



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT images

Ju Zhang^a, Lunduan Yu^b, Decheng Chen^b, Weidong Pan^b, Chao Shi^c, Yan Niu^b, Xinwei Yao^b, Xiaobin Xu^d, Yun Cheng^{d,*}

^a Zhijiang College of Zhejiang University of Technology, Shaoxing 312030, China

^b College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

^c College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

^d Department of Medical Imaging, Zhejiang Hospital, Hangzhou 310013, China

ARTICLE INFO

Keywords:

COVID-19

Attention

Medical image segmentation

Generative Adversarial Network

Deep learning

ABSTRACT

As the COVID-19 virus spreads around the world, testing and screening of patients have become a headache for governments. With the accumulation of clinical diagnostic data, the imaging big data features of COVID-19 are gradually clear, and CT imaging diagnosis results become more important. To obtain clear lesion information from the CT images of patients' lungs is helpful for doctors to adopt effective medical methods, and at the same time, is helpful to screen the patients with real infection. Deep learning image segmentation is widely used in the field of medical image segmentation. However, there are some challenges in using deep learning to segment the lung lesions of COVID-19 patients. Since image segmentation requires the labeling of lesion information on a pixel by pixel basis, most professional radiologists need to screen and diagnose patients on the front line, and they do not have enough energy to label a large amount of image data. In this paper, an improved Dense GAN to expand data set is developed, and a multi-layer attention mechanism method, combined with U-Net's COVID-19 pulmonary CT image segmentation, is proposed. The experimental results showed that the segmentation method proposed in this paper improved the segmentation accuracy of COVID-19 pulmonary medical CT image by comparing with other image segmentation methods.

1. Introduction

COVID-19 virus, which has been rapidly spreading since the end of 2019, has brought severe pneumonia of pandemic severity to the world [1,2]. According to official data from WHO, the internationally recognized mortality of COVID-19 is about 2–3%, and that of influenza is 0.1%. That is to say, novel Coronavirus infection has a mortality 20 times or higher than that of influenza. By November 25, 2020, the cumulative global death toll has reached 1,410,451, enough to illustrate the horror of the pandemic.

The lesson of this pandemic is that early detection of new pneumonia can not only give patients a chance of survival, but also prevent hospitals from overworking. In the latest guidelines issued by the Chinese government for the diagnosis and treatment of COVID-19 pneumonia (seventh edition), the diagnosis of COVID-19 must be confirmed by reverse transcription polymerase chain reaction (RT-PCR) or gene sequencing, but RT-PCR detection has disadvantages such as high false

negative rate and slow detection speed [3]. Therefore, there is an urgent need for a way to quickly detect lung abnormalities in medical judgment. The application of deep learning technology in medical diagnosis can improve the detection rate and efficiency of disease, and it has achieved great success in the field of medical image recognition. In order to diagnose lung cancer, lung tumor and lung nodules, many scholars have studied the deep learning based recognition methods for lung CT image, and it has been proved that CT image recognition is very useful for the diagnosis of lung diseases [4–6]. For today's large-scale infection COVID-19, there have been classification methods based on the deep learning and lung CT data analysis [7,8], and clinical practice has shown that chest CT is an effective screening strategy because it can detect the standard features between most COVID-19 cases, showing ground glass opacity in the lungs in early patients [9].

Available data suggest that 30–70% of early infected patients have negative RT-PCR results, which can be caused by a number of factors, including the fact that the viral load in the blood is not sufficient to

* Corresponding author.

E-mail address: 962110508@qq.com (Y. Cheng).

<https://doi.org/10.1016/j.bspc.2021.102901>

Received 15 March 2021; Received in revised form 11 June 2021; Accepted 19 June 2021

Available online 23 June 2021

1746-8094/© 2021 Published by Elsevier Ltd.

detect the virus in the early stage of infection and the use of the wrong detection kit. However, to date, a large study with more than 1000 cases has shown that the initial sensitivity of chest CT is over 95%, which means that at most 1 out of 20 cases is missed [8]. Applying the deep learning based lung CT imaging methods to large-scale screening for COVID-19 can effectively improve doctors' detection efficiency, and thus improve patient survival rate.

Although the deep learning based lung CT image screening method has the advantages of high efficiency and high accuracy, the following two challenges still exist to affect its performance:

1. Since pixel-level annotations are required in the segmentation of lung CT images, obtaining such labels often requires professional radiologists to spend a lot of time marking pixel by pixel. In this special period, few doctors have the time to do so. Therefore, it is difficult for us to collect enough marker data for deep learning model training.
2. Changes in the texture, size and location of lesions in lung CT images pose a certain challenge to the detection of COVID-19. For example, ground-glass boundaries as shown in Fig. 1 have a blurry appearance and low contrast, making them difficult to recognize and segment.

Compared with prediction or classification, segmentation helps doctors locate the lesions and observe the pulmonary lesions of patients infected with Covid19, laying a better foundation for subsequent pathological analysis.

The contributions of the proposed method are as follows:

- A novel Dense GAN was developed to generate higher quality images to expand the limited medical image segmentation data set of COVID-19, so as to improve the generalization and accuracy of the model.
- The U-NET network combined with multi-layer attention mechanism was proposed to improve the accuracy of lesion segmentation.
- Experimental and Ablation studies are made for the Dense GAN and multi-layer attention based method, and comparisons with the current-state of the art are presented, showing the effectiveness and strength of the proposed method.

The code and dataset of this manuscript have been released at: <https://github.com/YLDJack/DenseGAN>.



(a) COVID-19 Lung CT image



(b) Ground Truth

Fig. 1. CT scan of the lung of COVID-19 patient.

2. Related work

2.1. Generative Adversarial Network (GAN) based medical image generation

In recent years, GAN based unsupervised generation framework has been widely studied in medical image processing. GAN can provide data enhancement methods to solve the problem of most available data set fragmentation. Han et al. [11] proposed a three-dimensional multi-condition GAN (MCGAN) that can generate real and diverse nodules and place them naturally on the lung CT image, which can reduce the influence of different lesion locations, sizes and attenuation. Existing studies have shown that unsupervised GAN reduces cumbersome data labeling tasks and has considerable application prospect. It has been applied in retinal images [12], skin lesion images [13], pulmonary nodules [3] and other scenes in the medical field. GAN image of pulmonary nodules can improve the diagnostic efficiency of radiologists. Zhao C et al. [14] proposed a patched 3D U-NET and context convolutional neural network (CNN), which can automatically segment and classify pulmonary nodules, and applied GAN enhancement model training to improve the performance of image extraction of nodules.

2.2. Deep learning detection methods for COVID-19

Existing work includes block-based methods [14], section-based methods [15] and 3D CT-based methods [17], and all of them aim to distinguish whether patients are infected with COVID-19 or not. Pathak et al. [18] proposed a method to classify COVID-19 infected patients using deep migration learning technology, and used the top-2 smoothing loss function with cost-sensitive attributes to deal with the problems of noise and imbalance in the COVID-19 data set. The model achieved efficient classification effect. In the rapid diagnostic method proposed by Ardakani et al. [19], infected patients were classified by ten well-known pre-trained CNNs, and the neural network of the data set was optimized by using transfer learning to distinguish COVID-19 from other pneumonia and viral pneumonia more accurately. The study proved that RESnet-101 and The Xception network were the best.

Although COVID-19 auxiliary detection algorithms based on deep learning CT classification have made great progress in classification, segmentation and quantification, existing methods still face some challenges that cannot be ignored, as described below.

- 1) More work focuses on 3D CT detection, but the 3D model is more complex than the 2D model, and the resolution lacks the interpretability of the results.

2) The existing deep learning methods mainly focus on the classification task of infected patients, while the segmentation and quantification of images require a lot of manual operation, which is not conducive to the application in the epidemic period [20].

2.3. Attention mechanism based CT detection

In recent years, most studies on the combination of deep learning and visual attention mechanism have focused on the use of a mask to form the attention mechanism. The principle of mask is to identify the key features in the image data through another layer of new weight. Through learning and training, the deep neural network can learn the areas that need to be paid attention to in each new image, thus forming attention. SE-Net [21] explicitly models the interdependencies between feature channels. Therefore, a certain number of parameters can be reduced and a small amount of improvement can be brought about. RAN [22] proposed residual attentional learning, in which not only the feature tensor after the mask is used as the input of the next layer, but also the feature tensor before the mask is used as the input of the next layer. In this case, more features can be obtained and key features can be better paid attention to. RAN adopts a bottom-up, top-down, forward attention mechanism to extract the attention areas of the model in just one forward process, making model training easier. In addition, Han [23] et al. have applied the attentional learning mechanism to the accurate screening of 3D lung CT slices of COVID-19 in their paper, which has improved the interpretation of the results and reduced the heavy workload of radiologists.

Previous studies have shown that the attention mechanism in computer vision can enable the system to focus on the area of interest, and its application in computer-aided detection can significantly improve performance. Therefore, it is of great significance to continue to study the application of attention mechanism in disease diagnosis. A targeted multi-layer attention mechanism aiming at the early onset characteristics of COVID-19 is proposed in this paper, which effectively improves the detection rate and accuracy of lesion segmentation in COVID-19

patients.

3. Dense GAN and multilayer attention based lesion segmentation

In this section, an improved Dense GAN to expand data set is developed, and a multi-layer attention mechanism method, combined with U-Net's for COVID-19 pulmonary CT image segmentation is proposed. The proposed network structure is shown in Fig. 2.

Main aspects of the proposed Dense GAN and multilayer attention based lesion segmentation method are as follows.

3.1. Dense Generative Adversarial Network (DGAN)

3.1.1. Generative Adversarial Network

The generating countermeasure network is actually composed of a generator and a discriminator. The generator tries to create a randomly generated output (such as a face image, a digital image, etc.), and the discriminator tries to distinguish the output from the real sample. The back propagation algorithm and Dropout algorithm are used to train the two models. The expectation is that as the two networks compete with each other, the network will continue to optimize and the output will get better and better, resulting in a generator network that can produce realistic output.

Generator network: It captures the distribution of sample data and generates a sample similar to the real training data with random noise Z obeying a certain distribution (uniform distribution, Gaussian distribution, etc.).

Discriminator network: mix real data and generated data and input them to the discriminator to estimate whether the sample belongs to real data or generated data. If it comes from real data, high probability will be output, otherwise low probability will be output.

The objective function of GAN is expressed by mathematical formula, as shown in Equation (3-1)

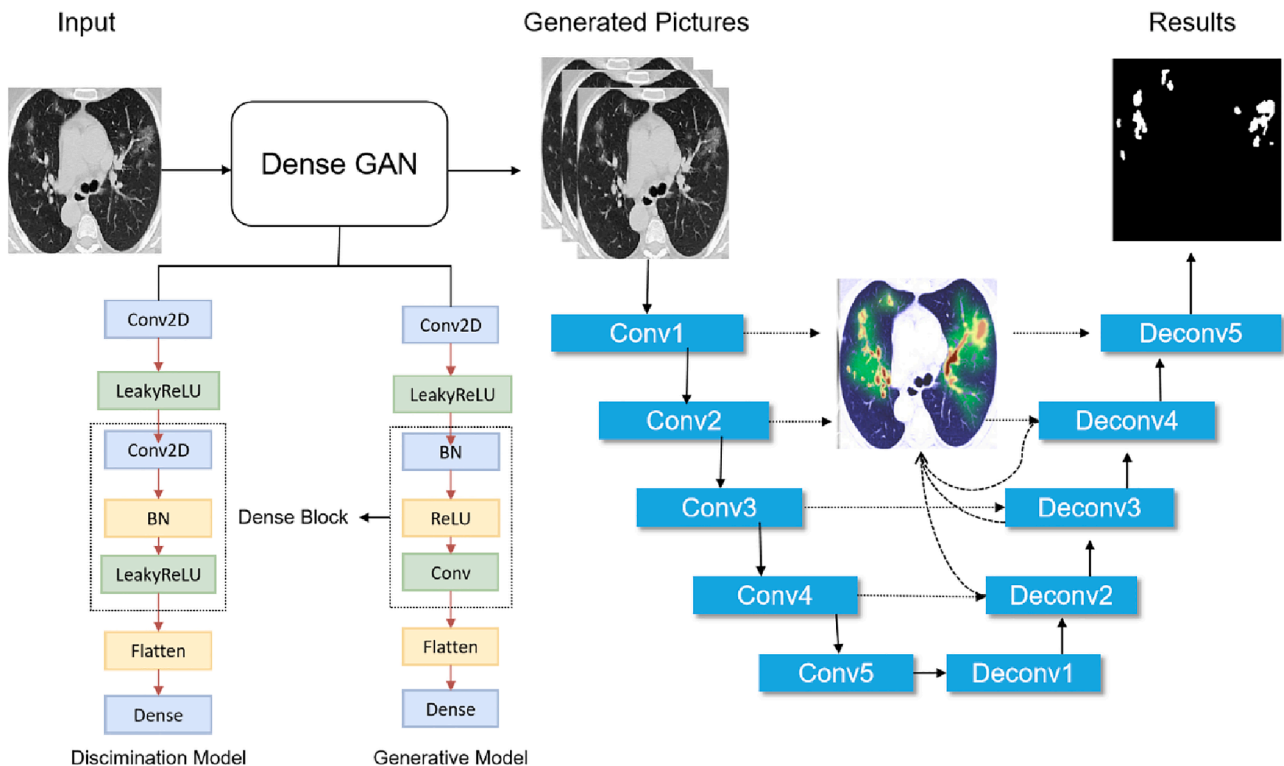


Fig. 2. The overall network structure.

$$\min_G \max_D V(D, G) = E_{x \sim P_{data(x)}} [\log(D(x))] + E_{z \sim P_{z(z)}} [\log(1 - D(G(z)))] \quad (3-1)$$

The above formula contains two optimization processes, and the disassembly is divided into two parts:

Optimize D, see Formula (3-2)

$$\max_D V(D, G) = E_{x \sim P_{data(x)}} [\log(D(x))] + E_{z \sim P_{z(z)}} [\log(1 - D(G(z)))] \quad (3-2)$$

When optimizing the discriminator D network, there is no need to operate on the generator. G(z) is a false sample from the generator. For the first term of the formula, if it is a real sample x then the output is close to 1 and the bigger the better, whereas if it is a generated sample G(z) the output is close to 0 and the smaller the better.

Optimize G, see Formula (3-3)

$$\min_G V(D, G) = E_{z \sim P_{z(z)}} [\log(1 - D(G(z)))] \quad (3-3)$$

There is no need to operate on the discriminator when optimizing the generator G network. In this case, you want the tag of the generated sample G(z) to be 1, so the bigger D(G(z)) the better.

According to the objective function, the discriminator's loss function can be obtained as shown in the formula (3-4)

$$-((1 - y)\log(1 - D(G(z))) + y\log D(x)) \quad (3-4)$$

The generator's loss function is shown in the formula (3-5)

$$(1 - y)\log(1 - D(G(z))) \quad (3-5)$$

In general, for the discriminator D network, this is a dichotomous problem, namely, the data is divided into true or false. For generator network G, in order to make the generated data G(z) as identical as possible to the real data, the two models need to be cross-trained in the way of D before G in the training process. Although GAN can generate more clear and real samples, it is difficult for GAN to reach Nash equilibrium in the actual training process. Therefore, GAN has problems such as training instability, mode collapse and gradient disappearance, and GAN is not suitable for processing discrete data such as text. The basic GAN network structure is shown in Fig. 3:

3.1.2. Deep convolution Generative Adversarial Network

In the training process of GAN, due to the instability of training, meaningless output often appears. Combining CNN and GAN, DCGAN uses full convolutional network to replace the multilayer perceptron in GAN, thus effectively stabilizing the training.

It basically did the following several optimizations:

1. Remove the pooling layer. In the generator, the transposed convolution is used for up-sampling, while in the discriminator, the convolution with stride is used instead of the pooling layer.

2. Batch normalization is used in both the generator and the discriminator to prevent the generator from converging all samples to the same point.
3. Remove the full connection layer.
4. The generator output layer use ReLU activation function and the last layer use Tanh activation function.
5. All layers in the discriminator use the Leaky ReLU activation function.

The modification of DCGAN greatly improves the stability of GAN training and the result of image generation. However, the disadvantage is that the generated image still lacks diversity and the image is relatively single.

The DCGAN generative model and discrimination model are shown in Fig. 4(a) and Fig. 4(b), respectively.

3.1.3. Dense Generative Adversarial Network (DGAN)

Considering the disadvantages of DCGAN, medical images have higher requirements on the generation results. We propose a five-layer Dense Block after the convolutional layer of the structure in the discriminator network of DCGAN. Each Dense Block consists of a 3x3 step size 2 average pooling layer and a number of 1*1 and 3*3 convolution kernels. The front of the 3*3 convolution of each dense block contains a 1*1 convolution operation, which is called the so-called "bottleneck layer", which aims to reduce the number of input feature maps, not only to reduce dimension and computation, but also to merge the features of each channel.

Joining the Dense Block gives our network the following advantages: The Dense Block allows the learned feature map to be received by all subsequent layers, increasing the reuse of features and making the model more compact. The advantage of this is that the images generated by the generation network have higher image generation quality and generalization.

In a nutshell, the Dense block is a situation where the input from each layer comes from the output of all previous layers.

Another advantage of the Dense block is that it makes the network narrower and has fewer parameters. In the Dense block, the number of output feature maps of each convolutive layer is small (less than 100), rather than hundreds and thousands of widths as in other networks. At the same time, this connection makes the transmission of features and gradients more effective, and the network is easier to train. The problem of gradient disappearance is more likely to occur when the network is deeper. The reason is that input information and gradient information are transferred between many layers. Now, this dense Connection is equivalent to each layer directly connecting input and loss, so the phenomenon of gradient disappearance can be alleviated, so the deeper network is not a problem

The formula representation of the Dense Block is shown in Formula

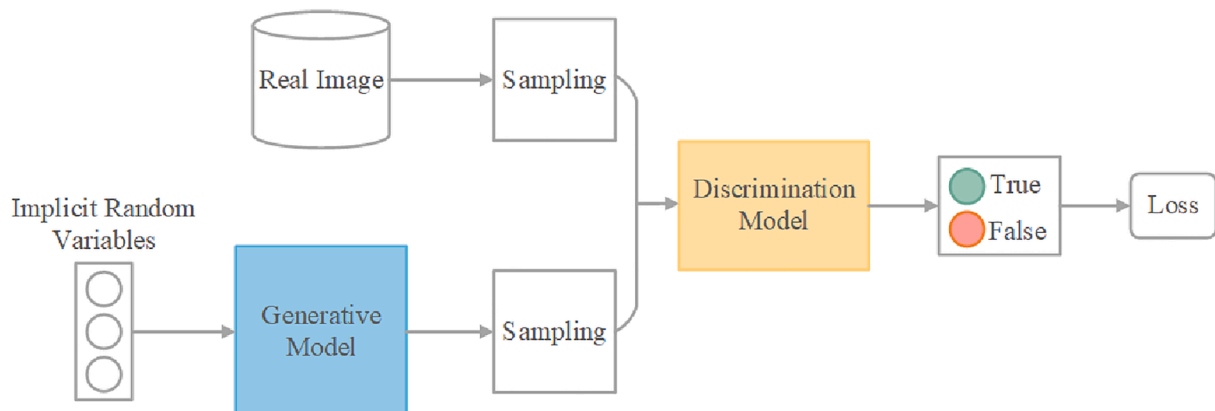


Fig. 3. Basic GAN network structure.

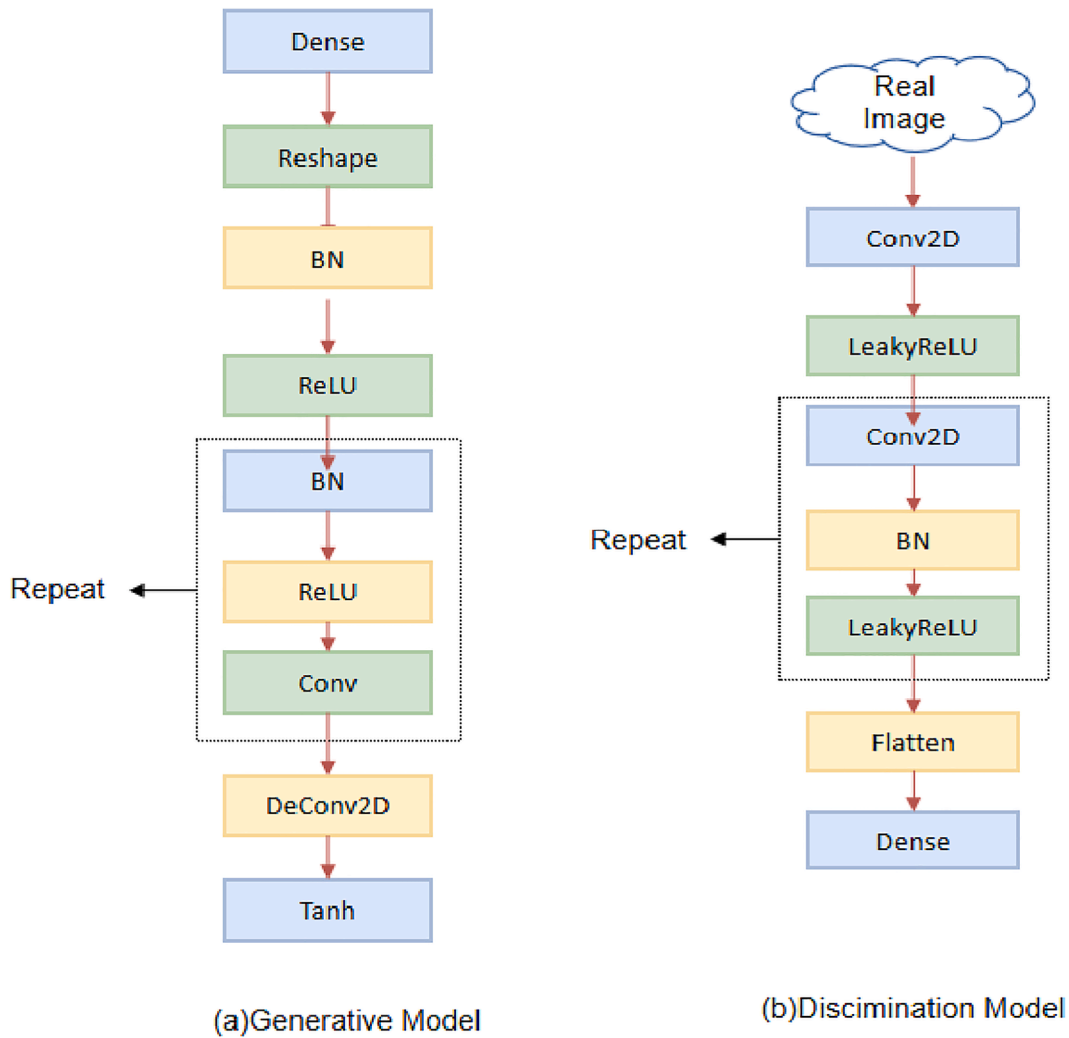


Fig. 4. DCGAN network structure.

(3–6):

$$X_l = H_{([x_0, x_1, \dots, x_{l-1}])} \quad (3-6)$$

$[x_0, x_1, \dots, x_{l-1}]$ means concatenation will be done from the output feature map of layer 0 to L-1. Concatenation is doing channel merging, just like Inception.

The novel generative network model of Dense GAN is shown in Fig. 5. The pseudocode for DenseGAN is shown in Table 1.

3.2. Multilayer attention mechanism U-Net

The COVID-19 virus comes from the air and is inhaled by the bronchi. When inhaled into the lungs, the virus will spread and present multiple clumps or smoke alterations. The diameter of the COVID-19 virus is 60 to 140 nm. The diameter of the virus is very small. The diameter of the alveoli is 200 μm , the size of the alveolar pores is 10 to 15 μm , and the viruses are much smaller. The amount of virus that we inhale from the air into our lungs is quite large, like a cloud of smoke. Once the virus gets into the lungs, it gets to the periphery of the alveolar. Inside the alveoli, the alveolar orifice is 10 to 15 μm , and the virus is nano-scale, so the virus is very easy to spread around through the alveolar orifice. We call it a smoke pattern, like a firework exploding in the air, spreading from one center to the outside, so it's a multiple ground glass lesion. In view of the above characteristics, we add multi-layer attention modules in the backbone U-Net and a novel network

model of multilayer attention mechanism U-Net is proposed, shown in Fig. 6.

3.2.1. Edge attention module

In U-Net architecture, the image information to be segmented is gradually input from the high level to the low level, and the position and overall feature information of the high level image will be gradually diluted by convolution kernel and pooling operation. The left side of the network is the down sampling side, which is also called the contracting path in the original text. That is to compress the image size and extract the features in the image. With the increase of the depth of the network, the extracted features tend to be more low-level features. In this paper, the edge attention module is added between the first layer and the second layer. In this module, firstly, the salient edge feature information of the segmented pathological tissue is modeled. In order to obtain more robust salient object features, we add three convolution layers, and add a relu layer after each convolution layer to ensure the nonlinear expression ability of the model. Conv2 preserves more abundant edge feature information of pathological tissue in the image. Therefore, this paper extracts local edge feature information from Conv2. Inspired by the mechanism of nonlocal attention, we know that in order to obtain more typical and rich edge features, only local feature information is not enough, and it must be combined with the advanced semantic information or location information of the lesion tissue. Therefore, the edge attention module designs a feature information fusion mechanism from top to bottom, and inputs the spatial position information of top level

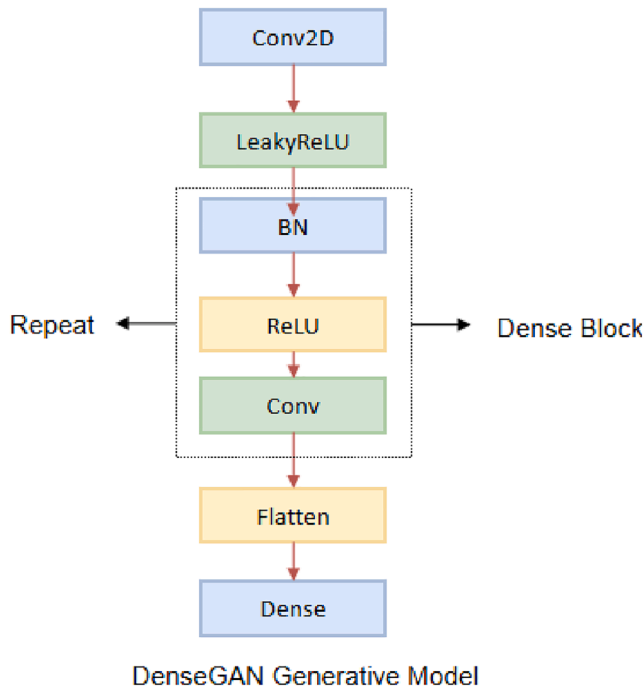


Fig. 5. Novel Generative network model of Dense GAN.

Table 1

The pseudocode for DenseGAN.

Dense GAN training algorithm with step size \times , using minibatch Adam for simplicity.
for number of training epochs do
for each batch do
● Sample batch of m real data distribution Z_i . (Generative network)
● Generate batch of m noise data distribution x_i . (Generative network)
● Input Z_i and x_i to the Discrimination network.
● Update the discriminator by ascending its stochastic gradient:
$\frac{1}{m} \sum_{i=1}^m [\log(D(x_i)) + \log(1 - D(G(z_i)))]$
● Update the generator by descending its stochastic gradient:
$\log(1 - D(G(z_i)))$
end for
end for

conv1 into the edge attention module, which can suppress the edge feature information of the region without spatial coherence, improve the weight ratio of the edge feature of the target region, and achieve the ultimate goal of improving the accuracy of network segmentation (Fig. 7).

3.2.2. Shape attention module

According to the shape characteristics of CT lesions in the lung of patients with COVID-19 (see Fig. 8 for details), two different filters are designed in this paper: \square considering that there are no obvious edges and corners in the shape of the lesions, but there are often round or quasi round pulmonary nodules. In this paper, a circular shape filter is introduced to enhance the attention ability of this kind of shape features. \square High gray-scale diffuse ground glass opacity and pulmonary nodule shadow often appear in the lung CT of patients with COVID-19. This is because COVID-19 virus will further invade the epithelial cells of alveoli after entering the lung bronchus, and make self-replication in them. The rapid replication of virus will lead to obvious swelling of epithelial cells, and high gray shadow with obvious gray difference from ordinary lung parenchyma will appear in CT images. Inspired by the traditional threshold method, this paper uses the gray threshold filter to extract the key pixels that exceed the gray threshold to improve the detection

efficiency of high gray shadow features.

3.2.3. Local attention module

In the clinical analysis of lung image lesions, experienced radiologists usually use two-step procedure to segment the lung infection area, roughly locate the infection area, and then accurately mark these areas by examining the local tissue structure. Inspired by the clinical experience, we believe that each pixel in the CT image to be segmented also needs to learn from the characteristics of its adjacent region, and determine whether it exists in the scope of the diseased tissue. Therefore, this paper introduces the local attention module, shown in Fig. 9.

After the addition of multi-layer attention module, the detection rate of ground glass lesions and the precision of lesion segmentation can be improved during the training of lung CT image input into the network. The thermodynamic distribution of the attention mechanism is shown in Fig. 10:

4. Experiments training and results analysis

4.1. Experimental training data set

In this paper, more than 100 CT images from different COVID-19 patients were used. All the CT images were collected by the Italian Society of Medical and Interventional Radiology, where a radiologist splits the CT images with different labels to identify lung infections. This is the first open access COVID-19 data set for segmentation of lung infections, and its sample size is small, with only 100 marker images available.

In order to improve the generalization ability of the model, we used Dense GAN to generate 1600 Lung CT images of COVID-19 patients. An unmarked training data set is constructed for effective semi-supervised segmentation. In the experimental training, the proportions of our training set, verification set and test set were 80%, 10% and 10% respectively. In order to ensure the authenticity of the trained segmentation neural network model, the test data set will be randomly selected from the original COVID-19 patient lung CT image dataset without participating in training. The segmentation results obtained by this method have good authenticity and can reflect the actual performance in practical application scenarios. The image data set of training is shown in Fig. 11:

4.2. Experimental environment

The computer hardware and software used in the experimental training were configured as two RTX2080 Ti (11 GB video memory) video CARDS, i9-9900 K CPU (3.60 GHz) and 64 GB RAM. Pytorch deep learning framework is used to carry out experimental training in Windows10 operating system.

4.3. Experimental training phase

4.3.1. Use of Dense GAN to extend a semi-supervised training data set

We used Dense GAN to train from the 100 identified images that have been opened to generate additional lung CT images. Our proposed Dense GAN uses a five-layer Dense Block after the convolutional layer in the network generated structure of DCGAN website. Each Dense Block is composed of a 3x3 step length 2 average pooling layer and a number of 1x1 and 3x3 convolution cores, enabling each layer of Dense Block to make full use of the features in the previous CNN layer. The batch size in the network is set to 8, and the resolution of the image is uniformly adjusted to 256x256.

During the training, the discriminator improves its ability to identify true data and false data as much as possible, while the generator generates true data as much as possible. The discriminator uses real data and false data for training to update weights, while the generator can only use the complete model (frozen discriminator, input false data, but marked as true) to feedback to the front through the discriminator error

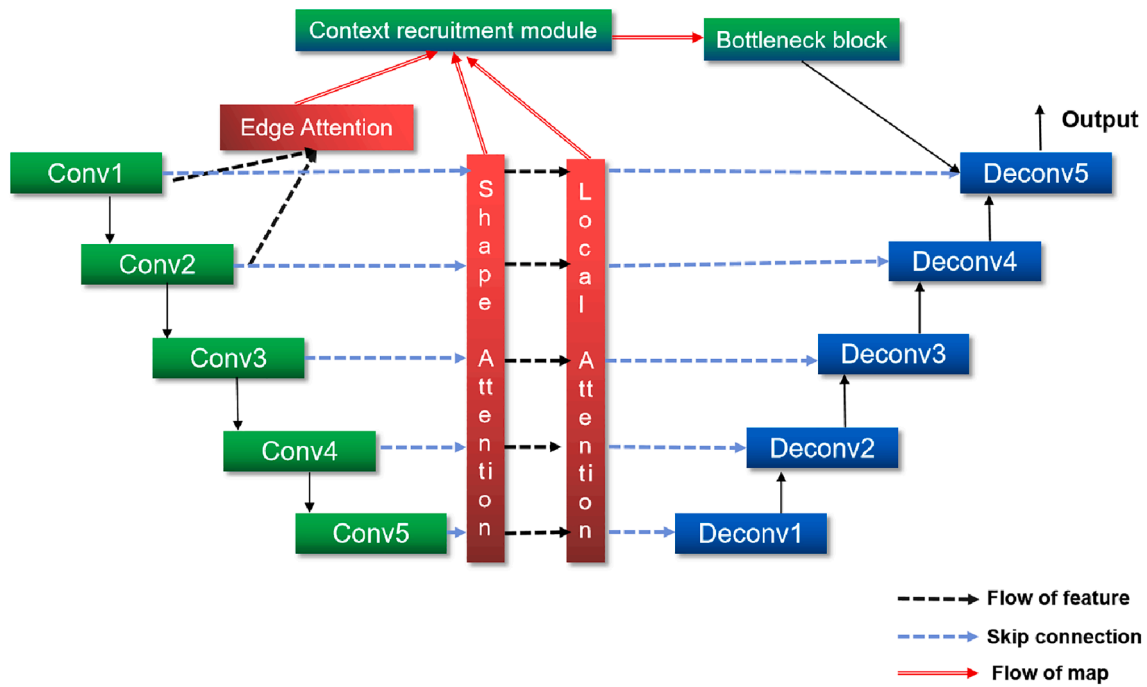


Fig. 6. A novel network model of multilayer attention mechanism U-Net.

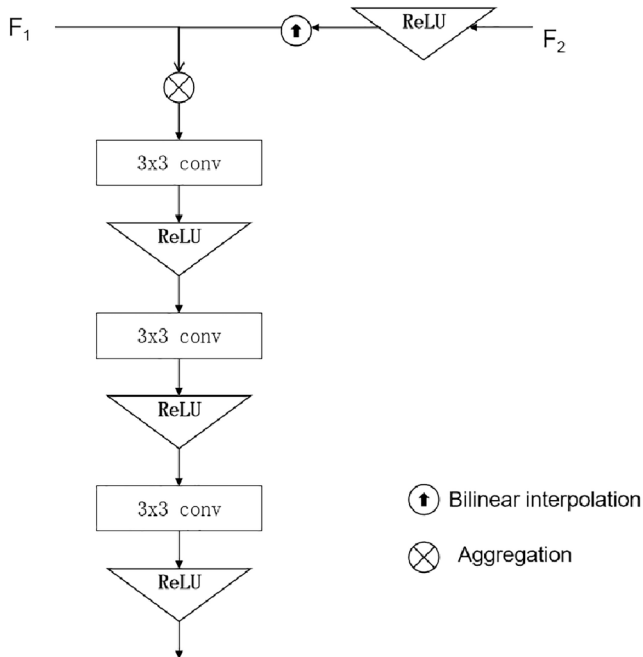


Fig. 7. The edge attention module.

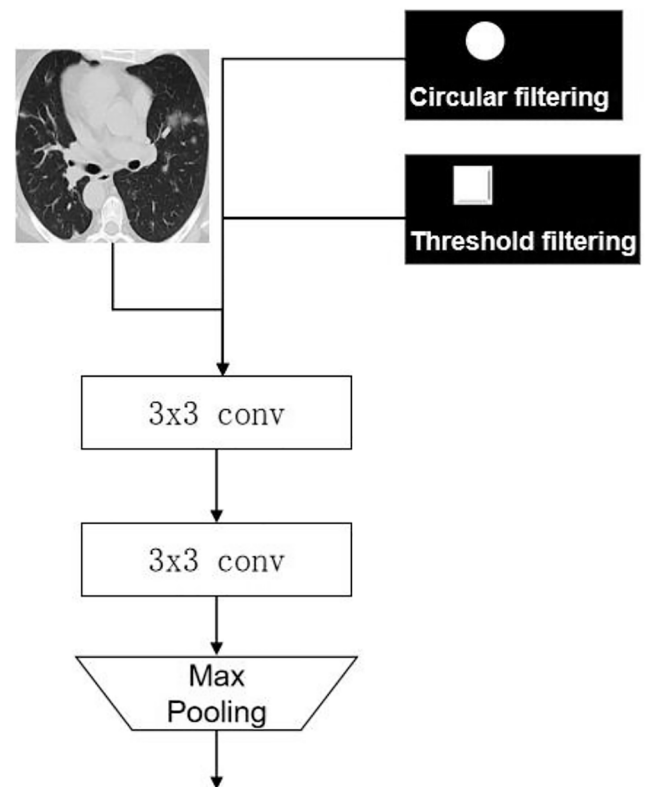


Fig. 8. The shape attention module.

to achieve weight update, so as to improve the authenticity of the generated image.

The image generated from our Dense Net is unlabeled at this time. To solve this problem, we use semi-supervised learning to generate pseudotags for subsequent segmentation network training. The generated CT image is shown in Fig. 12:

We introduce inception Score to evaluate the quality of DGAN generated images. Inception Score is referred to as IS, which can be used to evaluate the clarity and diversity of images generated by the generating networks of the generation model. IS sends the generated image \times into the trained Inception model, such as Inception Net-V3, which is a

classifier and outputs a 1000-dimensional label vector y for each input image. Each dimension of the vector represents the probability showing which category the input sample belongs to. Assuming that our Inception Net-V3 is trained well enough, for a high quality generated image \times , Inception Net-V3 can classify it into a class with high probability. The pictures generated by the Dense GAN model in this paper obtained 8.3 IS

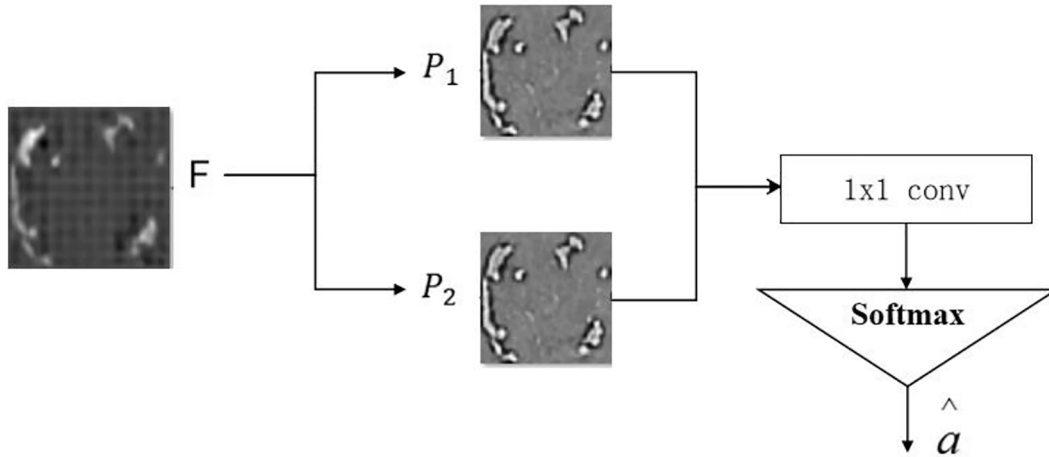


Fig. 9. The local attention module.

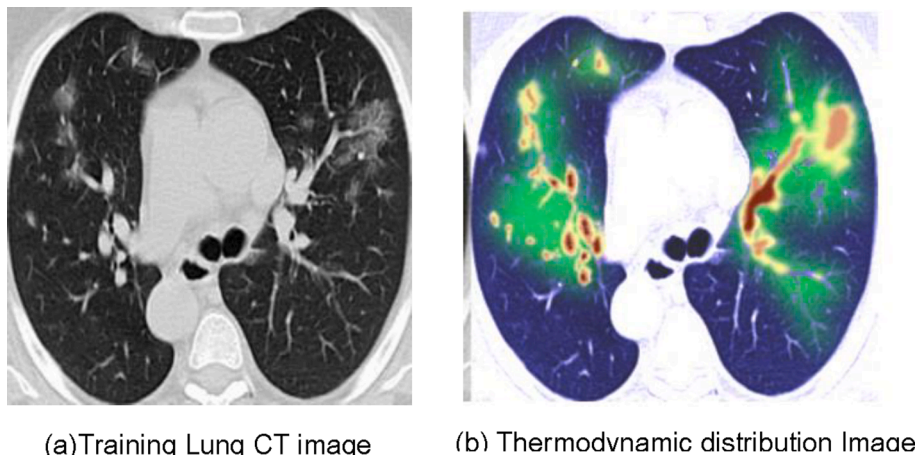


Fig. 10. Thermodynamic distribution of the attention mechanism.

score, indicating that the clarity and diversity of lung CT images generated by the Dense Gan model are very good, which can effectively expand our data set.

4.3.2. U-NET network with multi-layer attention mechanism was used for segmentation

On the basis of traditional U-Net, we introduced a multi-layer attention mechanism targeting the characteristics of lung lesions in COVID-19 patients. In the U-NET encoder, when the feature on each pixel and the corresponding feature in the decoder are splicing, a multi-layer attention module is added. Each attention module consists of two branches: the Trunk Branch and the Mask branch. The trunk branch can use any of the current SOTA convolutional neural network models. Through the processing of the feature map, the mask branch outputs the attention feature map with consistent dimensions, and at the same time increases the weight of the feature map of the bronchus. Then, the dot product operation is used to combine the feature map of the two branches to obtain the final output feature map.

The batch size in the network is set to 18, which is equivalent to randomly selecting 18 images from the medical CT image data set for each iteration. The resolution of the images is 256x256. Adam [23] optimization algorithm was used to calculate the loss function in the network, where Momentum was 0.9 and the learning rate was 0.0003. When the epoch was 600, the network performance obtained by training was the best.

4.4. Comparative analysis of experimental results

4.4.1. Evaluation indexes

We used five widely adopted measures, namely, Dice similarity coefficient, sensitivity, accuracy, MAE and Structure Measure. Their definitions are as follows:

1. DICE (value range: [0,1])

Represents the ratio of the area of the intersection of two objects to the total area. The value of perfect segmentation is 1.

$$DICE = \frac{2 * (R_{seg} \cap R_{gt})}{R_{seg} + R_{gt}} \quad (4-1)$$

R_{gt} represents the area of Ground Truth, R_{seg} represents the area of the image obtained after segmentation.

2. Sensitive

Represents the proportion of all positive examples that are correctly classified, and measures the classifier's ability to identify the positive examples.

$$Sen = TP / (TP + FN) \quad (4-2)$$

Where TP stands for true positive (in this case, patients diagnosed) and FN for false negative (in this case, patients misclassified as normal).

3. Precision

Precision is a measure of accuracy, representing the proportion of positive examples that are actually positive examples in an example divided into positive examples:

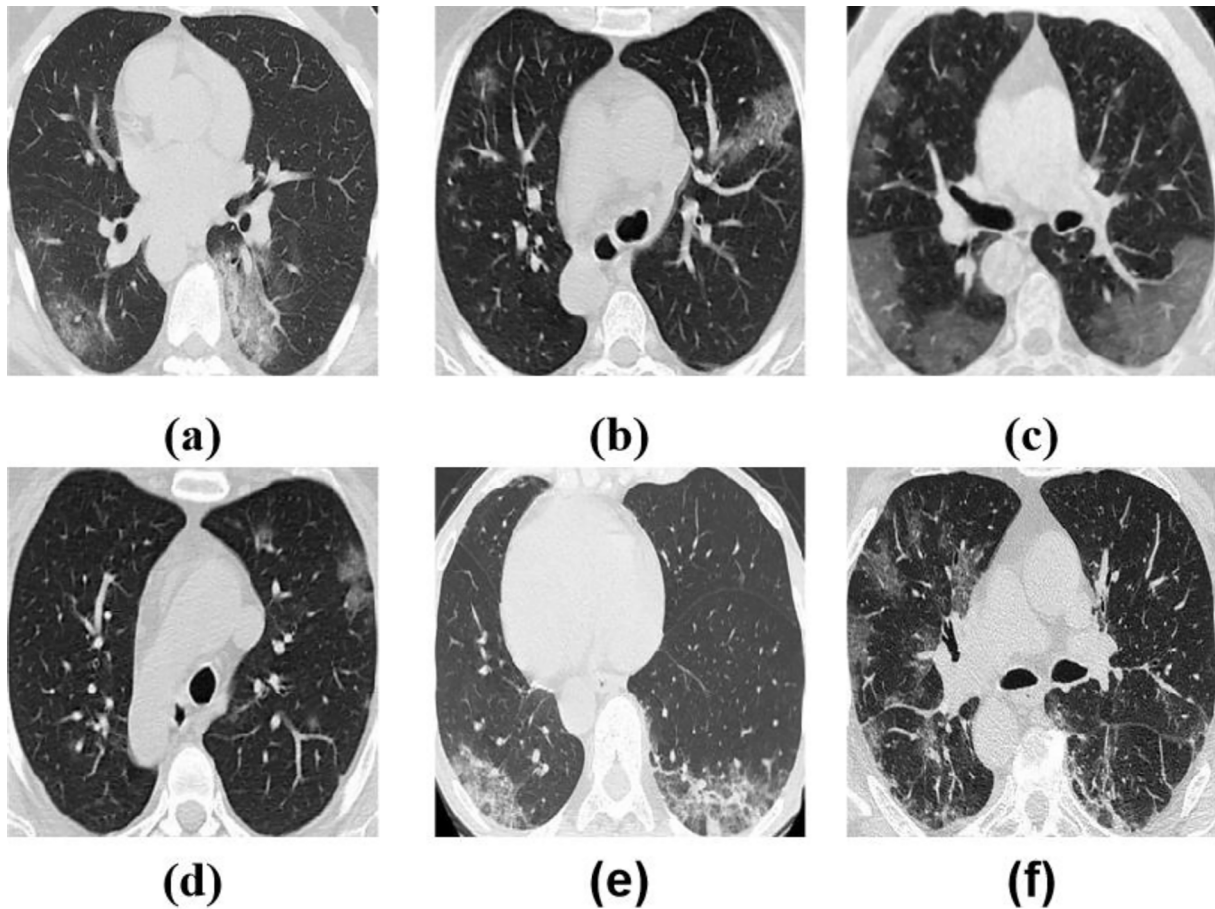


Fig. 11. Training images.

$$Pre = TP / (TP + FP) \quad (4-3)$$

Where FP stands for false positive (in this case, for ordinary people who have been misdiagnosed)

4. MAE

MAE is the average of the absolute error. It can better reflect the actual situation of the predicted value error. The smaller the value of MAE means the higher the segmentation accuracy:

$$MAE(X, h) = \frac{1}{m} \sum_{i=1}^m |h(x_i) - y_i| \quad (4-4)$$

5. Structure Measure (S_α)

This is a method proposed to measure the structural similarity between the predicted image and the ground-truth, which is more in line with the human visual system.

$$S_\alpha = (1 - \alpha) * S_o(S_p, G) + \alpha * S_r(S_p, G) \quad (4-5)$$

where α is a balance factor between object-aware similarity and region-aware similarity. We report using the default setting ($\alpha = 0.5$) suggested in the original paper.

4.4.2. Comparison of experimental results

Experimental studies are conducted for the Dense GAN and multi-layer attention based method, and comparisons with the current-state of the art are presented in this section to show the effectiveness and strength of the proposed method.

The proposed method is compared with the U-Net, Gated U-Net, Dense U-Net, U-Net and the Inf-Net networks. The change of the loss function and accuracy of the training set and the validation data set during the model training in this paper is shown in Fig. 13(a) (b), and the

specific segmentation results are shown in Fig. 14. Table 2 shows the parameters of each network and the quantitative results of experiments.

The backbone network is a network for feature extraction. It is used to extract image information features and generate feature map for the following segmentation network. The six different methods in Table 2 use three mainstream backbone networks for feature extraction. The backbones used by U-Net [25], Gated U-Net [26], U-Net++ [27] and our proposed method are VGG16. Compared with DenseNet161 and Res2Net used by Dense U-Net [28] and Inf-Net [21] respectively, VGG16 has the least network layers and the most simple structure, and thus its basic parameters are the least. In addition, due to the special gating structure used by Gated U-Net [26], the number of parameters is greatly reduced. Similarly, the multi-layer attention mechanism used in our method also effectively reduces a certain amount of parameters.

From the segmentation results shown in Fig. 14, we can see that, for other recently developed four networks, U-Net [25], Gated U-Net [26], U-Net++ [27], Dense U-Net [28], their boundaries are relatively fuzzy and the gap between them and the ground truth is large. However, the results of our proposed method and the Inf-Net segmentation are very close to ground truth, and the effect is very similar.

According to the quantitative indexes in Table 2, Inf-Net [21] and the method in this paper make use of the attention mechanism to achieve better segmentation effect than the other four models without a significant increase in the number of parameters. At the same time, we notice that the number of parameters of this network is less than that of Inf-Net, but the evaluation result of segmentation effect is better.

4.4.3. Ablation studies

In this section, we will further verify the effectiveness of the Dense GAN and multilayer attention mechanism through a variety of

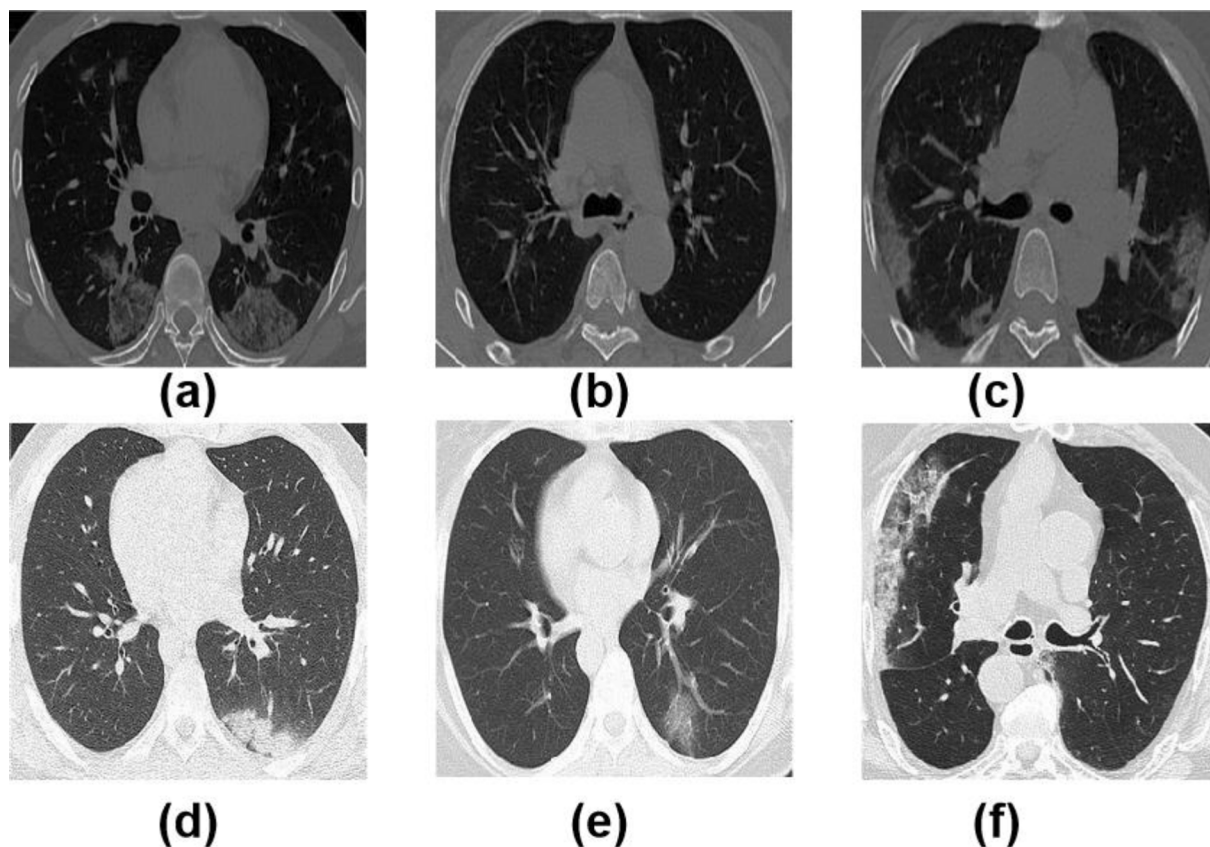


Fig. 12. Generated Pictures.

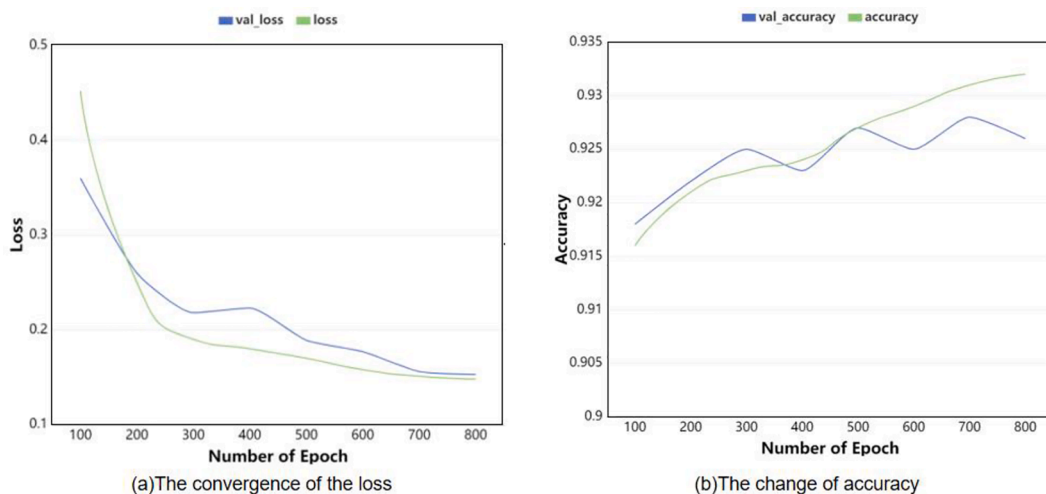


Fig. 13. Loss and accuracy during the model training.

experiments of Ablation studies.

a. Effectiveness of Dense GAN

In order to show the validity of the Dense GAN network for improving the segmentation accuracy of our model, we also trained CT images generated by GAN and DCGAN networks. The detailed comparison results are shown in Table 3. The results (a) and (b) show that Dense GAN can indeed help us improve the accuracy of image segmentation.

b. Effectiveness of multilayer attention mechanism

In order to show the importance of multilayer attention mechanism, we carried out two experiments (c) and (d) in Table 3. It can be seen that,

with the addition of multilayer attention mechanism, the segmentation quantization indexes have been improved to some extent, indicating that multilayer attention mechanism effectively helps us to segment lesions more accurately.

5. Conclusions

In this paper, a novel Dense GAN and multi-layer attention based lesion segmentation method for COVID-19 CT Images is presented. An improved Dense GAN was developed to generate higher quality images to expand the limited medical image segmentation data set of COVID-

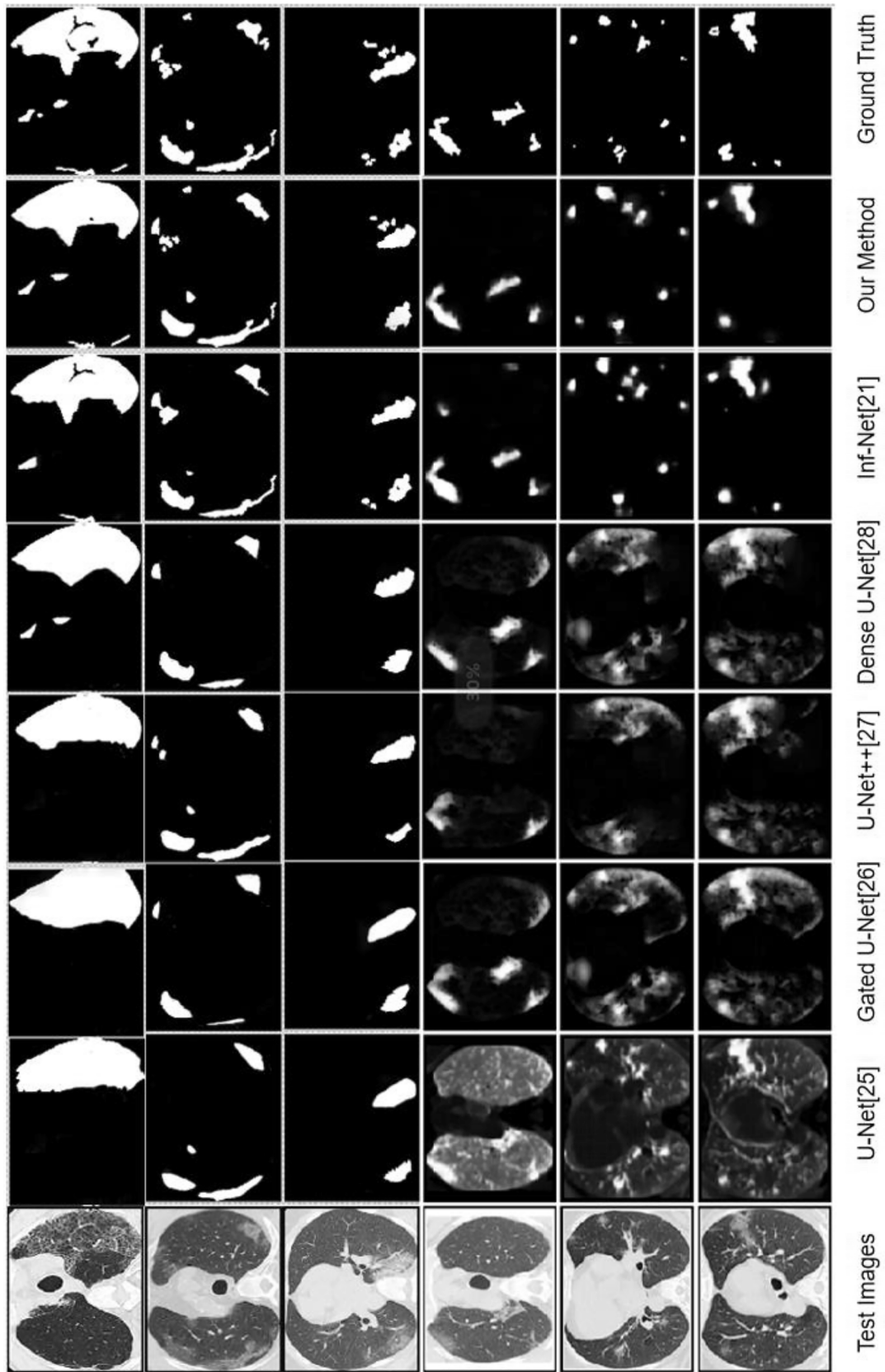


Fig. 14. The segmentation effects of experiments.

Table 2
Evaluation results of each network.

Methods	Backbone	Parameters	Dice	Sensitive	Precision	MAE	S_{α}
U-Net [25]	VGG16	7.853 M	0.439	0.534	0.858	0.186	0.622
Gated U-Net [26]	VGG16	175.093 K	0.623	0.658	0.926	0.102	0.725
U-Net++ [27]	VGG16	9.163 M	0.581	0.672	0.902	0.120	0.722
Dense U-Net [28]	DenseNet161	45.082 M	0.515	0.594	0.840	0.184	0.655
Inf-Net [20]	Res2Net	33.122 M	0.682	0.692	0.943	0.082	0.781
Our Method	VGG16	28.538 M	0.683	0.698	0.946	0.075	0.792

Table 3
Evaluation results of Ablation Studies.

Methods	Dice	Sensitive	Precision	MAE	S_{α}
(a)GAN	0.468	0.572	0.820	0.173	0.628
(b)DCGAN	0.537	0.613	0.871	0.136	0.683
(c)GAN + MA	0.563	0.638	0.893	0.129	0.703
(d)DCGAN + MA	0.642	0.663	0.928	0.098	0.758
(e)Our Method	0.683	0.698	0.946	0.075	0.792

MA: Multilayer attention mechanism.

19. By adding multilayer attention modules in the backbone U-Net, a novel network model of multilayer attention mechanism U-Net is proposed.

Experimental studies are conducted for the Dense GAN and multilayer attention based method, and comparisons with the current-state of the art, i.e. with recently developed methods such as the U-Net, Gated U-Net, Dense U-Net, U-Net and the Inf-Net networks, are presented. The experimental results showed that the segmentation method proposed in this paper improved the segmentation accuracy of COVID-19 pulmonary medical CT images, and achieved better segmentation effect. At the same time, the number of parameters of this network is less than that of Inf-Net, and the evaluation result of segmentation effect is better than that of Inf-Net. Moreover, the effectiveness of the proposed method are also verified through a variety of experiments of Ablation studies.

CRedit authorship contribution statement

Ju Zhang: Supervision, Conceptualization, Writing - review & editing. **Lunduan Yu:** Investigation, Methodology, Software, Writing - original draft. **Decheng Chen:** Validation, Software. **Weidong Pan:** Software, Validation. **Chao Shi:** Software, Validation. **Yan Niu:** Software, Validation. **Xinwei Yao:** Validation. **Xiaobin Xu:** Validation. **Yun Cheng:** Project administration, Resources, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Chen, L. Wu, J. Zhang, L. Zhang, D. Gong, Y. Zhao, H. Yu, Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography, *Sci. Rep.* 10 (1) (2020) 19196.
- [2] D. Singh, Kumar, V. Vaishali, M. Kaur, Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks, *Eur. J. Clin. Microbiol. Inf. Dis.* 39 (7) (2020) 1379–1389.
- [3] M. Kassimi, R. Filali, N. Chikhaoui, M. Mahi, COVID-19 with false negative RT-PCR test in 4 patients: When Chest CT becomes a need, *Int. J. Sci. Res. Publ.* (2020), <https://doi.org/10.29322/ijrsp.10.06.2020.p10284>.
- [4] Y. Xie, Y. Xia, J. Zhang, et al., Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT, *IEEE Trans Med Imaging* (2019).
- [5] D. Mhaske, K. Rajeswari, R. Tekade, Deep Learning Algorithm for Classification and Prediction of Lung Cancer using CT Scan Images, 2019 5th International Conference on Computing, Communication, Control and Automation (ICCCUBEA), Pune, India, 2019, pp. 1–5.
- [6] N. Mohanapriya, B. Kalaavathi, T.S. Kuamr, Lung Tumor Classification and Detection from CT Scan Images using Deep Convolutional Neural Networks (DCNN), 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 2019, pp. 800–805.
- [7] S. Hu, et al., Weakly supervised deep learning for COVID-19 infection detection and classification from CT images, *IEEE Access* 8 (2020) 118869–118883.
- [8] Z. Han, et al., Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning, *IEEE Trans. Med. Imaging* 39 (8) (Aug. 2020) 2584–2594.
- [9] C. Huang, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet* 395 (10223) (2020) 497–506.
- [11] C. Han et al., Synthesizing diverse lung nodules wherever massively: 3D multi-conditional GAN-based CT image augmentation for object detection, 2019 International Conference on 3D Vision (3DV), Quebec City, QC, Canada, 2019, pp. 729–737.
- [12] K.-B. Park, S.H. Choi, J.Y. Lee, M-GAN: retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks, *IEEE Access* 8 (2020) 146308–146322.
- [13] H. Rashid, M.A. Tanveer, H.A. Khan, Skin Lesion Classification Using GAN based Data Augmentation, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2019: 916–919.
- [14] C. Zhao, J. Han, Y. Jia, F. Gou, Lung Nodule Detection via 3D U-Net and Contextual Convolutional Neural Network, 2018 International Conference on Networking and Network Applications (NaNA), Xi'an, China, 2018, pp. 356–361.
- [15] Y. Lai, et al., 2019 Novel Coronavirus-infected pneumonia on CT: a feasibility study of few-shot learning for computerized diagnosis of emergency diseases, *IEEE Access* 8 (2020) 194158–194165.
- [17] C. Zheng, X. Deng, Q. Fu, et al. Deep learning-based detection for COVID-19 from chest CT using weak label. medRxiv, 2020.
- [18] Y. Pathak, P.K. Shukla, A. Tiwari, et al., Deep transfer learning based classification model for COVID-19 disease, *IRBM* (2020).
- [19] A.A. Ardakani, A.R. Kanafi, U.R. Acharya, et al., Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks, *Comput. Biol. Med.* (2020), 103795.
- [20] L. Zhou, et al., A rapid, accurate and machine-agnostic segmentation and quantification method for CT-based COVID-19 diagnosis, *IEEE Trans. Med. Imaging* 39 (8) (2020) 2638–2652.
- [21] D.-P. Fan, et al., Inf-Net: automatic COVID-19 lung infection segmentation from CT images, *IEEE Trans. Med. Imaging* 39 (8) (2020) 2626–2637.
- [22] S. Chen, X. Tan, B. Wang, H. Lu, X. Hu, Y. Fu, Reverse attention-based residual network for salient object detection, *IEEE Trans. Image Process.* 29 (2020) 3763–3776.
- [23] Z. Han, B. Wei, Y. Hong, T. Li, J. Cong, X. Zhu, W. Zhang, Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning, *IEEE Trans. Med. Imaging* 39 (8) (2020) 2584–2594.
- [25] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: *MICCAI*, Springer, 2015, pp. 234–241.
- [26] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, D. Rueckert, Attention gated networks: Learning to leverage salient regions in medical images, *Med. Image Anal.* 53 (2019) 197–207.
- [27] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, UNet++: a nested U-Net architecture for medical image segmentation, *IEEE Trans. Med. Imaging* (2019) 3–11.
- [28] X. Li, H. Chen, X. Qi, Q. Dou, C. Fu, P. Heng, H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes, *IEEE Trans. Med. Imaging* 37 (12) (2018.8.) 2663–2674.