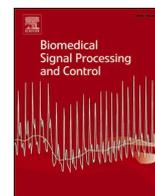




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

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



## Weakly supervised segmentation of COVID-19 infection with local lesion coherence on CT images

Wanchun Sun<sup>a</sup>, Xin Feng<sup>a,b,\*</sup>, Jingyao Liu<sup>a,c</sup>, Hui Ma<sup>d</sup>

<sup>a</sup> School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China

<sup>b</sup> Chongqing Research Institute, Changchun University of Science and Technology, Chongqing 401122, China

<sup>c</sup> School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China

<sup>d</sup> Computer Basic Teaching and Research Department, Anhui Vocational College of Police Officers, Hefei 232001, China

### ARTICLE INFO

#### Keywords:

Weakly Supervised  
COVID-19  
Segmentation  
Local Coherence

### ABSTRACT

At the end of 2019, a novel coronavirus, COVID-19, was ravaging the world, wreaking havoc on public health and the global economy. Today, although Reverse Transcription-Polymerase Chain Reaction (RT-PCR) is the gold standard for COVID-19 clinical diagnosis, it is a time-consuming and labor-intensive procedure. Simultaneously, an increasing number of individuals are seeking for better alternatives to RT-PCR. As a result, automated identification of COVID-19 lung infection in computed tomography (CT) images may help traditional diagnostic approaches in determining the severity of the disease. Unfortunately, a shortage of labeled training sets makes using AI deep learning algorithms to accurately segregate diseased regions in CT scan challenging. We design a simple and effective weakly supervised learning strategy for COVID-19 CT image segmentation to overcome the segmentation issue in the absence of adequate labeled data, namely LLC-Net. Unlike others weakly supervised work that uses a complex training procedure, our LLC-Net is relatively easy and repeatable. We propose a Local Self-Coherence Mechanism to accomplish label propagation based on lesion area labeling characteristics for weak labels that cannot offer comprehensive lesion areas, hence forecasting a more complete lesion area. Secondly, when the COVID-19 training samples are insufficient, the Scale Transform for Self-Correlation is designed to optimize the robustness of the model to ensure that the CT images are consistent in the prediction results from different angles. Finally, in order to constrain the segmentation accuracy of the lesion area, the Lesion Infection Edge Attention Module is used to improve the information expression ability of edge modeling. Experiments on public datasets demonstrate that our method is more effective than other weakly supervised methods and achieves a new state-of-the-art performance.

### 1. Introduction

As of April 17, 2022, the COVID-19 pandemic has already claimed over 6 million lives and infected over 500 million people [1]. COVID-19's estimated [2] incubation period is longer than that indicated by earlier studies on SARS, MERS, and COVID-19. On January 30, 2020, the World Health Organization classified the novel coronavirus outbreak (2019-nCov) as a Public Health Emergencies of International Concern (PHEIC). On March 11, 2020, COVID-19 was declared a pandemic by the World Health Organization. As of April 5, 2022, the latest news from the World Health Organization (WHO) warns of a new variant of the recombinant Omicron strains BA.1 and BA.2, "XE", which may be more transmissible than any new strain of Coronavirus previously seen. At

present, Shanghai and Changchun from China are the hardest hit areas of this round of new variants, which proves that the social danger of this variant is great.

Most patients with COVID-19 have mainly lower respiratory symptoms, and clinical trial data show that common symptoms of COVID-19 are fever, cough, sputum, weakness of the extremities, and headache. Some patients only show diarrhea, low fever, slight weakness, drowsiness, etc., without pneumonia manifestations, some patients even have no clinical manifestations [3,4]. The current primary medical screening tool utilizes real-time reverse transcription polymerase chain reaction (PCR) for the determination. In addition to this, X-rays, computed tomography (CT) and ultrasound, have all proven their effectiveness in detecting infectious diseases, tracking and assessing and estimating

\* Corresponding author.

E-mail address: [xinfeng\\_cust@163.com](mailto:xinfeng_cust@163.com) (X. Feng).

<https://doi.org/10.1016/j.bspc.2022.104099>

Received 26 May 2022; Received in revised form 30 July 2022; Accepted 15 August 2022

Available online 18 August 2022

1746-8094/© 2022 Elsevier Ltd. All rights reserved.

disease evolution. Among them, CT is more commonly used as an adjunct to detection because of its ability to unfold a three-dimensional view of the lung lesion area and its high contrast analysis. In performing CT processing for COVID-19 applications, performing segmentation characterization of focal areas of lung infections is a critical operation. Investigating the statistical analysis, it usually takes 7 h to mark and process the one CT scan with 250 slices using manual segmentation [5]. Therefore, there is currently a huge challenge in combating the outbreak both in terms of symptom presentation and medical analysis of the pathogen. Thus, there is an urgent need to develop smarter, more accurate and efficient detection methods to combat variant viruses.

In recent times, the use of AI technology in medical disease diagnosis has become increasingly common, particularly in the highly visible field of COVID-19 research. In [6], the patient's chest imaging data information was attempted to be applied to the COVID-19 diagnosis with the help of a deep network model, and the method was demonstrated to be feasible. In [7], considering the limited training samples of the current COVID-19 data, the authors used multiple random sampling of lesion regions to increase the training sample size in disguise to improve the training sample accuracy. In [8], the network model is divided into encoder and decoder, with the encoder stage enhancing the extraction of feature information of COVID-19 focal areas and the decoder stage using different feature information to further optimize the fusion to improve the judgment accuracy. In [9], model training at point-level supervision is attempted to improve the accuracy by comparing the gap between ground truth and point-level segmented images. In [10], training attempts are made using Scribble weakly supervised Annotation, and the training samples are divided into teacher-student models after different scale transformations to learn the optimized model by comparison. Despite the fact that various complementary strategies for diagnosing novel COVID-19 have arisen, almost no research has been expended in the domain of minimally supervised investigations.

Given that the majority of medical segmentation deep learning researchers continue to rely on a wide variety of sample labeling methods for network feature training, the fundamental reason remains that the accuracy of weakly supervised and unsupervised approaches is low and cannot meet clinical standards. To address the issue of unsatisfactory weak supervision segmentation accuracy, we first design the Local Lesion Coherence solution in the field of medical segmentation, which we call Local Lesion Coherence Network (LLC-Net) for this innovative solution of 2D CT weakly supervised labeling image COVID-19 lung infection segmentation. Weakly supervised research work in the process of in-depth deep learning techniques rely on strong training samples. Therefore, the most difficult problem at the moment is how to mine more effective information in the training samples without precise labeling to compensate for the lack of precise labeling of training samples. To begin, a simple point-level method is employed to identify the affected zone. Second, for a specific calibration region, there must be some non-correlation between the junction of the infected and non-infected regions, and we believe that this local correlation is useful for region mining. Following that, we discuss how to expand the training effect in a limited number of weakly labeled samples, and the multi-scale transformation strategy has demonstrated its usefulness numerous times in recent years at various top computer vision conferences. Finally, the infected region's border information is the only effective characteristic that can be reused to prevent the prediction of the focal area from growing, and the boundary information is used to further suppress the non-effective infected region to increase accuracy.

## 2. Relate work

### 2.1. Deep learning in medical image segmentation

Deep learning technology is currently a prominent and effective tool in Medical Image Segmentation research. Deep learning is used in clinical practice to identify distinct lesions from Medical Image

Segmentation to provide vital information for clinicians to make preliminary diagnoses. Many relevant algorithms have been developed in recent years to get good results. For example in [11], X-ray and CT scan imaging modalities are described in terms of their advantages and problems for COVID-19 detection, and the application of AI technology to COVID-19 detection is fully recognized as a valuable research. Beyond that, [12] fully acknowledges the role of medical images in the field of image segmentation, while it considers the technique as the initial and most important component of the diagnostic and therapeutic pipeline. It is often used to distinguish homogeneous regions. In [13], in the face of the difficulties of medical image segmentation such as limited labeling data, propose to add the a priori medical knowledge to the deep learning model with the loss function to improve the accuracy of segmentation. The loss function is based on the deep atlas prior loss and the likelihood loss, which contains a priori information for accurate segmentation such as organ location and shape. In [14] presents a fully automated deep learning approach to obtain robust medical image segmentation, where first the encoding-decoding forward system predicts the segmentation result from the input image, which is then encoded by a full convolutional network (FCN) based context feedback system, and this encoded feature space is fused back into the forward learning process of the forward system. In [15] proposes a multilayer boundary perception-self attention mechanism deep learning model of image segmentation algorithm R-UNET segmentation model, focusing on solving the boundary localization problem in model segmentation, while deep mining network structure to build global information features to reverse the accuracy of segmentation. In [16] major contribution is to embed the attention residual network into the deep learning framework structure to build the COVID-19 CT image detection model and prove its effectiveness. In [17] adopting deep learning network AlexNet as the backbone is used. During the training process, in order to improve the training efficiency and prevent the internal covariance shift, its data set is normalized and the classifier of AlexNet fully connected layer is redesigned. In [18] study proposed a novel deep rank-based average pooling network, whose main contribution is the design of an  $n$ -conv rank-based average pooling module to avoid overfitting phenomenon in classification.

### 2.2. Semantic segmentation for COVID-19 infected area

For COVID-19 disease, semantic segmentation techniques are used in applications for patient diagnosis, especially in the CT image layer. In [19], COVID-19 lung infection region segmentation process has challenging problems such as irregular shapes and different sizes. Co-supervision and attention fusion strategy is used to guide the network to learn edge and semantic features by enhancing the information of supervision and fusing different levels of multi-scale feature maps. In [20], adopted a novel pixel-precision attention model that can extract cues such as boundaries and shapes from CT contours and use these features to refine the infected region, and also added cross-context attention fusion upsampling to robustly reconstruct deep semantic features back into the high-resolution segmentation map. In particular [21], a novel deep learning framework structure with multi-view slice decomposition is proposed for fine-grained CT lesion area segmentation evaluation, and prior knowledge is integrated into the model training to effectively improve the model performance. In [22], a multilayer segmentation network learning method, CHS-Net, was designed capable of automatic deep learning of COVID-19 from CT images, with shrinking and expanding phases of depth-separable convolution and hybrid pooling to efficiently encode and decode semantic and focal region feature information. In [23], Lung Ultrasound data and deep learning were proposed as a deep fusion artificial neural network framework to assess COVID-19 severity using Lung Ultrasound data. In [24], it is pointed out that the convolutional neural network based on CT images used to extract the image mask process does not do enough amount of optimization work, and multi-agent deep reinforcement learning

approach is able to extract effective visual features of the lesion area and enhance the performance of the segmentation network. In [25], a multi-task regression network is proposed to segment COVID-19 lesions in a framework MT-nCov-Net, which customizes lesion segmentation based on multi-task regression to cope with segmentation of task low-level, intermediate-level, and high-level features. Multi-scale feature learning module is used to learn lesion when reducing semantic divide between different scales, the fine-grained lesion localization module adaptive dual attention mechanism to explore lesion infection. In [26], an expanded dual-attention U-Net network structure D2A U-Net based on attention mechanism and hybrid null convolution is proposed to cope with the problem of fuzzy boundaries and low sensitivity in the accuracy and automatic segmentation of COVID-19 lung infection, and its main contribution is to refine the feature maps using the dual-attention mechanism to reduce the semantic gaps between different levels of feature maps, which is introduced into the decoding model by expanded null convolution to obtain a larger perceptual field of view.

### 2.3. Weakly supervised semantic segmentation for COVID-19 infected area

Despite the value of COVID-19 focal zone segmentation for disease assessment, segmentation of COVID-19 infection has not been extensively studied due to difficulties in data labeling and the limited availability of data samples. Recently, researchers have gradually shifted their focus to weakly supervised semantic segmentation for COVID-19 Infected Area. In [27], proposes a Weak Variational Autoencoder for Localisation and Enhancement framework, in order to solve the problem of effective anomaly localisation without pixel-level annotations, with a new gradient-based technique for variational autoencoders in localisation of COVID-19 lung infection regions, and use of post-hoc attention maps to generate pseudo segmentation datasets for images. In [28] propose a novel attention framework to estimate weakly annotated CT COVID-19 dataset, a non-locality approach that correlates ground-glass opacities and consolidations features across different parts and spatial scales of the 3D Scan. In [29], the main contribution is to build a deep learning model using normal CT images and then perform focal zone labeling on unlabeled images using a focal zone feature recognition approach to demonstrate the validity of his proposed integration of COVID-19 positive diagnosis and lesion analysis into a unified framework, and that the approach can be extended to other needle-detection applications for chest diseases. In [10], weakly supervised learning using scribble annotation method, an uncertainty-aware mean teacher to enhance different perturbations for different CT images, while introducing pixel-level uncertainty methods in to guide the student model to get reliable predictions. In [30], a weakly supervised segmentation method based on image-level labeling of the generative adversarial network (GAN) is proposed to develop guiding the generator to capture feature matching strategies for complex texture feature information in CT images chests. In [31], a weakly supervised COVID-19 framework based on the U-Net structure was designed to segment the patient's lung region first using the U-Net pre-trained model, then 3D deep neural network to predict the probability of COVID-19 infectious, The activation regions in the classification network and the unsupervised connected components are combined to locate COVID-19 lesions.

### 2.4. Local coherence for weakly supervised

In our research work in the area of weak supervision, we aim to quantify and evaluate problems using as limited a set of information as possible. In [32] focuses its main efforts on multi-scale to change the network structure, thus, weakly supervised single datasets are improved, and the training effect is advanced in reverse. [33] separating low-level and high-level local feature information for feature interweaved aggregation. The low-level appearance features and high-level semantic features are retained to fully improve the utilization of

limited local information. In [34], a new set of loss function is designed for weakly supervised labeling, and cross-entropy loss is performed for locally labeled information, and unlabeled regions are evaluated using gated CRF loss. In [35] argues that most research efforts have focused on the localization information ignoring rule-based appearance information. In addition, the article attempts to establish guidance mining semantic affinities between pixels consistent with the image's local and global consistency. In [36] generates affine transform in a limited data set to form a siamese network to improve the accuracy of performance, the essential core of which is still to enhance the base of training samples. In addition, [37] also performs aggregation after deeply mining the model information of different layers, and monitors the detail changes between different local intervals to continuously optimize the objective.

### 2.5. Research gaps

Medical image segmentation has been studied in the field of COVID-19 so far, and although many effective methods have been proposed, some problems have been found to be in urgent need of solution during the research process, and solving such problems is an excellent contribution to enhance the research results in this field.

- (1) Despite the fact that the global outbreak region is extensive, data collection and tagging are difficult to complete in a short period of time due to high labor expenses and time constraints [11,38]. Weakly supervised learning is of particular research interest that are less labor and time compared to traditional training labeling methods.
- (2) Medical images differ substantially from natural images in terms of class correlation, there is a greater probability of no correlation between lesion regions [39], and it is difficult to mine deeper valid information in a weakly supervised domain using only semantic correlation between network structure layers [33,34]. For weakly supervised learning, studying local autocorrelation is a better entry point to break through the lack of label information.
- (3) Research on weakly supervised learning of medical images has been reported rarely [9,10,28], and it is in the initial stage of development. Therefore, the need to propose an effective and simple segmentation network model appears to be of great research value.

### 2.6. Research contributions

In summary, the key contributions of this research are reported as follows:

- (1) Presented a point-level weakly supervised COVID-19 CT segmentation network is proposed. The main innovation contains three parts: Local Self-Coherence (LSC), Scale Transform Mechanism (ST), and Lesion Infection Edge Attention Module (EAM).
- (2) LLC-Net demonstrated the effectiveness of the COVID-19 focal zone segmentation method on diverse medical datasets that were weakly supervised, and its performance was state-of-the-art when compared to other weakly supervised methods.
- (3) To the best of our knowledge, Local Lesion Coherence solution is the first optimization solution for the general problem in the field of medical weakly supervised images, which does not care about the original network structure model and has a certain plug-and-play nature, and also has a certain improvement effect for other problems of medical image classification and segmentation.

## 3. Methodology

In this subsection, we will focus on the LLC-Net structure in terms of its design ideas, core overview of the three core components.

### 3.1. Motivation

Since deep learning has been applied to the cross-field of medicine, the effect has been outstanding. In the past two years, moreover, due to the explosion of COVID-19, various related disease diagnosis solutions relying on deep learning technology have been emerging, but the problem of large number of training samples annotated with cost has also become a pressing challenge in the research process [40,41]. In [5], it was mentioned that it could take up to 7 h to manually segment the lesion from the one CT scan with 250 slices. It is the specialized nature of the markers and the high time costs that make the use of weakly supervised medical images in disease diagnosis all the more important. Medical images are different from natural images in that successive image lesion areas have certain similarity but no regularity. Only their own characteristics can be combined to mine effective information in the research process. In [31] is a relatively early proposal to use the weakly supervised approach for COVID-19 disease diagnosis solution, but limited by the short time of research in this field, there are few published articles in this category. Therefore, it is of great interest to find a flexible solution.

### 3.2. Local lesion coherence for COVID-19 infection segmentation network architecture

The structure of the LLC-Net network framework is shown in Fig. 1. The images generated by CT scan are transmitted to the backbone network layer for extracting low-level semantic features and high-level resolution information. Firstly, the feature information after encoder processing is local self-coherence roughly to localize the lung infection region. Secondly, the scale transform is enhanced in the decoder stage to enhance the data information. It is of interest that these modules are constructed using a cascade approach between them. Finally, the lesion Infection Edge information is guided at the very end stage to further optimize the prediction map of the lung infection region. In the following, the design logic for the construction of each key component is

specified.

### 3.3. Local self-coherence mechanism

Weakly supervised images cannot be evaluated by point annotations alone for effective pixels of the image lesion region. The lack of information about the extent, size, etc. of the lesion region makes learning structural features more complicated. Therefore, it is necessary to improve its structure learning model by digging into local autocorrelation information. In this case, the most similar learning approach says to find local relevance, so we propose the Local Self-Coherence Mechanism (LSC), For the CT image  $I$ , define  $I \in \mathbb{R}^{B \times H \times W}$ , pixels  $k, j$ , which belong to  $k, j \in H \times W$ , the local relationships between pixels are as follows:

$$\mathfrak{S} = \text{Norm} |I_k - I_j| \quad (1)$$

Where  $I_k$  and  $I_j$  denote the weights of the pixel  $k$  and  $j$  prediction lesion, respectively.  $L1 - \text{Norm}$  is used to calculate the distance between two pixel points. The corresponding Gaussian kernel method is invoked here to solve the local self-consistency problem:

$$L(k, j) = \exp \frac{\|d(k) - d(j)\|^2 - \|m(k) - m(j)\|^2}{\omega(\sigma_d^2 - \sigma_m^2)} \quad (2)$$

Where  $L(k, j)$  is the pixel, and Gaussian evaluation of the similarity of the two locations' proximity,  $d(\cdot)$  is the distance between the two pixels,  $m(\cdot)$  pixels color difference value, Gaussian kernel hyper parameter values are represented by  $\sigma$ .

### 3.4. Scale transform for self-correlation

The feature representation capability of scale transformation for segment [42] has proved its importance. Inspired by this, this paper argues that it is a very good solution for the case of insufficient samples of effective training data, and in addition the prediction models of different scales should have consistent results to be a good guiding

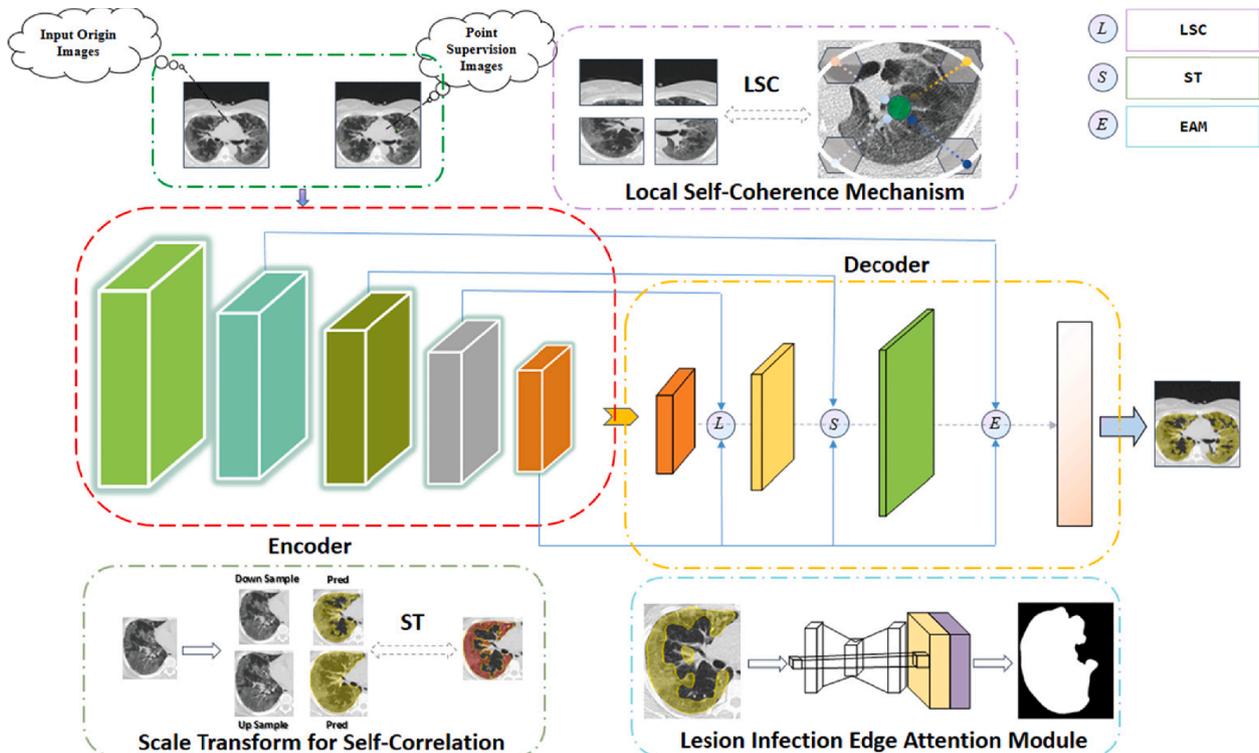


Fig. 1. Overall architecture of the proposed LLC-Net framework. LLC-Net consists of an Local Self-Coherence Mechanism (LSC), Scale Transform for Self-Correlation component (SC), and a Lesion Infection Edge Attention Module (EAM), respectively.

direction. CT image after scale variation over CNN to get low-level feature information  $f \in R^{B \times C \times H \times W}$ . In this case, do the following:

$$f^{(\omega)} = Up(\zeta(Trans(f; \theta)); \omega) \quad (3)$$

Where  $Up(\cdot)$  is a bilinear interpolation operation with a sampling size parameter  $\omega$ ,  $Trans(f; \theta)$  is a convolutional transformation operation with parameters  $\theta$  on the features  $f$  of the CT image.  $\zeta(\cdot)$  is a linear transformation.

However, it is difficult to monitor the different scales effectively, so scale transform for self-correlation is designed to monitor the prediction effect more effectively:

$$\hat{f}^{(\omega_i, \omega_j)} = AvgPool(f^{(\omega_i)} \times f^{(\omega_j)}) - (AvgPool(f^{(\omega_i)} \times f^{(\omega_j)}))^2 \quad (4)$$

$AvgPool$  is to perform average pooling operations on it, while doing the following:

$$ST^{up} = (2 \times AvgPool(f^{(\omega_i)} \times f^{(\omega_j)}) + \alpha) \times (2 \times \hat{f}^{(\omega_i, \omega_j)} + \beta) \quad (5)$$

$\alpha, \beta$  is the hyper parameter, in order to avoid the unstable fluctuation of the value:

$$ST^L = (AvgPool(f^{(\omega_i)})^2 + AvgPool(f^{(\omega_j)})^2 + \alpha) \quad (6)$$

$$ST^R = (AvgPool(f^{(\omega_i)})^2 + AvgPool(f^{(\omega_j)})^2 + \alpha) \times (\hat{f}^{(\omega_i, \omega_i)} + \hat{f}^{(\omega_j, \omega_j)} + \beta) \quad (7)$$

For the supervised information value of the scale transformation, that is:

$$ST = \rho \left( \frac{1}{2} \left( 1 - \frac{ST^{UP}}{ST^L + ST^R} \right); \theta \right) \quad (8)$$

The  $\rho$  define linear value constraint, scale correlation of the network can be improved based on the hyper parameter and continuously learn to scale transform for self-correlation according to Eq. (8).

### 3.5. Lesion infection edge attention module

The supervised role of the edge information in the lesion region is strongly guiding for the learning of the network model, and it has been shown in [43] that effective edge information is effectively constraining for feature segmentation. The given CT image has more accurate edge information in the lower layer of the CNN, and the second layer in Fig. 1,  $I_{Conv2}$  is selected to establish significant edge feature information in the specific operation. To summarize:

$$f_{pred}^{(\omega)} = \rho(f^{(\omega)}; \zeta) \quad (9)$$

$\zeta$  denotes the activation function, sigmoid is used here, refer to Eq. (8) expression, guiding the network's handling of lesion infection edge concerns:

$$f_{EAM}^{(\omega)} = \left| \log \left( 1 - f_{pred}^{(\omega)} \right) \cdot ME(I_{Conv2}; \theta) \right| + \left| \left( 1 - f_{pred}^{(\omega)} \right) \cdot \log \left( 1 - ME(I_{Conv2}; \theta) \right) \right| \quad (10)$$

$ME(I_{Conv2}; \theta)$  denotes the second convolutional layer of the network mask edge feature extraction. With the lesion infection edge attention module, the network is able to feed a more accurate edge outline without any additional information.

## 4. Experiments

### 4.1. Experimental datasets standard

To more fairly evaluate the performance of weakly supervised networks, the first publicly available and very widely used medical segmentation dataset [44] is cited here, contains mostly of 9 volumetric COVID-19 Chest CT, DICOM format is the medical industry standard, which contain 829 slices infection. In addition, considering the small sample size of this sample data set, in order to demonstrate the effectiveness of our method, increase the sample data set more rich [45], two radiologists designated the lungs and infection sites, which were then verified by an experienced radiologist, 3520 slices in a volume of 20 COVID-19 chest CT.

### 4.2. Data augmentation

Data augmentation is a routine pre-processing operation in medical images, and in this paper, we only do simple data augmentation to enhance generalization without adding too much performance overhead on the basis of a limited data set, and to prevent overfitting.

As shown in Fig. 2, only three types of processing, Rotate, Crop and noise, were done on the images during the experimental training phase. Owing to the different standards of the acquired image sizes, they are adjusted here according to standard  $352 \times 352$ , for original images  $X$  processed as shown in Eq. (11).

$$Y = Resize(X, [352, 352]) = \{x_1, x_2, x_3, \dots, x_n\} \quad (11)$$

Rotation is a common method of data augmentation, the rotation angle  $\theta = (0^\circ, 90^\circ)$  was applied.

to the images:

$$\overrightarrow{z^{Ro}(i)} = F_{Ro}[Y(i)] = [Y^{Ro}(i, \theta_1), \dots, Y^{Ro}(i, \theta_n)] \quad (12)$$

Where the value of  $\theta_i$  has a certain randomness, and  $F_{Ro}$  denotes the rotation function. In addition to the rotation method, the pre-processing process of cropping only does a certain amount of edge cropping to avoid the loss of valid data information.

$$\overrightarrow{z^{Crop}(i)} = F_{Crop}[Y(i)] = [Y^{Crop}(i, r_1), \dots, Y^{Crop}(i, r_n)] \quad (13)$$

Where the value of  $r_n$  is the intersection ratio between the cropped image and the original image, generally not more than 10 % here. In addition to the traditional methods of rotating and cropping augmentation data, injecting noise is also an effective way to enhance learning. Here use both simple Pretzel noise, and Gaussian noise for random noise generation.

$$\overrightarrow{z^{Noise}(i)} = F_{Noise}[Y^{Noise}(i)] = [Y^{Noise}(i), \dots, Y^{Noise}(i)] \quad (14)$$

### 4.3. Experimental results

**Implementation Details.** The experiment is deployed in the Ubuntu 20.04 environment, LLC-Net model is implemented in Pytorch, using transfer learning techniques with VGG-16 as the base backbone [9], pre-trained on ImageNet as baseline [46]. The experiments were carried out using an NVIDIA DGX Station deep learning workstation equipped with a 24 GB Tesla P40 graphics card. The tuned hyper-parameters of the model were the Learning Rate (LR), Batch Size (BS), Epochs, Optimizer, and Dropout Rate (DR). The BS, Epochs, DR, and LR initial values were set at 16, 120, 0.4, and 0.01, then optimized by SGD to obtain the best experimental results.

In the model parameter setting stage, the BS needs to be set first. According to the performance of Tesla P40 graphics card, 8, 16 and 24 parameter values were set during the experiment for experimental comparison. As shown in Fig. 3, when the BS parameter is set to 16, the loss reduction is very stable and the performance is optimal. At the same

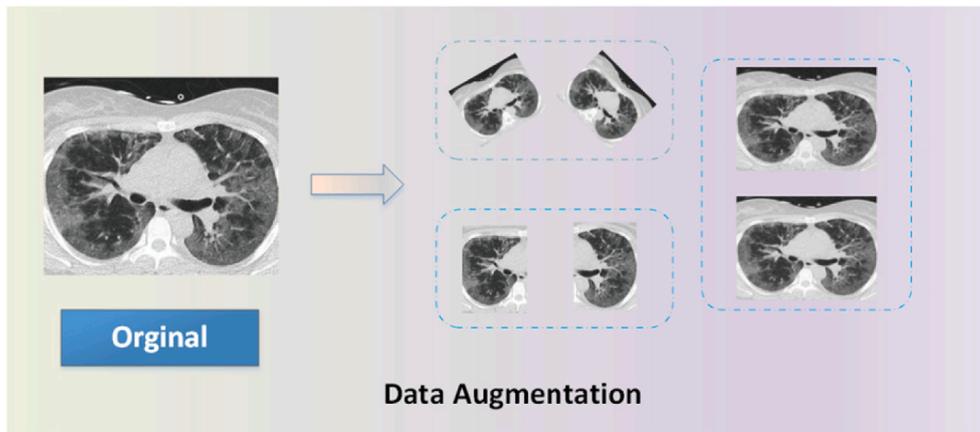


Fig. 2. Visualizing Data Augmentation effects.

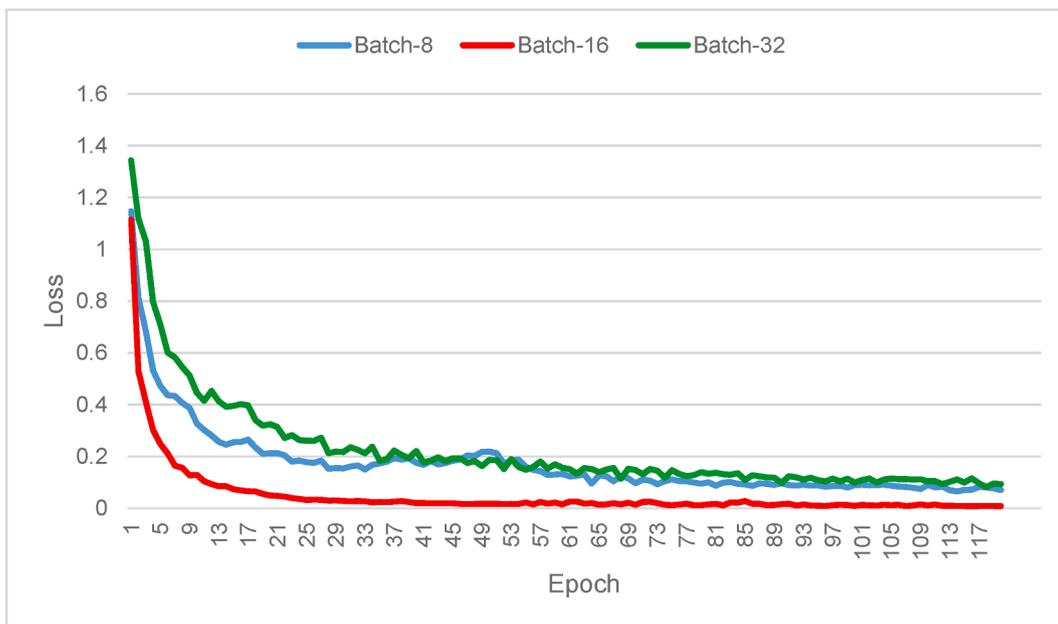


Fig. 3. Impact of Batch Size parameters on loss.

time, it can be seen that the loss converges when Epoch reaches 120, so the value of Epoch parameter is set to it values.

The importance of LR is usually higher than that of DR. Under normal circumstances, the value of LR is set at 0.001, 0.01 or 0.1. In Fig. 4, the loss works best when the LR parameter value is set to 0.01.

DR selection standard is generally around the value of 0.5, according to loss and Precision to select the evaluation parameters. The values of the DR parameter are set to 0.4, 0.5 and 0.6 in the same epoch training cycle for evaluation. As shown in Fig. 5, when DR is set to 0.4, the loss convergence and the Precision are optimal compared to others.

**Metrics of Performance Evaluation.** The process of selecting the performance evaluation metrics for our proposed LLC-Net segmentation network is consistent with other papers, six widely adopted metrics to quantitatively evaluate the method, this makes it easier to evaluate, including Dice similarity Coefficient (Dice), Intersection Over Union (Iou), Sensitivity (Recall), Specificity (SP), Accuracy (Acc) and Precision (Prec). The corresponding Equations are expressed below:

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (15)$$

Among them, TP represents true positive, FP and FN denote false

positive, false negative, respectively. Dice is a popular segmentation evaluation criterion, and the higher the value of similarity between samples, the more similar they are, with values often ranging from 0 to 1.

$$Recall = \frac{TP}{TP + FN} \quad (16)$$

Recall, also known as Sensitivity, is used to evaluate the proportion of positive sample evaluations to the total positive sample of the model, and it mainly examines the positive sample detection rate of the model.

$$Iou = \frac{TP}{TP + FP + FN} \quad (17)$$

Iou is used to calculate the ratio of the intersection and union of a category's prediction result and the true value.

$$SP = \frac{TN}{FP + TN} \quad (18)$$

TN indicates true negative in Eq. (18), exhibiting the same premise as Eq. (17), which primarily evaluates the negative sample completion rate.

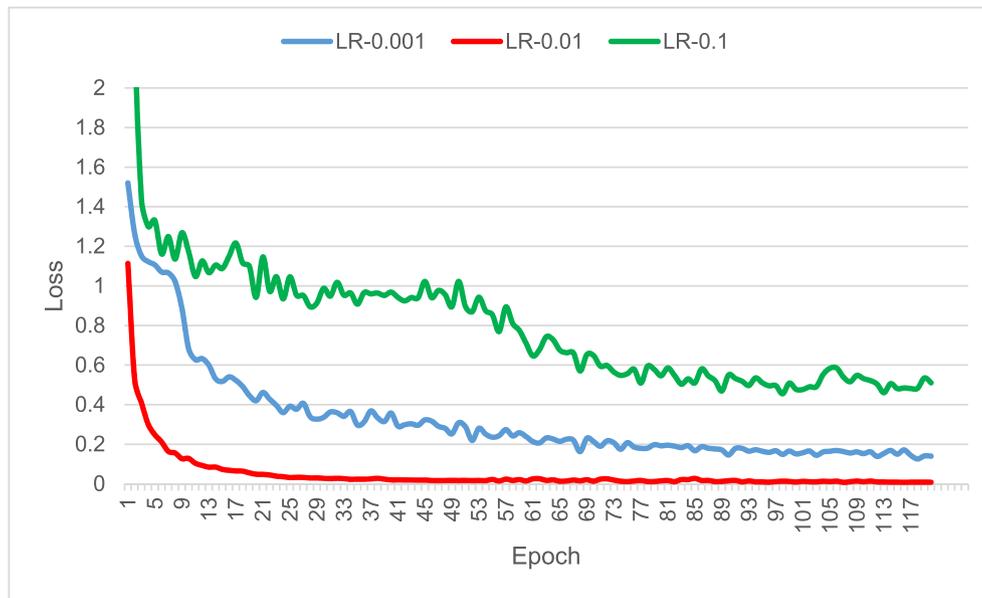


Fig. 4. Impact of Learning Rate parameters on Loss.

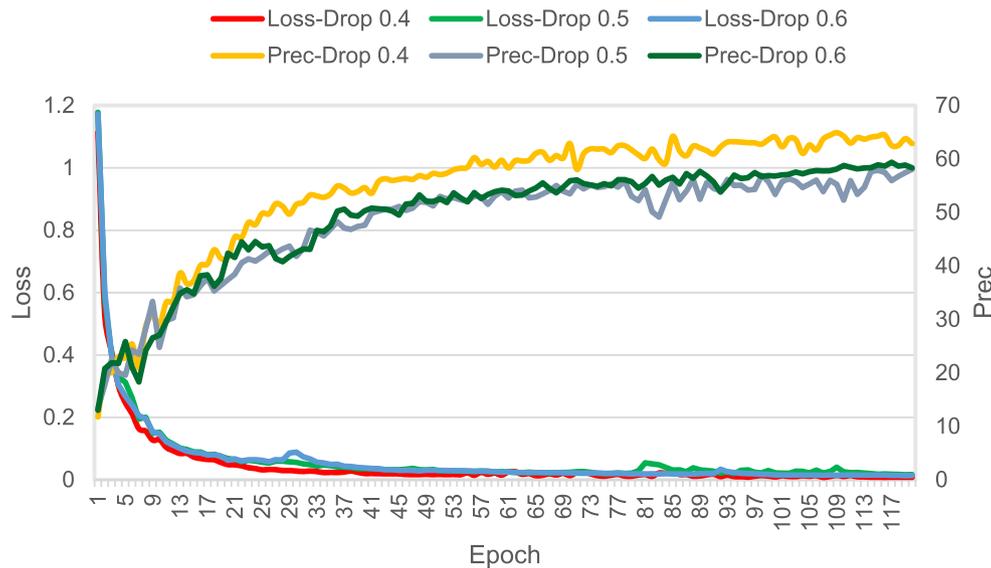


Fig. 5. Impact of Dropout Rate parameters on Loss and Precision.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

In Eq. (19), the Accuracy rate is an important indicator of a good or bad classifier, and the higher the value, the more valuable the classification model is.

$$Prec = \frac{TP}{TP + FP} \quad (20)$$

The model classification process, for a certain class of prediction evaluation criteria is represented by Precision.

#### 4.4. Comparison with State-of-the-arts approaches

To research the planned LLC-Net segmentation performance, we compared our methods with other six state-of-the-art approaches, including FCN8s [47], U-Net [48], PSPNet [49], DeepLabV3+ [50], WSCL [9] and Inf-Net [51] in Fig. 6. To more fairly compare the

superiority of our method, according to the other six methods, we conduct the approaches are implemented on dataset [44].

Table 1 shows the weakly supervised segmentation performance of different network models on COVID-19 in the [44] dataset, all experimental results in Table 1 are averages based on 5-fold cross-validation [52]. In the metrics, Dice is the one that most commonly used. With the help of local lesion coherence on CT images, our method significantly outperforms other weakly supervised segmentation of COVID-19 infection methods.

In addition, we evaluated performance on the dataset [45], and as shown in Table 2, according to the WSCL experimental standards, the data set is separated into separate and mixed evaluations. Reference [9] for more information on the relevant data set division standard. Local lesion consistency on CT images significantly increases performance when compared to the WSCL technique. We believe that local lesion consistency is particularly effective for weakly supervised COVID-19 infection segmentation so far.

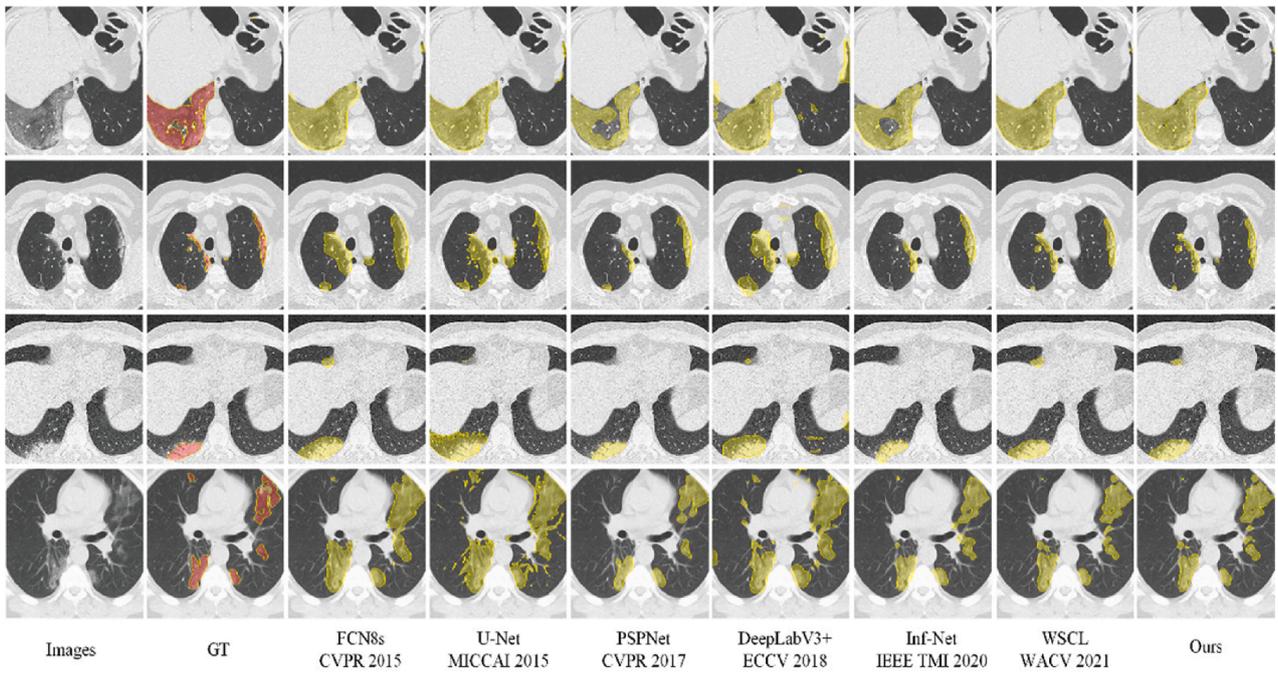


Fig. 6. Qualitative segmentation results comparison by ours methods and others state-of-the-art methods. From the left to right, Weakly Supervised approaches contains FCN8s [47], U-Net [48], PSPNet [49], DeeplabV3+ [50], Inf-Net [51], WSCL [9] and Ours.

Table 1

Comparison with other state-of-the-art methods of different methods on [44] datasets.

Methods	Publication	Dice	Iou	Recall	SP
FCN8s [47]	CVPR 2015	66.45	49.82	81.97	98.56
U-Net [48]	MICCAI 2015	66.58	49.93	82.12	98.52
PSPNet [49]	CVPR 2017	68.79	52.43	81.67	98.70
DeepLabV3+ [50]	ECCV 2018	58.50	43.35	72.33	98.25
Inf-Net [51]	IEEE TMI 2020	73.01	57.50	83.55	98.80
WSCL [9]	WACV 2021	74.13	58.92	77.93	99.07
Ours		<b>76.31</b>	<b>61.71</b>	<b>89.55</b>	98.79

Table 2

On [45] datasets, compare performance with WSCL [9].

Standard	Methods	Dice	Iou	Recall	SP
Separate	WSCL[9]	0.75	0.59	0.86	0.97
	Ours	<b>0.77</b>	<b>0.62</b>	<b>0.91</b>	<b>0.99</b>
Mixed Split	WSCL[9]	0.68	0.51	<b>0.85</b>	0.99
	Ours	<b>0.70</b>	<b>0.54</b>	0.82	<b>0.99</b>

#### 4.5. Ablation study

In this subsection, the performance of each key point optimization of LLC-Net is performed by performing on the experimental dataset [44], which are LSC, SC and EAM. Refer to Table 3 for detailed data.

To demonstrate the LLC-Net training process, added Fig. 7. As shown in the figure, the horizontal axis coordinates represent Epoch, the left

Table 3

Ablation experiments with all the blocks results.

Base	LSC	ST	EAM	MedSeg [44] Datasets Metrics (%)			
				Dice	Iou	Recall	SP
✓				75.43	60.55	<b>96.44</b>	98.57
✓	✓			75.77	60.99	95.71	98.62
✓	✓	✓		76.37	61.77	95.99	98.66
✓	✓	✓	✓	<b>77.24</b>	<b>62.92</b>	86.44	<b>99.09</b>

vertical axis corresponds to the training Loss, and the right vertical axis represents the Acc value, Prec value of the validation datasets during the training process. Along with the reduction of training Loss, the Acc value tends to be stable and almost close to 100 %, Prec reaches the highest value of 64.91 at epoch 109 stage, which is almost 2 percentage points higher compared to WSCL under the same testing conditions.

**Qualitative Ablation Visualization.** In Fig. 8, to demonstrate the effect of different stages of the ablation experiment, we randomly selected 4 example figures from the [44] dataset for presentation. The leftmost column is the sampling information of the CT scan lung images, the second column is the ground truth with COVID-19 lung infection after labeling, and the third to the sixth columns are the evaluation effects of different stages of the ablation experiment, which can be referred to Table 3 for better argumentation analysis. In the third column our method can effectively capture the infected region, although there are obvious misjudgments. In the fourth column, thanks to the Local Self-Coherence Mechanism, the irrelevant judgment regions are removed, but some of the prediction maps mask the important focal areas. In the fifth column, the effective focal areas are re-labeled by the Scale Transformation for Self-Correlation Mechanism. In the last column, the edge of the focal area is further modified and enhanced by the lesion infection Edge Attention Module.

## 5. Conclusions

In this research, we propose LLC-Net, a simple and effective weakly supervised CT infection region segmentation network for COVID-19. Our method is particularly useful since it has a minimal overall dependency on the network topology and has some reference and portability. For lesion area detection, propose a Local Self-Coherence Mechanism. In addition, the Scale Transform for Self-Correlation solves the problem of insufficient sample size. More crucially, we apply the Lesion Infection Edge Attention Module to improve the edge information of the lesion region, decreasing the unfavorable influence of the Local Self-Coherence Mechanism. We successfully demonstrate that our strategy beats state-of-the-art weakly supervised segmentation models and enhances the performance of state-of-the-art in public data tests.

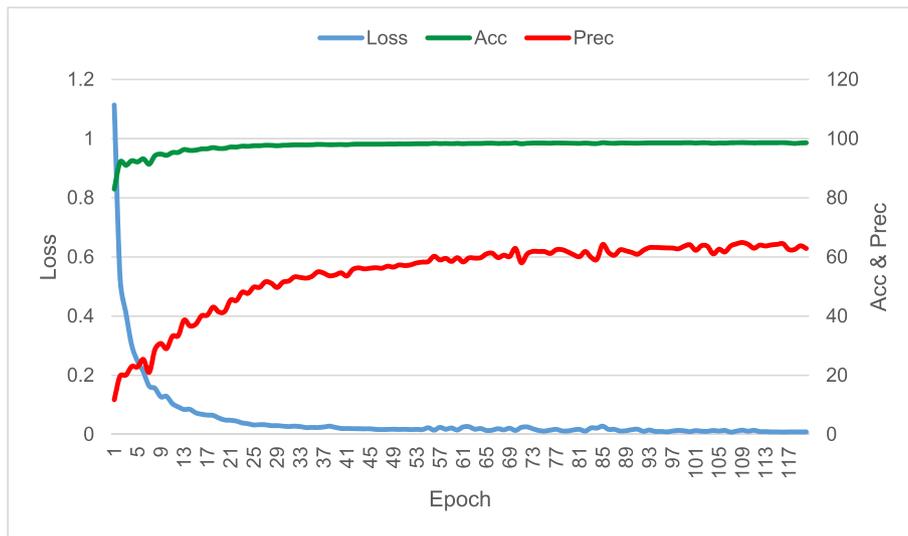


Fig. 7. LLC-Net model training performance fluctuation chart.

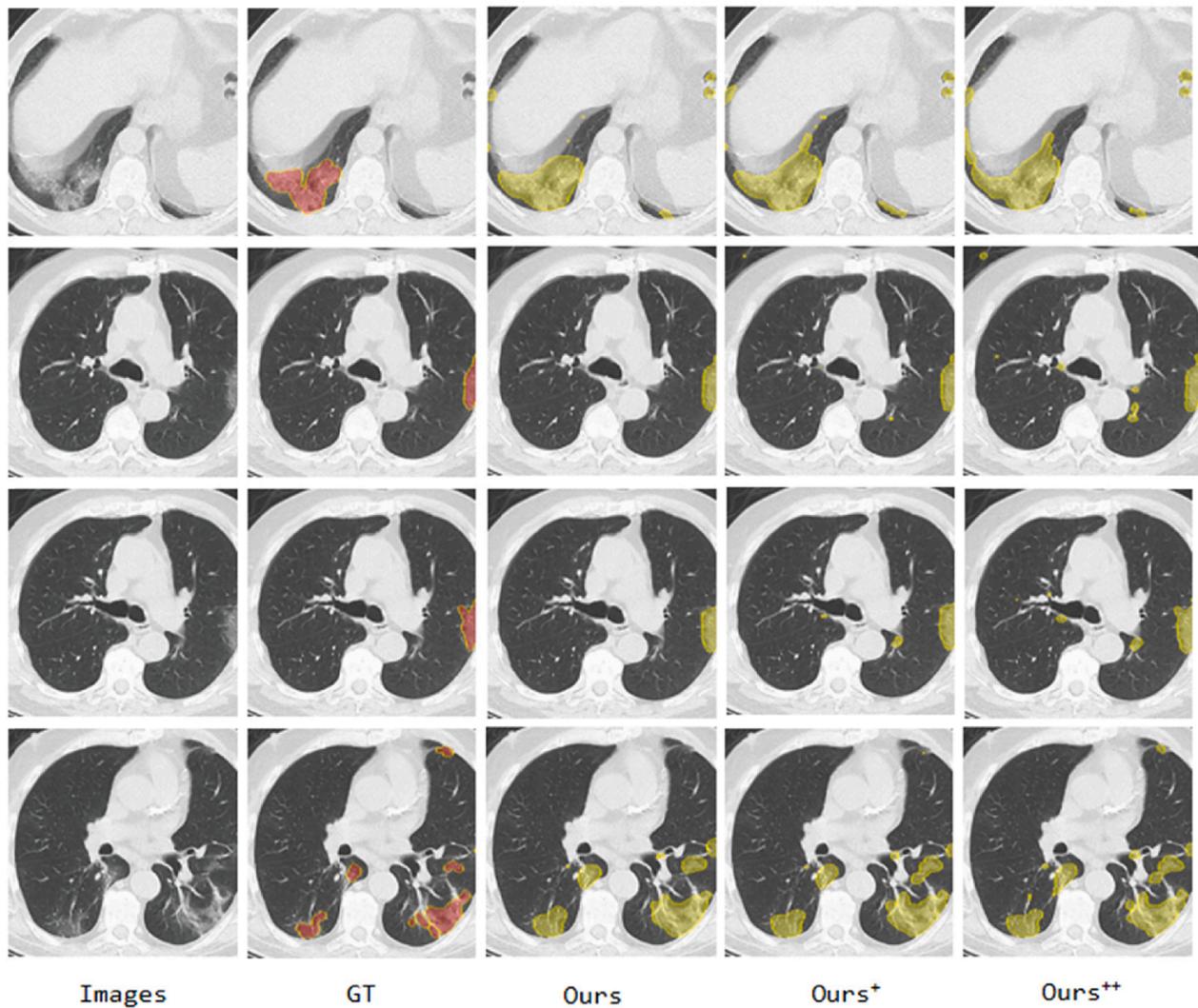


Fig. 8. Qualitative COVID-19 infection segmentation results presented by our different component approach. Ours, Ours<sup>+</sup>, and Ours<sup>++</sup> in the table are shown with reference to the model settings in the second, third, and fourth rows of the Table 3.

## CRedit authorship contribution statement

**Wanchun Sun:** Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Xin Feng:** Methodology, Supervision, Project administration. **Jingyao Liu:** Data curation, Visualization, Formal analysis. **Hui Ma:** Resources, Investigation.

## Declaration of Competing Interest

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

## Data availability

No data was used for the research described in the article.

## References

- [1] W.H. Organization. WHO coronavirus disease (COVID-19) Dashboard. 2022.
- [2] J. Qin, C. You, Q. Lin, T. Hu, S. Yu, X.-H. Zhou, Estimation of incubation period distribution of COVID-19 using disease onset forward time: a novel cross-sectional and forward follow-up study, *Sci. Adv.* 6 (2020) eabc1202.
- [3] C. Qin, L. Zhou, Z. Hu, S. Zhang, S. Yang, Y. Tao, et al., Dysregulation of immune response in patients with coronavirus 2019 (COVID-19) in Wuhan, China, *Clin. Infect. Dis.* 71 (2020) (2019) 762–768.
- [4] W.J. Wiersinga, A. Rhodes, A.C. Cheng, S.J. Peacock, H.C. Prescott, Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review, *JAMA* 324 (2020) (2019) 782–793.
- [5] J. Ma, Y. Wang, X. An, C. Ge, Z. Yu, J. Chen, et al. Towards efficient COVID-19 CT annotation: A benchmark for lung and infection segmentation. (2020).
- [6] M.E. Karar, E.-E.-D. Hemdan, M.A. Shouman, Cascaded deep learning classifiers for computer-aided diagnosis of COVID-19 and pneumonia diseases in X-ray scans, *Complex & Intelligent Systems.* 7 (2021) 235–247.
- [7] Y. Oh, S. Park, J.C. Ye, Deep learning COVID-19 features on CXR using limited training data sets, *IEEE Trans. Med. Imaging* 39 (2020) 2688–2700.
- [8] Y.-H. Wu, S.-H. Gao, J. Mei, J. Xu, D.-P. Fan, R.-G. Zhang, et al., Jcs: An explainable covid-19 diagnosis system by joint classification and segmentation, *IEEE Trans. Image Process.* 30 (2021) 3113–3126.
- [9] I. Laradji, P. Rodriguez, O. Manas, K. Lensink, M. Law, L. Kurzman, et al., A weakly supervised consistency-based learning method for covid-19 segmentation in ct images, in: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021*, pp. 2453–2462.
- [10] X. Liu, Q. Yuan, Y. Gao, K. He, S. Wang, X. Tang, et al., Weakly supervised segmentation of covid19 infection with scribble annotation on ct images, *Pattern Recogn.* 122 (2022), 108341.
- [11] R. Karthik, R. Menaka, M. Hariharan, G. Kathiresan, Ai for COVID-19 detection from radiographs: Inclusive analysis of state of the art techniques, key challenges and future directions, *IRBM.* (2021).
- [12] M.H. Hesamian, W. Jia, X. He, P. Kennedy, Deep learning techniques for medical image segmentation: achievements and challenges, *J. Digit. Imaging* 32 (2019) 582–596.
- [13] H. Huang, H. Zheng, L. Lin, M. Cai, H. Hu, Q. Zhang, et al., Medical image segmentation with deep atlas prior, *IEEE Trans. Med. Imaging* 40 (2021) 3519–3530.
- [14] K.B. Girum, G. Créhange, A. Lalande, Learning with context feedback loop for robust medical image segmentation, *IEEE Trans. Med. Imaging* 40 (2021) 1542–1554.
- [15] F.-P. An, J.-E. Liu, Medical image segmentation algorithm based on multilayer boundary perception-self attention deep learning model, *Multimedia Tools and Applications.* 80 (2021) 15017–15039.
- [16] R. Karthik, R. Menaka, S. Anand, A. Johnson, K. Srilakshmi, Attention-Based Residual Learning Network for COVID-19 Detection Using Chest CT Images, *Decision Sciences for COVID-19.* Springer (2022) 367–391.
- [17] X. Zhang, S. Lu, S.-H. Wang, X. Yu, S.-J. Wang, L. Yao, et al., Diagnosis of COVID-19 pneumonia via a novel deep learning architecture, *Journal of computer science and technology.* 37 (2022) 330–343.
- [18] W. Shui-Hua, M.A. Khan, V. Govindaraj, S.L. Fernandes, Z. Zhu, Z. Yu-Dong, Deep rank-based average pooling network for COVID-19 recognition, *Computers, Materials, & Continua.* (2022) 2797–2813.
- [19] H. Hu, L. Shen, Q. Guan, X. Li, Q. Zhou, S. Ruan, Deep co-supervision and attention fusion strategy for automatic COVID-19 lung infection segmentation on CT images, *Pattern Recogn.* 124 (2022), 108452.
- [20] R. Karthik, R. Menaka, M. Hariharan, D. Won, Contour-enhanced attention CNN for CT-based COVID-19 segmentation, *Pattern Recogn.* 125 (2022), 108538.
- [21] Z. Li, S. Zhao, Y. Chen, F. Luo, Z. Kang, S. Cai, et al., A deep-learning-based framework for severity assessment of COVID-19 with CT images, *Expert Syst. Appl.* 185 (2021), 115616.
- [22] N.S. Punn, S. Agarwal, CHS-Net: A Deep Learning Approach for Hierarchical Segmentation of COVID-19 via CT Images, *Neural Process. Lett.* 1–22 (2022).
- [23] O. Frank, N. Schipper, M. Vaturi, G. Soldati, A. Smargiassi, R. Inchingolo, et al., Integrating domain knowledge into deep networks for lung ultrasound with applications to covid-19, *IEEE Trans. Med. Imaging* 41 (2021) 571–581.
- [24] H. Alliou, M.A. Mohammed, N. Benamer, B. Al-Khateeb, K.H. Abdulkareem, B. Garcia-Zapirain, et al., A multi-agent deep reinforcement learning approach for enhancement of COVID-19 CT image segmentation, *Journal of personalized medicine.* 12 (2022) 309.
- [25] W. Ding, M. Abdel-Basset, H. Hawash, O.M. Elkomy, MT-nCov-Net: A Multitask Deep-Learning Framework for Efficient Diagnosis of COVID-19 Using Tomography Scans, *IEEE Transactions on Cybernetics.* (2021).
- [26] X. Zhao, P. Zhang, F. Song, G. Fan, Y. Sun, Y. Wang, et al., D2A U-Net: Automatic segmentation of COVID-19 CT slices based on dual attention and hybrid dilated convolution, *Comput. Biol. Med.* 135 (2021), 104526.
- [27] Q. Zhou, S. Wang, X. Zhang, Y.-D. Zhang, WVALE: Weak variational autoencoder for localisation and enhancement of COVID-19 lung infections, *Comput. Methods Programs Biomed.* (2022) 106883.
- [28] R. Karthik, R. Menaka, M. Hariharan, D. Won, CT-based severity assessment for COVID-19 using weakly supervised non-local CNN, *Appl. Soft Comput.* 121 (2022), 108765.
- [29] K. Wu, B. Jelfs, X. Ma, R. Ke, X. Tan, Q. Fang, Weakly-supervised lesion analysis with a CNN-based framework for COVID-19, *Phys. Med. Biol.* 66 (2021), 245027.
- [30] Z. Yang, L. Zhao, S. Wu, C.-Y.-C. Chen, Lung lesion localization of COVID-19 from chest CT image: a novel weakly supervised learning method, *IEEE J. Biomed. Health. Inf.* 25 (2021) 1864–1872.
- [31] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, et al., A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, *IEEE Trans. Med. Imaging* 39 (2020) 2615–2625.
- [32] W. Yang, T. Zhang, Z. Mao, Y. Zhang, Q. Tian, F. Wu, Multi-scale structure-aware network for weakly supervised temporal action detection, *IEEE Trans. Image Process.* 30 (2021) 5848–5861.
- [33] Z. Chen, Q. Xu, R. Cong, Q. Huang, Global context-aware progressive aggregation network for salient object detection, in: *Proceedings of the AAAI Conference on Artificial Intelligence, 2020*, pp. 10599–10606.
- [34] A. Obukhov, S. Georgoulis, D. Dai, L. Van Gool. Gated CRF loss for weakly supervised semantic image segmentation. arXiv preprint arXiv:190604651. (2019).
- [35] S. Yi, H. Ma, X. Wang, T. Hu, X. Li, Y. Wang, Weakly-supervised semantic segmentation with superpixel guided local and global consistency, *Pattern Recogn.* 124 (2022), 108504.
- [36] Y. Wang, J. Zhang, M. Kan, S. Shan, X. Chen, Self-supervised equivariant attention mechanism for weakly supervised semantic segmentation, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020*, pp. 12275–12284.
- [37] S. Yu, B. Zhang, J. Xiao, E.G. Lim, Structure-consistent weakly supervised salient object detection with local saliency coherence, in: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). AAAI Palo Alto, CA, USA, 2021.*
- [38] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, A.K. Nandi, Medical image segmentation using deep learning: A survey, *IET Image Proc.* 16 (2022) 1243–1267.
- [39] S. Asgari Taghanaki, K. Abhishek, J.P. Cohen, J. Cohen-Adad, G. Hamarneh, Deep semantic segmentation of natural and medical images: a review, *Artif. Intell. Rev.* 54 (2021) 137–178.
- [40] Z.-H. Zhou, A brief introduction to weakly supervised learning, *Natl. Sci. Rev.* 5 (2018) 44–53.
- [41] E. Granger, P. Cardinal, Weakly Supervised Learning for Facial Behavior Analysis, A Review. arXiv preprint arXiv:210109858. (2021).
- [42] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, P. Torr, Res2net: A new multi-scale backbone architecture, *IEEE Trans. Pattern Anal. Mach. Intell.* 43 (2019) 652–662.
- [43] Y. Liu, M.-M. Cheng, X. Hu, K. Wang, X. Bai, Richer convolutional features for edge detection, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017*, pp. 3000–3009.
- [44] MedSeg. Covid-19 CT Segmentation Dataset, URL <https://medicalsegmentation.com/COVID19>. (2020).
- [45] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, P. Xie, COVID-CT-dataset: a CT scan dataset about COVID-19. arXiv preprint arXiv:200313865. (2020).
- [46] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., Imagenet large scale visual recognition challenge, *Int. J. Comput. Vision* 115 (2015) 211–252.
- [47] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015*, pp. 3431–3440.
- [48] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention.* Springer, 2015. p. 234-41.
- [49] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017*, pp. 2881–2890.
- [50] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: *Proceedings of the European Conference on Computer Vision (ECCV), 2018*, pp. 801–818.
- [51] D.P. Fan, T. Zhou, G.P. Ji, Y. Zhou, L. Shao, Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Transactions on Medical Imaging.* PP (2020) 1-.
- [52] T.-T. Wong, Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation, *Pattern Recogn.* 48 (2015) 2839–2846.