Image Classification of Covid-19 Pneumonia Based on Mask-EfficientNet

Xiao Xu[1], Wu Wang[1], Quanfeng Xu[1,2]
[1]School of Mathematics and Computer Science, Yunnan Minzu University, 2929 Yuehua Street, Kunming, 650500, China
[2] Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory,
Chinese Academy of Sciences, 80 Nandan Rd., Shanghai 200030, China

Novel coronavirus is a serious disease-causing virus which spreads through the air, such a highly contagious virus will cause great harm to the body after disease. After the Novel coronavirus infects someone, viruses hidden in the body will spread rapidly and widely in the population as the carrier moves, that cause catastrophic consequences. Therefore, how to quickly detect the infection of novel coronary pneumonia has become an urgent issue. Analysing the lung image of Computed Tomography (CT) is an important method to accurately detect whether people is infected by novel coronavirus in medical practice. In this paper, firstly, we use the binarized features of the novel coronary pneumonia image, and then use the features of histogram and mask as additional features, finally we design an improved network based on Efficient-Net. Through comparative experiments with other mainstream Convolutional Neural Network(CNN) networks, it is found that the model proposed in this paper reduces the parameters of the model and improves the detection accuracy.

***Index Terms*—EfficientNet, image classification, Convolutional Neural Network, Covid-19.**

I. INTRODCTION

SINCE 2020, the novel coronavirus epidemic has ravaged the world, posing a serious threat to the lives of people around the world. As of December 2022, more than 644 million people have been diagnosed with COVID-19 worldwide. With the rapid increase in the number of infected people, how to quickly diagnose the new coronary pneumonia has become very urgent. At present, the detection of novel coronaviruses is mostly based on reverse transcription polymerase chain reaction (RT-PCR). However, it still takes 4-6 hours to get accurate results, and there is also a risk of virus spreading during the process of queuing up for testing.

Analysing lung imaging is also one of the effective detection methods for COVID-19 pneumonia [1] (A lung CT image of a COVID-19 patient shown in Fig. 1. However, there is still a high threshold for diagnosing whether there is new coronary pneumonia based on CT images. Artificial intelligence can help screen the virus and COVID-19 pneumonia more efficiently [2]. In order to reduce the burden of medical staff, this paper introduces deep learning into the diagnosis of COVID-19, and proposes the Mask-EfficientNet network, which can efficiently and quickly classify whether patients have new coronary pneumonia or not.

Nowadays, the processing of medical images based on deep learning has become an important branch of computer vision. For example, picking up red blood cells images from background based on U-Net (U-Net is a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg), and tumor pathological images classification based on VGG networks (Visual Geometry Group network) and other networks. In addition, medical images are mostly black and white images with low definition. Therefore, the feature extraction network has become an important factor affecting the classification results. With the introduction of Efficient-Net, the accuracy of image classification has been greatly improved. In this paper, the Efficient-Net model will be improved based on lung CT image characteristics.

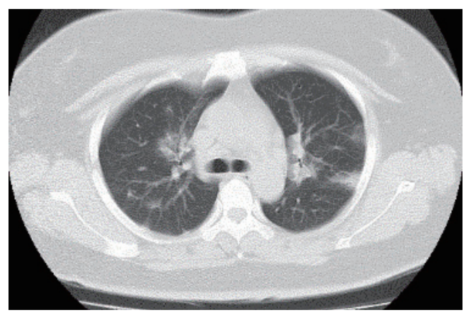

Fig. 1: A lung CT image of a COVID-19 patient

However, personal lung CT images related with personal privacy information, it is difficult get collections of these images lead to there are very few datasets containing lung CT images of COVID-19 positive patients. In addition, labeling this type of dataset requires professional medical personnel. Today, when the novel coronavirus epidemic is raging, medical personnel are busy helping novel coronavirus patients, resulting in medical personnel barely have enough time to label the lung images. Preceding problems greatly limit the development of deep learning in the field of novel coronavirus diagnosis.

In the face of extremely insufficient training data, transfer learning and semi-supervised learning can improve the situation of insufficient samples. In this paper, we adopt transfer learning and self-supervised learning to process this task, so that the pre-training network can adapt to the characteristics

Manuscript received October 26, 2022; revised December 17, 2022. Corresponding author: Wu Wang (email: wwkmyn@139.com).

of CT images, and can also reduce the overfitting caused by the small sample size.

The contributions of this paper summarized as follows:

1) Adopting CT image masks as additional features to enhance the performance of the feature extraction network and improved the classification results.
2) According to the characteristics of medical CT images, Efficient-Net was improved to adapt to lung CT images with low resolution and inconspicuous features.
3) Different transfer learning approaches were designed and applied in Mask-EfficientNet network.
4) Extensive experiments were conducted on the lung CT datasets, results showed that our model improve classification effect compare with current mainstream CNN networks.

II. Related Technologies and Related Works

With the improvement of computer computing power, computer vision has become a research hotspot today. Due to the excellent performance of deep learning, it has been widely used in people's daily life. Particularly, the deep learning network represented by the convolutional neural network [3] has achieved great success due to its natural applicability to images. A standard convolutional neural network is shown in Fig. 2. With the continuous expansion of its application field, medical image processing based on deep learning has gradually become a research hotspot.

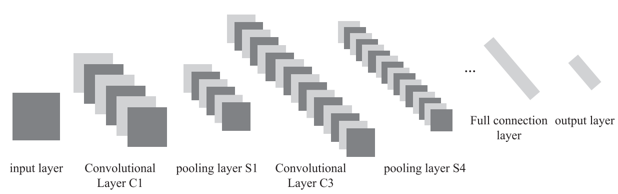

Fig. 2: The convolutional neural network

In order to effectively control the spread of the new coronavirus, doctors need to confirm whether the patient is infected with the new coronavirus or not as soon as possible. Thus, deep learning diagnostic methods based on CT image of lung are receiving widespread attention. Chaddad et al. in [4] established multiple convolutional neural network models to classify COVID-19 CT images; Ramin in [5] proposed a two-way CNN to extracting global and local features to detect and classify COVID-19 infection from CT images. Milletari et al. in [6] proposed a 3D image segmentation method neural network based on volumetric, fully convolutional. Skourt et al. in [7] proposed a high segmentation accuracy algorithm for lung CT image segmentation. Song et al. in [8] designed three neural networks and applied them to the CT image classification task and achieved high accuracy.

The above are all the contributions that scientists have made to medicine with artificial intelligence in recent years. There is no doubt that convolutional neural networks have made outstanding contributions to medicine. The following are brief introductions to the related technologies involved in this article.

A. Convolutional Neural Network and Its Convolution Processes

Convolutional Neural Network is a feedforward neural network, can be used for large-scale image processing. It contains a feature extractor composed of convolution layer and subsampling layer. In the convolutional layer of the convolutional neural network, one neuron is only connected with some neighboring neurons. A convolution layer of CNN usually contains several feature planes, and each feature plane is composed of some rectangular arranged neurons. Neurons in the same feature plane share weight parameters, and the shared weight parameter here is the convolution kernel.

Generally, the convolution kernel is initialized by the form of random decimal matrix. With the continuous training of neural network, the convolution kernel will learn more reasonable weights. Shared weights (convolution kernels) can reduce connections between layers of the network while reducing the risk of overfitting.

The convolution process is the most important feature of convolutional neural network. Convolution operation is actually a simple mathematical operation, which has two steps: one is matrix inner product, and the other is total addition of the results of inner product. The inner product of the matrix is to multiply the elements in the same position of the two multiplied matrices, and a new matrix will be obtained. The total sum is the sum of all the values of this new matrix obtained, which is the result of the convolution operation. The convolution calculation is shown in Fig. 3.

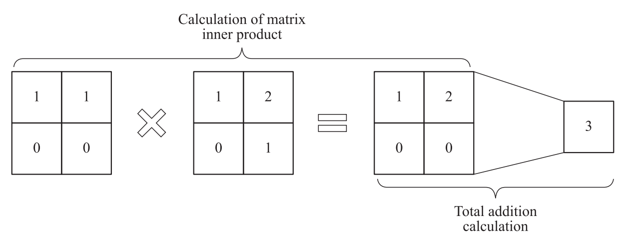

Fig. 3: Illustration of convolution calculation

In the actual convolution process, not only need an input image but also a convolution kernel. The value of the convolution kernel is artificially set and can also be adjusted. In a specific convolutional neural network, the value of the convolution kernel refers to the parameters of the network. Generally, training the convolution network refers to the adjustment the parameters of these convolution kernels. In practice, convolution kernel chose are 3*3, 5*5, and 1*1.

The input of convolution calculation can be not only image, but also other two-dimensional matrix information. Because both feature images and convolution kernels can be represented by matrices, they can perform convolution calculations. The convolution kernel first performs a convolution operation on the first area of the feature image, and the result of this convolution calculation will be used as a point on the output feature image, as shown in the Fig. 4.

The sliding process of convolution kernel on feature image operates as follows: the result of each convolution calculation be used as a point of the output feature image, and after the

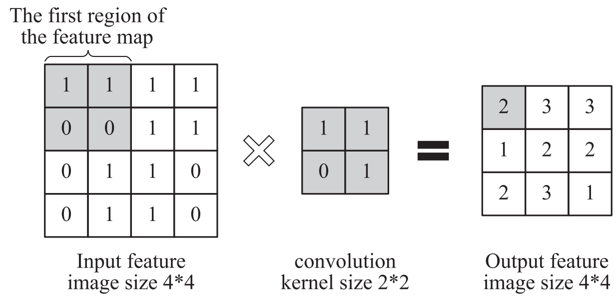

Fig. 4: The first convolution calculation of feature image

convolution process of the feature image, a new feature image will be output, which requires multiple convolution calculations between the convolution kernel. The Fig. 5 illustrates the sliding process with step 1.

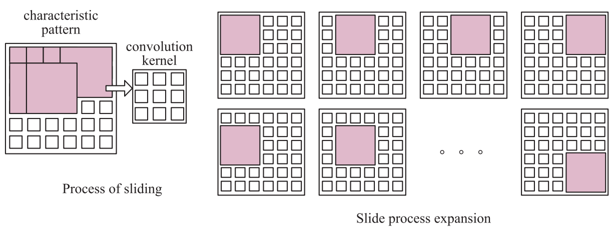

Fig. 5: The sliding process of convolution kernel in feature image with step 1

B. Classification Network Model

To better understand the classification networks and prove the advantages of the model presented here, I refer to several excellent classification networks in recent years, namely VGG, ResNet, DenseNet and EfficientNet. After understanding and summarizing the characteristics of each network, finally, the Mask-EffieceitNet is designed according to the particularity of CT image based on the EfficientNet network. In turn, we present an overview of several classic classification models.

1) VGG Network

Due to the Visual Geometry Group (VGG) network [9] won the second place in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 competition classification project, VGG network has received extensive attention and has become the current mainstream feature extraction network.

In order to avoid the feature map becoming smaller and smaller during the convolution process, VGG used a large number of 3*3 convolution kernels instead of large convolution kernels, and fills the edges, so that the volume of product does not reduce the image size. The authors believe that continuous use of 3*3 convolution kernels can replace a largesized convolution kernel under the condition that the receptive field is similar, and the continuous using of small convolution kernels can significantly reduce network parameters compared to using a single large convolution kernel.

The structure of CNN network is shown in Fig. 6.

The characteristic of this network is that its model has excellent generalization ability. It could express complex data with less depth, so that it could to achieve good generalization.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Fig. 6: The structure of CNN network

2) ResNet Network

In 2015, He et al. proposed the ResNet network [10] and won the first place in the ImageNet classification task. Due to the excellent performance of ResNet and its excellent convergence speed, there has been an uproar in the field of neural networks. ResNet is still one of the backbones of most tasks so far.

Previous studies indicated that as the number of neural network layers increases, the capability of feature extraction also increases. However, the gradient back-propagation mechanism caused the gradient disappears and the gradient explodes, resulting in a situation where the more the number of network layers, the worse the effect. Therefore, He et al. proposed a residual mechanism to solve the problems that the number of network layers cannot be increased and the gradient disappears and the gradient explodes. The residual network model they proposed was composed of several residual blocks.

The single residual block function mapping formula and structure diagram as shown in Equation (1) and Fig. 7, respectively.

$$g(x) = x + f(x) \tag{1}$$

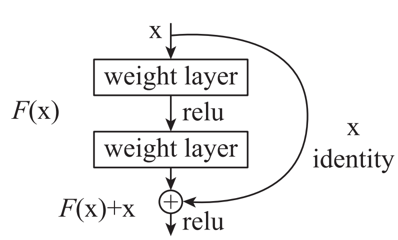

Fig. 7: Residual block

ResNet 18, ResNet 50, and ResNet 152 with different network depths are obtained by continuously stacking several residual blocks.

Compared with the VGG network presented in II A, the ResNet network replaces VGG-16 with ResNet 101, which improves the $\mathrm{mAP}(\%)$ of the COCO dataset and PASCAL VOC dataset. Furthermore, the ResNet network has good generalization ability on other classification tasks.

3) DenseNet

Inspired by the idea of ResNet algorithm, Huang et al. in [11] proposed DenseNet. Commonly, the input of the last layer of the traditional neural network often only depends on the output of the previous layer. Whereas, the authors introduced a dense connection blocks in DenseNet, in such dense connection block, the output of the first N layers of the network in the block was input to the $\mathrm{N}+1$ layer, so that the connections between layers were established. Design of dense connections could enhance the information transmission in the network, and the reusability of features was greatly improved, the amount of parameters were reduced, too.

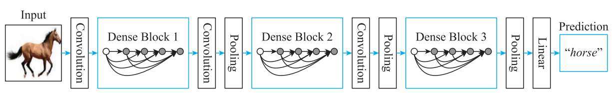

Fig. 8: DenseNet

As shown in Fig. 8, the input of the lth layer is not only related to the $(l-1)$th layer, the lth layer receives all the feature maps of the previous layers as input, namely

$$x_l = H_l\left(\left[x_0, x_1, \cdots, x_{l-1}\right]\right) \tag{2}$$

The network introduces direct connections between any two layers with the same feature map size. While following simple connection rules, DenseNets naturally introduced properties of identity mapping, deep supervision, and diverse depths. They allow features to be reused across the network and thus could learn more compact, more accurate models.

4) Efficient-Net

Tan et al. in [12] found that deepening the network depth, widening the network width, and increasing the resolution of the input image can improve the performance of the neural network. In contrastly, if the above three methods are mixed, namely, deepening the network depth, widening the network width, and increasing the resolution of the image, simultaneously. The performance of the neural network would be ineffective. Thus, the authors tried to find a balance between the depth, width and image resolution of the convolutional network. Finally, the authors proposed a compound scaling method, known as Efficient-Net, which find this balance and greatly improved the performance of the neural network.

Efficient-Net uses several MBConv modules in series to build the network. Its architecture is shown in Fig. 9.

The MBConv module enhances its cross-channel feature fusion capability by combining residual idea and attention mechanism. The network structure of MBConv module is shown in Fig. 10.

The most contribution of Efficient-Net is adopting an efficient compound scaling method that enables it to easily scale the baseline ConvNet to any target resource constraints in a more principled manner while maintaining model efficiency, solving the trade-off between network width, depth and image resolution.

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3x3	224×224	32	1
2	MBConv1,k3x3	112×112	16	1
3	MBConv6,k3x3	112×112	24	2
4	MBConv6,k5x5	56×56	40	2
5	MBConv6,k3x3	28×28	80	3
6	MBConv6,k5x5	14×14	112	3
7	MBConv6,k5x5	14×14	192	4
8	MBConv6,k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

Fig. 9: Efficient-Net

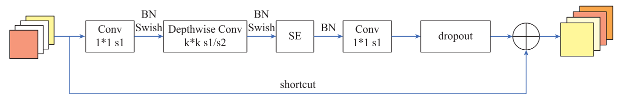

Fig. 10: MBConv block

Efficient-net's greatest contribution is the use of an efficient composite scaling method that enables it to easily scale baseline Conv Net to any target resource constraint in a more principled manner while maintaining model efficiency, resolving tradeoff resolutions between network width, depth, and image.

EfficientNet shows the best performance among the corresponding classification networks, but it does not achieve the desired accuracy due to the small data set of COVID-19 patient lung CT images and their relatively fuzzy features. Therefore, this paper optimizes and innovates the fuzzy problem of small data sets and CT images.

C. COVID-19 Lung CT Image Classification

1) COVID-19 CT Images of Lung

The main manifestation of COVID-19 in lung CT is ground-glass opacity scattered throughout the lung. This means that the alveoli are filled with fluid and appear as gray shading on CT images. In more severe or rapidly progressing lung infections, more and more fluid accumulates in the lobes, causing the ground-glass degeneration of the lungs to gradually progress to solid white "lung consolidations ". As shown in Fig. 11. Eventually, CT showed a "gravel road sign" due to swelling of the interstitial wall of the lung lobules. The alveolar wall thickens, splitting the blurred ground-glass shadows like white lines, and accumulating in the lungs like many irregularly shaped stones are laid on the road, hence the name "gravel road sign".

2) Medical CT Image Classification

Image classification is that given an image, the computer runs a deep learning algorithm to determine the category of the image.

Image classification includes classifying natural images and classifying unnatural images. The natural image is the photo taken by the photosensitive unit of the camera, and the unnatural image is the radiographic imaging, functional imaging, magnetic resonance imaging, ultrasound imaging, etc. CT images are unnatural images.

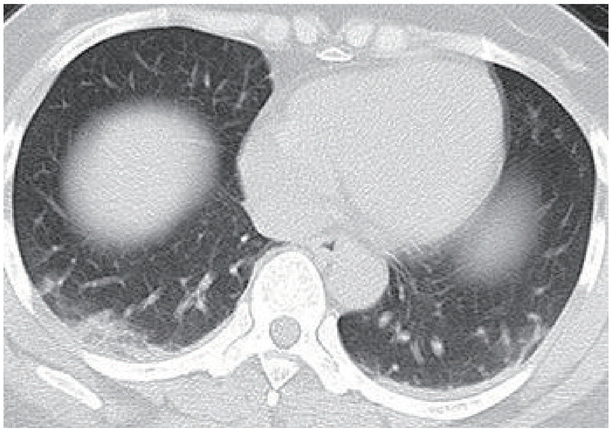

Fig. 11: Consolidation of the lung

Most of the traditional image classification is to classify natural images, as the continuous development of machine learning, the classification of unnatural images is also feasible. Hiroyuk et al. studied the relationship between the number of CT images [13] (including contrast-enhanced data used to create a classification model) and achieved high accuracy in classifying. Cho et al. also adopted deep learning to classify CT imaging.

D. Lung CT Image Diagnosis of COVID-19 Based on Deep Learning

The basic steps of the CNN-based image classification algorithm are as follows:

1) The input image is preprocessed to meet the input requirements for the convolutional neural network.
2) Feature extraction is performed using a convolutional neural network as the backbone. Common feature extraction networks such as AlexNet, VGG, ResNet.
3) Using the features output by the feature extraction network as input, use the fully connected layer for classification prediction.

E. Transfer Learning

The learning of convolutional neural network requires a large amount of sample data, and the number of COVID-19 lung CT images is small. To solve this problem, this paper introduces transfer learning to train the network.

Hussain et al. achieved high accuracy and high efficiency through the study of transfer learning on a new image dataset [14]. Transfer learning usually uses standard network architecture for training on large-scale datasets (such as ImageNet), and then fine-tunes the trained network structure and pre-training weights to transfer to the target task, expecting to enhance the model feature extraction ability and generalization. In the medical field, transfer learning is also widely used in medical image classification and recognition tasks, such as tumor classification and skin disease diagnosis using transfer learning.

F. Attention Mechanism

Currently, deep learning methods are used to analyze and process medical CT images. Due to the high homogeneity of medical images of human organs, the pathology of medical images can only be distinguished by small differences, so classification is a challenge. The attention mechanism can solve this problem very well which focuses on local information. The basic idea of this mechanism is to ignore irrelevant information and focus on key information. Give a larger weight to the objects of interest, and give a smaller weight to the objects of no concern. For example, when recognizing an image, instead of giving the same degree of attention to all the information in the image, it fixes more attention on a local area of image. Wang et al. in [15] designed a collaborative network with a backbone attention mechanism to extract medical image features and achieved good results. This paper will also introduce an attention mechanism for lung medical CT images to achieve better results.

III. Mask-EfficientNet

In order to improve the accuracy of diagnosing COVID-19 based on lung CT images, in view of the particularity of lung CT images of COVID-19, we have done a series of data preprocessing for theses images, including normalization, data expansion and enhancement, and image feature extraction. Aftering preceding processing, we modify the EfficientNet, one of the best performing networks in the current classification network, Mask-EfficientNet is finally proposed. The following is a detailed introduction.

A. Data Preprocessing

For medical imaging, different scanning equipment and the physique of the scanned person could impact on the image. These factors can conduct a huge impact on the neural network, making it difficult to converge. To this end, this paper firstly normalize the input image to eliminate the influence of uneven data distribution while retaining the grayscale difference with judgment value. Commonly used normalization methods are: maximum and minimum normalization, zero mean normalization. This paper will use the zero-mean normalization method.

Due to reasons such as personal privacy mentioned in section I, there are too few patient's lungs CT images. This paper will consider using a variety of data augmentation methods to increase the amount of data. The specific method is as follows:

1) Apply affine transformation to the image, and perform a certain range of translation, rotation, and left-right mirror transformation on the image.
2) On the premise of ensuring the quality of the input image, artificial random disturbance is added to the image, such as superimposing a certain range of Gaussian noise on the image.

In addition, in this paper, a variety of network transfer learning methods are used to evaluate the network, the purpose is to pre-train on other datasets so that our model can be trained on the same number of datasets it came out better. Since ImageNet has a large number of images and good target diversity, so the current mainstream pre-training model is based on the ImageNet dataset. However, because CT images are far from traditional images, our experiment first uses the ImageNet dataset pre-training model and the C2L proposed by Zhou in [16] as the pre-training model. Two kinds of pre-training models were used for experiments on the ResNet 18 network. The results found that in the training process, the convergence speed of the pre-training model using C2L was significantly better than that based on the ImageNet-based pre-training model. Therefore, this paper finally uses the model C2L proposed by Zhou as a pretrained model.

Because medical CT images are relatively blurry, in order to better extract features from medical lung CT images, this paper expands the input dimension of the data and supplements the original input images. As described in the section II-C, CT images of the lungs with COVID-19 clearly show localized white solids or gray shading, which contrasts sharply with the black background of the CT images. Considering this characteristic of the image, if the image binarization method is adopted, the image can be more focused on the diseased area, thereby implicitly providing supervision information to the network. Due to the influence of various factors, the pure medical image has a lot of interference information, which will cause adverse effects on the results. If the image mask information is added to the neural network, that is, the similarity image matching method is used to detect and extract the structural features similar to the mask in the image, which will enable the neural network to pay more attention to the characteristics of the lesion and reduce the disturbance caused by local pixel values.

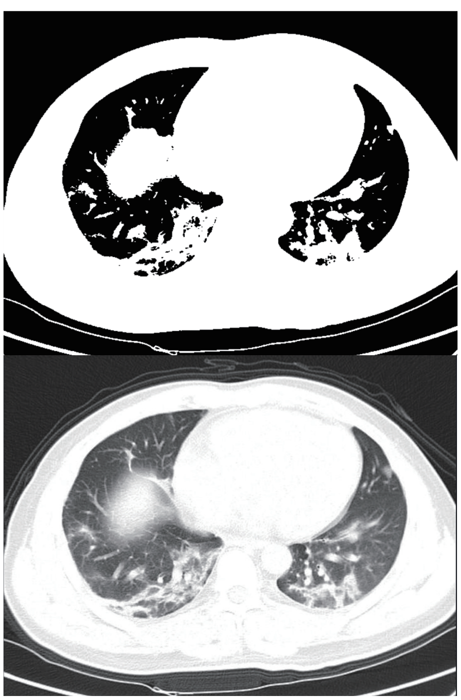

Fig. 12: COVID-19 lung CT mask image vs original image

Through analysis, in order to optimize the extraction of structural features, this paper will use image binarization to generate image masks as additional supplementary information of the original image. The mask image and original image are shown in Fig. 12.

B. The Mask-EfficientNet network and its framework

EfficientNet is one of the best performing neural networks, since there is not much data on CT images of the lungs of COVID-19 patients, and these CT images are relatively blurry, Efficient-Net cannot achieve good accuracy. Therefore, this paper optimizes and innovates the problem of small datasets and CT image blurring. Mask-EfficientNet is proposed by taking the mask image as the input image and incorporating the attention mechanism.

Mask-EfficientNet features the introduction of the attentional mechanism according to the special features of CT images, which differ from natural images. In this paper, an attention mechanism, extrusion activation module, is added to the make-layer, so that the lung image part of CT can receive more attention weight. It works by the principle that the original image is EfficientNet model after the make-layer, and the mask image is splicing together with the branches of the original image after the first make-layer.

The specific structure of its network is shown in Fig. 13.

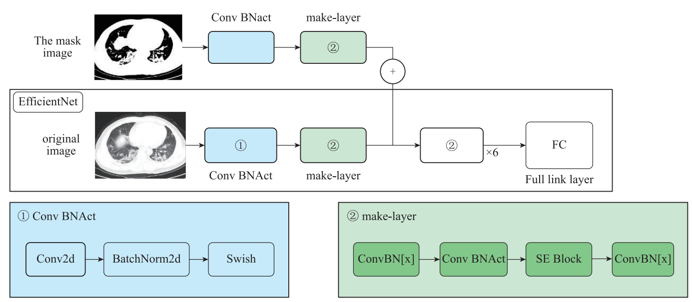

Fig. 13: Mask-EfficientNet framework

This way not only improves the characteristics of lung CT images, but also retains the advantages of EfficientNet, so that the classification can achieve the same high performance as EfficientNet while accurately identifying.

IV. EXPRIMENTS

Extensive experiments were conducted to verify the high efficiency of our proposed Mask-EfficientNet. The experiments used the COVID-19 CT image dataset of the University of California, USA, which contains 349 positive new coronary pneumonia lung images and 397 negative images. These images were classified by VGG, ResNet, DenseNet, EfficientNet introduced in section II-B, and Mask-EfficientNet proposed in this paper, that is, to identify which of these images are COVID-19 positive and which are COVID-19 negative.

To evaluate the performance of our network model, this paper adopts the following metrics,

$$P = \frac{TP}{TP + FP} \tag{3}$$

The precision rate P indicates the accuracy rate of predicting whether or not to be infected with pneumonia at the time of testing, TP is total number of samples predicted to be infected with pneumonia and predicted correctly, FP is total number of samples in which pneumonia was predicted but incorrectly predicted.

$$R = \frac{TP}{TP + FN} \tag{4}$$

The recall rate R indicates all the samples infected with pneumonia have been predicted, FN is the number of samples infected with pneumonia was not predicted.

$$F1 = \frac{2PR}{R + P} \tag{5}$$

F1 comprehensively represents the value of P and R.

$$acc = \frac{TP + TN}{TP + FN + TN + FP} \tag{6}$$

The accuracy rate acc indicates how many of the total samples are correctly predicted, TP+FN+TN+FP is the number of samples in the population.

As acc is the most valuable data, Mask-EfficientNet is calculated and compared with acc for other mainstream network models. Experimental results show that the accuracy of Mask-EfficientNet is indeed the best.

The classification accuracy of current mainstream and Mask-EfficientNet is summarized in Table I.

TABLE I: Experimental performance of baseline

Method	Accuracy
VGG	76.3
ResNet 50	84.6
DenseNet-121	89.4
EfficientNet	91.2
Mask-EfficientNet	92.1

In addition, we also calculated the P, R, and F1 of Mask-EfficientNet respectively, the results are: the precision rate P is 92.2 %, the recall rate R is 79 %, and the F1 rate is 85%.

V. CONCLUSION

In order for computer to diagnose whether the novel coronary pneumonia is positive or negative based on lung CT images, we conducted a series of data preprocessing for lung CT medical images, and improved EfficientNet to make it more accurate for lung CT medical images, proposed Mask-EfficientNet image classification network. A large number of experimental results showed that Mask-EfficientNet had higher resolution and accuracy for COVID-19 infection diagnosing through lung CT image than other classification networks. In the fight against COVID-19, the model proposed in this paper can be used for rapid and accurate detection of COVID-19 infection.

REFERENCES

[1] Jeffrey P Kanne. Chest ct findings in 2019 novel coronavirus (2019-ncov) infections from wuhan, china: key points for the radiologist. *Radiology*, 2020.

[2] Muhammad EH Chowdhury, Tawsifur Rahman, Amith Khandakar, Rashid Mazhar, Muhammad Abdul Kadir, Zaid Bin Mahbub, Khandakar Reajul Islam, Muhammad Salman Khan, Atif Iqbal, and Nasser Al Emadi. Can ai help in screening viral and covid-19 pneumonia? *IEEE Access*, 8:132665–132676, 2020.

[3] Sakshi Indolia, Anil Kumar Goswami, Surya Prakesh Mishra, and Pooja Asopa. Conceptual understanding of convolutional neural network-a deep learning approach. *Procedia computer science*, 132:679–688, 2018.

[4] Ahmad Chaddad, Lama Hassan, and Christian Desrosiers. Deep cnn models for predicting covid-19 in ct and x-ray images. *Journal of medical imaging*, 8(S1):014502, 2021.

[5] Ramin Ranjbarzadeh, Saeid Jafarzadeh Ghoushchi, Malika Bendechache, Amir Amirabadi, Mohd Nizam Ab Rahman, Soroush Baseri Saadi, Amirhossein Aghamohammadi, and Mersedeh Kooshki Forooshani. Lung infection segmentation for covid-19 pneumonia based on a cascade convolutional network from ct images. *BioMed Research International*, 2021, 2021.

[6] Milletari Fausto, Navab Nassir, and Ahmadi Seyed-Ahmad. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016.

[7] Skourt Brahim Ait, El Hassani Abdelhamid, and Aicha Majda. Lung ct image segmentation using deep neural networks. *Procedia Computer Science*, 127:109–113, 2018.

[8] Qingzeng Song, Lei Zhao, and Xingke Luo. Using deep learning for classification of lung nodules on computed tomography images. *Journal of healthcare engineering*, 2017, 2017.

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

[10] 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.

[11] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[12] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.

[13] Sugimori and Hiroyuki. Classification of computed tomography images in different slice positions using deep learning. *Journal of healthcare engineering*, 2018, 2018.

[14] Hussain Mahbub, Bird Jordan J, and Faria Diego R. A study on cnn transfer learning for image classification. In *UK Workshop on computational Intelligence*, pages 191–202. Springer, 2018.

[15] Shanshan Wang, Tao Zhang, and Fei Li. A synergic neural network for medical image classification based on attention mechanism. In *2022 Asia Conference on Algorithms, Computing and Machine Learning (CACML)*, pages 82–87. IEEE, 2022.

[16] Hongyu Zhou, Shuang Yu, Bian Cheng, Yifan Hu, Kai Ma, and Yefeng Zheng. Comparing to learn: Surpassing imagenet pretraining on radiographs by comparing image representations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 398–407. Springer, 2020.

Xiao Xu is Master's program at Yunnan Minzu University. His research interests include machine learning and computer vision.

Wu Wang received his B.S. and M.S. degrees from Yunnan University, China , in 2003 and 2007, respectively. Received his Ph.D. degree from Future University Hakodate, Japan, in 2018. He is currently an associate professor at the School of Mathematics and Computer Science at Yunnan Minzu University. His research interest includes ad hoc networks, neural network, and network security.

Quanfeng Xu is under joint Master's program at Yunnan Minzu University and Shanghai Astronomical Observatory of the Chinese Academy of Sciences. His research interests include machine learning and computer vision.