



SMA-Net: Sobel Operator Combined with Multi-attention Networks for COVID-19 Lesion Segmentation

Fangfang Lu^(✉), Chi Tang^(✉), Tianxiang Liu, and Zhihao Zhang

College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai, China

lufangfang@shiep.edu.cn, tangchi97100163.com,
{liutianxiang,zhangzhihao}@mail.shiep.edu.cn

Abstract. Coronavirus disease 2019 (COVID-19) has been spreading since late 2019, leading the world into a serious health crisis. To control the spread rate of infection, identifying patients accurately and quickly is the most crucial step. Computed tomography (CT) images of the chest are an important basis for diagnosing COVID-19. They also allow doctors to understand the details of the lung infection. However, manual segmentation of infected areas in CT images is time-consuming and laborious. With its excellent feature extraction capabilities, deep learning-based method has been widely used for automatic lesion segmentation of COVID-19 CT images. But, the segmentation accuracy of these methods is still limited. To effectively quantify the severity of lung infections, we propose a Sobel operator combined with Multi-Attention networks for COVID-19 lesion segmentation (SMA-Net). In our SMA-Net, an edge feature fusion module uses Sobel operator to add edge detail information to the input image. To guide the network to focus on key regions, the SMA-Net introduces a self-attentive channel attention mechanism and a spatial linear attention mechanism. In addition, Tversky loss function is adopted for the segmentation network for small size of lesions. Comparative experiments on COVID-19 public datasets show that the average Dice similarity coefficient (DSC) and joint intersection over Union (IOU) of proposed SMA-Net are 86.1% and 77.8%, respectively, which are better than most existing neural networks used for COVID-19 lesion segmentation.

Keywords: Lesion segmentation · Deep learning · COVID-19 · Multi-attention

1 Introduction

Coronavirus disease 2019 (COVID-19) is an epidemic disease caused by a new coronavirus (formerly known as 2019 nCoV). This new coronavirus has strong adaptability, so far, it has produced eleven different mutant strains [1]. According to the latest statistics from the Johns Hopkins Center for Systems Science and Engineering (CSSE) (updated October 8, 2022), the number of confirmed COVID-19 cases worldwide has reached 621 million, including 6.56 million

deaths. Currently, reverse transcription-polymerase chain reaction (RT-PCR) is the standard test for diagnosing COVID-19 disease [2]. However, the RT-PCR test has the possibility of false negatives when the nucleic acid content of the new coronavirus is too low in the test sample. The missed diagnosis cases caused by false negatives will lead to more widespread transmission, which is extremely unfavorable for the prevention and control of the epidemic [3].

In order to better suppress the spread of the coronavirus, chest computed tomography (CT) images have become an important tool for diagnosing COVID-19. Studies in [4, 5] show that CT scan has high sensitivity, and abnormal features such as ground-glass opacity (GGO), consolidation and rare features in CT images can reflect the severity of infection in patients. However, it takes a lot of time to manually segment the lesion area in CT images, and for an experienced radiologist, it needs about 21.5 min to get the diagnostic results of each case by analyzing CT images [6]. Therefore, it is necessary to propose an automatic lesion segmentation method to assist doctors in diagnosis. Recently, with the powerful feature extraction capability of deep convolutional neural networks, deep learning-based method has been widely used in medical image processing [1, 7]. Wang et al. [8] developed a deep learning method combined with CT classification and segmentation that can extract CT image features of COVID-19 patients and provide medical diagnosis for doctors. Matteo et al. [9] proposed a lightweight convolutional neural network for distinguishing CT images of COVID-19 patients from healthy CT images.

It is worth noting that the encoder-decoder structure is the most common used in lesion segmentation models. Many studies [10, 11] have confirmed that this structure has good segmentation performance and robustness. As a result, a number of studies have been conducted on segmentation of COVID-19 lesions by using encoder-decoder structures. FCN [12], SegNet [13], Unet [14] and deeplav3 [15] are applied to the COVID-19 segmentation task. In addition, Unet and its variants have also been applied to the COVID-19 segmentation task. Chen et al. [16] used Unet combined with residual network to achieve automatic segmentation of COVID-19 lesions. Bhatia et al. [17] proposed a U-Net++-based segmentation model for identifying 2019-nCoV pneumonia lesion on chest CT images of patients. A large number of deep learning-based methods [18, 19] have been applied to the lesion segmentation of COVID-19. Although these methods are more efficient than manual segmentation, they still have shortcomings in segmentation accuracy. They tend to have the following problems. (1) Although the encoder-decoder structure can extract high-level features with rich semantics, it will lose spatial detail information such as the edge information of the lesion area when the encoder performs down sampling. (2) These networks lack an effective mechanism to learn the channel information and spatial information of features. (3) The previous loss function of semantic segmentation is not suitable for the lesion segmentation task of COVID-19, which makes the network insensitive to small lesion areas.

To solve the above problems, we propose a Sobel operator combined with Multi-Attention networks (SMA-Net) to segment the lesions of COVID-19. Different from previous methods, we pay more attention to the edge information

of images. We propose self-attentive channel attention mechanism and spatial attention mechanism to guide the network in the concatenation of low-level and high-level features for feature extraction. The Tversky loss function adopted by SMA-Net can take into account the small lesion area and improve its sensitivity.

Our contributions are summarized as follows:

- 1) We propose a module for fusing COVID-19 CT images and their edge features to provide more detailed information for the network. This module uses the Sobel edge detection operator to obtain edge information.
- 2) We propose a self-attentive channel attention mechanism with a spatial linear attention mechanism module that is independent of the resolution size of the feature map, which we apply to each layer of the low-level feature and high-level feature splicing. It enables the network to focus on important semantic information, thereby improving the segmentation performance of the network.
- 3) SMA-Net has a suitable loss function for the small lesion area of COVID-19. Compared with other segmentation methods, SMA-Net has better segmentation accuracy in small lesion.

2 Dataset and Data Enhancement

Public COVID-19 segmentation dataset: The public dataset used in this paper is from zendo [20]. The dataset contains 20 COVID-19 CT scans, including lung and lesion segmentation labels. The dataset was annotated by two radiologists and examined by an experienced radiologist. In this study, 2237 CT images were selected for the experiment. To speed up the convergence of the network and improve efficiency, some preprocessing operations were performed on this dataset. We cropped the CT images to a resolution of 512*512 size to reduce the amount of calculation in the training process. The CT images are then normalized. Image normalization is the process of centering the data by de-meaning, which can improve the generalization of the network.

3 The Proposed Method

In this section, we first propose the overall structure of SMA-Net. Then the core modules of the network are introduced in detail, including edge feature fusion, self-attentive channel attention mechanism and spatial linear attention mechanism. Finally, the loss function used for training is described.

3.1 Network Structure

The structure of proposed SMA-Net is shown in Fig. 1. It can be seen that the original CT image is first fused with its corresponding edge features to obtain the input tensor of the network. Then the input feature map is divided into two directions after convolution and activation operations. Feature map is sent to the SCAM module, and it is also sent to the next layer by pooling for further

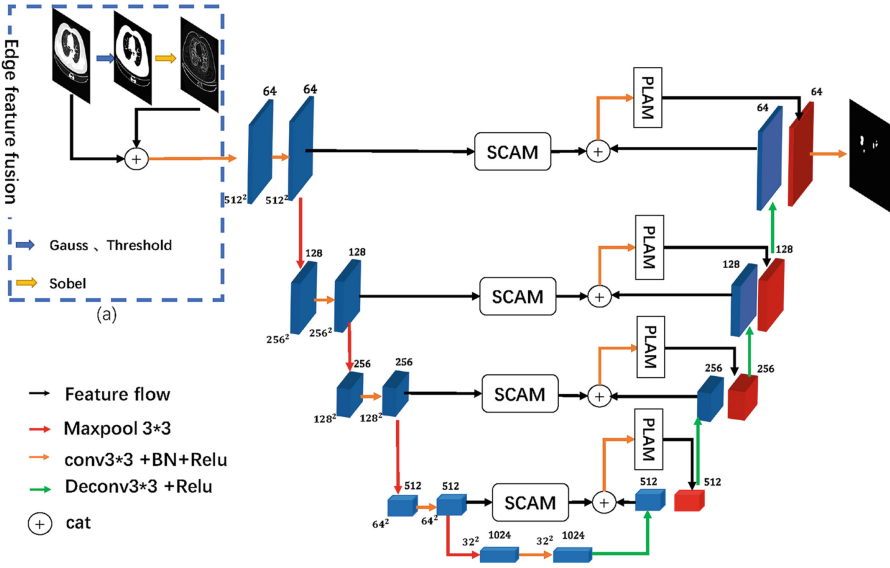


Fig. 1. The network structure of SMA-Net. The part (a) in the blue dashed box is the edge feature fusion. After the edge feature fusion is completed, the image is fed into the segmentation network. The segmentation network has four layers. Each layer has a corresponding channel attention mechanism and a spatial linear attention mechanism. The input image and the segmentation result output by the network have the same resolution size. (Color figure online)

feature extraction. SMA-Net performs 4 times down-sampling for the features. The feature map is reduced from a resolution size of 512×512 to 32×32 . Then, the feature map is begun up-sampling. After up-sampling, the feature map is concatenated with the encoder feature map in the same layer after it passing through SCAM module, and then the obtained feature map is sent to the PLAM module to get a feature map with rich semantic information. Next, the feature map is further upsampling, and then repeat the above operation to upsample the feature map to the original image size. Finally, compress the channel to get the final segmentation result.

3.2 Edge Feature Fusion

In the semantic segmentation models applied to medical images, most of them use encoding-decoding structure as the overall architecture. The encoder extracts the feature maps from the images through convolution and pooling operations. The low-level feature maps are often containing more edge information of the lesions in the CT images. But in the process of downsampling, the edge details in the feature map will be partially lost. To solve the loss of edge information, we propose to fuse CT images with their edge features to add spatial detail information from the source of the model input. As shown in Fig 1(a), we first

do a Gaussian filtering process on the CT image. The idea of Gaussian filtering is to suppress noise and retain detail information by weighted average of pixels. Then, a thresholding process is done to obtain a binary map U .

$$U = T(G(X, k = 3), t = 127) \tag{1}$$

where G denotes Gaussian filtering operation, k is the filter size and X is the grey scale map of the input CT image. t denotes image thresholding and it is set to 127 in this paper. The Sobel operator is then used to calculate the gradient in the X and Y directions for the binary map, and the two gradients are combined to obtain the edge feature map. Finally, the model input Z is obtained by fusing the extracted edge feature map with its CT image.

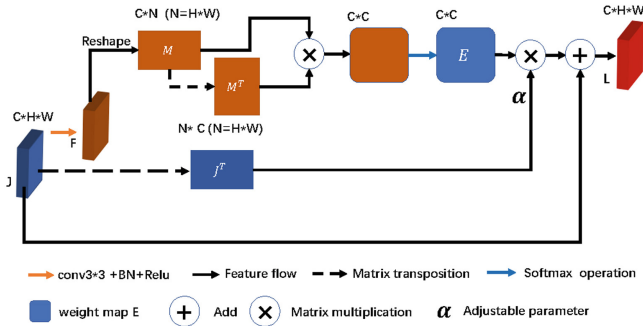


Fig. 2. SCAM: Self-attentive channel attention mechanism

3.3 Self-attentive Channel Attention Mechanism (SCAM)

To improve semantic segmentation performance, U-shaped networks concatenate high-level features with low-level features to obtain richer semantic information. In the process of concatenation, a redundancy channel of feature map often occurs. So, channel attention modules (such as the classical SE module) are usually added to the network in order to emphasize the meaningful features of channel axis. SE module obtains the compressed feature vectors by global average pooling of the feature maps, and then the obtained compressed feature vectors go through the full connection layer to generate the weight of each channel of the feature map. The SE module is simple and easy to apply it to the model. However, the global averaging pooling operation in the SE module results in a loss of semantic information.

In order to solve this problem, we propose a self-attentive channel attention mechanism (SCAM) module as shown in Fig. 2. Instead of compressing the feature map by global average pooling, the module first performs a convolution operation on the feature map J input to the SCAM module to obtain the feature map $F \in R^{C \times H \times W}$ as shown in Eq. 2. The feature map F is then reshaped to obtain the matrix $M \in R^{C \times N}$, and then the transpose of matrix M and M are calculated as matrix product. Finally, using softmax to activate the matrix product yields the channel attention weight map E .

$$F = f(J, k = 3) \quad (2)$$

$$E_{ji} = \frac{\exp(M_i \cdot M_j)}{\sum_{i=1}^C \exp(M_i \cdot M_j)} \quad (3)$$

where E_{ji} denotes the effect of the i -th channel on the j -th channel. After obtaining the weight map E , E and the transpose of matrix J are calculated as matrix product. This assigns the values in the weight map to each channel of J . Given on the idea of residual networks, the result of the product is multiplied by the adaptive coefficient α and then summed with J to obtain the final output L of the SCAM module.

$$L = \alpha \sum_{i=1}^c (E_{ji} J^T) + J \quad (4)$$

where the initial value of α is set to 0 and it can be changed with the needs of the network during the training process. L is used as the output of the input J passing through the SCAM module. L is then connected in series with its corresponding high-level features in the decoder.

3.4 Spatial Linear Attention Mechanism (PLAM)

After completing the concatenation of low-level features with high-level features, the decoder obtains a semantic rich feature map. However not all regions of this rich semantic information are equally important for lesion segmentation. To enhance the representation of key regions, we introduce the spatial linear attention module as shown in Fig. 3. Before introducing the spatial linear attention module, we first review the principle of the compressed dot product attention mechanism (Scaled-Dot Attention, SDA), as given in Eq. 5.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

where Q , K and V denote the query matrix, the key matrix and the value matrix respectively. These three matrices are obtained by convolving the input feature map by compressing the number of channels and then reshaped the feature map. $\sqrt{d_k}$ denotes the scaling factor. The overall dot product attention

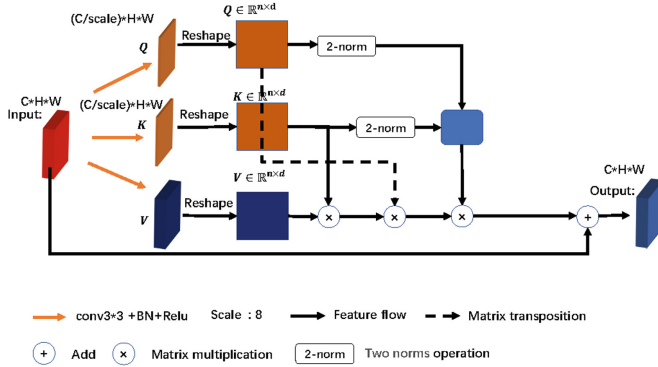


Fig. 3. PLAM Spatial linear attention mechanism structure

mechanism can be summarized as modeling the similarity between pixel points by matrix multiplication and the softmax function is used to activate the matrix multiplication result.

However, since $Q \in \mathbb{R}^{n \times d}$, $K \in \mathbb{R}^{n \times d}$, $V \in \mathbb{R}^{n \times d}$ where $n = W * H$, W and H represent the width and height of the feature map respectively. The complexity of the dot product attention mechanism is $\mathcal{O}(n^2)$, which makes SDA limited by the image resolution. Moreover, the resolution of CT images is usually large and if SDA is used directly, it will exceed the computational power of the computer. If the resolution of the CT image is scaled by scaling, much detailed information is lost in the image.

In order to improve SDA, we propose a spatial linear attention mechanism module. The complexity of the module is reduced from $\mathcal{O}(n^2)$ to $\mathcal{O}(n)$, which allows the module to be flexibly applied to segmentation networks. We start by equivalently rewriting Eq. 5 as Eq. 6. Because PLAM does not use a scaling factor, $\sqrt{d_k}$ is removed from Eq. 6. Equation 6 represents the result of the i -th row of the matrix obtained from the feature map after feeding into the dot product attention mechanism.

$$\text{Attention}(Q, K, V)_i = \frac{\sum_{j=1}^n e^{q_i^\top k_j} v_j}{\sum_{j=1}^n e^{q_i^\top k_j}} \tag{6}$$

where $e^{q_i^\top k_j}$ is essentially a weighted average over v_j , so Eq. 6 can be generalized to a general form by replacing the softmax function with the general function as given in Eq. 7.

$$\text{Attention}(Q, K, V)_i = \frac{\sum_{j=1}^n \text{sim}(q_i, k_j) v_j}{\sum_{j=1}^n \text{sim}(q_i, k_j)} \tag{7}$$

where $\text{sim}(q_i, k_j) \geq 0$. In order to reduce the complexity of Eq. 7, the order of concatenation of q_i, k_j, v_j needs to be changed and the normalization of q_i, k_j

needs to be solved. In the construction of linear attention mechanism, we start with Taylor expansion. Turn $e^{q_i^T k_j}$ into $1 + q_i^T k_j$.

$$e^{q_i^T k_j} \approx 1 + q_i^T k_j. \tag{8}$$

According to the Taylor expansion of Eq. 7, $\text{sim}(q_i, k_j) = 1 + q_i^T k_j$. Since we need to normalize q_i, k_j and ensure that $\text{sim}(q_i, k_j) > 0$. We can use the two norms of the matrix for normalization. Following this Eq. 6 can be equated to Eq. 9.

$$\text{Attention}(Q, K, V)_i = \frac{\sum_{j=1}^n \left(1 + \left(\frac{q_i}{\|q_i\|_2} \right)^T \left(\frac{k_j}{\|k_j\|_2} \right) \right) v_j}{\sum_{j=1}^n \left(1 + \left(\frac{q_i}{\|q_i\|_2} \right)^T \left(\frac{k_j}{\|k_j\|_2} \right) \right)} \tag{9}$$

By modifying the original form of the attention mechanism, we have completed the construction of a spatial linear attention mechanism.

3.5 Loss Function

Due to the exist of small lesions in the CT images of COVID-19, and the early clinical manifestations of COVID-19 are not obvious. The small lesion part of the CT images can be used as a basis for the early diagnosis of COVID-19. When the proportion of pixels in the target region is small, network training becomes more difficult, so small lesions are easily ignored in the network training process. Therefore, after the network has been built, it is important to choose a suitable loss function that is appropriate for the segmentation task. The Dice Loss function, which is often used in segmentation tasks, cannot meet the segmentation needs of small lesions in COVID-19. In order to fit the segmentation task, we chose the Tversky loss. As given in Eq. 10.

$$TL(\alpha, \beta) = \frac{\sum_{i=1}^N p_{0i} g_{0i}}{\sum_{i=1}^N p_{0i} g_{0i} + \alpha \sum_{i=1}^N p_{0i} g_{1i} + \beta \sum_{i=1}^N p_{1i} g_{0i}} \tag{10}$$

where α, β are hyper parameters, set to 0.3 and 0.7 respectively in this paper. p_{0i} represents the probability of a pixel point being diseased. g_{0i} is 1 and g_{1i} is 0 when the pixel point is diseased. p_{1i} represents the probability of a pixel point being non-diseased. When the pixel point is non-lesioned, g_{0i} is 0 and g_{1i} is 1. As can be seen from the Eq. 10, the trade-off between false negatives and false positives can be controlled when adjusting the values of α, β . The value of β is taken to be 0.7 greater than α , improving sensitivity by emphasizing false negatives. This allows the network to focus on small lesion areas during training, thus addressing the problem of data imbalance in CT images of patients with neocoronary pneumonia.

4 Experiment

4.1 Experimental Setup

Baseline: In the lesion segmentation experiments, our proposed SMA-Net is compared with the classical network Unet and Unet++. In addition, we also refer to the advanced semantic segmentation networks Deeplabv3, FCN, SegNet. Moreover, we also compare three newly proposed COVID-19 lesion segmentation networks: AnamNet, JCS, and Inf-Net.

AnamNet [21]: A lightweight CNN based on deformation depth embedding for segmentation network of COVID-19 chest CT image anomalies, can be deployed to mobile terminals.

JCS [22]: A novel combined classification and segmentation system for real-time and interpreted COVID -19 chest CT diagnosis.

Inf-Net [23]: A semi-supervised segmentation framework based on a random selection propagation strategy for a network with a fully supervised form, which we have selected for its fully supervised approach.

Table 1. Comparison of lesion segmentation performance

Methods	DSC	IOU	SEN	SPE
Unet	0.797	0.701	0.869	0.998
Unet++	0.754	0.687	0.836	0.993
Deeplabv3	0.773	0.646	0.861	0.997
FCN	0.689	0.612	0.795	0.876
Segnet	0.731	0.634	0.854	0.993
AnamNet	0.808	0.71	0.846	0.979
JCS	0.847	0.754	0.852	0.989
Inf-Net	0.818	0.723	0.871	0.985
Ours	0.861	0.778	0.915	0.997

4.2 Segmentation Results

To compare the segmentation performance of SMA-Net, we refer to the classical medical image segmentation network Unet and its variant Unet++. In addition we also refer to the advanced semantic segmentation networks Deeplabv3, FCN, SegNet. For the three recently proposed COVID-19 lesion segmentation networks (AnamaNet, JCS, Inf-Net), we have also conducted comparative experiments. The quantitative results are shown in Table 1. It can be seen that for the other methods our proposed SMA-Net achieves a significant improvement in IOU metric, with a 7.8% improvement compared to Unet. The DSC coefficient also achieves the best. We attribute this improvement to our edge feature fusion

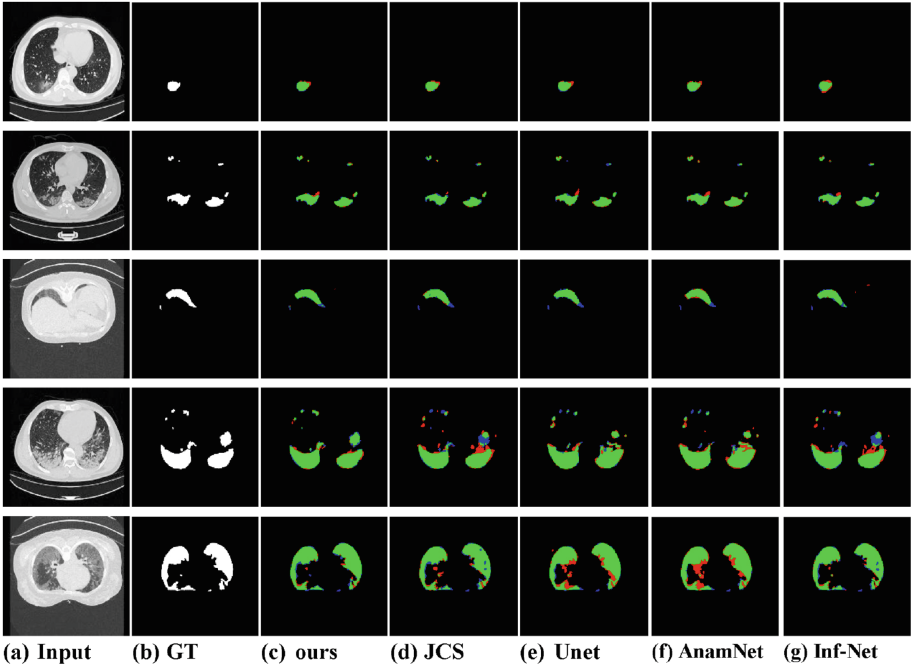


Fig. 4. Visual comparison of lesion segmentation results using different networks. (a) represents CT images (b) represents ground truth. (c), (d), (e), (f) and (g) represent the segmentation results of SMA-Net, JCS, Unet, AnamNet and Inf-Net, respectively. The green, blue, and red regions refer to true positive, false negative and false positive pixels, respectively. (Color figure online)

module as well as to self-attentive channel attention mechanism and spatial linear attention mechanism. Thanks to the two attention mechanisms guiding SMA-Net, SMA-Net can sample a richer feature map of semantic information during feature extraction.

Figure 4 shows a visual comparison of SMA-Net with Unet and the three newly proposed COVID-19 lesion segmentation networks (AnamaNet, JCS, Inf-Net). The green, blue, and red regions refer to true positive, false negative and false positive pixels, respectively. It can be seen that SMA-Net is closest to the ground truth. In contrast, many false positive pixels appear in the Unet and AnamNet segmentation results. Thanks to our choice of Tversky loss function, SMA-Net achieves good results in the segmentation of small lesions. Compared to the other networks, our increased sensitivity of the loss function to the small lesion region allows the network to segment the small lesion region well.

4.3 Ablation Studies

In this section, we experimentally demonstrate the performance of key components of SMA-Net, including the edge feature fusing module, the self-attention channel attention mechanism module (SCAM), and the spatial linear attention mechanism module (PLAM). In Table 2, A is the SMA-Net without the SCAM module, B is the SMA-Net without the PLAM module, C is the SMA-Net with the edge feature fusion module removed and D is the complete SMA-Net. E is the SMA-Net without SCAM, PLAM and the feature fusion module.

Table 2. Results of the ablation experiment

Model	DSC	IOU	SEN	SPE
A	0.846	0.756	0.903	0.997
B	0.837	0.732	0.897	0.983
C	0.783	0.697	0.884	0.997
D	0.861	0.778	0.915	0.997
E	0.797	0.701	0.869	0.998

- Effectiveness of **SCAM**: To explore the SMA-Net’s self-attentive channel attention module, we propose two benchmarks: as shown in Table 2, A (SMA-Net without SCAM), D (SMA-Net), and the results show that SCAM is effective in improving network performance.
- Effectiveness of **PLAM**: From Table 2, it can be observed that the IOU values decrease more between B (SMA-Net without PLAM) compared to D. This indicates that the spatial linear attention mechanism has an important role in guiding the network to learn to segment the lesion area, allowing the network SMA-Net to focus more on the pixels in the lesion area.
- Effectiveness of edge feature stitching: After the fusing of edge features is completed, the encoder obtains richer semantic information. As can be seen from Table 2, C has the lowest IOU metric compared to A, B and D, which indicates that edge features are important for the detail complement of CT images.

5 Comparison of Loss Function

5.1 Selection of Loss Function

After the construction of the network SMA-Net is completed, the selection of the loss function has a great impact on the performance of the network. Therefore, for different semantic segmentation tasks, the selection of the loss function is based on the characteristics of the task. Commonly used loss functions such as Dice loss (DL) function, Balanced cross-entropy loss function BCE for binary classification task, Weighted cross-entropy loss function WCE. In addition, we

also selected excellent loss functions that have been used for semantic segmentation in recent years, namely Asymmetric Loss Functions (AL), Tversky Loss (TL), PenaltyGDiceLoss (PL).

- 1) Asymmetric Loss Functions (AL): A novel loss function is designed to address the problem of positive and negative sample imbalance in classification tasks. Adaptive methods are proposed to control the asymmetric rank.
- 2) Tversky Loss (TL): In order to solve the problem of data imbalance, a new loss function is proposed to improve the sensitivity of small lesion areas by adjusting the hyper parameters of the tversky index.
- 3) PenaltyGDiceLoss (PL): improves network segmentation performance by adding false negative and false positive penalty terms to the generalized dice coefficients (GD).

Table 3. Comparison of SMA net results under different loss functions

Loss	DSC	IOU	SEN	SPE
BCE	0.834	0.743	0.979	0.896
WCE	0.823	0.726	0.983	0.852
DC	0.783	0.697	0.997	0.884
GD	0.799	0.689	0.973	0.854
AL	0.824	0.727	0.998	0.897
TL	0.861	0.778	0.997	0.915

5.2 Comparison Results

As can be seen from Table 3, Tversky Loss (TL) performed the best among the three indicators of IOU, DSC and SPE. Compared to the BCE loss function IOU and DSC coefficients improved by 6.8% and 7% respectively. Among them, AL performed the best in sensitivity. We also made a visual comparison of the output results of SMA-Net with different loss functions. The results from TL are more sensitive for small lesion regions and can do well in segmenting small lesions. In contrast, the lack of segmentation for small lesion regions can be observed from the segmentation results of BCE as well as AL.

6 Conclusion

To improve the efficiency of diagnosis of COVID-19, we have developed a COVID-19 lesion segmentation network. In our network, we propose the first edge feature fusion module, which allows the network to capture more edge feature information. In addition, we introduce a self-attentive channel attention mechanism and a spatial linear attention mechanism to improve the network performance. Two attention mechanisms guide SMA-Net, which captures lesion

areas more accurately during feature extraction. Compared with the classical medical image segmentation network U-net, the DSC and IOU of SMA-Net are improved by 7% and 7.8% respectively. Although our method achieves good results in terms of performance, it still has the following shortcomings. (1) the network has a high computational complexity, and (2) the network does not perform the classification task simultaneously. Therefore, our future work will try to start with the light weighting of the model and to achieve simultaneous network classification and segmentation as a way to improve the diagnosis of neocrown pneumonia.

References

1. Shi, F., et al.: Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE Rev. Biomed. Eng.* **14**, 4–15 (2020)
2. Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lu, W., et al.: Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in china: a report of 1014 cases. *Radiology* **296**(2), E32–40 (2020)
3. Fang, Y., et al.: Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* **296**(2), E115–E117 (2020)
4. Ankur, G.-W., et al.: False-negative RT-PCR for COVID-19 and a diagnostic risk score: a retrospective cohort study among patients admitted to hospital. *BMJ Open* **11**(2), e047110 (2021)
5. Swapnarekha, H., Behera, H.S., Nayak, J., Naik, B.: Role of intelligent computing in COVID-19 prognosis: a state-of-the-art review. *Chaos Solitons Fractals* **138**, 109947 (2020)
6. Wang, Y., Hou, H., Wang, W., Wang, W.: Combination of CT and RT-PCR in the screening or diagnosis of COVID-19 (2020)
7. Rajinikanth, V., Dey, N., Raj, A.N.J., Hassanien, A.E., Santosh, K.C., Raja, N.: Harmony-search and otsu based system for coronavirus disease (COVID-19) detection using lung ct scan images. arXiv preprint [arXiv:2004.03431](https://arxiv.org/abs/2004.03431) (2020)
8. Wang, B., et al.: Ai-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system. *Appl. Soft Comput.* **98**, 106897 (2021)
9. Polsinelli, M., Cinque, L., Placidi, G.: A light CNN for detecting COVID-19 from CT scans of the chest. *Pattern Recogn. Lett.* **140**, 95–100 (2020)
10. Chen, J., Qin, F., Lu, F., et al.: CSPP-IQA: a multi-scale spatial pyramid pooling-based approach for blind image quality assessment. *Neural Comput. Appl.* pp. 1–12, (2022). <https://doi.org/10.1007/s00521-022-07874-2>
11. Xiaoxin, W., et al.: Fam: focal attention module for lesion segmentation of COVID-19 CT images. *J. Real-Time Image Proc.* **19**(6), 1091–1104 (2022)
12. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015)
13. Badrinarayanan, V., Kendall, A., Cipolla, R.: SEgNET: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017)
14. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

15. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11211, pp. 833–851. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_49
16. Chen, J., et al.: Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. *Sci. Rep.* **10**(1), 1–11 (2020)
17. Bhatia, P., Sinha, A., Joshi, S.P., Sarkar, R., Ghosh, R., Jana, S.: Automated quantification of inflamed lung regions in chest CT by UNET++ and SegCaps: a comparative analysis in COVID-19 cases. In: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 3785–3788. IEEE (2022)
18. Tang, S., et al.: Release and demand of public health information in social media during the outbreak of COVID-19 in China. *Front. Pub. Health* **9**, 2433 (2021)
19. Gao, K., et al.: Dual-branch combination network (DCN): towards accurate diagnosis and lesion segmentation of COVID-19 using CT images. *Med. Image Anal.* **67**, 101836 (2021)
20. Ma, J., et al.: Toward data-efficient learning: a benchmark for COVID-19 CT lung and infection segmentation. *Med. Phys.* **48**(3), 1197–1210 (2021)
21. Paluru, N., et al.: Anam-net: anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images. *IEEE Trans. Neural Netw. Learn. Syst.* **32**(3), 932–946 (2021)
22. Kimura, K., et al.: JCS 2018 guideline on diagnosis and treatment of acute coronary syndrome. *Circ. J.* **83**(5), 1085–1196 (2019)
23. Fan, D.-P., et al.: INF-net: Automatic COVID-19 lung infection segmentation from CT images. *IEEE Trans. Med. Imaging* **39**(8), 2626–2637 (2020)