



DUDA-Net: a double U-shaped dilated attention network for automatic infection area segmentation in COVID-19 lung CT images

Feng Xie^{1,2,3} · Zheng Huang^{2,3,4} · Zhengjin Shi¹ · Tianyu Wang^{1,2,3} · Guoli Song^{2,3} · Bolun Wang^{1,2,3} · Zihong Liu^{1,2,3}

Received: 13 January 2021 / Accepted: 21 May 2021 / Published online: 5 June 2021
© CARS 2021

Abstract

Purpose The global health crisis caused by coronavirus disease 2019 (COVID-19) is a common threat facing all humankind. In the process of diagnosing COVID-19 and treating patients, automatic COVID-19 lesion segmentation from computed tomography images helps doctors and patients intuitively understand lung infection. To effectively quantify lung infections, a convolutional neural network for automatic lung infection segmentation based on deep learning is proposed.

Method This new type of COVID-19 lesion segmentation network is based on a U-Net backbone. First, a coarse segmentation network is constructed to extract the lung areas. Second, in the encoding and decoding process of the fine segmentation network, a new soft attention mechanism, namely the dilated convolutional attention (DCA) mechanism, is introduced to enable the network to focus on better quantitative information to strengthen the network's segmentation ability in the subtle areas of the lesions.

Results The experimental results show that the average Dice similarity coefficient (DSC), sensitivity (SEN), specificity (SPE) and area under the curve of DUDA-Net are 87.06%, 90.85%, 99.59% and 0.965, respectively. In addition, the introduction of a cascade U-shaped network scheme and DCA mechanism can improve the DSC by 24.46% and 14.33%, respectively.

Conclusion The proposed DUDA-Net approach can automatically segment COVID-19 lesions with excellent performance, which indicates that the proposed method is of great clinical significance. In addition, the introduction of a coarse segmentation network and DCA mechanism can improve the COVID-19 segmentation performance.

Keywords Medical image analysis · Deep learning · Lesion segmentation · U-Net · Attention mechanism

Introduction

Due to the infectivity of the new coronavirus disease 2019 (COVID-19) and the shortage of medical resources since the outbreak of COVID-19 in 2020, a large number of COVID-19 automatic prediction and diagnosis systems based on deep learning technology have been proposed, such as multistep autoregression methods [1] and convolutional neural

network approaches [2]. Although the existing automatic diagnosis system can improve the diagnostic efficiency and relieve pressure on medical systems, most of the existing COVID-19 automatic diagnosis systems directly diagnose entire computed tomography (CT) images [3]. Normal lung tissues and other diseased tissues will greatly interfere with the diagnosis system, which greatly affects the diagnostic accuracy [4]. To avoid this problem, it is necessary to extract the diseased tissues in the CT images and apply the automatic diagnosis system to analyse the COVID-19-diseased tissues [2]. At present, most hospitals extract lesions by time-consuming and labour-intensive manual segmentation methods. To improve the efficiency of lesion extraction, it is necessary to propose an automatic segmentation system for COVID-19 lesions.

Since a fully convolutional neural network was proposed in 2015, a large number of studies have verified that deep neural networks can achieve state-of-the-art performance in medical image segmentation tasks [5, 6]. Due to their

✉ Guoli Song
songgl@sia.cn

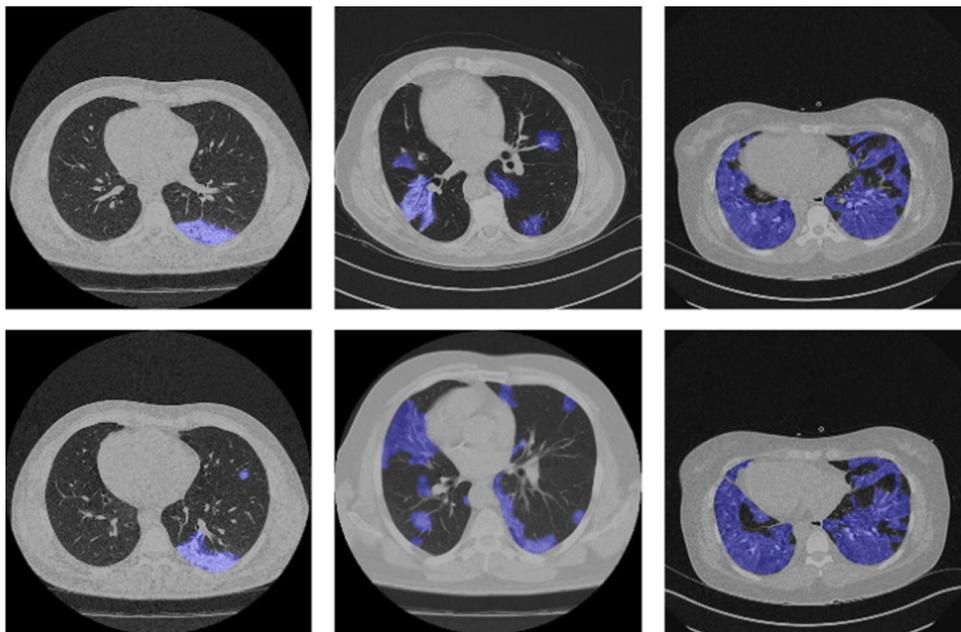
¹ School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang, China

² State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China

³ Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, China

⁴ University of Chinese Academy of Sciences, Beijing, China

Fig. 1 Images in the CT dataset. Lung consolidation is marked in purple



efficiency and excellent generalization, numerous deep learning-based methods have been proposed for COVID-19 lesion segmentation [7–10]. Although these methods have better segmentation accuracy than the direct use of U-Net, they still have the following problems. (1) The CT image input into the network contains nonpulmonary regions, which will cause the trained model to overfit. (2) The neural network lacks the spatial and channel information learning of CT images, and there is a large error in the small segmentation area. (3) The choice of a single loss function is difficult. For the COVID-19 lesion segmentation task, to effectively control the balance between false negatives and false positives, it is necessary to select a suitable loss function to train the network.

A double U-shaped dilated attention network (DUDA-Net) is proposed for automatic infection area segmentation in COVID-19 lung CT images to solve these problems. Our contributions mainly pertain to the following three areas. (1) A COVID-19 coarse segmentation method is proposed for the first time. The coarse segmentation network eliminates the interference of the nonpulmonary areas and improves the learning efficiency for the fine segmentation network. (2) A designed dilated convolutional attention (DCA) mechanism, which acquires multiscale context information and focuses on channel information, is proposed to improve the ability of the network to segment small COVID-19 lesions. (3) DUDA-Net with a suitable loss function for COVID-19 lesion segmentation has certain clinical value. In addition to improving the segmentation accuracy, it can reduce the segmentation time compared with manual segmentation methods.

Materials and Method

Dataset

In this work, a public database¹ obtained on March 30, 2020, from Radiopaedia [11] is employed to evaluate the performance of the proposed system. The public dataset contains CT images of more than 40 COVID-19 patients, with an average of 300 axial CT slices per patient, and infections are labelled by two radiologists and verified by an experienced radiologist. In this work, CT slices are employed to automatically segment the lesions. However, most of the data do not contain lesions, which easily causes a class imbalance problem. To avoid this issue, 557 CT slices are extracted from the public database. Figure 1 shows some CT samples of the dataset, which are utilized to train the neural network; the lung consolidations are marked in purple.

Image Preprocessing and Data Augmentation

To emphasize CT image characteristics and improve image quality, global histogram equalization [12] is applied to enhance the image contrast. The main idea of the global histogram equalization method is to equally redistribute each pixel value. By using this method, the COVID-19 infection area in a CT image becomes more obvious.

Deep neural networks are a kind of data-driven model. Small datasets can lead to overfitting. To avoid overfitting

¹ <https://academictorrents.com/details/136ffddd0959108becb2b3a86630bec049fcb0ff/>

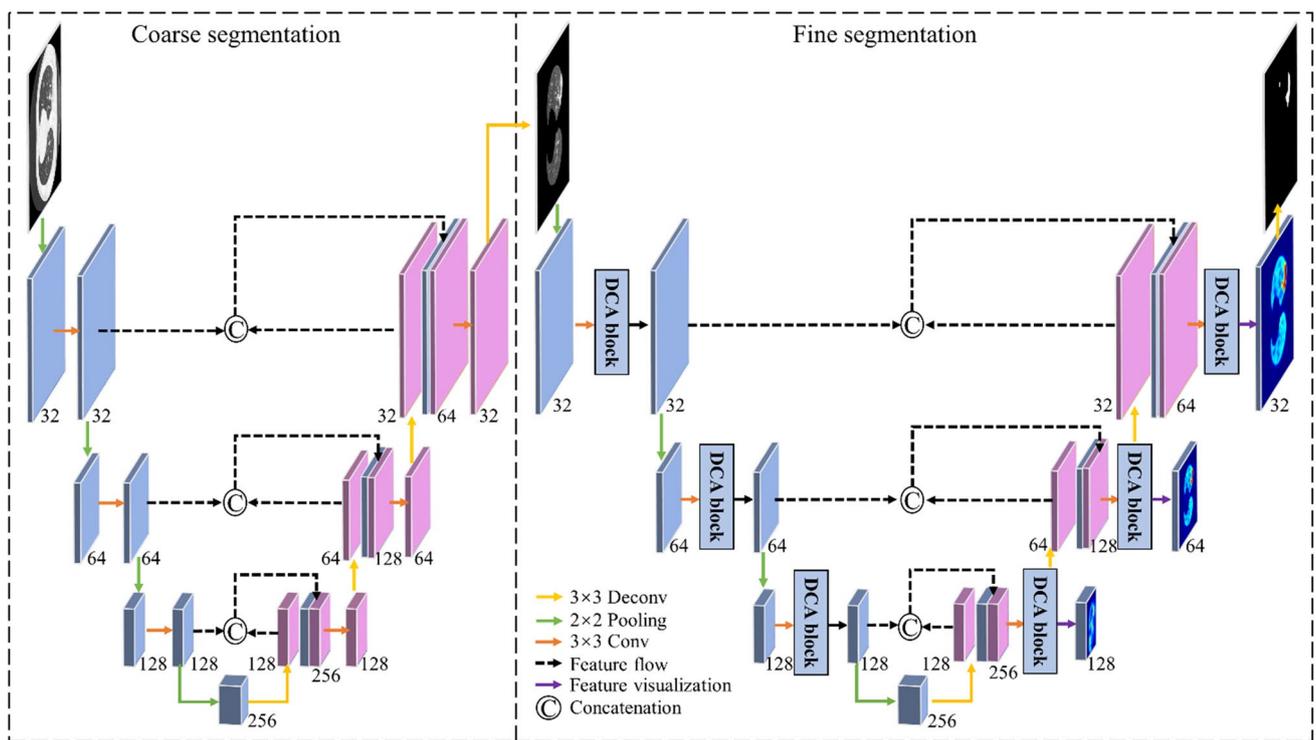


Fig. 2 DUDA-Net structure

and improve the generalizability of the proposed system, data augmentation techniques are implemented. In this work, data augmentation techniques, namely Gaussian noise [13] addition and image rotation by 90° , 180° and 270° , are implemented to enlarge the training dataset fivefold. After data augmentation, the training set contains 2628 CT slices, and the test set contains 157 CT slices. In addition, 10% of CT slices in the training set are randomly selected as the validation set.

Network Structure

Recently, a large number of studies have shown that U-shaped convolutional neural networks perform better than traditional machine learning methods in medical image segmentation. Since COVID-19 lesions appear only in the lung regions, using U-Net directly to segment the lesions will cause a high false-negative rate [14]. A U-shaped coarse-to-fine segmentation network is proposed to improve the segmentation performance. The network structure is shown in Fig. 2.

In this work, the coarse segmentation network contains 6 convolutional layers, 4 pooling layers and 4 transpose convolutional layers. First, CT images with a size of 256×256 are fed into the coarse segmentation network. Then, through 4 iterations of 2×2 max-pooling layers and 3×3 convolutional layers with strides of 1 in the encoder,

multilevel semantic features with sizes of 128×128 , 64×64 , 32×32 and 16×16 are acquired. Moreover, to iteratively recover the image resolutions, 3×3 transpose convolutional layers with a stride of 2 are introduced in the decoder. Furthermore, the high-level semantic feature maps in the decoder are densely concatenated with the low-level detail feature maps in the encoder to recover the details of the lung regions. In addition, batch normalization is added after each convolutional and transpose convolutional layer so that the input feature maps of each layer maintain the same distribution as the input images, and the training convergence is accelerated [15].

The fine segmentation network contains 6 convolutional layers, 4 transpose convolutional layers, 4 max-pooling layers and 6 DCA blocks, and it is the same as the coarse segmentation network on the backbone, which is a U-shaped network. However, segmentation of the lesions is more difficult than segmentation of the lung areas. The lesions are unevenly distributed and have different sizes. The U-shaped network used alone performs poorly. To improve the lesion segmentation performance, a channel attention mechanism, namely a DCA block, is proposed to force the network to focus on the key regions and channels. In this work, a DCA block is added after the ordinary convolution operation of each layer in the fine segmentation network. The DCA block can obtain multilevel context

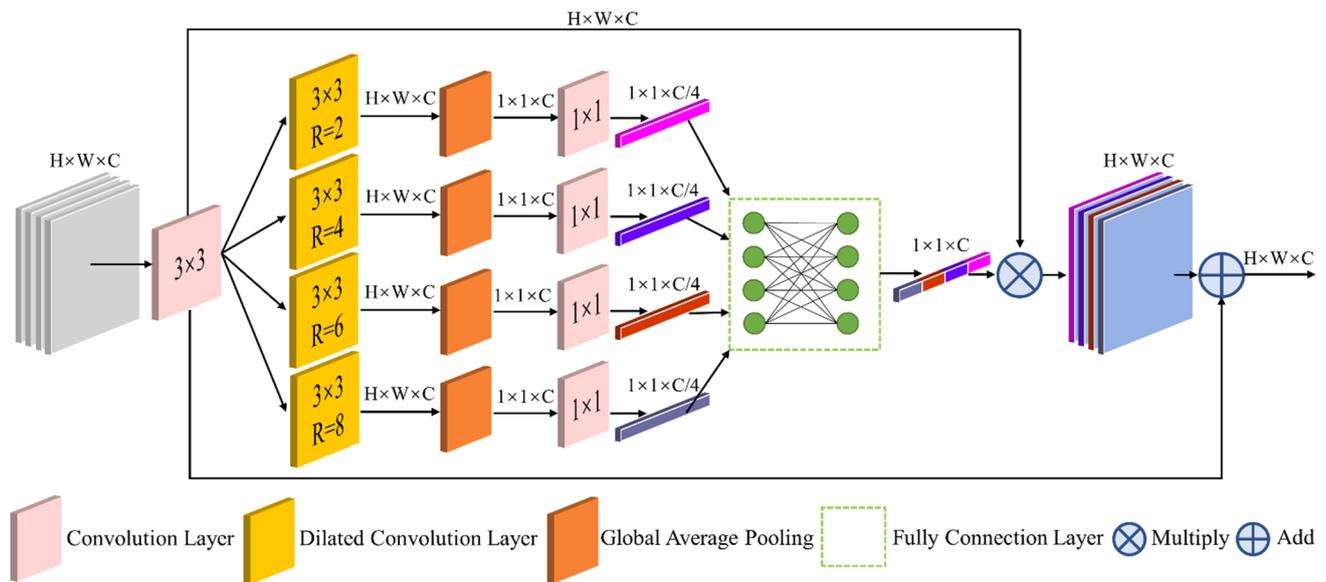


Fig. 3 Structure diagram of the DCA blocks

information to reduce the error rate of the segmentation boundaries and improve the accuracy.

In addition, the activation function of the last layer of the coarse segmentation network and the fine segmentation network are sigmoid functions, and the other layers all use the rectified linear unit (ReLU) activation function. The sigmoid and ReLU functions are defined as follows:

$$\text{ReLU} : x \rightarrow \max\{0, x\} \quad (1)$$

$$\text{Sigmoid} : x \rightarrow \frac{1}{1 + e^{-x}} \quad (2)$$

DCA Mechanism

Concatenation of high-level and low-level features in the U-shaped network can lead to feature channel redundancy. Therefore, it is necessary to propose a channel attention mechanism to suppress redundant channels and focus on key feature channels. Generally, the squeeze-and-excitation (SE) mechanism is one of the most typical cross-attention modules. The main procedure of the SE mechanism is to acquire the global distributions of feature maps by applying global average pooling and obtain the channel weights by introducing a two-layer dense neural network. Due to their simplicity, SE blocks are widely used in current methods. However, global average pooling in SE blocks can lead to information loss. To avoid the loss of information and introduce multiscale context information, a DCA module is proposed in this paper. The DCA mechanism not only focuses on channel information but

also introduces parallel dilated convolution with different dilation rates to acquire multiscale receptive fields, which is conducive to learning scale-invariant features without information loss. The overall structure of the DCA block is shown in Fig. 3. The height, width and number of channels of the input features are H , W and C , respectively. The size of the output feature maps is still $H \times W \times C$. The main procedures of the DCA blocks are as follows:

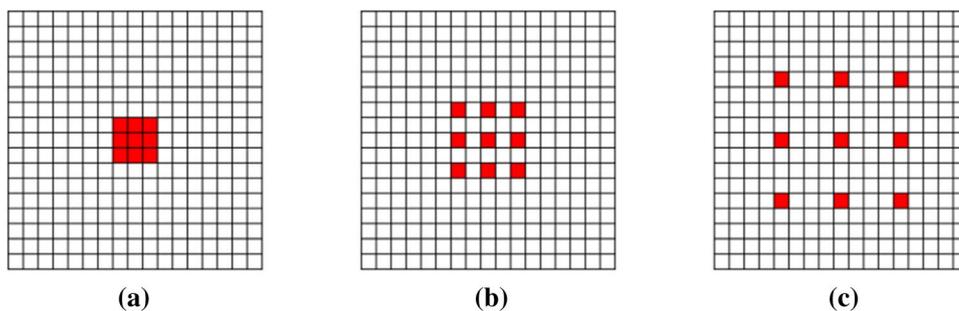
Step 1: Implement a 3×3 convolution on each input feature map to extract the low-level features. The convolution operation is defined by Eqs. (3) and (4), in which I is the input, V is the output, v_n is the convolution output of the n th convolution kernel, k_n is the n th convolution kernel, and I_s is the s th input.

$$F_{cov} : I \rightarrow V, I, V \in \mathbb{R}^{H \times W \times C} \quad (3)$$

$$v_n = k_n * I = \sum_{s=1}^C k_n * I_s \quad (4)$$

Step 2: Feed the initially extracted features into parallel dilated convolutional layers with rates of 2, 4, 6 and 8 to obtain multiscale context information. A dilated convolution is designed to insert holes into the standard convolution to expand the receptive fields. The dilated convolution can enlarge the receptive fields without information loss. Therefore, it is adopted in numerous semantic segmentation networks to replace the pooling layers. A schematic diagram of the dilated convolution receptive fields is shown in Fig. 4. The mapping relationship of the dilated convolution can be expressed by Eq. (5), where D is the

Fig. 4 Schematic diagram of the convolution receptive fields: **a** 3×3 convolution; **b** 3×3 dilated convolution, rate = 2; and **c** 3×3 dilated convolution, with rate = 4



dilated convolution output, v_n^d is the dilated convolution output of the n th dilated convolution kernel, k_n^d is the n th dilated convolution kernel, and v_{ns} is the s th input.

$$F_{cov}^d : V \rightarrow D, V, D \in R^{H \times W \times C} \tag{5}$$

$$v_n^d = k_n^d * v_n = \sum_{s=1}^C k_n^d * v_{ns}$$

Step 3: Perform global average pooling on the output feature maps of each dilated convolutional layer (Eq. (7), in which g_n represents the output of the n th global average pooling layer). By implementing global average pooling, the feature maps are squeezed into 4 vectors with C channels.

$$g_n = F_{gap}(v_n) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W v_n(i, j) \tag{7}$$

Step 4: Apply a 1×1 convolution to these 4 feature vectors for dimension reduction (Eq. (8), in which $G \in R^{1 \times 1 \times C}$ is the input of the 1×1 convolution and $L \in F_{cov}(G, w)$ is the output of the 1×1 convolution).

$$L = F_{cov}(G, w) \tag{8}$$

Step 5: Introduce a 2-layer dense neural network to acquire the channel weights of the initial feature maps. First, the 4 feature vectors are concatenated to form a feature vector with C channels. Second, the concatenated feature vector is fed into the dense neural network. Finally, the output of the fully connected neural network is generated by Eq. (9), in which the input is defined as x and the output is defined as a .

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_C \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdots & w_{1C} \\ w_{21} & w_{22} & w_{23} & \cdots & w_{2C} \\ w_{31} & w_{32} & w_{33} & \cdots & w_{3C} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{C1} & w_{C2} & w_{C3} & \cdots & w_{CC} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_C \end{bmatrix} \tag{9}$$

Step 6: Multiply the feature vector obtained in step 5 by the initial feature maps obtained in step 1 to generate weighted

feature maps (Eq. (10), in which $M \in R^{H \times W \times C}$ is the result of multiplication).

$$M(:, :, c) = V(:, :, c) \times a(c) \tag{10}$$

Step 7: Apply a residual connection to prevent information loss and network degradation (Eq. (11), in which $O \in R^{H \times W \times C}$ is the output of a DCA block).

$$O = M + V \tag{11}$$

Hyperparameters

Furthermore, the selection of hyperparameters is essential. In this work, an Adam [16] optimizer with an initial learning rate of 0.001 is used to train the network. When the loss value does not decrease after training for 3 consecutive epochs, the learning rate is reduced by half. In addition, early stopping is used to prevent overfitting; that is, when the loss value has not decreased for 10 consecutive epochs, training is stopped. In addition, the batch size and epoch number are set to 16 and 50, respectively.

Experimental Results and Discussion

DUDA-Net is programmed in Keras, and all the experiments are carried out on a server with 4 NVIDIA RTX 2080 Ti GPUs. In this work, the DSC, intersection over union (IoU), accuracy (ACC), sensitivity (SEN) and specificity (SPE) are introduced to verify the network performance (Eqs. (12) to (16)), where FN, FP, TN and TP are the numbers of false-negative, false-positive, true-negative and true-positive samples, respectively [17].

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{12}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{13}$$

$$ACC = \frac{TP + TN}{P + N} \tag{14}$$

Table 1 Results of different loss function experiments

Loss	DSC	IoU	ACC	SEN	SPE
WCE	79.33%	65.75%	99.14%	79.32%	99.56%
BCE	79.51%	66.04%	99.07%	88.55%	99.21%
DL	86.58%	76.35%	99.05%	86.56%	99.72%
GDL	85.00%	74.26%	98.98%	79.27%	99.85%
TL	87.06%	77.09%	99.06%	90.85%	99.59%

The bold figures in Table are the optimal performance

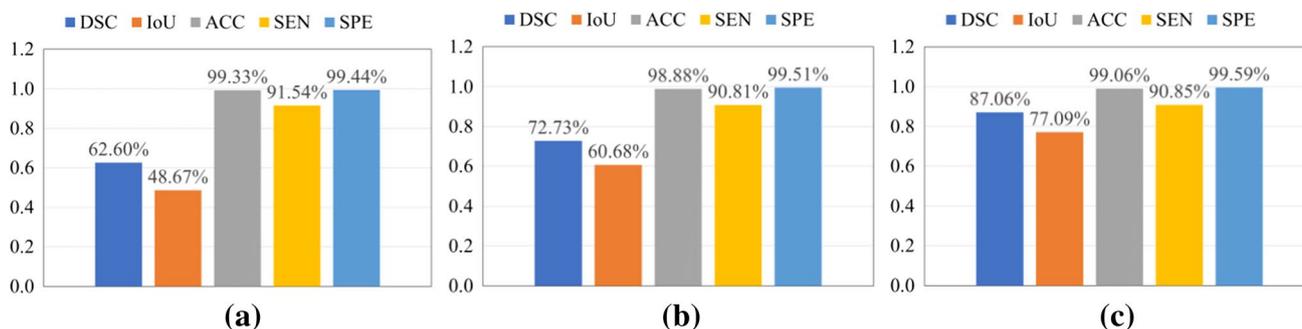


Fig. 5 Results of the ablation experiment: **a** DUDA-Net without coarse segmentation, **b** DUDA-Net without DCA blocks and **c** DUDA-Net

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (15)$$

$$\text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (16)$$

Loss Function Comparison

The selection of an appropriate loss function is significant after the construction of DUDA-Net. Generally, Dice loss (DL) is commonly applied in most image segmentation networks. However, in the COVID-19 lesion segmentation task, the proportion of lesions in the CT images is small, which can cause class imbalance problems. To avoid this problem, weighted cross-entropy (WCE) loss, balanced cross-entropy (BCE) loss, generalized DL (GDL) and Tversky loss (TL) are introduced. To determine the optimal loss function for COVID-19 segmentation tasks, the performances of DUDA-Net with different loss functions, namely the WCE loss, BCE loss, DL, GDL and TL, are compared. As indicated in Table 1, the accuracy of DUDA-Net with WCE loss is the best, as the accuracy can reach 99.14%. The GDL outperforms other loss functions in terms of the SPE, which reaches 99.85%. Moreover, the TL outperforms the other loss functions in terms of the DSC (87.06%), IoU (77.09%) and SEN (90.85%), and compared with those of the suboptimal loss function, the DSC, IoU and SEN of the TL are

improved by 0.48%, 0.74% and 2.3%, respectively. Since the ACC and SPE obtained by DUDA-Net with the TL are only 0.08% lower and 0.26% lower than those of DUDA-Net with the WCE and GDL, respectively, the TL is the optimal loss function for COVID-19 segmentation tasks.

Model Comparison

Furthermore, to verify that the coarse segmentation network and DCA blocks in the fine segmentation network can

improve the segmentation performance, two kinds of networks, namely DUDA-Net without coarse segmentation and DUDA-Net without DCA blocks, are constructed, and their performances are compared. As indicated in Fig. 5, the DSC, IoU, ACC, SEN and SPE of DUDA-Net without coarse segmentation reach 62.60%, 48.47%, 99.33%, 91.54% and 99.44%, respectively. By introducing coarse segmentation, the DSC, IoU and SPE are improved by 24.46%, 28.42% and 0.15%, respectively. In addition, the DSC, IoU, ACC, SEN and SPE of DUDA-Net without DCA blocks reach 72.73%, 60.68%, 98.88%, 90.81% and 99.51%, respectively. By introducing DCA blocks, these metrics are improved by 14.33%, 16.41%, 0.18%, 0.04% and 0.08%. As indicated by Fig. 6, the largest area under the receiver operating characteristic (ROC) curve (AUC) obtained by DUDA-Net reached 0.965. Compared with those of DUDA-Net without coarse segmentation and DUDA-Net with DCA blocks, the AUCs of DUDA-Net are improved by 0.238 and 0.051, respectively. Obviously, coarse segmentation can significantly improve the performance of the network, and when both coarse segmentation and DCA blocks are used at the same time, the network achieves the best segmentation performance.

Moreover, the segmentation results of the lesions are indicated in Fig. 7. Although DUDA-Net without the coarse segmentation network can segment some small lesions, there are disturbances from the nonpulmonary areas, and misjudgement occurs in some areas. In the case of DUDA-Net

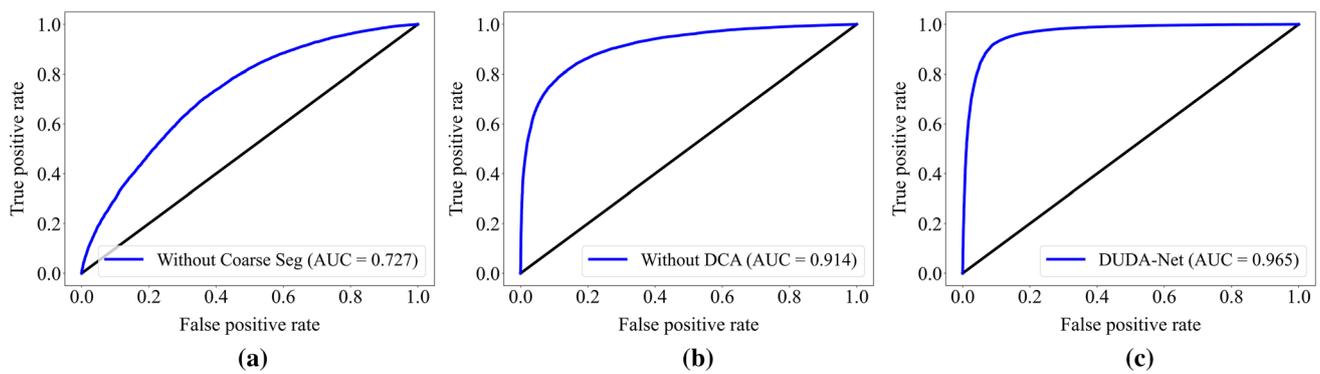
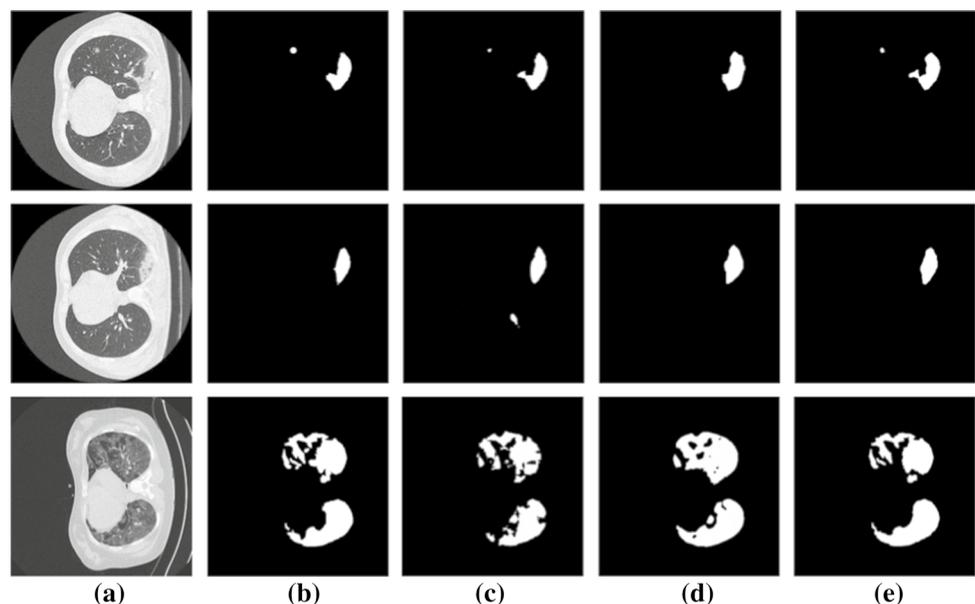


Fig. 6 ROC curve of the ablation experiment and the AUC indicator: **a** DUDA-Net without coarse segmentation, **b** DUDA-Net without DCA blocks and **c** DUDA-Net

Fig. 7 Prediction results of the ablation experiment: **a** the CT image, **b** the ground truth, **c** the results of DUDA-Net without coarse segmentation, **d** the results of DUDA-Net without DCA blocks and **e** the results of DUDA-Net



without using DCA blocks, the segmentation error of small lesions and boundaries is large. Moreover, compared with DUDA-Net without the coarse segmentation network and DCA blocks, DUDA-Net locates the lesion more accurately. The results indicate that the introduction of a coarse segmentation network and DCA blocks can contribute to removing the disturbances of the nonpulmonary areas and improving the segmentation performance of the small lesions.

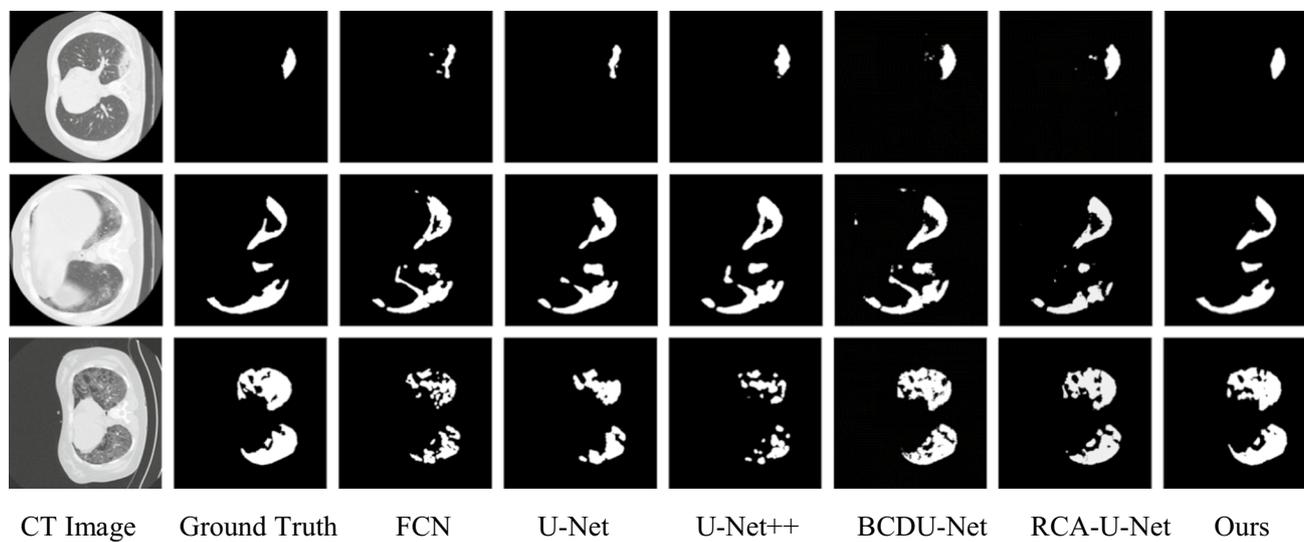
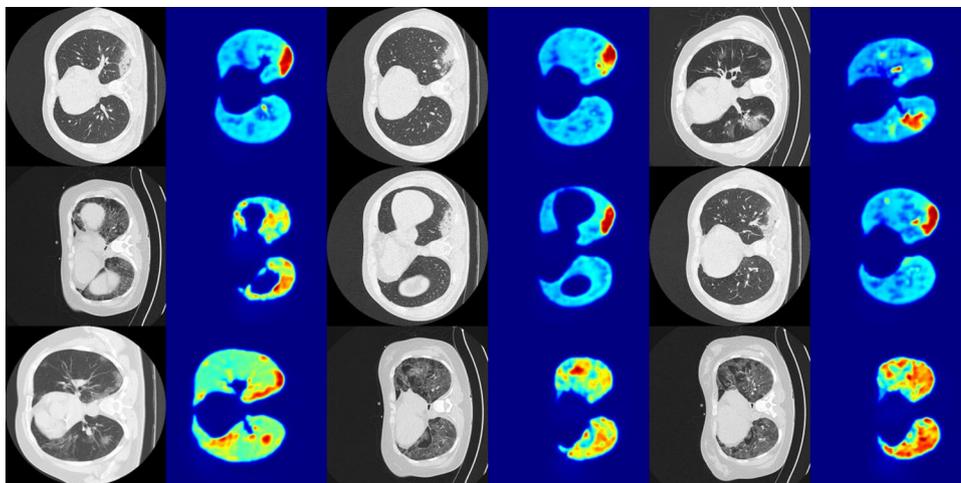
To further illustrate the superior performance of DUDA-Net, the performance of DUDA-Net is compared with that of several typical medical segmentation networks: a fully convolutional network (FCN), U-Net, U-Net + +, bidirectional convolutional long short-term memory U-Net with densely connected convolutions (BCDU-Net) and residual channel attention U-Net (RCAU-Net). As indicated in Table 2, DUDA-Net outperforms 5 other kinds of typical models in DSC, IoU, ACC and

SEN. In addition, compared with the suboptimal model, DUDA-Net can improve the DSC, IoU, ACC and SEN by 4.46%, 6.67%, 0.03% and 0.07%, respectively. Moreover, the prediction samples of these segmentation networks are shown in Fig. 8, and the results further verify that DUDA-Net outperforms other networks. The FCN and U-Net can precisely segment large lesions. However, the performance of these two models on small lesions is not ideal. Furthermore, the segmentation performance of U-Net + +, BCDU-Net and RCAU-Net is better than that of FCN and U-Net, but the error rates of these 3 models on boundaries are still very high. Compared with that of these five typical models, the overall performance of DUDA-Net on small lesions is better. In addition, the testing time of these methods is also provided. It takes 16.51 s for DUDA-Net to generate the prediction results for 55 testing samples. This indicates that the introduction of the coarse-to-fine scheme

Table 2 Results of different typical models

Method	DSC	IoU	ACC	SEN	SPE	Testing time
FCN [18]	52.96%	36.55%	98.12%	44.71%	99.62%	6.85 s
U-Net [19]	59.81%	42.90%	98.29%	51.97%	99.65%	7.54 s
U-Net+ + [20]	69.98%	54.51%	98.46%	65.74%	99.74%	9.99 s
BCDU-Net [21]	79.29%	65.69%	98.94%	90.78%	99.64%	11.29 s
RCA-U-Net [22]	82.60%	70.42%	99.03%	77.79%	99.77%	10.42 s
Ours	87.06%	77.09%	99.06%	90.85%	99.59%	16.51 s

The bold figures in Table are the optimal performance

**Fig. 8** Prediction results of each model**Fig. 9** Heat map of the DUDA-Net results

can cause an increase in computational complexity. In fact, compared with the efficiency, the proposed method focuses more on the segmentation precision. Therefore, DUDA-Net is still regarded as the optimal model with reasonable computational complexity.

Gradient-weighted class activation mapping (Grad-CAM) is applied to acquire the class activation maps of DUDA-Net. As shown in Fig. 9, the network model is more inclined to learn the features from the lesions during the training process.

Table 3 Comparison of DUDA-Net and several existing works

References	Years	Method	DSC	IoU	SEN	SPE
Zhou et al. [23]	2020	U-Net using an attention mechanism	83.1%	–	86.7%	99.3%
Omar et al. [24]	2020	Region of interest extraction segmentation network	78.6%	–	71.1%	99.3%
Qiu et al. [9]	2020	MiniSeg segmentation network	77.28%	82.12%	83.62%	97.42%
Ours	2021	DUDA-Net	87.06%	77.09%	90.84%	99.58%

The bold figures in Table are the optimal performance

In addition, the proposed DUDA-Net model is compared with several existing works on the same dataset. As indicated in Table 3, the proposed network outperforms the existing works in terms of the DSC, SEN and SPE. By introducing DUDA-Net, the DSC, SEN and SPE are improved by 8.46%, 4.14% and 0.28%, respectively. The results indicate that the proposed method can better achieve state-of-the-art segmentation performance. Zhou et al. [23] applied a single U-Net model with SE blocks as a channel attention mechanism. In fact, the SE blocks learn the channel weights by implementing global average pooling, which can lead to information loss; as a result, the channel weights learned by SE blocks are inaccurate. Compared with those of the original SE blocks, the channel weights learned by the DCA mechanism are more accurate, as multiscale context information is introduced by implementing parallel dilated convolution. In addition, Zhou et al. [23] directly segmented whole CT images, and disturbances from unrelated regions can result in poor segmentation performance. To address this issue, a coarse segmentation model is proposed in DUDA-Net to segment the lungs. Omar et al. [24] proposed a network to segment the lungs, which was followed by fine segmentation. However, the original images are concatenated with the lung images, and the disturbances from unrelated regions are preserved; as a result, the generalizability of the method in [24] is poor. Qiu et al. [9] proposed an attentive hierarchical spatial pyramid (AHSP) module for effective lightweight multiscale learning, but the lack of network parameters leads to low accuracy. Therefore, compared with that of current methods, the performance of DUDA-Net is better.

Conclusion

An automatic lesion segmentation system was developed for COVID-19 in this study. The highlights of the proposed system are as follows. (1) A coarse-to-fine segmentation scheme is introduced. To prevent disturbances from unrelated regions, lung areas are segmented by a coarse segmentation network, which is followed by a fine network to obtain the fine details of COVID-19 lesions. The experimental results indicate that the coarse-to-fine scheme can improve the DSC by 24.46%. (2) A DCA module is proposed, and parallel

dilated convolution layers are introduced to determine the significant channels with a multiscale receptive field; as a result, the accuracy of small lesions and boundaries is further improved. The experimental results indicate that the DCA mechanism can improve the DSC by approximately 14.33%. (3) DUDA-Net can achieve state-of-the-art performance, which indicates that the proposed method is of great clinical significance.

Although the proposed method can achieve precise segmentation, there are still some weaknesses, as follows. (1) The complex structure of DUDA-Net results in high computational complexity and low efficiency. (2) Accurate quantification of lung infection results requires further segmentation, such as ground glass shadows and pleural effusions. Therefore, our future work will reduce the computational complexity of DUDA-Net and collect more data to realize multicategory segmentation for COVID-19 lesions. For further research, we made the source code available at <https://github.com/AaronXieSY/DUDANet-for-COVID-19-lesions-Segmentation.git>.

Author's contribution All authors contributed equally to the revisions of this manuscript and approved the submission of this final version.

Funding This study is funded by the National Key R&D Program of China under grant 2017YFB1303003, the National Natural Science Foundation of China under grants 62073314 and 61821005, Youth Innovation Promotion Association of the Chinese Academy of Sciences under Grant 2019205, Program GQRC-19-20 and the Special Fund for High-level Talents (Shizhen Zhong Team) of the People's Government of Luzhou Southwestern Medical University, China Postdoctoral Science Foundation: 2020M670815.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval The article uses open-source datasets. All procedures in studies involving human participants were performed in accordance with the ethical standards of the institutional and/or national research committee and the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Nasirpour MH, Sharifi A, Ahmadi M, Ghouschi SJ (2021) Revealing the relationship between solar activity and COVID-19 and forecasting of possible future viruses using multi-step autoregression (MSAR). *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-021-13249-2>
- Hassantabar S, Ahmadi M, Sharifi A (2020) Diagnosis and detection of infected tissue of COVID-19 patients based on lung X-ray image using convolutional neural network approaches. *Chaos, Solitons Fractals* 140:110170. <https://doi.org/10.1016/j.chaos.2020.110170>
- Ucar F, Korkmaz D (2020) COVIDiagnosis-Net: Deep Bayes-SqueezeNet based Diagnostic of the Coronavirus Disease 2019 (COVID-19) from X-Ray Images. *Medical Hypotheses*. 140:109761
- Hofmann HS, Hansen G, Burdach S, Bartling B, Silber RE, Simm A (2004) Discrimination of human lung neoplasm from normal lung by two target genes. *Am J Respir Crit Care Med* 170(5):516–519. <https://doi.org/10.1164/rccm.200407-127OC>
- Ahmadi M, Sharifi A, Jafarian Fard M, Soleimani N (2021) Detection of brain lesion location in MRI images using convolutional neural network and robust PCA. *Int J Neurosci*. <https://doi.org/10.1080/00207454.2021.1883602>
- Ahmadi M, Sharifi A, Hassantabar S, Enayati S (2021) QAIS-DSNN: Tumor Area Segmentation of MRI Image with Optimized Quantum Matched-Filter Technique and Deep Spiking Neural Network. *Biomed Res Int*. <https://doi.org/10.1155/2021/6653879>
- Wu YH, Gao SH, Mei J, Xu J, Fan DP, Zhao CW, Cheng MM (2020) JCS: An explainable COVID-19 diagnosis system by joint classification and segmentation. *arXiv preprint arxiv:2004.07054*
- Chen X, Yao L, Zhang Y, (2020). Residual attention U-Net for automated multi-class segmentation of COVID-19 Chest CT Images. *arXiv preprint arXiv:2004.05645*
- Qiu Y, Liu Y, Xu J (2020) MiniSeg: An extremely minimum network for efficient COVID-19 Segmentation. *arXiv preprint arXiv:2004.09750*
- Amyar A, Modzelewski R, Li H, Ruan S (2020) Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. *Comput Biol Med* 126:104037. <https://doi.org/10.1016/j.compbimed.2020.104037>
- Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M (2020) Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:2006.11988*
- Senthikumar N, Thimmiraja J (2014) Histogram equalization for image enhancement using MRI brain images. In 2014 world congress on computing and communication technologies, IEEE. pp 80–83. <https://doi.org/10.1109/WCCCT.2014.45>
- Gowen AA, Downey G, Esquerre C, O'Donnell CP (2011) Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients. *J Chemom* 25(7):375–381. <https://doi.org/10.1002/cem.1349>
- Roelofs R, Shankar V, Recht B, Fridovich-Keil S, Hardt M, Miller J, Schmidt L (2019) A meta-analysis of overfitting in machine learning. In *Adv Neural Inf Process Syst* pp 9179–9189
- Ioffe S, Szegedy C (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*. PMLR. pp 448–456
- Jais IKM, Ismail AR, Nisa SQ (2019) Adam optimization algorithm for wide and deep neural network. *Knowl Eng Data Sci* 2(1):41–46. <https://doi.org/10.17977/um017v2i12019p41-46>
- Huang Z, Zhao Y, Li X, Zhao X, Liu Y, Song G, Luo Y (2020) Application of innovative image processing methods and Ada-Bound-SE-DenseNet to optimize the diagnosis performance of meningiomas and gliomas. *Biomed Signal Process Control* 59:101926. <https://doi.org/10.1016/j.bspc.2020.101926>
- Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp 3431–3440
- Ronneberger O, Fischer P, Brox T (2015) October. U-net: Convolutional networks for biomedical image segmentation. In: *International conference on medical image computing and computer-assisted intervention*, Springer, Cham. pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
- Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J (2018). Unet++: A nested u-net architecture for medical image segmentation. In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*, Springer, Cham. pp 3–11. https://doi.org/10.1007/978-3-030-00889-5_1
- Azad R, Asadi-Aghbolaghi M, Fathy M, Escalera S (2019) Bi-directional ConvLSTM U-net with Densley connected convolutions. In: *Proceedings of the IEEE international conference on computer vision workshops*. 0–0
- Fang Z, Chen Y, Nie D, Lin W, Shen D (2019) RCA-U-Net: Residual Channel Attention U-Net for Fast Tissue Quantification in Magnetic Resonance Fingerprinting. In: *International conference on medical image computing and computer-assisted intervention*, Springer, Cham. pp 101–109 https://doi.org/10.1007/978-3-030-32248-9_12
- Zhou T, Canu S, Ruan S (2020) An automatic covid-19 ct segmentation network using spatial and channel attention mechanism. *arXiv preprint arXiv:2004.06673*
- Elharrouss, O., Subramanian, N., Al-Maadeed, S., 2020. An encoder-decoder-based method for covid-19 lung infection segmentation. *arXiv preprint arXiv:2007.00861*

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.