019(Covid-19) has made people around the world suffer. And there are many researchers make efforts on deep learning methods based on CT imgaes, but the limitation of this work is the lackage of the dataset, which is not easy to obtain. In this study, we try to use data augmentation to compensate this weakness. In the first part, we use traditional DenseNet-169, and the result shows that data augmentation can help improve the calculating speed and the accuracy. In the second part, we combine Self-trans and DenseNet-169, and the result shows that when doing data augmentation, many model performance metrics have been improved. In the third part, we use UNet++, which reaches accuracy of 0.8645. Apart from this, we think GAN and CNN may also make difference.

***Keywords**:* Covid-19; DenseNet-169; Data Augmentation; CT Images

1. Introduction

The outbreak of covid-19 is a big challenge for the whole world. It has infected more than 240 million individuals and cause almost 5 million deaths. Nowadays, the most popular way to detect covid-19 is nucleic acid testing，but it needs about 3-6 hours and costs tons of money. So people are seeking for ways to diagnose covid-19 efficiently for some clinical resource limited areas, etc. Deep learning has made a lot of extraordinary results these years and has been proved that is a good way to help detect potential patients when the result of nucleic acid testing has not come out. In the field of deep learning, Convolutional Neural Network makes a big difference in image classification. After the groundbreaking work of AlexNet[1], we have witnessed many milestones that CNN achieved. With GPU, it's possible to train a CNN model very fast and with great accuracy. When training a CNN model, the model can probably be overfitted if the dataset is too small[2]. To overcome this obstacle, data augmentation is an efficient way. In our study, we use CT images of lung as our dataset and DenseNet-168[3] as our model. We also use horizontal flipped, vertical flipped and diagonal flipped for data augmentation. And also we train UNet++ to do segmentation to detect focus of infection.

2. Literature review

Recently, artificial intelligence has been widely studied and used in medical field, like diagnosing diseases, surgical robot, virtual nursing. Inside, there are various DL framework like RNNs, AEs, GANs. One of the hottest research topics is about COVID-19. A mass of essays have been published during this epidemic, which fall into three categories: detection of COVID-19, severity assessment and infection segmentation. In this paper we will discuss the study of the COVID-19 binary classification issue using CT images.

When searching for information and technology of detecting COVID-19, we find that an army of paper with keywords of "detecting COVID-19 with CT scans" and "machine learning with covid-19 diagnosis". In their work, the majority of them used machine learning technology and deep learning or convolution neural network (CNN), which truly attained certain results. Typically, Ardakani [4] gave a detailed view of ten representative CNNs with comparing their performances in detection of COVID-19 : VGG-16, GoogleNet, ResNet-18, Xception, ResNet-101, AlexNet, MobileNet-V2 and so on. Among them, Xception and ResNet-101both reached effective results- an AUC of 0.994, which is better than the radiologist's AUC of 0.873.

Nevertheless, a large quantity of data is needed in deep learning methodology which is impossible to gain right now. The limited and unbalance data influence the performance of deep learning hugely, so many first-class scientists and doctors are trying to figure out this drawn backs-limited above. Currently, for refining the diagnose accuracy, there are lots of ways to be explored.

Firstly, using data augmentation technique, like affine transformation (rotation, translation, scaling, reflection, shearing), image mirroring. Recently Li et al. [5] reduce data scarcity by decomposing the 3D CT scan into multi-view slices as input data and integrate prior domain learning into their model. In the end, they achieved an obviously improved accuracy from 0.867 to 0.966. Zhou [6] combined several 2D models and Taylor et al. [7] utilized photometric and geometric to deal with the data-scarcity issue to enhance the effect of the model. Sameena et al. [8] used AdaBoost of decision stump trees to diminish the degree of overfitting and generalization and achieved an accuracy of 0.96, whose classifier is dynamically chosen depend on test sample' s characteristics. Kamrul et al.[9] proposed a 3D-CNN structure, which is integrated with segmentation, class-rebalancing, progressive resizing, augmentation and can expand training data through being trained on the 3D-CT patches to study the inter- and intra-slice spatial voxel information. Ozturk et al.[10] two-stage data enhancement approaches- a shallow image augmentation and the Synthetic minority over-sampling technique algorithm to solve the deficient and unbalanced data problem, which contributes to next to 10% performance.

In addition to the data, the suitable hyperparameter is also vital to determine the classification performance. Sameena et al. [8] used WOA-BAT optimization to choose hyperparameters of CNN and proved that using WOA-BAT optimized CNN performed superior to the standard CNN architecture. Priya et al.[11] offered a SqueezeNet structure network based ResNet-50, which is used for lung infection segmentation of CT and can be automatically optimized.

Transfer learning also has a satisfying performance in detecting COVID-19, which can reduce the dependence on data while achieving better accuracy. A sort model to diagnose COVID-19 is proposed by Ilyas LAHSAINI et al. [12] based on deep transfer learning and the DenseNet-201 architecture with 0.988 of accuracy. Tuan D Pham [13] has conducted a comprehensive study on the use of pre-trained CNNs for COVID-19 diagnosed. He investigated 16 pre-trained CNNs and concluded that CNNs can performed well after using several epochs training and DenseNet-201 did the best work, which can reach the highest average specificity of 0.9667. Additionally, he proved the transfer learning with using image slices not data augmentation can do better classification. Contrastive self-supervised learning and transfer learning are combined and utilized by He et al.[14] to study unbiased and useful feature representations and this framework achieved an AUC of 0.94 under limited training CTs. Maghdid et al.[15] created an image dataset containing a mass of CT and X-ray images and utilizede a pre-trained AlexNet structure on the dataset based on transfer learning and deep learning, resulting accuracy up to 0.98.

There is one more point that design fresh neural network framework. A sequence of new-style neural networks is being proposed all the time. Hong et al.[16] put forward a lightweight convolutional neural network model derived from the attention mechanism and depth-wise separable convolutions named MGMADS-CNN, which achieved accuracy of 0.9825 on CT images. M.Polsinelli et al.[17] presented a light oriented capsule network derived from the SqueezeNet and achieved 0.830 of accuracy. The light of it is to achieve a satisfying result on medium power computers, alleviate the requirement for hardware. Hu et al.[18] put forward one kind of new weakly supervised deep learning structure instead of commonly used supervised learning framework to learning from image-level label, which can reduce the dependence of manual labelling of images

Another important way is to combine several models to make diagnose. Rohit et al.[19] raised a method of integrating four pre-trained models with Sugeno fuzzy integral and achieving 98.93% accuracy. Sameena et al.[18] built their architecture by utilizing features selected from five CNN architectures. Ardakani et al.[20] nurtured twenty radiological features extracted using CT scans into five classifiers to develop the best CAD system performing in COVID-19 diagnose, which has an AUC of 0.965.

In addition to the above methods, there are still many other ones, like the “deep domain adaptation “which is used to deal with the shortage of labeled data, removing images from the majority classed to down sample.

In conclusion, they explored and optimized each process of deep learning classification by proposing new ways of their own. One same limitation of the majority studies is that their models’ performances don’t be compared with radiologists. It will be a better framework by combination of methods.

3. Research methods

This study, we used the dataset collected by[14], which reported the biggest public dataset so far of CT images for detection of Covid-19. As for the model, we chose DenseNet-169 as it has the best accuracy in the same article above. We also used Self-trans[14] like what they do because in their article, it’s proved that this is a practical way to improve the accuracy. And data augmentation is what our work wants to research if it is useful for improve the efficiency of our model.

3.1 DenseNet-169

We use pre-trained DenseNet-169 model as our model and Stochastic Gradient Descent (SGD) as our optimizer. DenseNet-169 is one of the DenseNet models families, which are designed to work image classification. There is a parameter **k** , called growth rate, which refers to the number of extra channels in each layer, or the convolution kernel of each layer. If the channel of the input feature graph is **K**, then the channel number of the **L** layer is **K+(L-1)k**, because each layer accepts the characteristic graphs of all the previous layers, so this **k** can be very small, usually 12 will have a good result. We should note that the actual meaning of this **k** is the newly extracted features. There are 1 x 1 bottleneck layers applied before each 3 x 3 convolution layer followed by transition layer. Diverse kind of Dense-Net has different **k**. So they have unequable sizes and accuracies .This is the main difference.

DenseNet exploit the potential of the network through feature re-use[1].This means that every layer can re-use the features produced by all previous layers. By doing so, it can deal with the degradation problem and vanishing gradient. DenseNet can be said to be an implicit strong supervision mode, because each layer establishes a connection with the previous layer, and the error signal can be effortlessly propagated to the former layer, so the earlier layer can obtain direct supervision from the last classification layer. At the same time, it has the characteristics of fewer parameters and higher computational efficiency. Besides, in DenseNet, it uses different levels of features, and it tends to give smoother decision boundaries. This also explains why DenseNet still performs well when training data is insufficient.

3.2 Self-trans + DenseNet-169

Transfer learning is a ideal way to mitigate data deficiency, which can use data rich source tasks to help learn target tasks with insufficient data through previously extracting useful features on big datasets and then finetuning the wights on the inefficient datasets. But there are some problems like the discrepancy in visual appearance and class labels between source data and target data, the over-parameterized pre-trained network.

For the sake of these problems, we integrate it with self-supervised learning(SSL), which is usually utilized to learn general representations without considering labels. But this time it is used to learn unbiased and powerful features without human annotations, which can get some intrinsic features and characteristics of the dataset. In other words, SSL is only based on data itself to learn meaningful results and satisfying performance. In this study, we construct some auxiliary tasks to provide self-supervision for the transfer learning process.

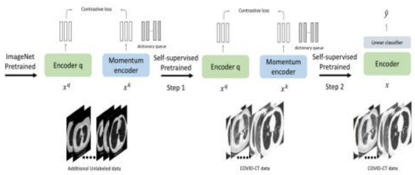

Figure 1 structure of Self-trans + DenseNet-169[31].

3.3 UNet++

3.3.1 Introduction to UNet++

UNet is a deep learning network using coder and encoder, which is widely used in medcal images segmentation problems.

To avoid the fusion of semantically dissimilar features of pure jump connections in UNet, UNet++ further strengthens these connections by introducing nested and dense jump connections, in order to reduce the semantics between encoder and decoder.

3.3.2 Our network Structure

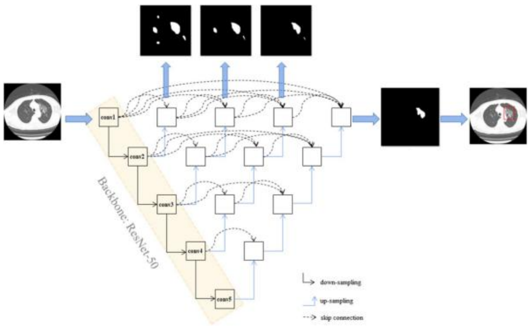

Figure 2 structure of our network

3.4 Data Augmentation

When the dataset is too small, it is likely that the model is overfitting. And data augmentation is a good way to solve this problem, because it helps the model extract more features from those images. We use horizontal flipped, vertical flipped and diagonal flipped as our methods to make the dataset larger. These ways do not change the nature of those CT images so the produced images share the same label with the original ones. We set 3 degrees as the flipped angle after lots of experiments that set different angles.

4. Result

4.1 DenseNet-169

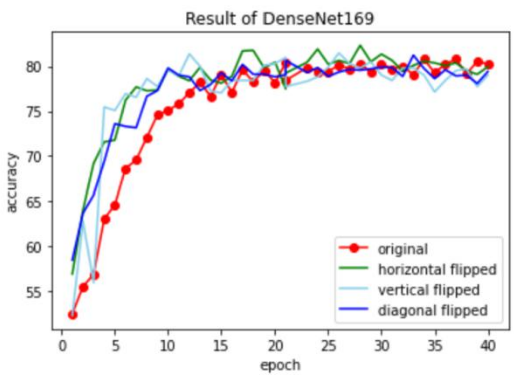

Figure 3 the result of DenseNet-169

The graph shows that data augmentation can improve before 15th epoch but its improvement is not evident after 15th epoch.

4.2 Self-trans+DenseNet-169

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1\,score = \frac{2 * precision * recall}{precision + recall}$$

We train the model with 200 epochs of train and 1 epoch of prediction. And the result is in TABLE 1. From the Table we can see that data augmentation did make a difference on improving the result, but not all data augmentation methods work. Self-trans+Horizontal flipped or Vertical flipped can improve the F1 score and accuracy apparently, but Self-trans + Diagonal flipped do not have an apparent effect, and the mix of 3 data augmentation methods make the result even worse than original results.

Table 1 The results after training the model with 200 epochs of train and 1 epoch of prediction.

	TP	TN	FN	FP	TP+FP
DenseNet-169 + self-trans	99	50	6	48	147
DenseNet-169 +self-trans + horizontal flipped	88	73	17	25	113
DenseNet-169 + self-trans+ vertical flipped	90	73	15	25	115
DenseNet-169 + self-trans+ diagonal flipped	72	85	33	13	85
DenseNet-169 + self-trans+ horizontal flipped + vertical fllipes + diagonal	82	61	23	37	119
	precision	recall	F1	acc	AUC
DenseNet-169 + self-trans	0.67	0.94	0.79	0.73	0.88
DenseNet-169 +self-trans + horizontal flipped	0.78	0.84	0.81	0.79	0.88
DenseNet-169 + self-trans+ vertical flipped	0.78	0.86	0.81	0.80	0.88
DenseNet-169 + self-trans+ diagonal flipped	0.84	0.69	0.76	0.77	0.87
DenseNet-169 + self-trans+ horizontal flipped + vertical fllipes + diagonal flipped	0.69	0.78	0.73	0.70	0.81

5. Discussion and conclusion

In our study, we chose DenseNet-169 as our model and used several data augmentation approaches to improve the accuracy of model. In the first experiment, the number of covid and noncovid images in training set are 234 and 191 respectively. Then we doubled the training set by using each of the 3 augmentation methods to produce one more image for each of the origin image in second, third and fourth experiment respectively. In the fifth experiment, we used all methods for origin images and enlarge the training set by 3 times. We also used self-trans to improve accuracy because this is a method that has been proved practical [19]. And we calculate the F1 score, test accuracy and AUC in each experiment.

As the results show, with F1 score and AUC almost the same, the test accuracy of second to fourth experiments are all 4% - 7% better compared with the benchmark accuracy, which means the accuracy became better when we only use one of the three methods of data augmentation. However, if we use all of these 3 methods in one (the fifth) experiment, all 3 measuring parameters became worse. This might because we use the same learning rate (0.001) in all 5 experiments, and it might be too large for the fifth experiment, the dataset in which is much bigger than the others. We noticed that during the training process in the fifth experiment, the train loss become oscillated after 7-8 epochs. According to previous articles [22], this is a sign of setting the learning rate too large.

The work published in [21] also used DenseNet-169 with self-trans and get an accuracy of 83% with the same dataset, which is 10% better than our first experiment (also used DenseNet-169 + self-trans), though their AUC and F1 score is almost the same as ours. We consider that one probable reason is that, according to the paper, they used unlabeled images from Lung Nodule Analysis (LUNA) for self-supervised learning, a learning form between supervised and unsupervised learning. However, we didn't do this due to the time limit.

One of the benefits of self-supervised learning is that it can solve the overfitting problem very well [23], which is also an advantage of data augmentation [24]. But our best accuracy, which was got in third experiment, is 3% worse than theirs, though we our AUC is slightly better and we share the same F1 score. To improve our result, utilizing more data augmentation methods is a potential way. We only used horizontal flipped, vertical flipped and diagonal flipped, which are some basic ways for data augmentation. In Pham's paper[25], some other methods, such as reflection, horizontal translation, vertical translation, horizontal scaling and vertical scaling are also worth trying. And for those flipping method, it's a time-consuming work to test the best flip angle. If the angle is too small, than the produced images will be too similar to the origin images, which is not good for model to extract more features. If the angle is too big, than some part of lung shadow will fade out from the produced image. We tried several angles range from 1 to 5 with horizontal flipped as the data augmentation method and the result is almost the same. Furthermore, we have also thought of using GAN[26] for data augmentation, but GAN may change the nature of images, so we are not able to know the label of generated images, which was the problem that prevented us from using GAN to enlarge our dataset.

When it comes to the model we used, we chose DenseNet-169[27] because this net has the best accuracy according to[21]. Apart from this, we also used swin transformer[28] and unet++[29] with a new data augmentation method, mixup[30], and got very good results, with the accuracy of 84% and 86.5% respectively. However, it took too much time for us to write and tune these 2 nets and so far we have not known the F1 score and AUC of these 2 nets. In future work, we will continue working on complete the test results.

For the hyperparameters in DenseNet-169, we tried to modify some of those in[21] but didn't have good results. For instance, the images are resized to 224*224 originally, we tried some other scale, like 300*300 but the result is even worse. Therefore, we used hyperparameters the same as orinal ones. And for the function transforms.Normalize we used ((0.485, 0.456, 0.406), (0.229, 0.224, 0.225)), which is calculated by the mean and standard deviation of images in ImageNet. According to an answer in Stackoverflow, if images are special, like medical images such as CT or CXR images, than it's recommended to calculate the mean and standard deviation of images in our dataset to normalize our dataset better. This is also a time-consuming work and we don't have enough time to do this. We consider doing this will probably improve our test results. What's more, we tested both SGD and Adam as our optimizer, with learning rate=0.001, momentum = 0.9 for SGD and learning rate = 0.001 for Adam. The results are almost the same. But we consider using a lower learning rate, especially for the fifth experiment as mentioned above.

References

[1] Technicolor, T. , Related, S. , Technicolor, T. , & Related, S.. ImageNet Classification with Deep Convolutional Neural Networks [50].

[2] Shorten, C., & Khoshgoftaar, T. M.. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1).

[3] Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., & Keutzer, K.. (2014). Densenet: implementing efficient convnet descriptor pyramids. Eprint Arxiv.

[4] Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., & Mohammadi, A.. (2020). Application of deep learning technique to manage covid-19 in routine clinical practice using ct images: results of 10 convolutional neural networks. Computers in Biology and Medicine, 121.

[5] Li, Z., Zhao, S., Chen, Y., Luo, F., Kang, Z., Cai, S., ... & Li, Y. (2021). A deep-learning-based framework for severity assessment of COVID-19 with CT images. Expert Systems with Applications, 185, 115616.

[6] Zhou, L., Li, Z., Zhou, J., Li, H., & Gao, X.. (2020). A rapid, accurate and machine-agnostic segmentation and quantification method for ct-based covid-19 diagnosis. IEEE Transactions on Medical Imaging, 39(8), 2638–2652.

[7] Taylor, L., & Nitschke, G. (2018). Improving Deep Learning with Generic Data Augmentation. In 2018 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1542–1547).

[8] Sameena Pathan, P.C. Siddalingaswamy, Preetham Kumar, Manohara Pai M M, Tanweer Ali, U. Rajendra Acharya, Novel ensemble of optimized CNN and dynamic selection techniques for accurate Covid-19 screening using chest CT images, Computers in Biology and Medicine, Volume 137,2021,104835, ISSN 0010-4825.

[9] Md. Kamrul Hasan, Md. Tasnim Jawad, Kazi Nasim Imtiaz Hasan, Sajal Basak Partha, Md. Masum Al Masba, Shumit Saha, Mohammad Ali Moni, COVID-19 identification from volumetric chest CT scans using a progressively resized 3D-CNN incorporating segmentation, augmentation, and class-rebalancing, Informatics in Medicine Unlocked, Volume 26,2021,100709.

[10] Ş. Öztürk, U. Özkaya, M. Barstuğan Classification of Coronavirus (COVID‐19) from X‐ray and CT images using shrunken features Int. J. Imag. Syst. Technol., 31 (2021), pp. 5-15.

[11] C. Priya, S.M.H. Sithi Shameem Fathima, N. Kirubanandasarathy, A. Valanarasid, M.H. Safana Begam, N. Aiswarya, Automatic optimized CNN based COVID-19 lung infection segmentation from CT image, Materials Today: Proceedings,2021.

[12] Ilyas LAHSAINI, Mostafa EL HABIB DAHO, Mohamed Amine CHIKH, Deep transfer learning based classification model for covid-19 using chest CT-scans, Pattern Recognition Letters, Volume 152,2021, Pages 122-128, ISSN 0167-8655.

[13] Pham TD. A comprehensive study on classification of COVID-19 on computed tomography with pretrained convolutional neural networks. Sci Rep. 2020 Oct 9;10(1):16942.

[14] He, XH., Yang, XY., Zhang, SH., Zhao, JY., Zhang, YC., Xing, E., Xie, PT., Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT ScansmedRxiv 2020.04.13.20063941.

[15] Halgurd S. Maghdid, Aras T. Asaad, Kayhan Zrar Ghafoor, Ali Safaa Sadiq, Seyedali Mirjalili, and Muhammad Khurram Khan "Diagnosing COVID-19 pneumonia from x-ray and CT images using deep learning and transfer learning algorithms", Proc. SPIE 11734, Multimodal Image Exploitation and Learning 2021, 117340E (12 April 2021).

[16] Hong, G., Chen, X., Chen, J. et al. A multi-scale gated multi-head attention depth-wise separable CNN model for recognizing COVID-19. Sci Rep 11, 18048 (2021).

[17] M. Polsinelli, L. Cinque, G. Placidi A Light CNN for Detecting COVID-19 from CT Scans of the Chest arXiv preprint arXiv:2004.12837 (2020).

[18] S. Hu et al., "Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification From CT

Images," in IEEE Access, vol. 8, pp. 118869-118883, 2020.
[19] Rohit Kundu, Pawan Kumar Singh, Seyedali Mirjalili, Ram Sarkar, COVID-19 detection from lung CT-Scans using a fuzzy integral-based CNN ensemble, Computers in Biology and Medicine, Volume 138, 2021, 104895.
[20] A.A. Ardakani, U.R. Acharya, S. Habibollahi, A. Mohammadi COVIDiag: a clinical CAD system to diagnose COVID-19 pneumonia based on CT findings Eur. Radiol., 31 (1) (2021), pp. 121-130
[21] He, XH., et al. "Sample-efficient deep learning for COVID-19 diagnosis based on CT scans." medrxiv (2020).
[22] Jason Brownlee. “How to Configure the Learning Rate When Training Deep Learning Neural Networks”.
[23] Yan, W., “A Summary of Semi Supervised Learning”.
[24] Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60 (2019).
[25] Pham, T.D. “A comprehensive study on classification of COVID-19 on computed tomography with pretrained convolutional neural networks”. Sci Rep 10, 16942 (2020).
[26] Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems 27 (2014).
[27] Iandola, F., et al. "Densenet: Implementing efficient convnet descriptor pyramids." arXiv preprint arXiv:1404.1869 (2014).
[28] Liu, Z., et al. "Swin transformer: Hierarchical vision transformer using shifted windows." arXiv preprint arXiv:2103.14030 (2021).
[29] Li, XM., et al. "H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes." IEEE transactions on medical imaging 37.12 (2018): 2663-2674.
[30] Zhang, HY., et al. "mixup: Beyond empirical risk minimization." arXiv preprint arXiv:1710.09412 (2017).
[31] A. Krizhevsky, I. Sutskever, G.E. Hinton Imagenet classification with deep convolutional neural networks Commun. ACM, 60 (6) (2017), pp. 84-90.

Note：Hu Ruiwen, Tianrun Wang, Jing Yaxing, Tangyu Xu, Feiyu Chen&These authors contributed equally to this work and should be considered co-first authors.