



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.



GFNet: Automatic segmentation of COVID-19 lung infection regions using CT images based on boundary features[☆]



Chaodong Fan^{a,b,c,d}, Zhenhuan Zeng^{b,*}, Leyi Xiao^{a,d,e,f,g,*}, Xilong Qu^d

^a School of Computer Science and Technology, Hainan University, Haikou 570228, China

^b School of Computer Science, Xiangtan University, Xiangtan 411100, China

^c Foshan Green Intelligent Manufacturing Research Institute of Xiangtan University, Foshan 528000, China

^d School of Information Technology and Management, Hunan University of Finance and Economics, Changsha 410205, China

^e AnHui Key Laboratory of Detection Technology and Energy Saving Devices, AnHui Polytechnic University, Wuhu 241000, China

^f Fujian Provincial Key Laboratory of Data Intensive Computing, Quanzhou Normal University, Quanzhou 362000 China

^g Vehicle Measurement, Control and Safety Key Laboratory of Sichuan Province, Xihua University, Chengdu 610039, China

ARTICLE INFO

Article history:

Received 25 September 2021

Revised 31 July 2022

Accepted 7 August 2022

Available online 8 August 2022

Keywords:

Image segmentation

COVID-19

Edge-guidance

Convolutional neural network

CT image

ABSTRACT

In early 2020, the global spread of the COVID-19 has presented the world with a serious health crisis. Due to the large number of infected patients, automatic segmentation of lung infections using computed tomography (CT) images has great potential to enhance traditional medical strategies. However, the segmentation of infected regions in CT slices still faces many challenges. Specially, the most core problem is the high variability of infection characteristics and the low contrast between the infected and the normal regions. This problem leads to fuzzy regions in lung CT segmentation. To address this problem, we have designed a novel global feature network (GFNet) for COVID-19 lung infections: VGG16 as backbone, we design a Edge-guidance module (Eg) that fuses the features of each layer. First, features are extracted by reverse attention module and Eg is combined with it. This series of steps enables each layer to fully extract boundary details that are difficult to be noticed by previous models, thus solving the fuzzy problem of infected regions. The multi-layer output features are fused into the final output to finally achieve automatic and accurate segmentation of infected areas. We compared the traditional medical segmentation networks, UNet, UNet++, the latest model Inf-Net, and methods of few shot learning field. Experiments show that our model is superior to the above models in Dice, Sensitivity, Specificity and other evaluation metrics, and our segmentation results are clear and accurate from the visual effect, which proves the effectiveness of GFNet. In addition, we verify the generalization ability of GFNet on another "never seen" dataset, and the results prove that our model still has better generalization ability than the above model. Our code has been shared at <https://github.com/zengzhenhuan/GFNet>.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Since December 2019, a large number of novel coronavirus cases have been reported in Wuhan, Hubei Province, China, and the number of infections is increasing. Novel coronavirus can cause

acute respiratory diseases in humans and may even cause fatal acute respiratory distress syndrome (ARDS). The International Committee on Taxonomy of Viruses (ICTV) named the novel coronavirus SARS-CoV-2, and the World Health Organization (WHO) named it COVID-19. The Novel Coronavirus has been confirmed as capable of human-to-human transmission. The Novel Coronavirus spreads rapidly in China and the world due to massive traffic and population movement during the Spring Festival. The Novel Coronavirus has extremely high rates of morbidity and mortality. According to the World Health Organization (WHO), as of 10 April 2020, there have been 1521, 252 confirmed cases globally [1]. As of 12 July 2021, there have been 180 million confirmed cases worldwide and 4 million cumulative deaths [2].

Reverse transcription polymerase chain reaction (RT-PCR) is considered by the industry to be the gold standard for screening

[☆] This work is supported by [Hunan Provincial Natural Science Foundation of China](#) (No. 2020JJ4587), Guangdong Basic and Applied Basic Research Foundation (No. 2019A1515110423), Degree & Postgraduate Education Reform Project of Hunan Province (No. 2019JGYB115), Scientific Research Project of Hunan Provincial Department of Education (No. 21C0922), Open Fund Project of Fujian Provincial Key Laboratory of Data Intensive Computing (No. BD202004), Open Research Fund of AnHui Key Laboratory of Detection Technology and Energy Saving Devices (No. JCKJ2021B05), Open Fund Project of Vehicle Measurement, Control and Safety Key Laboratory of Sichuan Province (No. QCCK2021-006).

* Corresponding authors.

E-mail addresses: zzh9573@hotmail.com (Z. Zeng), xiaolyttkx@163.com (L. Xiao).

for COVID-19. However, this has limited rapid and accurate detection. In addition, RT-PCR testing has been reported to have a high rate of false negatives. As an important complement to RT-PCR testing, radiological imaging techniques such as X-ray and computed tomography (CT) have also played a role in current diagnosis, including follow-up evaluation and assessment of disease progression [7,8]. A clinical study of 1014 patients from Wuhan, China, showed that chest CT analysis could achieve a sensitivity of 0.97, a specificity of 0.25, and an accuracy of 0.68 in detecting cases of neocoronary pneumonia, accompanied by RT-PCR test results as a reference [23]. Computed tomography (CT) imaging plays a crucial role in detecting the pulmonary manifestations of COVID-19 [3,4], and the segmentation of infected lesions in CT scan is very important for quantitative measurement of disease progression [5,6]. Lung CT image segmentation is a necessary initial step for lung image analysis [37]. The segmentation of lesions can remove unnecessary background areas and assist doctors in diagnosis, which is an important step in CT image analysis. In contrast to common pneumonia, COVID-19 presents with pulmonary ground-glass opacity (GGO) and signs of pulmonary solidity. CT imaging of COVID-19 in its early stages usually appears as one or more GGO nodular, patchy or lamellar shadows and is generally distributed along the field 1/3 of the lung and under the pleura. As the disease progresses, most patients with COVID-19 will develop solid lung lesions. During the COVID-19 outbreak in Wuhan, a large number of patients, including suspected cases, confirmed cases and follow-up cases, required chest CT examinations to observe the changes and severity of pneumonia. The qualitative evaluation of infection and longitudinal changes in CT slices could thus provide useful and important information in fighting against COVID-19. So we need to segment the lesion regions separately. Existing papers [36] have shown that considering a real-world application, segmentation is an important step since it removes background information, reduces the chance of data leak, and forces the model to focus only on important image areas. Under the current circumstances, any missed cases will continue to lead to COVID-19 transmission. Therefore, a large amount of work and high diagnostic accuracy pose great challenges to radiologists. In addition, the radiologist's eyestrain increases the potential risk of missing some small lesions. In the face of such a serious epidemic, it is very necessary to apply deep learning to disease diagnosis. Early, accurate and rapid diagnosis of suspected cases is crucial to timely isolation and medical treatment, and is of great significance to patient treatment, epidemic control and public health security. So, developing an artificial intelligence (AI) method for COVID-19 computer-aided diagnosis could be very helpful to radiologists.

In Section 2, we will introduce some recent examples of deep learning applied to lung CT segmentation, explain their shortcomings, and then introduce the ideas and solutions of this paper. In Section 3, we will introduce the network structure and its core modules and loss functions in detail. In Section 4, we will introduce the experimental environment, dataset sources, evaluation metrics, and illustrate how to determine the optimal output of the network. Then we will present our experimental results qualitatively and quantitatively and compare them with other methods, and add ablation experiments to verify the effectiveness of the core modules. Finally, the conclusion is given in Section 5.

2. Related work

Recently, deep learning systems have been designed to examine patients infected with COVID-19 via radiological imaging [41,42]. In clinical practice, automatic segmentation of lesions is highly desirable [6]. Although CT scans are important for diagnosis and treatment decisions, automatically segmenting COVID-19 pneumonia lesions from CT scans is challenging for several reasons. First of all,

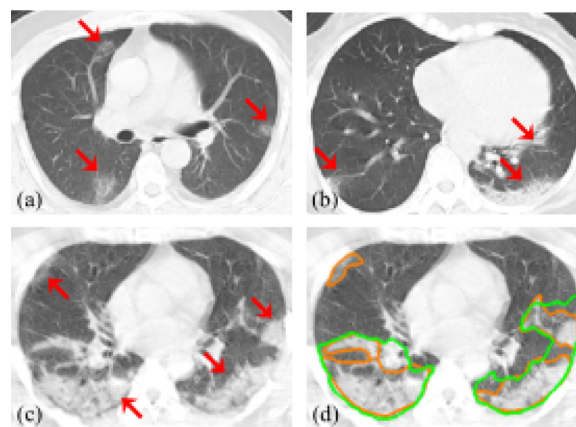


Fig. 1. CT scan of COVID-19 patients with complex findings of pneumonia lesions (a-c) from three different patients, with some lesions highlighted by arrows.(d) Display the labels made by different observers on item (c). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the infected lesions have a variety of complex manifestations, such as ground glass shadow (GGO), consolidation, etc. [9]. Second, the size and location of pneumonia lesions vary greatly between different stages of infection and different patients. As shown in Fig. 1(a-c) [10], the lesions were irregular in shape and blurred in boundary. Some lesions, such as GGO, had low contrast with the surrounding region. In addition, radiologists' labeling of infected region is a highly subjective task, often subject to personal biases and influenced by clinical experience. As shown in Fig. 1(d). Due to the outbreak of COVID-19, it is difficult to collect enough labeled data for deep model training in a short period of time. In addition, obtaining high quality pixel-level lung infection annotations on CT slices is expensive and time consuming. These challenges make the task of automatically segmenting lesions more difficult.

Li et al. [4] used U-Net [12] to segment lung from CT scan. UNet++ [13] is also used to detect [14] and segment [15] infected lesions from CT scan. Liu et al. [21] propose a weakly supervised COVID-19 infection segmentation method with scribble supervision. A et al. [37] design and evaluate an automatic tool for automatic COVID-19 Lung Infection segmentation and measurement using chest CT images. An uncertainty-aware mean teacher framework is incorporated into the proposed method to guide the model training, encouraging the segmentation predictions to be consistent under different perturbations for an input image. With the pixel level uncertainty measure on the predictions of the teacher model, the student model is guided with reliable supervision. Zhao et al. [22] proposed a new deep-learning-based method that integrates a 3D V-Net with shape priors for medical image segmentation. The shape prior was used to optimize the model weights in both V-Net input and output, which significantly improved the model performance. Wang et al. [16] trained the lung segmentation network by using ground-truth mask obtained by unsupervised method, designed an effective lightweight 3D residual network and proposed a weakly supervised COVID-19 lesion detection method. Wang et al. [10] proposed a novel noise-robust-DICE loss function and a noise robust learning framework based on CNNs self-integration, which is robust to noise labels and Dice losses. Fan et al. [17] proposed an Inf-Net automatic segmentation of infected region from lung CT images and a semi-supervised segmentation framework based on random selection propagation strategy, which only required a small number of labeled images and mainly used unlabeled data to alleviate the shortage of high-quality labeled data. Maghdid et al. [38] build a comprehensive dataset of X-rays and CT scan images from multiple sources as well as pro-

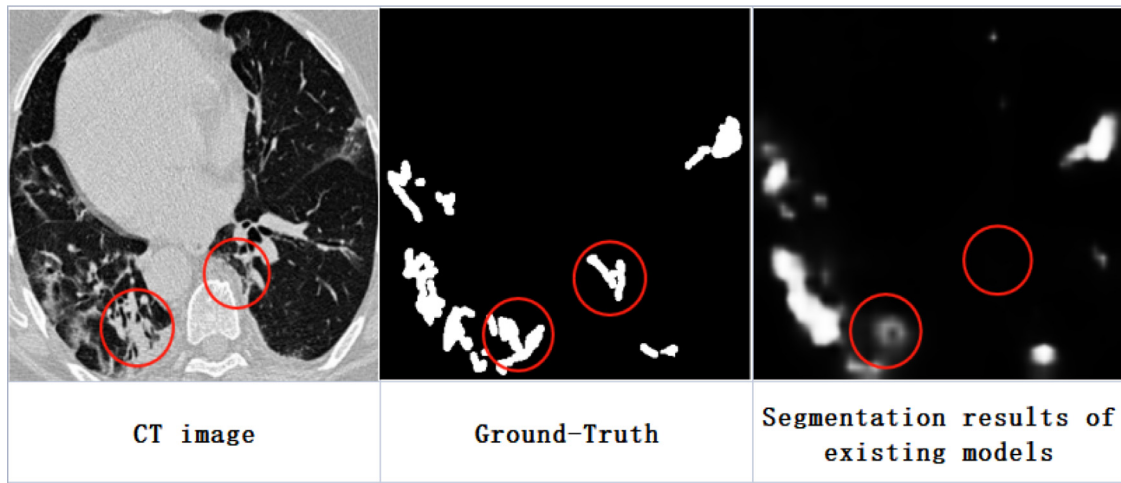


Fig. 2. The position marked by the circle in the figure points out the problem of fuzzy boundary of segmentation results in an existing model [17]. In the two regions marked, it may be difficult for health care workers to determine from the naked eye whether these regions are infected or not. In image segmentation, these regions are difficult to identify by previous methods because of their fuzzy boundaries and low contrast with normal tissues. Such areas of ambiguity may interfere with a health care provider's perception of the infected region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

vides a simple but an effective COVID-19 detection technique using deep learning and transfer learning algorithms. Zhou et al. [43] proposed an ensemble deep learning model for novel COVID-19 detection from CT images. Mu et al. [45] proposed a multi-scale multi-level feature recursive aggregation (mmFRA) network which is used to integrate multi-scale features with multi-level features. Katsamenis et al. [46] proposed a deep learning framework that can detect COVID-19 pneumonia in thoracic radiographs, as well as differentiate it from bacterial pneumonia infection. Voulodimos et al. [25,44] presented a few-shot learning paradigm for segmenting COVID-19 infectious regions. The main difference of the proposed algorithm compared with the traditional approaches is that it is an online learning paradigm, not the static supervised learning of U-Net.

However, all the above methods have a common deficiency, that is, the fuzzy boundary problem of lesion segmentation is not solved completely, in other words, the fuzzy boundary is not exactly segmented. So the segmentation results are still kind of ambiguous. As shown in Fig. 2. If ambiguous, for segmenting the lesion regions will affect the medical staff of judgment, if the normal tissue segmentation as the lesion regions, can make the patient's condition mistaken for serious, if the lesion regions segmentation to normal tissue, may let the patient do not get proper treatment cases lead to misdiagnosis and even produce COVID-19 further spread.

Therefore, the method proposed in this paper is mainly aimed at solving the problem of fuzzy boundary and the problem that normal organizations are difficult to distinguish correctly. In clinical practice, doctors first determine the location of the lesion and then judge the results according to the characteristics of the diseased region when judging the patient's condition according to CT images. According to the above steps, we should focus on the location and boundary of the lesion region when we use the deep neural network for learning.

To solve the above problems, we will carry out the following two main operations successively: 1. Locate the lesion. 2. Accurately extract contour. As we all know, in the convolutional neural network, the resolution of low-layer features is higher and contains more detailed information, but because of less convolution, it has lower semantics and more noise. High-layer features have stronger semantic information, but the resolution is low and the perception of details is poor [11]. How to integrate them efficiently, take their

advantages and discard their faults is the key to improving the segmentation model. Therefore, we designed a kind of framework. First, aggregation high-level feature extracting rough estimate of the lesion region. We also designed a Edge-guidance module, used to provide guidance in the use of reverse attention modules for accurate contour extraction of lesions after localization. Finally, by fusing feature at each layer, the accurate segmentation of the lesion region was obtained.

In short, our contributions in this paper are threefold:

1. A novel deep learning network framework, GFNet, was proposed for the segmentation of infected region of COVID-19 in two-dimensional CT images of lungs. By aggregating the high-layer features using the aggregation module, the aggregated features can capture context information and generate a global location map as an initial boot region for subsequent steps. In order to further dig the boundary information of the target, we use the reverse attention module step by step from the high-layer to the low-layer, then further extract the hidden details of each layer, and finally fuse the features of each layer, so that the network can fully extract the details that are difficult to be noticed by the previous model.

2. We design a Edge-guidance map that contains the boundary features of each layer to further extract the boundary information when the features of each layer are extracted. The experiment proves that this design is very effective.

3. We applied the GFNet framework to VGG16 and used our method on two different datasets. One data set was "seen" to verify learning ability, and the other was "not seen" to verify generalization ability. Experimental results show that GFNet has better learning ability and generalization ability than existing models.

4. We conducted experiments with each model on training datasets of different sizes. Our model can achieve good performance when the training set is relatively small. In real-world case decision making, our GFNet is fully capable of such tasks if it is put into application under time constraints or with few training samples. Our GFNet can also be sufficiently trained to achieve maximum performance if there is sufficient time or a large training sample.

3. Proposed method

In this section, we will give a detailed introduction to the GFNet network structure, core network components, and loss functions.

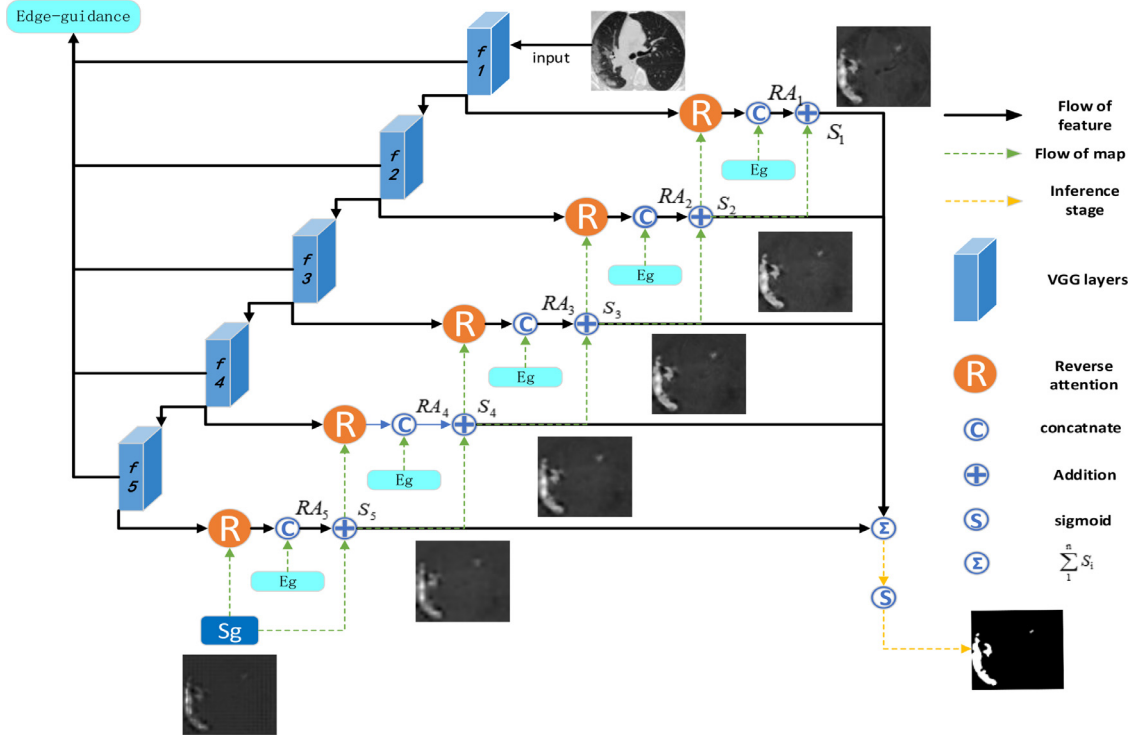


Fig. 3. Structure diagram of GFNet model, f_1 to f_5 are the backbone of VGG16. GFNet contains one Eg, one aggregation module S_g , and five RA modules connected. The input image goes from F_1 to F_5 , then from S_5 back to S_1 via S_g , Eg, RA modules. The final output result is the sum of each lateral output S_i . See III. A for more details.

3.1. Global-feature network(GFNet)

Network overview: The framework of GFNet is shown in Fig. 3. With VGG16 as the backbone network, the whole network is divided into five layers, and the high-layers features are layer 3, layer 4, and layer 5. We use the aggregation module to aggregate the high-layers features (f_3 , f_4 , f_5) to generate global location map S_g [17]. Then, starting from the S_g , importing S_g and f_5 into a reverse attention module (RA), meanwhile, under the guidance of Edge-guidance module(Eg). The output S_5 and f_4 go to the next RA. Repeating the above operation to get S_1 finally, fuse n S_i and get the prediction graph through sigmoid output. Next, we will introduce the core modules in this network and the loss function in detail.

3.2. Edge guidance module

We know that the deeper the feature, the coarser the information, while the low-layer feature contains a lot of useful details. In order to make good use of the features of each layer, based on the idea of edge detection [18], we designed a Edge-guidance module(Eg) fusing the features of each layer, it contains the boundary information of each layer and has rich edge features, which helps us to extract the boundary of the lesion region more effectively and accurately in the subsequent feature extraction, so as to solve the problem of fuzzy boundary. The specific approach is as follows: For the VGG16, VGG16 network composes of 13 conv layers and 3 fully connected layers. We cut all the fully connected layers and the pool5 layer. On the one hand, we remove the fully connected layers due to the fact that they do not align with our design of a fully convolutional network. On the other hand, adding pool5 layer will increase the stride by two times, and it's harmful for edge localization. Its conv layers are divided into five stages, in which a pooling layer is connected after each stage. The useful information captured by each conv layer becomes coarser with its receptive field size increasing. Detailed receptive field sizes of different lay-

Table 1

Detailed receptive field and stride sizes of standard VGG16 net.

layer	conv1 ₁	conv1 ₂	pool ₁	conv2 ₁	conv2 ₂	pool ₂
rf size	3	5	6	10	14	16
stride	1	1	2	2	2	4
layer	conv3 ₁	conv3 ₂	conv3 ₃	pool ₃	conv4 ₁	conv4 ₂
rf size	24	32	40	44	60	76
stride	4	4	4	8	8	8
layer	conv4 ₃	pool ₄	conv5 ₁	conv5 ₂	conv5 ₃	pool ₅
rf size	92	100	132	164	196	212
stride	8	16	16	16	16	32

ers can be seen in Table 1. Hence, we combine hierarchical features from all the conv layers into a holistic framework, in which all of the parameters are learned automatically. Since receptive field sizes of conv layers in VGG16 are different from each other, our network can learn multi-scale, including low-level and objectlevel, information that is helpful to extract more boundary information. Since our boundary guidance module combines all the accessible conv layers to employ richer features, it is expected to achieve a boost in accuracy. At the same time of obtaining the Edge-guidance map, we compute the gradient of Ground-truth of the input image to obtain edge-Ground-truth of the boundary [17]. Then we conduct deep supervision of the Edge-guidance map and let it further learn the boundary features from the rich features. In Fig. 4, we show the Edge-guidance results after combination. The edge obtains strong response at the COVID-19 infection region. Details of this module are shown in Fig. 4.

We use the standard Binary Cross Entropy(BCE) loss function to measure the dissimilarity between the generated Edge-guidance map (E_g) and the edge-Ground-truth map (G_e) calculated by using the gradient of the ground-truth (GT):

$$L_{E_g} = - \sum_{x=1}^w \sum_{y=1}^h [G_e \log(E_g) + (1 - G_e) \log(1 - E_g)] \quad (1)$$

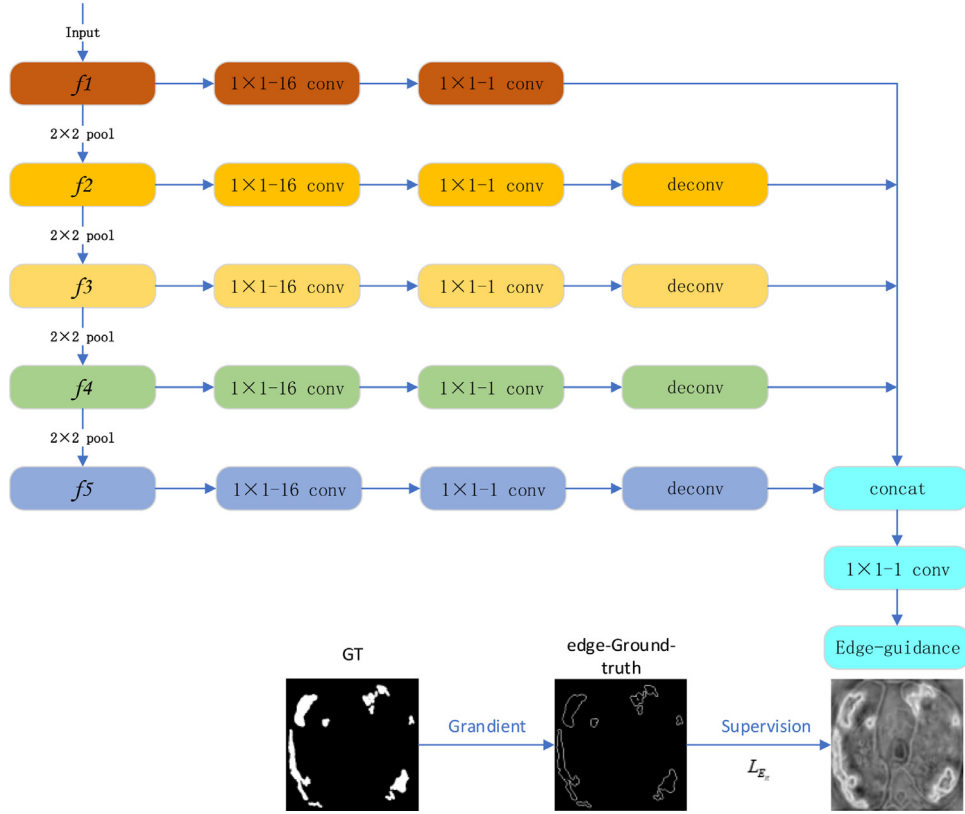


Fig. 4. A Edge-guidance map is generated by fusing every layer feature to rich edge features. See III. B for more details.

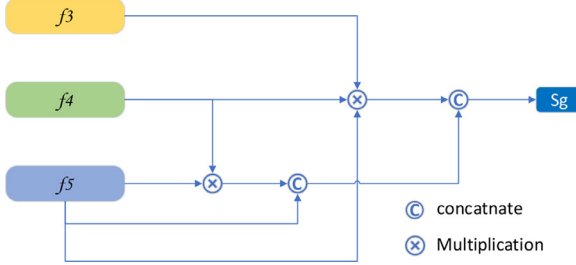


Fig. 5. Aggregated high-layers features are used to generate a global location map.

Where, (x, y) the coordinates of each pixel in the Edge-guidance map E_g and edge-ground-truth map E_e . In addition, w and h respectively represent the width and height of the corresponding feature map.

3.3. Aggregation module

Many existing medical image segmentation networks [19,20], use high and low level features in encoders to segment target organs and lesion regions. However, Wu et al. [35] pointed out that, compared with high-level features, low-level features require more computing resources due to their greater spatial resolution, but contribute less to performance. Inspired by this, we using a *Parallel Partial Decoder (PPD)* modules to aggregate high-layers features. Specifically, for the input 2d CT images, we first obtain three high-layers features $f_i, i=3, 4, 5$ of the VGG16. Then, we use a novel partial decoder $pd()$ [17,35] to aggregate the high-layers features in a paralleled connection. The decoder produces a global location map $S_g = pd(f_3, f_4, f_5)$, which then serves as an initial guide in the RA module. The details are shown in Fig. 5.

3.4. Reverse attention module

In clinical practice, clinicians usually use a two-step operation to segment the lung infection region, first roughly locate the infected region, and then accurately mark these region by examining the local tissue structure. Inspired by this process and [24], we designed GFNet using five reverse attention modules(RA).

First, the *PPD* module acts as a rough locator and generates a high-level global map S_g with unstructured details to provide rough location information for the lung infection region.

Secondly, our structure mine discriminative infection regions in an erasing manner [26,27]. Specifically, compared with simply aggregating features from all layers [27], we propose to adaptively learn the reverse attention in all five layers features. Importing S_g and f_5 into a RA, meanwhile, under the guidance of E_g . The specific approach is as follows: S_{i+1} goes through *Sigmoid* and reverse to get A_i with reverse attention weight, and then A_i is multiplied with the feature f_i of the lower layer. Before getting the RA feature, we concatenate E_g with the result multiplied, so as to get RA_i feature under the guidance of E_g . Finally, we add RA_i with S_{i+1} to get lateral output S_i of layer i . For each lateral output S_i , we conduct deep supervised learning. The output S_5 and f_4 go to the next RA. Repeating the above operation to get S_1 finally. Formula is as follows:

$$RA_i = C(f_i \odot A_i, \text{down}(E_g)) \quad (2)$$

Where $C(\cdot)$ represents concatenate, $\text{down}(\cdot)$ represents down sampling, followed by two two-dimensional convolution layers with 64 filters. In fact, RA weights A_i widely adopted design in the salient object detection in computer vision [27], it is defined as:

$$A_i = \varepsilon(\Theta(\text{Sig}(\text{up}(S_{i+1}))) \quad (3)$$

Among them up represents the upsampling operation, Sig represents activate function for Sigmoid, Θ represents the reverse op-

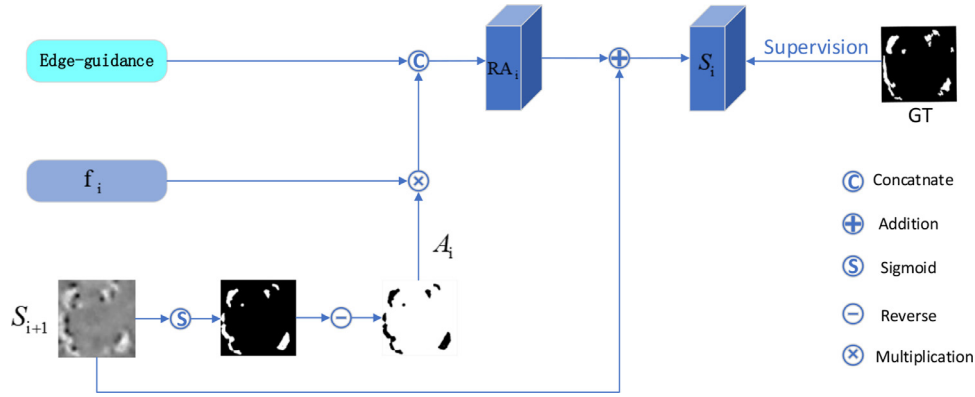


Fig. 6. The reverse attention module is used to implicitly learn edge features.

eration of subtracting the input from a matrix of all 1s. The symbol ε represents the extension of the single-channel feature map to multi-channels, including the reversion of each channel of the candidate tensors in Eq. (2). The final output of this layer S_i is:

$$S_i = RA_i + S_{i+1} \quad (4)$$

The details of this process are shown in Fig. 6. It is worth noting that erasure strategies in RA can ultimately refine inaccurate and coarse prediction region into an accurate and complete prediction map.

In [17], RA module is also used in Res2Net to extract features from layer 5 to layer 3. We differ from them in that here we apply the RA module to each layer and demonstrate that our approach is better in subsequent ablation experiments.

3.5. Loss function

We define the loss function L_{seg} as the combination of the weighted IoU loss function L_{seg}^{IoU} and the weighted binary cross entropy (BCE) loss function L_{BCE}^{w} , namely:

$$L_{seg} = L_{IoU}^{w} + \lambda L_{BCE}^{w} \quad (5)$$

Where λ represents the weight, which was set to 1 in our experiment. The two parts of L_{seg} provide effective global (image-level) and local (pixel-level) supervision to obtain accurate segmentation effects. Unlike the standard IoU loss function, which is widely used in the segmentation task, the weighted IoU loss function increases the weights at the sample points of the difficult pixel to highlight its importance. In addition, compared with the standard BCE loss function, L_{BCE}^{w} pays more attention to the difficult pixel sample points rather than assigning the same weight to all pixels. The definitions of these losses are the same as those in [31,32], and their effectiveness has been validated in target detection domain. We adopt deep supervision for the five lateral outputs ($S_i, i = 1, 2, 3, 4, 5$) and the global location map S_g , each of which is upsampled to the same size as the object-level segmented ground-truth map G . Therefore, the total loss function is:

$$L_{total} = L_{seg}(G, S_g) + \sum_{i=1}^5 L_{seg}(G, S_i) + L_{E_g} \quad (6)$$

4. Experiment

4.1. Experimental environment and parameter setting

Our model was configured with the PyTorch 1.10.0, cuda 11.6 framework and trained on a single NVIDIA RTX 3080 laptop GPU. Before the training, we uniformly resize all the inputs to 352×352 , and a multi-scale training method with a scaling ratio

of [0.75, 1, 1.25] was used to train the datasets, so as to improve the generalization performance of the model. The network is trained using the Adam optimizer. Set the Epoch to 100 with a batch size of 4 and the learning rate to $1e-4$.

4.2. Baselines

For the infected region experiment, we applied the GFNet framework of this paper to VGG16. We first identify the best value for n of $S_{out} = \sum_{i=1}^n S_i$. And compared with two classic segmentation models in the medical field, including U-Net [12], U-Net++ [13], the latest model Inf-Net [17], and two segmentation models “Few-shot UNet” [25] and “SG-One” [39] based on few-shot learning.

4.3. Datasets

Currently, the number of CT images with segmentation labeling is very limited because manually segmenting areas of lung infection is a difficult and time-consuming task, and the outbreak of the disease is in its early stages. To solve this problem, we use a semi-supervised learning strategy to improve GFNet, using a large number of unlabeled CT images to effectively expand the training data. This strategy refers to the method in [17] and is based on the random sampling strategy to gradually expand the unlabeled data of the training dataset. See Algorithm 1 [17] for details. Specific-

Algorithm 1

Require: Labeled training data $D_{Labeled}$ and unlabeled training data $D_{Unlabeled}$
Ensure: Trained GFNet
1: Construct a training dataset $D_{Training}$ using all the labeled CT images from $D_{Labeled}$
2: Training our model GFNet using $D_{Training}$
3: **repeat**
4: Perform testing using the trained model GFNet and K CT images randomly selected from $D_{Unlabeled}$, which yields network-labeled data $D_{Net-labeled}$, consisting of K CT images with pseudo labels
5: Enlarge the training dataset using $D_{Net-labeled}$, i.e., $D_{Training} = D_{Training} \cup D_{Net-labeled}$
6: Remove the K testing CT images from $D_{Unlabeled}$
7: Fine-tune GFNet using $D_{Training}$
8: **until** $D_{Unlabeled}$ is empty
9: **return** Trained model GFNet

cally, we generate the pseudo labels for unlabeled CT images using the procedure described in Algorithm 1. The resulting CT images

with pseudo labels are then utilized to train our model. This semi-supervised approach is simple, it requires no measures to evaluate the labels of predictions, and it has no threshold. Secondly, this strategy can provide better performance than other semi-supervised learning methods and prevent network overfitting. Recent research work [34] has confirmed this conclusion. We use the same training settings as in [17]. We generate pseudo labels for unlabeled CT images using the protocol described in Algorithm 1 [17]. The number of randomly selected CT images is set to 5, i.e., $K = 5$. Now there are 1600 unlabeled images, we need to perform 320 iterations with a batch size of 4. The unlabeled CT images are extracted from the COVID-19 CT Collection [33] dataset, which consists of 20 3D CT volumes from different COVID-19 patients. Fan et al. [17] had extracted 1600 2d CT axial slices from 3D volumes, removed non-lung regions, and constructed an unlabeled training dataset $D_{Unlabeled}$ for effective semi supervised segmentation. After obtaining a semi-supervised dataset with pseudo labels, our next training phase consists of two steps: (i) Pre-training on 1600 CT images with pseudo labels, which takes 7.5 h to converge over 100 epochs with a batch size of 4. (ii) Fine tuning on 60 CT images with the ground-truth labels, which takes about 15 min to converge over 100 epochs with a batch size of 4. These 60 CT images were taken from COVID-19 CT Segmentation dataset [29], which consists of 100 axial CT images collected by the Italian Society of Medical and Interventional Radiology. A radiologist segmented the CT images using different labels to identify lung infections. Although this is the first open-access COVID-19 dataset for lung infection segmentation, it suffers from a small sample size, i.e., only 100 labeled images are available. We employ them as the labeled data $D_{Labeled}$, which consists of 60 CT images randomly selected as training samples, 20 CT images for validation, and the remaining 20 images for testing. We verify the model's learning ability on this testing dataset. In addition, we found a public dataset "COVID-CS" [30] and tested the trained model directly on this dataset to verify our model's generalization ability to "never seen" data. For a fair comparison, we adopt the same training mode and setting for the baseline model.

4.4. Evaluation metrics

According to Han et al. [5], Huang et al. [26], we used three widely used metrics: Dice similarity coefficient, Sensitivity (Sen.), Specificity (Spec.) and Precision (Prec.). We also introduce three golden indicators from the object detection field, including Structure Measure [27], Enhanced-alignment Measure [28], and Mean Absolute Error. In our evaluation, we choose S_{out} with Sigmoid function as the final prediction Seg . The seven metrics used to measure the performance of a model can be expressed as:

4.4.1. Dice coefficient

Dice is mainly used to calculate the similarity of two sets, and its definition is as follows:

$$Dice = \frac{2 * |Seg \cap GT|}{|Seg| + |GT|} \quad (7)$$

4.4.2. Sensitivity (Sen.)

Sen. reflects the percentage of lung infections that were correctly segmented. It is defined as follows:

$$Sen. = \frac{|Seg \cap GT|}{|GT|} \quad (8)$$

4.4.3. Specificity (Spec.)

Spec. reflects the percentage of lung non-infections that were correctly segmented. It is defined as follows:

$$Spec. = \frac{|I - Seg \cap GT|}{|I - GT|} \quad (9)$$

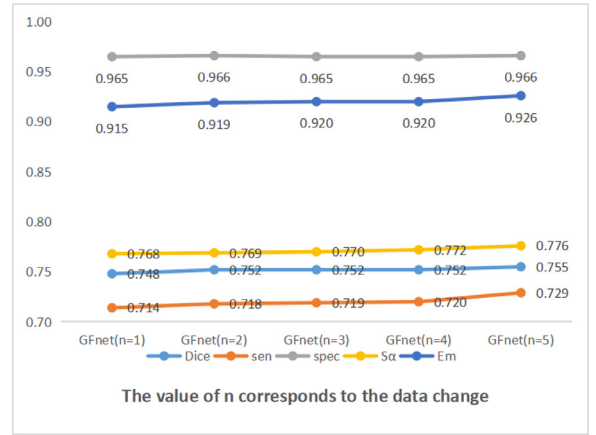


Fig. 7. The value of n corresponds to the data change of [Dice, Sen., Spe., S α , Em].

Among them is the pixel set of the entire CT.

4.4.4. Structural measure (S_α):

This metric is more consistent with the human visual system and is used to measure the structural similarity between the prediction map and the ground-truth mask:

$$S_\alpha = (1 - \alpha) \times S_0(\text{Seg}, G) + \alpha \times S_r(\text{Seg}, G) \quad (10)$$

Where α is the balance factor used to control object-aware similarity S_0 and region-aware similarity S_r . In this paper, we set the same metric value as the original text with the default setting ($\alpha=0.5$).

4.4.5. Enhanced-alignment measure (E_ϕ^{mean})

This is a recently proposed metric that measures both local and global similarity between two binary graphs. The formula is as follows:

$$E_\phi = \frac{1}{w \times h} \sum_x \sum_y \phi(S_p(x, y), G(x, y)) \quad (11)$$

Where w and h represent the width and height of ground-truth map G , and (x, y) represents the coordinates of each pixel in G . The symbol ϕ is an enhanced alignment matrix. We get a set of E_ϕ value by thresholding the prediction Seg with thresholds 0 to 255 to obtain a binary mask. In our experiment, we reported the mean of E_ϕ under all thresholds.

4.4.6. Mean absolute error (MAE)

This metric measures the pixel-level error between Seg and G , and is defined as:

$$MAE = \frac{1}{w \times h} \sum_x \sum_y |Seg(x, y) - G(x, y)| \quad (12)$$

4.5. The determination of the optimal value of n

We first conducted experiments in different n on the "COVID-19 CT Segmentation" dataset [29]. The results can be seen from Fig. 7. Some metrics, such as $Dice$, Sen , S_α and E_ϕ^{mean} , will increase with the number of fusion feature layers increasing. So we set the best value of n to 5, that is, the final output is $S_{out} = S_1 + S_2 + S_3 + S_4 + S_5$. Through this step experiment, we proved that the results combined by all the five lateral output will have the best performance.

Table 2

1600 training sets with pseudo-labels were trained and tested directly on the “COVID-19 CT Segmentation” dataset [29] to verify the validity of the dataset.

Method	Dice	Sen.	Spe.	S_a	E_m	MAE
U-Net	0.682	0.607	0.972	0.710	0.842	0.084
U-Net+	0.584	0.608	0.845	0.625	0.796	0.117
Inf-Net	0.689	0.630	0.970	0.733	0.853	0.070
GFnet(ours)	0.697	0.613	0.976	0.723	0.858	0.069

4.6. Verify the validity of data sets with pseudo labels

Our training process is divided into two steps: first, pre-training on 1600 training sets with pseudo labels, and then fine-tuning with 60 images with ground-truth labels. In this part, we first verify the validity of the pseudo-label dataset to ensure that it is helpful to the learning process of the model. The results are shown in Table 2. It can be seen from the above results that there is a certain gap between the test result after training with only the training set with pseudo-label and the test result after fine-tuning with ground-truth label, indicating that training with only the pseudo label data set is not enough. However, if only ground-truth labels are used for training (“0+60”), due to the small sample size, the final result is inferior to pseudo-label data set in some evaluation metrics such as Spe.. Therefore, we refer to the semi-supervised training method in Inf-Net [17], which firstly used pseudo-label data set for pre-training, and then used images with ground-truth labels for fine-tuning, so as to ensure sufficient training sample size and achieve the best performance.

4.7. Experimental results and analysis

4.7.1. Quantitative analysis

In order to compare the performance of infected region segmentation, the existing segmentation models as U-Net, U-Net++ and Inf-Net were added. In addition, we have added two algorithms for few-shot learning field: Few-shot U-Net and SG-One. The quantitative results in datasets [29] and [30] are shown in Tables 3 and 4. For each model, we used the same training method. We first conducted experiments to observe the influence of training sets of different sizes on the performance of the models. The specific approach is: we divide 1600 pseudo-label training sets into different sizes: [0,400,800,1200,1600]. We train the model on training sets of different sizes. Then we fine-tune models on 60 CT images with the ground-truth labels. In general, the experimental results of the traditional medical segmentation models U-Net and U-Net++ are mediocre. The GFNet proposed in this paper performs the best in many metrics. For the few-shot U-Net, which is based on U-Net and uses an online learning paradigm to further improve the segmentation ability of traditional U-Net for COVID-19. Its performance is improved over traditional U-Net. But for SG-One, another few-shot learning method, it was only moderately effective. Compared with the latest COVID-19 segmentation model and our GFNet, the experimental results of these two methods are not outstanding. Under the influence of training sets of different sizes, the performance of existing methods is gradually enhanced with the increase of data sets. We think this is because as the training set increases, more and more different knowledge is learned by the model. For GFNet, with the increase of training set, the performance of GFNet is slightly improved. This shows that GFNet not only has a good learning ability, but also can achieve good performance when training samples are small. This is because “Edge-guidance” in GFNet is a module with strong robustness and universality. It can learn the boundary features of the target region at

a fast speed and with a small sample size, and expand it to the segmentation result. In the dataset2 [30], our GFNet continues to perform best on almost every metric. This means that our GFNet performs better on “never seen” dataset. The universality of Eg to boundary features in different data sets is verified again. Indicating that it has greater generalization ability overall.

In addition, we also show the training and computational costs of each model in Table 5. In real-world applications, the accuracy of a deep learning model is important. Not only that, its training cost, test speed and other performance are also taken seriously. As can be seen from the table, on a training set of 1600 images, our GFNet is slightly slower than U-Net and Inf-Net in terms of training time. This is due to the relative complexity of our model. However, it is worth mentioning that GFNet can achieve good performance only by relying on a small number of training sets, which greatly saves the training time of the model and has great potential. U-Net++ has a very long training time, because when we were training U-Net++, on the same GPU, the batchsize was set to be smaller in case of insufficient video memory. In terms of test speed, we counted the time taken for a test set of 20 images. Our GFNet is second only to U-Net in testing speed by a small margin. To sum up, in real world case decision making, our GFNet is fully capable of such tasks if it is put into application under time constraints or with few training samples. If there is enough time or a large training sample, our GFNet can be sufficiently trained to achieve the best results.

Finally, we also examine the effect of the number of the parameters on the GFNet performance. To be specific: we changed the number of parameters by adding and removing one layer from the original GFNet (5 layers). We named the changed models GFNet (6) and GFNet (4). We examine the performance of these two models, and the detailed results are shown in Table 6. From the perspective of model performance, when we add or reduce the number of GFNet layers, the performance of the model is degraded. Generally speaking, the higher the number of parameters of the model, the better the performance will be, but the performance of GFNet (6) is slightly reduced. We suspect that this is because the boundary details of COVID-19 are already very difficult to extract. As the number of layers increases, many details are lost. For deeper structures, simple stacking of layers will lead to network degradation [40]. The last, we show the training and test time of GFNet (6) and GFNet (4) in Table 5.

4.7.2. Quantitative analysis

The segmentation results in different datasets of lung infection are shown in Figs. 8 and 9. It is easy to see that our GFNet is significantly better than the baselines. Specifically, the segmentation results generated by it are close to the ground-truth map, and there are fewer tissue regions wrongly segmented. In contrast, the results given by U-Net, U-Net++ are not satisfactory because some results have a large amount of normal tissue that has not been segmented. Although Inf-Net is much better than both, there are still many regions where the boundaries are blurred. Our GFNet demarcates the region with very clear and precise boundaries.

4.8. Ablation experiments

In this Section, we conducted some experiments to verify the core module and steps of GFNet (Edge-guidance module, number of RA module usage (i.e., the best value of i in A_i and S_i ($i = 1, 2, 3, 4, 5$)), as shown in Table 7. The best result of each metric is indicated in red. (The results of the fusion module are given in Fig. 7.

Table 3

Quantitative results of segmentation of infected region on the “COVID-19 CT Segmentation” dataset [29]. The table shows the effect of training sets of different sizes on the performance of each model. The best of each evaluation metric is marked in **bold** and the second best in *italic*.

Method	Training set size (Pseudo label training + Ground-truth label Finetuning)	Dice	Sen.	Spec.	S_{α}	E_{ϕ}^{mean}	MAE
U-Net	0 + 60	0.632	0.653	0.917	0.762	0.880	0.083
	400 + 60	0.651	0.702	0.937	0.765	0.892	0.076
	800 + 60	0.675	0.693	0.930	0.764	0.894	0.078
	1200 + 60	0.682	0.687	0.949	0.763	0.892	0.078
	1600 + 60	0.691	0.720	<i>0.965</i>	0.775	0.891	0.076
Few-shot U-Net	0 + 60	0.645	0.659	0.921	0.759	0.881	0.079
	400 + 60	0.662	0.673	0.934	0.769	0.899	0.078
	800+60	0.683	0.695	0.945	0.772	0.901	0.075
	1200 + 60	0.701	0.713	0.946	0.762	0.898	0.076
	1600 + 60	0.717	0.720	0.952	0.775	0.904	0.073
SG-One	0 + 60	0.594	0.589	0.882	0.643	0.793	0.129
	400 + 60	0.611	0.603	0.890	0.658	0.812	0.114
	800 + 60	0.625	0.616	0.913	0.671	0.824	0.098
	1200 + 60	0.632	0.639	0.927	0.685	0.821	0.087
	1600 + 60	0.640	0.635	0.936	0.693	0.834	0.079
U-Net+	0 + 60	0.607	0.692	0.898	0.681	0.767	0.125
	400 + 60	0.603	0.642	0.894	0.633	0.726	0.145
	800 + 60	0.611	0.618	0.940	0.685	0.827	0.090
	1200 + 60	0.625	0.640	0.930	0.702	0.830	0.098
	1600 + 60	0.623	0.638	0.935	0.693	0.835	0.076
Inf-Net	0 + 60	0.699	0.715	0.928	0.782	0.862	0.078
	400 + 60	0.727	0.721	0.960	0.793	0.901	0.065
	800 + 60	0.739	0.716	0.961	0.781	0.896	0.064
	1200 + 60	0.741	0.728	0.960	0.788	0.906	0.064
	1600 + 60	0.743	0.739	0.963	0.779	0.908	0.062
GFNet(ours)	0 + 60	0.739	0.731	0.944	0.767	0.919	0.064
	400 + 60	0.745	0.725	0.959	0.768	0.921	0.061
	800 + 60	0.748	0.724	0.964	0.766	0.921	0.061
	1200 + 60	0.750	0.721	0.966	0.764	0.922	0.062
	1600 + 60	0.755	0.729	0.966	0.776	0.926	0.059

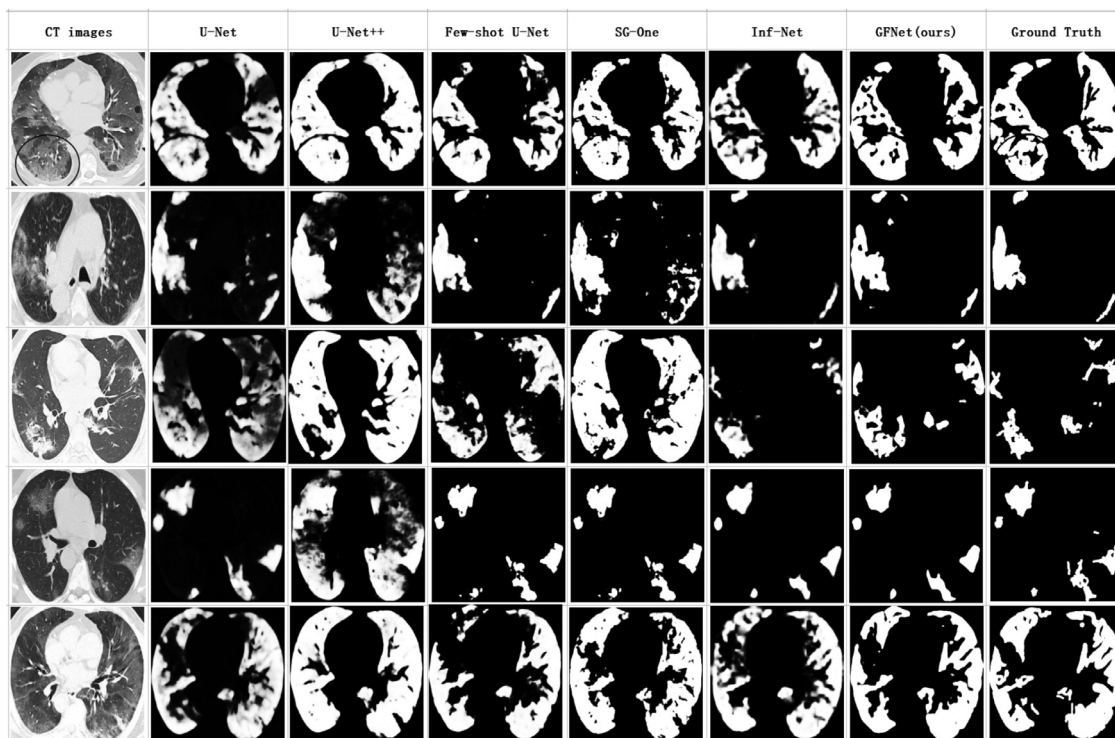


Fig. 8. Comparison of visual effects of segmentation of infected regions on the “COVID-19 CT Segmentation” dataset [29].

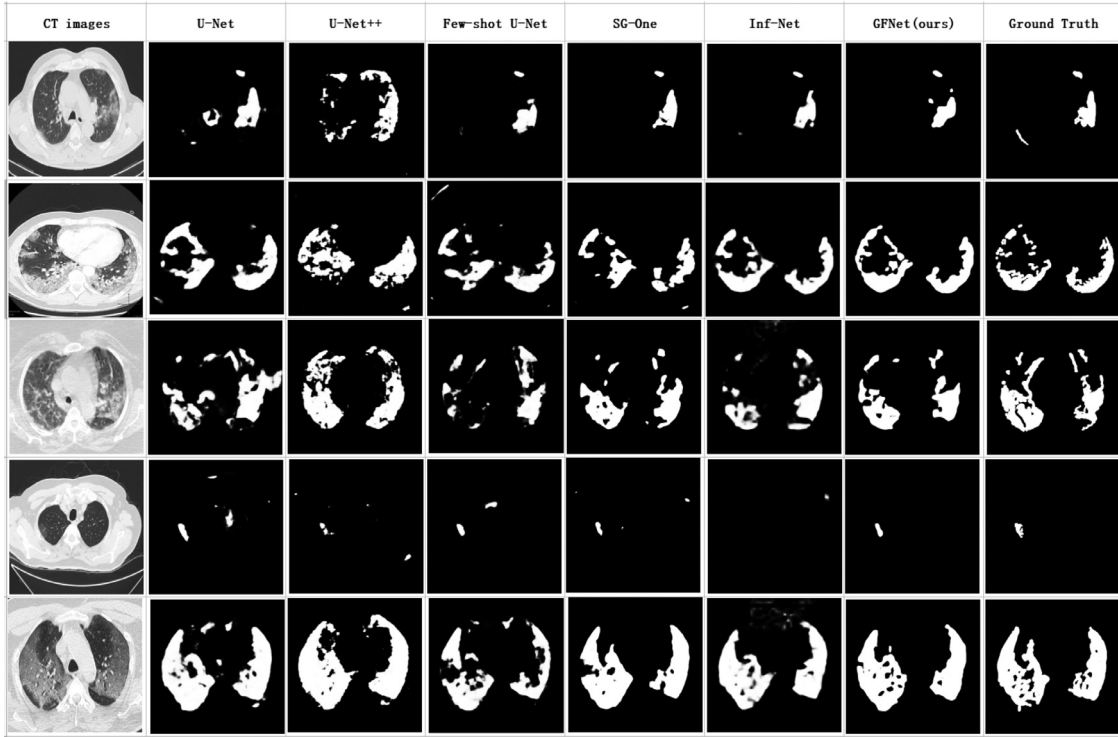


Fig. 9. Comparison of visual effects of segmentation of infected regions on the "COVID-CS" dataset [30].

Table 4

Quantitative results of segmentation of infected region on the dataset of "COVID-CS" [30]. The best segmentation results are shown in **bold**.

Method	Dice	Sen.	Spec.	S_{α}	E_{ϕ}^{mean}	MAE
U-Net	0.602	0.665	0.964	0.719	0.883	0.050
Few-shot U-Net	0.611	0.668	0.964	0.725	0.870	0.051
SG-One	0.596	0.581	0.979	0.721	0.892	0.056
U-Net+	0.485	0.617	0.946	0.650	0.801	0.077
Inf-Net	0.615	0.565	0.963	0.732	0.849	0.038
GFNet(ours)	0.663	0.605	0.981	0.743	0.895	0.033

4.8.1. Effectiveness of edge-guidance module

As can be seen from No.1 and No.2, No.3 and No.4, after the addition of Edge-guidance module, many metrics have significantly improved.

4.8.2. Effectiveness of RA module

As can be seen from No.1 and No.3, after the RA module is added, many metrics have greatly improved.

4.8.3. Effectiveness of the combination of edge-guidance module and RA module

As can be seen from No. 3 and No. 4, when the combination of Eg and RA module is used, many indicators are improved again.

Table 5

The table shows the size of each model and the computational cost for training and testing. The training time was conducted on the training set of 1600 images. The testing time was conducted on a test set of 20 images.

Method	Backbone	Param	FLOP	Training time	Testing speed
U-Net	VGG16	7.9M	38.116G	7.2h	1.51s
U-Net+	VGG16	9.2M	65.938G	26.7h	3.93s
SG-One	VGG16	19.0M	45.916G	7.3h	1.77s
Inf-Net	Res2Net	33.1M	13.922G	5.3h	1.92s
GFNet(ours)	VGG16	18.1M	50.827G	7.5h	1.63s
GFNet(6)(ours)	VGG16	27.4M	49.537G	7.1h	1.69s
GFNet(4)(ours)	VGG16	10.9M	59.391G	9.8h	1.55s

Table 6

Three GFNets which have different number of parameters.their performance in dataset "1600+60".

Method	Dice	Sen.	Spec.	S_{α}	E_{ϕ}^{mean}	MAE
GFNet(6)	0.741	0.713	0.961	0.758	0.912	0.064
GFNet(4)	0.723	0.700	0.958	0.747	0.910	0.067
GFNet	0.755	0.729	0.966	0.776	0.926	0.059

4.8.4. Usage times of RA module

In the paper [17], RA module was used three times, it is not rigorous enough, because we think RA module is a process of fusing high-level information with low-level information. So, as the features start at the deepest S_5 and backward to the S_1 , the information at each layer can be utilized. From No. 4 to No. 8, we can see that with the increase of RA module, the experimental results become better and better.

4.9. Deficiencies and future prospects

Although our GFNet is significantly better than the traditional methods in two datasets, there are still some deficiencies. For example, in terms of training speed and testing speed, our GFNet is not the fastest. We will continue to improve GFNet in this area. In

Table 7Ablation experiments of GFNet. The best segmentation results are shown in **bold**.

No.	BLOCKS							Dice	Sen	Spec	S_α	E_ϕ^{mean}	MAE
	Backbone (VGG16)	E_g	S_5	S_4	S_3	S_2	S_1						
1	✓							0.601	0.612	0.837	0.695	0.720	0.151
2	✓	✓						0.643	0.654	0.863	0.725	0.759	0.127
3	✓		✓					0.639	0.635	0.878	0.734	0.792	0.085
4	✓	✓	✓					0.681	0.667	0.921	0.757	0.856	0.076
5	✓	✓		✓				0.710	0.672	0.944	0.776	0.872	0.069
6	✓	✓			✓			0.725	0.698	0.951	0.791	0.890	0.065
7	✓	✓				✓		0.737	0.715	0.962	0.792	0.899	0.063
8	✓	✓					✓	0.748	0.714	0.965	0.798	0.915	0.061

the future, we will combine GFNet with more advanced models, and improve the existing modules such as Edge-guidance module and the final fusion module.

5. Conclusion

In this paper, we propose a COVID-19 lung CT region segmentation network framework named GFNet. The edge guidance module is added on the basis of the reverse attention module and the aggregation module, which can help the model to capture the boundary information on the fuzzy focus boundary effectively, and then make the boundary segmentation result very accurate. Then, based on the idea of feature fusion, the obtained features are fused. Finally, the model can produce an optimal output. Meanwhile, we proved the problem of the usage times of RA module. In Inf-Net, RA module was used in three deep convolution layers. We proved that the experimental results would be better if RA module was used for each layer. In order to solve the problem of the dataset is too small, we use a semi-supervised approach to extend the training dataset to train GFNet and to prevent overfitting. On the existing two datasets, we verify GFNet's learning ability and generalization ability respectively, and the experiment results show that GFNet is superior to previous models. The qualitative results also show that our segmentation results are more accurate and clear than other models. At the same time, we conducted experiments with each model on training datasets of different sizes. Experimental results show that for other models, increasing the size of the training set will improve the performance of the model. But our GFNet works well when the training set is small. This is because "Edge-guidance" in GFNet is a module with strong robustness and universality. It can learn the boundary features of the target region at a fast speed and with a small sample size, and expand it to the segmentation result. This shows that our GFNet has great potential in image segmentation with fuzzy boundaries facing other similar targets. Finally, we also show the training and computational costs of each model. In our experiment, the training and testing speed of our model is not the fastest, but our model can also achieve good performance when the training set is relatively small. In real world case decision making, our GFNet is fully capable of such tasks if it is put into application under time constraints or with few training samples. Our GFNet can also be sufficiently trained to achieve maximum performance if there is sufficient time or a large training sample. In the future, we will apply the GFNet framework to other medical image segmentation tasks, such as colonoscopic polyps, cells, etc., or image segmentation in other fields. Therefore, our GFNet has great potential to assist healthcare professionals in medical images segmentation.

Declaration of Competing Interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest

of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

Data availability

Our code has been shared at <https://github.com/zengzhenhuan/GFNet>.

References

- [1] WHO, Coronavirus disease 2019 (COVID-19) situation report-81, 2020, (April 10, 2020)[online] Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200410-sitrep-81-covid19.pdf?sfvrsn=ca96eb84_2.
- [2] WHO, Weekly operational update on COVID-19, 2021, 12 July 2021 <https://www.who.int/publications/m/item/weekly-operational-update-on-covid-19-12-july-2021>.
- [3] J. Lei, J. Li, X. Li, X. Qi, CT imaging of the 2019 novel coronavirus (2019-NCoV) pneumonia, *Radiology* (2020) 200236.
- [4] L. Li, et al., Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT, *Radiology* (2020) 200905.
- [5] F.S. Han, et al., Lung infection quantification of COVID-19 in CT images with deep learning, (2020). [arXiv:2003.04655](https://arxiv.org/abs/2003.04655).
- [6] C. Huang, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet* 395 (10223) (2020) 497–506.
- [7] G.D. Rubin, et al., The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner society, *Radiology* (2020) 201365.
- [8] F. Shi, J. Wang, et al., Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19, *IEEE Rev. Biomed. Eng.* (2020).
- [9] M.Y. Ng, et al., Imaging profile of the COVID-19 infection: radiologic findings and literature review, *Radiol. Cardiothorac. Imaging* 2 (1) (2020) E200034.
- [10] G. Wang, et al., A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images, *IEEE Trans. Med. Imaging* (2020). PP(99):1-1
- [11] Y. Liu, et al., Richer convolutional features for edge detection, *CVPR*, 2016.
- [12] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer International Publishing, 2015.
- [13] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, U-Net++: a remote U-Net architecture for medical image segmentation, *MICCAI* 11045 LNCS (2018) 3–11.
- [14] S. Jin, et al., AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical AI system in four weeks, *medRxiv* (2020).
- [15] J. Chen, et al., Deep learning-based model for detecting 2019 novel coronavirus pneumonia high-resolution computed tomography, *medRxiv* (2020).
- [16] X.G. Wang, et al., A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT, *IEEE Trans. Med. Imaging* (2020). PP(99):1-1
- [17] D.P. Fan, et al., Inf-Net: automatic COVID-19 lung infection segmentation from CT images, *IEEE Trans. Med. Imaging* (2020). PP(99):1-1
- [18] H. Fu, J. Cheng, Y. Xu, D.W.K. Wong, J. Liu, X. Cao, Joint optic disc and optical structure. cup segmentation based on multi-label deep network and polar transformation, *IEEE Trans. Med. Imaging* 37 (7) (2018) 1597–1605.
- [19] Z. Gu, et al., CE-NET: context encoder network for 2D medical image segmentation, *IEEE Trans. Med. Imaging* 38 (10) (2019) 2281–2292.
- [20] S. Zhang, et al., Attention guided network for retinal image segmentation, in: *MICCAI*, 2019, pp. 797–805.
- [21] X. Liu, et al., Weakly supervised segmentation of COVID19 infection with scribble annotation on CT images, *Pattern Recognit.* 122 (2022).
- [22] C. Zhao, et al., Lung segmentation and automatic detection of COVID-19 using radiomic features from chest CT images, *Pattern Recognit.* 119 (2021).
- [23] T. Ai, Z. Yang, et al., Correlation of chest CT and RT-PCR testing in coron-

- avirus disease 2019 (COVID-19) in China: a report of 1014 cases, *Radiology* 2019 (2020) 200642.
- [24] S. Chen, X. Tan, B. Wang, X. Hu, Salient attention for salient object detection, in: *ECCV*, 2018, pp. 234–250.
- [25] A. Voulodimos, et al., A few-shot U-Net deep learning model for COVID-19 infected area segmentation in CT images, *Sensors* 21 (6) (2021) 2215.
- [26] F. Shi, L. Xia, et al., Large-scale screening of COVID-19 from the community acquired pneumonia using infection size aware classification, *arXiv* (2020).
- [27] D.-P. Fan, M.-M. Cheng, Y. Liu, T. Li, A. Borji, Structure-measure: a new approach way to evaluate foreground maps, in: *ICCV*, 2017, pp. 4548–4557.
- [28] D.P. Fan, et al., Enhanced alignment measure for binary foreground map evaluation, in: *IJCAI*, 2018, pp. 698–704.
- [29] COVID-19 CT segmentation dataset, 2020, <https://medicalsegmentation.com/covid19>, accessed: 2020-04-11.
- [30] Y.H. Wu, et al., JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation, *IEEE Trans. Image Process.* (2021). PP(99):1-1
- [31] X. Qin, et al., BASNet: boundary-aware salient object detection, in: *CVPR*, 2019, pp. 7479–7489.
- [32] J. Wei, S. Wang, Q. Huang, F3Net: fusion, feedback and focus for salient object detection, *AAAI*, 2020.
- [33] J.P. Cohen, P. Morrison, L. Dao, COVID-19 image data collection, *arXiv* (2020).
- [34] S. Mittal, M. Tatarchenko, O. Çiçek, T. Brox, Parting with illusions about deep active learning, *arXiv* (2019).
- [35] Z. Wu, L. Su, Q. Huang, Cascaded partial decoder for fast and accurate salient object detection, in: *CVPR*, 2019, pp. 3907–3916.
- [36] L. Teixeira, et al., Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest X-ray images, *Sensors* 21 (21) (2021) 7116.
- [37] A. Oulefki, et al., Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images, *Pattern Recognit.* 114 (2021) 107747.
- [38] H.S. Maghdid, et al., Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms, *Multimodal Image Exploitation and Learning 2021*, vol. 11734, International Society for Optics and Photonics, 2021.
- [39] X. Zhang, Y. Wei, Y. Yang, et al., SG-One: similarity guidance network for one-shot semantic segmentation, *IEEE Trans. Cybern.* (2020) 1–11. PP(99)
- [40] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *CVPR* (2016) 770–778.
- [41] W. Zhao, W. Jiang, X. Qiu, Deep learning for COVID-19 detection based on CT images, *Sci. Rep.* 11 (1) (2021) 1–12.
- [42] M.Z. Alom, et al., COVID_MNet: COVID-19 detection with multi-task deep learning approaches, (2020). *arXiv:2004.03747*.
- [43] T. Zhou, et al., The ensemble deep learning model for novel COVID-19 on CT images, *Appl. Soft Comput.* 98 (2021) 106885.
- [44] A. Voulodimos, et al., Deep learning models for COVID-19 infected area segmentation in CT images, in: *The 14th PErvasive Technologies Related to Assistive Environments Conference*, 2021.
- [45] N. Mu, et al., Progressive global perception and local polishing network for lung infection segmentation of COVID-19 CT images, *Pattern Recognit.* 120 (2021) 108168.
- [46] I. Katsamenis, et al., Transfer learning for COVID-19 pneumonia detection and classification in chest X-ray images, in: *24th Pan-Hellenic Conference on Informatics*, 2020.

Chaodong Fan received his master's degree and doctor's degree from Hunan University in 2011 and 2014 respectively. Mainly engaged in intelligent computing, pattern recognition and intelligent systems, smart grid related fields. Some new intelligent optimization algorithms and image processing technologies, such as molecular dynamics optimization algorithm, spatial section projection histogram and post-processing strategy based on threshold, are proposed. His research interests include artificial intelligence, image processing and model recognition.

Zhenhuan Zeng was born in 1995 in Changsha, Hunan Province, China. He is studying in Xiangtan University for his master's degree. His main research interests are evolutionary algorithms, and computer vision.

Leyi Xiao was born in 1986 in Changde, Hunan, China. She graduated from Hunan University with a doctorate in engineering in 2020. Her research interest include image processing and model recognition, which relate to both the field of computer vision and pattern recognition.

Qu Xilong was born in January 1978, graduated from Southwest Jiaotong University with a doctor's degree in 2006. His research interests include manufacturing informatization, distributed system integration technology, information security technology.