International Journal of Innovative Computing, Information and Control
Volume **19**, Number **4**, August **2023**

ICIC International ©2023 ISSN 1349-4198
pp. **1209–1220**

ESR-NET: A NETWORK FOR SEGMENTING COVID-19 LUNG INFECTION REGIONS IN CT IMAGES

Jianfei Zhang[1,2,*], Shutao Wu[1,2] and Bo Wang[1,2]

[1]College of Computer and Control Engineering
[2]Heilongjiang Key Laboratory of Big Data Network Security Detection and Analysis
Qiqihar University
No. 42, Wenhua Street, Qiqihar 161006, P. R. China
solocica@126.com; bowanggdr@qqhru.edu.cn
*Corresponding author: zhangjianfei@qqhru.edu.cn

Received October 2022; revised February 2023

Abstract. *Since the outbreak of coronavirus disease 2019 (COVID-19), it has spread rapidly around the world, posing a huge challenge to public health. Computed tomography (CT) has played an extremely important role in the diagnosis of COVID-19. The automatic segmentation of infected areas in CT images using computer-aided diagnosis techniques is of great significance for the diagnosis of COVID-19. This paper proposes a novel segmentation deep network, ESR-Net, for segmenting COVID-19 infected regions in chest CT images. In ESR-Net, a pre-trained ResNet is used to enhance the feature extracting process. Then edge attention helps improve the boundary of infected region, side output filter and reverse attention work together aimed to improve the detail performance in the segmentation results. Our method is tested on a dataset which contains 1939 chest CT slices. Among them, 1551 CT volumes are used for training and 388 CT volumes are used for testing. The results show ESR-Net achieving 83.1% Dice coefficient, 83.8% sensitivity, and 84.3% F-measure, outperforms other existing methods.*
Keywords: COVID-19, Computed tomography, Infection segmentation, Deep learning

1. **Introduction.** Since the outbreak of COVID-19 in December 2019, the number of infected people around the world has risen rapidly due to the strong contagiousness and variability of the virus. According to statistics from Johns Hopkins University, 473.95 million people have been infected and 6.09 million people have died as of 10:00 on March 23, 2022. In recent studies, it has also been demonstrated that the disease can still have sequelae with different symptoms even after cure. As the COVID-19 virus continues to mutate, its contagiousness has increased rapidly, which has brought unprecedented challenges to the global epidemic prevention and control work.

At present, the main method to control the spread of the epidemic is to diagnose the disease in time. The widely used detection method is reverse transcription polymerase chain reaction (RT-PCR) [1], which is the gold standard for COVID-19 detection. However, this method is easily affected by the quality of test kits, and due to the huge number of patients, the stock of test kits in some regions is not enough to support large-scale testing. Recent studies have shown that even when the quality of the test kit is guaranteed, this method has a high false-positive rate in the detection of patients with early COVID-19 infection. However, the imaging features of COVID-19 are easy to identify even in the early stage of the disease, so using CT technology can diagnose the disease accurately even in the early stage of COVID-19 infection. Therefore, using CT technology as a complement to RT-PCR can greatly improve the accuracy of COVID-19 detection.

DOI: 10.24507/ijicic.19.04.1209

However, there are also many problems with the manual methods to detect COVID-19 based on CT technology. For example, a complete CT scan will generate a large workload due to the hundreds of CT slice images; the shortage of radiologists and the subjective factors of the interpreters will easily affect the final detection results. Thanks to the development of computer-aided diagnosis technology, especially the application of deep learning methods in the medical field, the above problems have been effectively solved. Recently, there have been many studies on the detection of COVID-19 based on medical images using deep learning methods. Shi et al. [2] have presented a system using ResNet-50 to identify COVID-19 automatically. Wang et al. [3] have proposed a novel CNN architecture named COVID-Net to classify COVID-19 among healthy people, COVID-19 patients and other pneumonia patients. Cao et al. [4] and Gozes et al. [5] have used U-Net to measure the progression of disease overtime. Jin et al. [6] have develop a system based on UNet+ and CNN which can highlight the lesion region. Because of the low interpretability of deep learning method, the traditional classification network is hard to apply in the clinical field. Using segmentation network to label the infection region is becoming a trend. The COVID-19 infection regions always have complex shape and the boundaries between lesions and normal tissues are hard to distinguish; these problems increase the difficulty of segmentation models.

In order to solve these problems, this paper presents a novel segment network named ESR-Net. This network is based on the structure of U-Net, which is widely used in the field of medical image segmentation. As we all know, U-Net consists of an encoder and a decoder, and ESR-Net's design strategy is to optimize these two parts and add new modules to improve the final segmentation performance. In the encoder part, we use the pre-trained ResNet-34 as the backbone, and we designed an edge attention module to take full advantage of the features extracted by the encoder to add edge supervision to the network to improve the boundary details, and in the decoder part, we designed side output filter module and reverse attention to further improve the infection region details. We test it on the latest open-source dataset, and design comparative experiments to verify the performance of ESR-Net in the task of COVID-19 infection region segmentation, and the results show ESR-Net performs well on this task.

2. **Methodology.**

2.1. **Overview of network.**

The encoding-decoding structure of U-Net [7] has shown good performance in the field of medical image segmentation. Therefore, the main structure of ESR-Net is based on U-Net, which consists of an encoder and a decoder. The overall structure of the network is shown in Figure 1.

The encoder part uses the ResNet-34 classification network [8] (without the average pooling layer and the fully connected layer) pre-trained on ImageNet [9] as the backbone. Transfer learning is used to improve the feature extraction ability of the encoder and to reduce the degradation of segmentation results due to insufficient samples in the dataset. After the feature extraction of CT images through the backbone, considering that the shallow features occupy more computing resources due to their larger spatial resolution and the performance improvement for the segmentation network is very small [10], we discard the shallow features. The last four layers of features are used as the initial input of the decoder and the edge attention module. The edge attention module is responsible for extracting the edge information of the infection region and the decoder extracts the infection regions information under the guidance of the side output filter and the reverse attention. The final result is obtained by modifying the output of the decoder with the output of the edge attention.

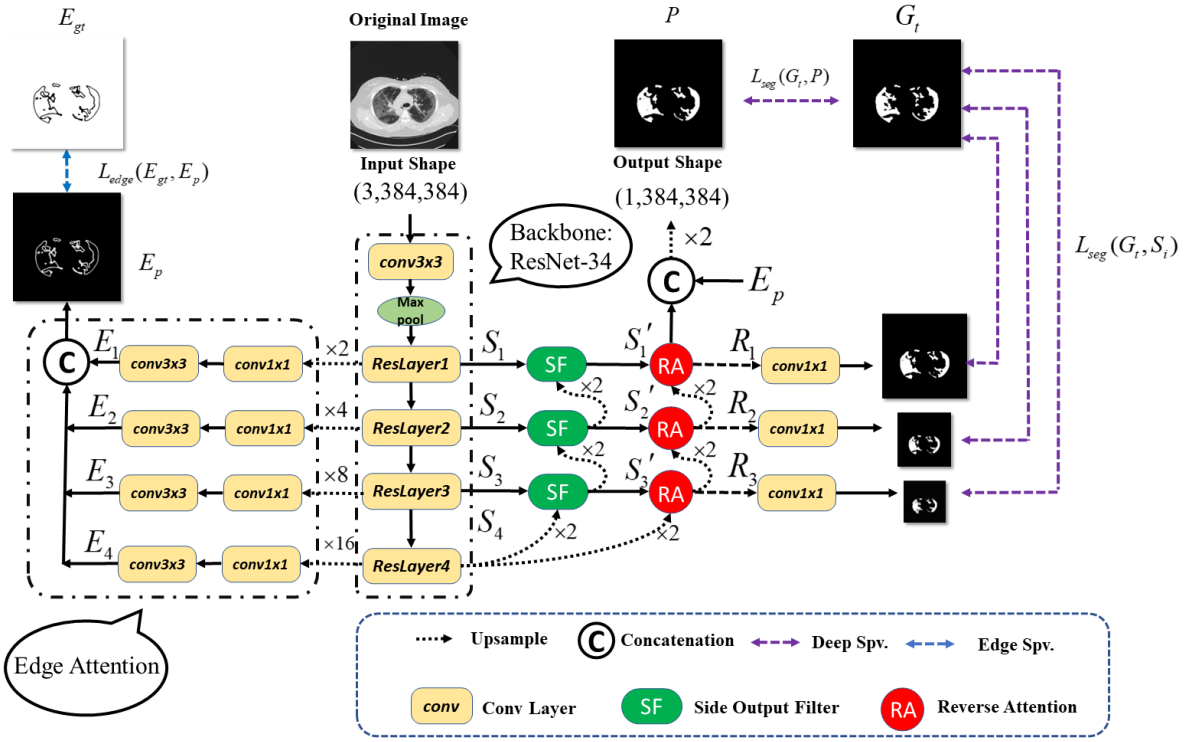

FIGURE 1. The architecture of ESR-Net

2.2. **Edge attention.** Compared with region-level based methods, object-level based salient object detection methods tend to show better performance. However, this will inevitably sacrifice the spatial coherence in the image, resulting in unsatisfactory performance of the edge segmentation of the target area. In recent years, a large number of studies [11,12] have shown that edge information can provide additional and effective constraints for image segmentation to guide the segmentation network to obtain better segmentation results. Inspired by this, we design an edge attention module to extract edge features. Different from the existing edge attention module that generates edge features in a specific layer in the backbone, the edge attention module designed in this paper extracts features from different depth networks in parallel. The edge features thus obtained can not only fully use the spatial information of the shallow features, but also improve the details under the guidance of the deep features. The edge attention works as this: first, the features extracted by the backbone network of each layer depth are up-sampled to the size of the ground truth; second, the edge features of each layer are obtained through the convolution layer. This process is shown in Equation (1).

$$E_i = conv_{3\times3}\left(conv_{1\times1}\left(up\left(S_i\right)\right)\right) \tag{1}$$

where $up(\cdot)$ represents the up-sampling operation, and then the features of each branch are concatenated. The parallel features obtained after concatenation are adjusted by a convolution layer with a convolution kernel size of 3×3, and finally the edge features are obtained. The process is shown in Equation (2).

$$E_p = conv_{3\times3}(Cat(E_i)) \tag{2}$$

where $Cat(\cdot)$ represents the concatenation operation and the edge feature E_p is supervised by the edge ground truth E_G. In this study, the edge ground truth is obtained from the ground truth after processing via the Sobel operator.

2.3. **Side output filter.** In the traditional segmentation model, the output of the decoder not only contains the information obtained by the up-sampling of the deep features, but also is often concatenated with the features obtained by the same-level encoder to enrich the diversity of the features. However, a simple concatenation operation will also produce a lot of redundant information, which will have a negative impact on the segmentation accuracy. The features are given trainable weights and the key features are screened out through the weights to provide guidance for the decoding work.

The side output filter works as this: side output S_i adds the deep feature S_{i+1} element by element with S_i after up-sampling, non-linearizes the ReLU function, going through a 1×1 convolution layer, and then uses the sigmoid function to obtain the weights. Finally, the weight is given to S_i to obtain the filtered side output S_i'. The specific implementation of this process is expressed by Equation (3).

$$S_i' = sigmoid\left(conv_{1\times1}\left(\sigma\left(conv_{1\times1}\left(up\left(S_{i+1}\right)\right) \oplus conv_{1\times1}\left(S_i\right)\right)\right)\right) \tag{3}$$

where $\oplus$ is the element addition operation corresponding to the pixel value and σ is the nonlinear function ReLU.

2.4. **Reverse attention.** In the field of object detection, in order to capture richer detailed features, additional supervision mechanisms are often introduced into the network. Reverse attention is an effective supervision method [13,14]. Since the backbone of conventional saliency detection networks is essentially fine-tuned from image classification networks, image classification networks often only respond to small and sparse regions, which often cannot meet the requirements of saliency detection tasks. The reverse attention mechanism can integrate deep features and side output features to gradually expand the target area, thus effectively solving this problem. The implementation process is that the deep features extracted from the backbone network are used as input, and the currently predicted regions are removed from the side output features to guide the entire network to discover complementary object regions and details at one time, thereby improving the segmentation performance of the network. This process can be summarized as Equation (4):

$$R_i = S_i' \otimes (E - sigmoid(up(S_{i+1}))) \tag{4}$$

where $\otimes$ is the multiplication operation of the corresponding elements of the pixel and E represents the identity matrix of the same shape as S_i.

The structure of this module is given in Figure 2.

2.5. **Loss function.** Because of the huge difference between the edge map (E_p) and the predict map (P) (see Figure 1), neither Dice loss nor BCE loss can lead segmentation model to a good performance, so we propose a combined loss function as follows:

$$L_{seg} = \alpha L_{Dice} + (1 - \alpha) L_{BCE} \tag{5}$$

where α is the weight coefficient, and in this research, α is set to 0.5. With the help of combined loss function, the segmentation task can be supervised both at the region level and at the pixel level. Furthermore, ESR-Net also adds deep supervision to three side outputs, which can improve the accuracy effectively. The final loss function is shown in Equation (6):

$$L_{total} = L_{seg}\left(G_t, P\right) + L_{seg}\left(E_{gt}, E_p\right) + \sum_{i=1}^{3} L_{seg}\left(G_t, S_i\right) \tag{6}$$

where $L_{seg}\left(G_t, P\right)$ represents the loss of the final result of supervision, $L_{seg}\left(E_{gt}, E_p\right)$ represents the loss of edge supervision and $\sum_{i=1}^{3} L_{seg}\left(G_t, S_i\right)$ represents the loss of three side outputs.

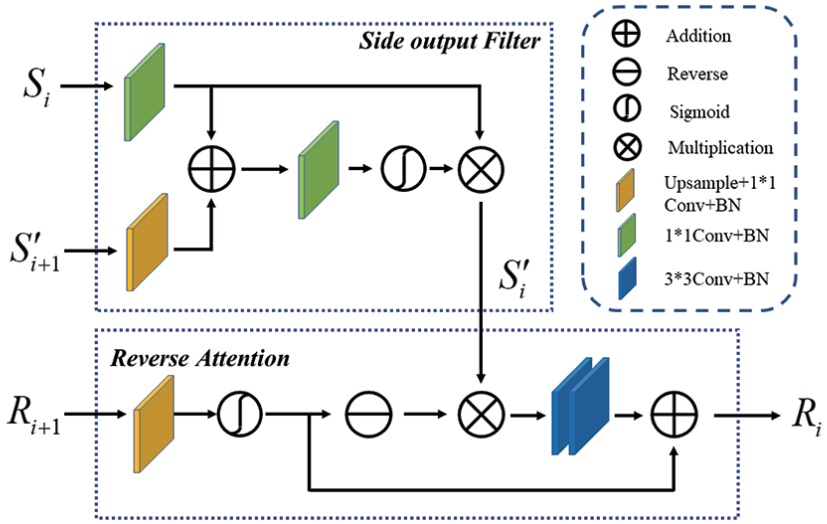

FIGURE 2. The architecture of the side output filter and the reverse attention

3. Experiments.

3.1. **Dataset.** The dataset used in this paper consists of two open-source datasets. The first dataset has a total of 100 2D CT images, all collected by the Italian Society of Medicine and Interventional Radiology [15]. The images in this dataset were marked by radiologists in the lung CT images infected with COVID-19, and the infection regions were divided into ground glass shadow areas and lung consolidation areas according to the degree of infection. This study focuses on the segmentation of COVID-19 infection areas, so these two types of infection regions in the dataset are regarded as the target area. The second dataset comes from [16], which contains 20 3D CT images of COVID-19 lesions. A total of 3520 2D CT images are obtained by slicing all 3D CT images, of which 1839 CT images contain COVID-19 infection. These images which contain infected regions are used in this study, and the rest of the CT images are not used.

3.2. **Implementation details.** This study is carried out on a platform equipped with a Tesla P100 graphics card. The model was built via PyTorch. All input images are rescaled to 384×384. Recent research [17] shows using suitable way to pre-preprocess the CT volumes can highly improve the accuracy of COVID-19 diagnosis. The data augmentation strategy used in this study is a multi-scale training strategy. By re-shaping the images in three different ratios, $\{0.75, 1, 1.25\}$, then training the model with the enhanced dataset, this increases the number of samples and reduces the risk of overfitting. The training process uses the Adam optimizer, the learning rate is set to 1e-4, the batch size is set to 8, and the number of training epochs is set to 60.

3.3. **Evaluation metrics.** The evaluation metrics include the Dice similarity coefficient, sensitivity, and F-measure widely used in segmentation models. The formula for the Dice similarity coefficient is as follows:

$$Dice(P, G) = \frac{2|P \cap G|}{|P| + |G|} \tag{7}$$

where P represents the model prediction result and G represents the ground truth. The sensitivity is defined as

$$Sensitivity = \frac{TP}{TP + FP} \tag{8}$$

where TP stands for true positives and FP stands for false positives. Generally speaking, the sensitivity represents the proportion of positive (COVID-19 infected area) pixels that are correctly predicted in the prediction map to the total positive pixels. The F-measure is defined as

$$F_\beta = \frac{(1+\beta^2)\,\text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}} \tag{9}$$

where β is set to 0.3 according to the recommendation in [18].

In addition, three evaluation metrics commonly used in object detection are also used in the research: Structural Measure [19], Enhanced-alignment Measure [20], and Mean Absolute Error.

The Structural Measure formula is as follows:

$$S_\alpha = (1-\alpha) \times S_o(P,G) + \alpha \times S_r(P,G) \tag{10}$$

This metric is used to measure the structural similarity between the prediction map and the ground-truth map, S_o represents the object-level similarity, S_r represents the region-level similarity, and α is the balance coefficient, which is set to 0.5 in this study.

The Enhanced-alignment Measure formula is as follows:

$$E_\emptyset = \frac{1}{w \times h} \sum_x^w \sum_y^h \emptyset(P(x,y), G(x,y)) \tag{11}$$

This metric is mainly used to measure the local and global similarity between two binary images, where w and h represent the width and height of the ground truth G, and x, y represent the coordinates of each pixel in G, and $\emptyset$ represent the enhanced alignment matrix.

The Mean Absolute Error formula is as follows:

$$\text{MAE} = \frac{1}{w \times h} \sum_x^w \sum_y^h |P(x,y) - G(x,y)| \tag{12}$$

This metric is mainly used to measure the pixel-level error between P and G.

4. **Results.**

4.1. **Segmentation results.**

The research first compares the performance of our method with other medical segmentation methods on the dataset used in this paper. The methods used in the comparison include U-Net [7], U-Net++ [21], Attention-UNet [22], and Dense-UNet [23], which are compared qualitatively and quantitatively.

4.1.1. *Qualitative results.* By comparing the infection regions segmentation results of each method in Figure 3, it is obvious that there are many false positive and false negative regions in the segmentation results of U-Net, U-Net++, and Attention-UNet. Although Dense-UNet has improved, it still has a large discrepancy with the ground-truth map in terms of details. It is clear that there are many false positive areas in these four methods, which greatly affects the segmentation results. Under the guidance of the edge attention of ESR-Net, the shape of the infection regions of the final segmentation result is very close to the ground-truth map. With the help of the side output filter and the reverse attention, the details inside the infection regions are also closer to the ground truth.

4.1.2. *Quantitative results.* It can be seen from Table 1 that the performance of using U-Net for the infection area segmentation task is not good, and only the sensitivity can reach a high level. The performance of U-Net++ and Attention-UNet is slightly improved

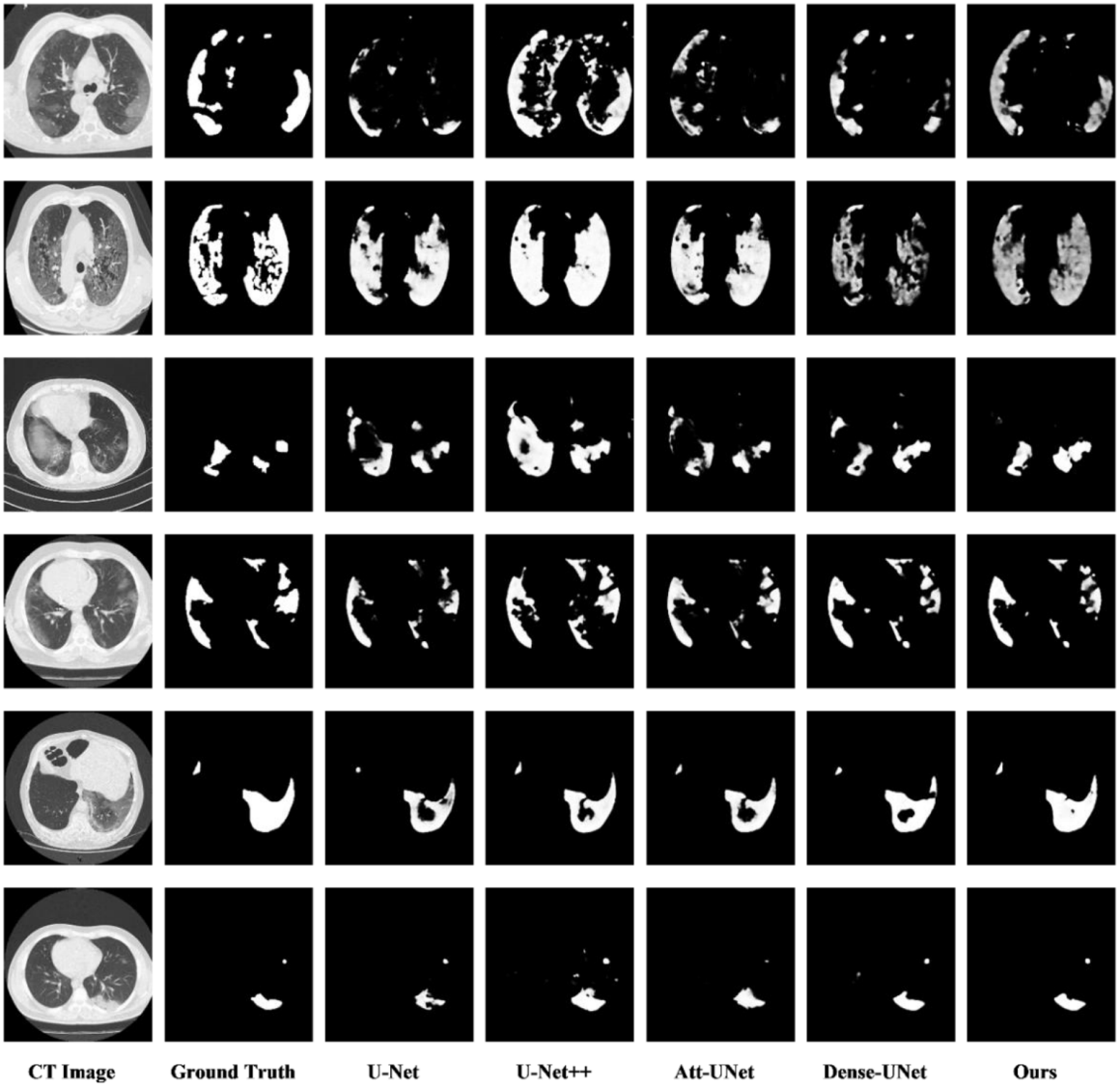

FIGURE 3. Visual comparison of the results of different methods for segmenting lung infection regions

TABLE 1. Quantitative segmentation results for different methods

Method	*Dice*	*Sen*	*mean-F*	S_α	$E_\emptyset$	*MAE*
U-Net	0.649	0.817	0.643	0.771	0.839	0.013
UNet++	0.751	0.801	0.774	0.822	0.915	0.011
Attention-UNet	0.759	0.755	0.793	0.836	0.932	0.009
Dense-UNet	0.814	0.815	0.827	0.861	0.957	0.007
Ours	0.831	0.838	0.843	0.882	0.964	0.006

but the Dice coefficient is still at a low level. The performance of Dense-UNet and ESR-Net has been greatly improved, which shows that the method of introducing transfer learning in the encoder can improve stably the feature extraction level, thereby improving the final segmentation quality. Compared with Dense-UNet, the ESR-Net approach proposed in this paper has a further improvement in various measures. From Figure 4 and Figure 5, it can also be seen that the overall performance of ESR-Net is better, which means that the sub-modules in ESR-Net are indeed effective.

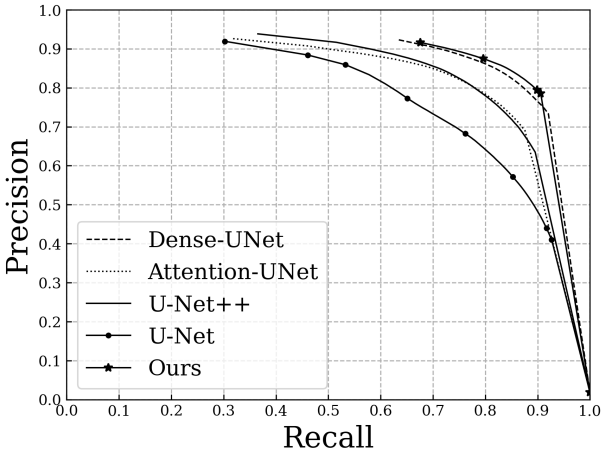

FIGURE 4. P-R curve of segmentation with different methods

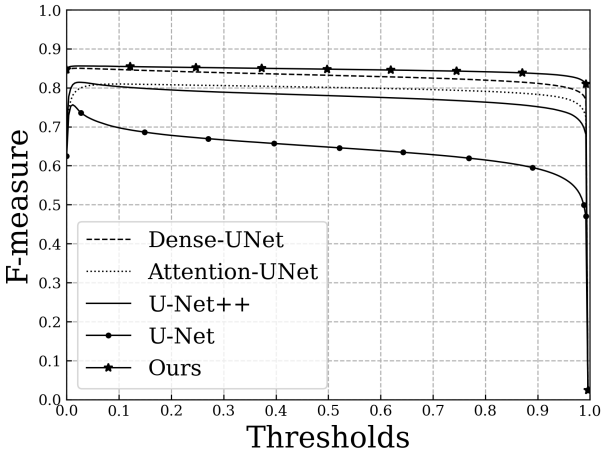

FIGURE 5. F-measure curve with different methods

4.2. **Ablation study.** In this section, we will verify the three key modules in the ESR-Net proposed in this paper: the edge attention, the side output filter, and the reverse attention. The effect of the combination of these modules on network performance is also investigated. The visualization and numerical results of the ablation experiment are shown in Figure 6, Figure 7, Figure 8 and Table 2.

4.2.1. *Effectiveness of edge attention.* From the results of experiments a and b in Table 2, it is clear that the characteristics and sensitivity of segmentation are greatly improved after the edge attention is introduced into the model. From experiments e and f, it can also be found that the edge attention cooperates with the side output filter and the reverse attention to improve the segmentation indicators to a certain extent. Observing the results of experiment b, it can be seen that the edge of the infection region is clearer and smoother than the segmentation results of the backbone after using the edge attention, which proves that the edge attention can effectively improve the performance of the network.

4.2.2. *Effectiveness of side output filter.* From experiments a and c in Table 2, it can be observed that the overall performance of the model is significantly improved by the side output filter, and the five indicators in the table have been greatly improved. According to the results of experiments d and e, it can be seen that the side output filter assisting the reverse attention can effectively improve the sensitivity of the model. However, after

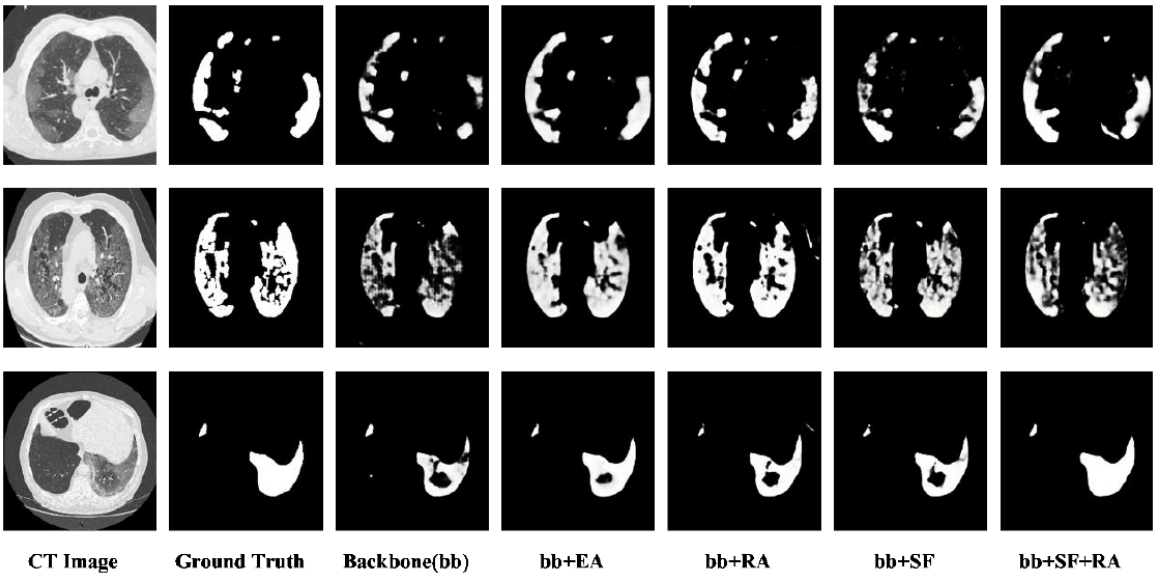

FIGURE 6. Visual comparison of the results with different modules for lung segmentation

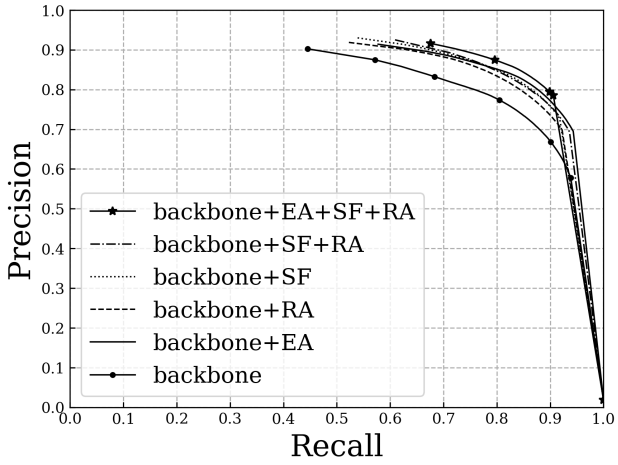

FIGURE 7. P-R curve of segmentation with different modules

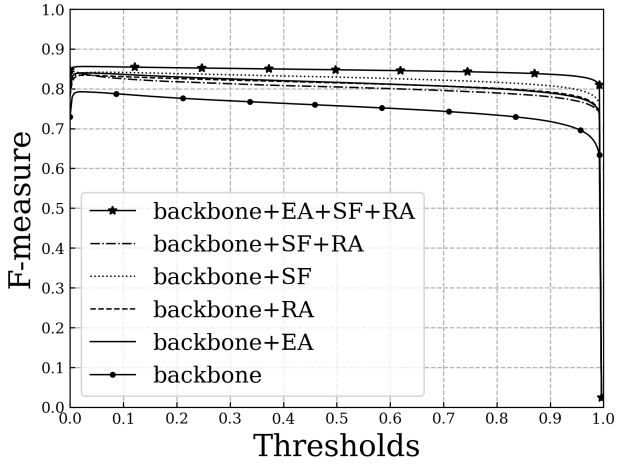

FIGURE 8. F-measure curve with different modules

TABLE 2. Quantitative segmentation results for different modules

Method	*Dice*	*Sen*	*mean-F*	S_α	$E_\emptyset$	*MAE*
a. backbone	0.761	0.818	0.751	0.837	0.929	0.009
b. backbone+EA	0.746	0.894	0.811	0.807	0.908	0.007
c. backbone+SF	0.812	0.827	0.823	0.868	0.957	0.007
d. backbone+RA	0.799	0.827	0.809	0.856	0.949	0.008
e. backbone+SF+RA	0.821	0.858	0.800	0.872	0.956	0.008
f. backbone+SF+RA+EA	0.830	0.843	0.843	0.880	0.963	0.007

using this module, the model shows that it is extremely insensitive to edge features. Even with the help of the reverse attention, it is difficult to overcome this shortcoming, which is also the motivation for designing the edge attention in the research.

4.2.3. *Effectiveness of reverse attention.* From the results of experiments a and d in Table 2, it is obvious that the reverse attention makes a large contribution to the improvement of various indicators of the model, especially the improvement of the Dice measure. From the visualization results of the segmentation, it can also be found that using this module can effectively improve the details of the segmentation, and the results are closer to the ground-truth map at the pixel level. This demonstrates that the reverse attention plays a role, but it can also be found that there are some false positive areas in the segmentation results after using the reverse attention module, which is also one of the inspirations for the improvement of the model structure in this study.

4.2.4. *Effectiveness of side output filter and reverse attention.* From experiments a and e in Table 2, it can be observed that adding a combination of the side output filtering and the reverse attention in the backbone can effectively improve the performance of the network. Compared with using these two modules alone, the improvement of using them at the same time is more significant. This can be seen in experiments c, d, and e, and it is also obvious from the visualization results that the results of using two modules in combination are much better than using these two modules alone.

5. **Conclusions.** This paper proposes an innovative segmentation deep network, ESR-Net, to provide visible evidence of COVID-19 diagnosis. With the help of side output filter, edge attention and reverse attention, ESR-Net can provide clear boundaries and plentiful details on COVID-19 infection region segmentation task. The model was applied to the latest COVID-19 lung CT dataset and achieved excellent segmentation results. There are still the following problems in the current research. First, this model cannot distinguish the features of COVID-19 infection tissues at different stages. Second, this model can only segment 2D images, while CT images themselves are a set of 3D images, spatial information is lost in the process of slicing it into a 2D image. Future research will focus on developing a segmentation network that can distinguish different infection characteristics to judge the stage of COVID-19 infection.

Acknowledgment. This work was supported by Disciplinary Collaborative Innovation Project (LJGXCG2022-124), Basic Scientific Research Business Fund Research Projects of Colleges and Universities in Heilongjiang Province (145209806).

REFERENCES

[1] T. Ai, Z. Yang, H. Hou et al., Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases, *Radiology*, 2020.

[2] F. Shi, J. Wang, J. Shi et al., Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19, *IEEE Reviews in Biomedical Engineering*, vol.14, pp.4-15, 2020.

[3] L. Wang, Z. Q. Lin and A. Wong, COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images, *Scientific Reports*, vol.10, no.1, pp.1-12, 2020.

[4] Y. Cao, Z. Xu, J. Feng et al., Longitudinal assessment of COVID-19 using a deep learning-based quantitative CT pipeline: Illustration of two cases, *Radiology Cardiothoracic Imaging*, vol.2, no.2, e200082, 2020.

[5] O. Gozes, M. Frid-Adar, H. Greenspan et al., Rapid AI development cycle for the coronavirus (COVID-19) pandemic: Initial results for automated detection & patient monitoring using deep learning CT image analysis, *arXiv Preprint*, arXiv: 2003.05037, 2020.

[6] B. Wang, S. Jin, Q. Yan et al., AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system, *Applied Soft Computing*, vol.98, 106897, 2021.

[7] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp.234-241, 2015.

[8] K. He, X. Zhang, S. Ren et al., Deep residual learning for image recognition, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.770-778, 2016.

[9] O. Russakovsky, J. Deng, H. Su et al., ImageNet large scale visual recognition challenge, *International Journal of Computer Vision*, vol.115, no.3, pp.211-252, 2015.

[10] Z. Wu, L. Su and Q. Huang, Cascaded partial decoder for fast and accurate salient object detection, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.3907-3916, 2019.

[11] Z. Zhang, H. Fu, H. Dai et al., ET-Net: A generic edge-attention guidance network for medical image segmentation, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp.442-450, 2019.

[12] J. X. Zhao, J. J. Liu, D. P. Fan et al., EGNet: Edge guidance network for salient object detection, *Proc. of the IEEE/CVF International Conference on Computer Vision*, pp.8779-8788, 2019.

[13] S. Chen, X. Tan, B. Wang et al., Reverse attention for salient object detection, *Proc. of the European Conference on Computer Vision (ECCV)*, pp.234-250, 2018.

[14] D. P. Fan, T. Zhou, G. P. Ji et al., INF-Net: Automatic COVID-19 lung infection segmentation from CT images, *IEEE Transactions on Medical Imaging*, vol.39, no.8, pp.2626-2637, 2020.

[15] *COVID-19 CT Segmentation Dataset*, https://medicalsegmentation.com/covid19/, Accessed on Apr. 11, 2020.

[16] J. Ma, Y. Wang, X. An et al., Toward data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation, *Medical Physics*, vol.48, no.3, pp.1197-1210, 2021.

[17] K. R. Ummah, T. Karlita, R. Sigit, E. M. Yuniarno, I K. E. Purnama and M. H. Purnomo, Effect of image pre-processing method on convolutional neural network classification of COVID-19 CT scan images, *International Journal of Innovative Computing, Information and Control*, vol.18, no.6, pp.1895-1912, 2022.

[18] A. Borji, M. M. Cheng, H. Jiang et al., Salient object detection: A benchmark, *IEEE Transactions on Image Processing*, vol.24, no.12, pp.5706-5722, 2015.

[19] D. P. Fan, M. M. Cheng, Y. Liu et al., Structure-measure: A new way to evaluate foreground maps, *Proc. of the IEEE International Conference on Computer Vision*, pp.4548-4557, 2017.

[20] D. P. Fan, C. Gong, Y. Cao et al., Enhanced-alignment measure for binary foreground map evaluation, *arXiv Preprint*, arXiv: 1805.10421, 2018.

[21] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh et al., UNet++: A nested U-Net architecture for medical image segmentation, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp.3-11, 2018.

[22] O. Oktay, J. Schlemper, L. L. Folgoc et al., Attention U-Net: Learning where to look for the pancreas, *arXiv Preprint*, arXiv: 1804.03999, 2018.

[23] X. Li, H. Chen, X. Qi et al., H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes, *IEEE Transactions on Medical Imaging*, vol.37, no.12, pp.2663-2674, 2018.

Author Biography

Jianfei Zhang received the Ph.D. degree in computer application technology from Yanshan University, Qinhuangdao, China, in 2015. He is currently a Professor with the College of Computer and Control Engineering, Qiqihar University. His research interests include Bayesian networks and deep learning.

Shutao Wu received the B.S. degree from Shanghai University, in 2013. He is currently the M.S. student with Qiqihar University. His research interests include deep learning and medical image processing.

Bo Wang received the Ph.D. degree in computer science and technology from Harbin Engineering University, in 2021. He is currently a Professor with the College of Computer and Control Engineering, Qiqihar University. His research interests include biomedical big data analysis and processing and LncRNA-disease association prediction.