COVID Edge-Net: Automated COVID-19 Lung Lesion Edge Detection in Chest CT Images*

Kang Wang(✉), Yang Zhao, Yong Dou, Dong Wen, and Zikai Gao

National Laboratory for Parallel and Distributed Processing,
School of Computer, National University of Defense Technology,
Changsha, Hunan, China
{wangkang,zhaoyang10,yongdou,wendong19,gaozk18}@nudt.edu.cn

Abstract. Coronavirus Disease 2019 (COVID-19) has been spreading rapidly, threatening global health. Computer-aided screening on chest computed tomography (CT) images using deep learning, especially, lesion segmentation, is an effective complement for COVID-19 diagnosis. Although edge detection highly benefits lesion segmentation, an independent COVID-19 edge detection task in CT scans has been unprecedented and faces several difficulties, e.g., ambiguous boundaries, noises and diverse edge shapes. To this end, we propose the first COVID-19 lesion edge detection model: COVID Edge-Net, containing one edge detection backbone and two new modules: the multi-scale residual dual attention (MSRDA) module and the Canny operator module. MSRDA module helps capture richer contextual relationships for obtaining better deep learning features, which are fused with Canny features from Canny operator module to extract more accurate, refined, clearer and sharper edges. Our approach achieves the state-of-the-art performance and can be a benchmark for COVID-19 edge detection. Code related to this paper is available at: https://github.com/Elephant-123/COVID-Edge-Net.

Keywords: COVID-19 · Edge detection · Canny operator · Multi-scale residual dual attention · Computed tomography (CT) images.

1 Introduction

As one of the most serious pandemics, Coronavirus Disease 2019 (COVID-19) [1–3] has been spreading violently around the world since December 2019, causing a devastating effect on global public health and economy. Because it has fast progression and infectious ability [4], it is necessary to develop effective tools or methods to accurately diagnose and evaluate COVID-19. Although the reverse transcription polymerase chain reaction (RT-PCR) [7, 8] becomes the gold standard for COVID-19 screening, it is time-consuming and suffers from high false negative rates [5, 6]. Computed tomography (CT) technique [9] is widely preferred due to its non-invasive imaging and three-dimensional view of the lung, which is regarded as a significant complement to RT-PCR tests.

* Supported by the National Key Research and Development Program of China (Grant No. 2018YFB0204301).

Recently, deep learning-based applications on COVID-19 CT images have demonstrated quite promising results in lesion segmentation [16–19, 15, 13] and COVID-19 diagnosis [10, 11], particularly, COVID-19 segmentation is an essential step for COVID-19 follow-up assessment. Several COVID-19 segmentation works [16–19, 15, 13] in CT scans have appeared, among which the Inf-Net model [13] achieves the state-of-the-art performance. Different from other segmentation approaches, Inf-Net utilizes the edge attention to model infection boundaries for better feature representations, effectively illustrating edges benefit segmentation performance. However, Inf-Net mainly exploits low-level features to represent edges and its edge extraction is a part of segmentation, thus, edge detection is insufficient. Moreover, there is no independent COVID-19 edge detection task currently. Based on above inspirations, we propose a new task of independent COVID-19 edge detection.

Up to now, to improve edge detection performance, some edge detection works [14, 12] using deep learning methods have existed in many applications, where [12] presents a dynamic feature fusion (DFF) model to produce weighted fusion features, acquiring excellent performance in semantic edge detection. However, DFF model is unable to enjoy same superiority in COVID-19 edge detection due to COVID-19 CT scans's characteristics: (1) ambiguous boundaries caused by the low contrast between infected regions and normal tissues, (2) noises (e.g., blood vessels in lungs), (3) the high variation in shapes, sizes and positions of infected edges. Furthermore, DFF model ignores capturing richer contextual relationships from original images in the initial feature extraction stage and merely considers deep learning features as edge features, which leads to limited COVID-19 edge representations.

To alleviate above problems, we come up with the first COVID-19 lung lesion edge detection model named COVID Edge-Net. Our COVID Edge-Net consists of the edge detection backbone, an effective multi-scale residual dual attention (MSRDA) module and the Canny operator module. The edge detection backbone is capable of detecting basic but coarse COVID-19 edges by extracting discriminative deep learning features. The MSRDA module added in the backbone mainly focuses on semantic features (i.e., edge shapes, sizes and positions), helping to capture richer contextual relationships from CT slices and boost edge identification with better deep feature representations. The Canny edge detection operator can highlight and locate thinner, more continuous and more refined boundaries, which is combined with the deep learning method to further enrich edge information. In a nutshell, our contributions in this paper are fourfold:

- It is the first time to propose a COVID-19 infection lesion edge detection model called COVID Edge-Net, which has the following clinical implications: (1) Due blurred edges of lesions make it difficult for doctors to identify boundaries accurately, this work can assist doctors to locate boundaries of lesions more intuitively and provide appropriate clinical guidance. (2) It is also able to raise COVID-19 infected segmentation performance in CT scans.

- An effective multi-scale residual dual attention (MSRDA) module is designed to help capture richer semantic relationships from multiple scales for detecting more accurate and clearer edges.
- The traditional Canny operator is considered and combined with the deep learning method to enhance feature representations for extracting more refined, more continuous and sharper edges.
- Our proposed method is superior to existing models and advances the state-of-the-art performance, which can be regarded as a benchmark for COVID-19 infection lesion edge detection.

2 Related Works

2.1 COVID-19 segmentation

Up to now, abundant COVID-19 segmentation researches have occurred. For instance, [17] exploited several preprocessing and data augmentation methods to generate random image patches for COVID-19 segmentation, reducing the overfitting risk. Zhou et al. [18] incorporated spatial and channel attention strategies to the U-Net model and introduced the focal tversky loss to solve small lesion segmentation. Aggregated residual transformations and soft attention mechanism were used by Chen et al. [19] to improve the ability of distinguishing various COVID-19 symptoms. Although these approaches overcome some problems caused by limited data and diverse lesion shapes in COVID-19 segmentation tasks, they ignore another important information, i.e., lesion edge information, which has the potential to improve segmentation performance. Later, Fan et al. [13] put forward the Inf-Net model, which achieves the state-of-the-art result in COVID-19 segmentation. It exploits a parallel partial decoder to aggregate high-level features and uses the reverse attention to enhance feature representations. Furthermore, the edge attention is utilized to model COVID-19 boundaries, providing more plentiful feature descriptions for the segmentation. Different from other works, Fan et al. considered COVID-19 edge information, which benefits COVID-19 segmentation. However, Inf-Net's edge extraction is merely used as a submodule of the segmentation task, and mainly low-level features are used as edge features, all of which lead to insufficient edge detection. Currently, there does not exist the sole edge detection task of COVID-19 infected regions. To this end, we propose an independent COVID-19 lesion edge detection task rather than as a submodule for the first time.

2.2 Edge detection

Several semantic edge detection works achieve fantastic results in other applications, such as natural scenes. Yu et al. proposed the CASENet model [14], which is a novel end-to-end deep semantic edge detection architecture via ResNet. In CASENet model, category-wise edge activations from the top layer are fused with the same set of bottom layer features. Its multi-scale feature fusion method

greatly benefits the semantic edge detection task, however, it adopts a fixed weight fusion strategy that forces images with different semantics to share the same weights. To better consider the heterogeneity in contributions made by different locations of feature maps, a new dynamic feature fusion (DFF) strategy was proposed by Hu et al. [12]. They designed a weight learner to assign different fusion weights for different feature maps and locations adaptively. Although Hu et al.'s DFF method reaches superior performance in edge detection of natural scenes, it is not completely suitable for COVID-19 edge detection due COVID-19 CT scans have their own traits (e.g., unclear infected edges, diverse lesion shapes of the same disease, noises) compared with natural images. Meanwhile, DFF's multi-layer features extracted from ResNet are directly applied to the feature fusion layer, ignoring capturing richer contextual relationships from original images. Merely deep learning features are considered in DFF, which are still unable to depict edge information fully. To solve above issues, we propose two new modules, where the multi-scale residual dual attention (MSRDA) module is designed to capture richer contextual features from original images and the Canny operator module is used to extract more edges and emphasize edge information.

3 Methodology

3.1 Task definition

COVID-19 infection lesion edge detection task aims to outline specific contours of infected areas. Particularly, given an one-channel grayscale input image X, the task outputs an edge map $\hat{Y}$ with single channel, which has the same size as X. Each value in $\hat{Y}$ is denoted as $\hat{Y}(p|X, W) \in [0, 1]$, indicating the computed COVID-19 edge probability at pixel p, where W stands for edge detection network's parameters, and $p \in \{1, 2 \cdots, |X|\}$.

3.2 Overview of COVID Edge-Net

To extract more accurate, more continuous and sharper edges, we address COVID Edge-Net model for COVID-19 edge detection (shown in Fig. 1). The input image is fed to residual blocks to generate multi-scale basic features. Then an effective multi-scale residual dual attention (MSRDA) module is used to generate semantically enhanced features, which are processed by feature normalization blocks and a concatenation operation. Later, the adaptive weight fusion block produces weighted deep learning features, which are combined with traditional Canny features from the Canny operator module to further highlight edge features. In the end, one 1×1 convolution operation and the sigmoid function follow to obtain predicted edges. The backbone consisting of residual blocks, feature normalization blocks and the adaptive weight fusion block is described in Sec. 3.3. The MSRDA module and the Canny operator module are specifically introduced in Sec. 3.4 and Sec. 3.5, respectively. And global loss function is displayed in Sec. 3.6.

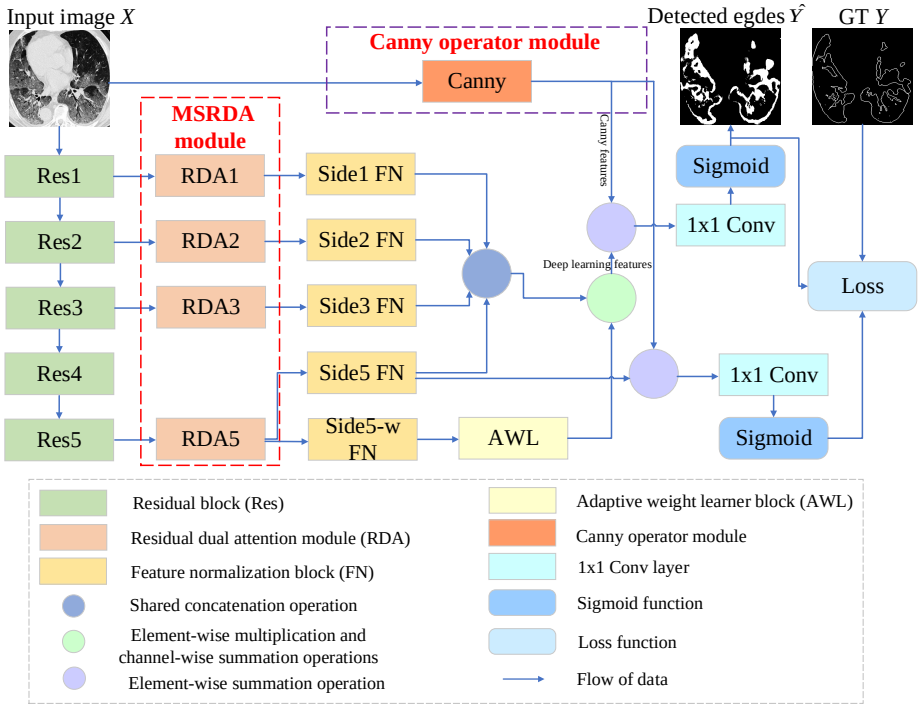

Fig. 1. Overall architecture of our COVID Edge-Net model.

3.3 The edge detection backbone

Our edge detection backbone adopts a dynamic feature fusion (DFF) model [12] via ResNet [21]. In detail, the input image X is transmitted into residual blocks to generate a set of features with different scales. The first three and the fifth stack of residual blocks are directly followed by feature normalization blocks, producing one-channel and K-channel response maps with original image size for Side1-3 and Side5 respectively. K is the number of categories of objects, and $K = 1$ in our task. These response maps are concatenated into a $4K$-channel feature map F_{cat} by shared concatenation, which replicates three one-channel response maps of Side1-3 for K times to separately concatenate with each map of the K-channel response maps in Side5. Another feature normalization block (Side5-w) is connected to the fifth stack of the residual block to generate a $4K$-channel feature map, which goes into the adaptive weight learner to predict dynamic fusion weights W_l. The final output $\hat{Y}_D$ of DFF is computed as $\hat{Y}_D = \sigma(f(F_{cat}, W_l))$, where f represents element-wise multiplication and channel-wise summation operations, σ is the sigmoid function.

Loss function. The backbone's loss function L is disassembled to two losses:

$$L = w_1 L_{fuse} + w_2 L_{side},$$

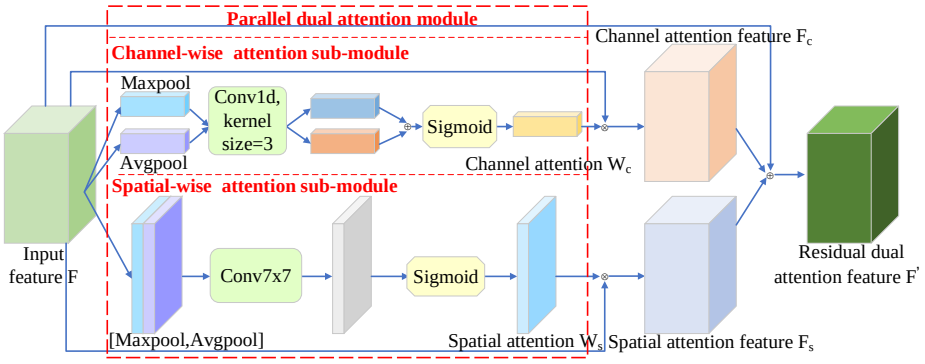

Fig. 2. Residual dual attention module.

$$
\begin{aligned}
L_{fuse} = \sum\nolimits_p \{ & -\alpha Y(p) log \hat{Y}_D(p|X, W_D) - (1-\alpha) \\
& (1 - Y(p)) log(1 - \hat{Y}_D(p|X, W_D))\}, \\
L_{side} = \sum\nolimits_p \{ & -\alpha Y(p) log \hat{Y}_{D\_s}(p|X, W_D) - (1-\alpha) \\
& (1 - Y(p)) log(1 - \hat{Y}_{D\_s}(p|X, W_D))\},
\end{aligned} \tag{1}
$$

where L_{fuse}, L_{side} denote its loss function for final output and Side5 output, respectively. w_1, w_2 are corresponding weighting factors to balance two losses. L_{fuse} and L_{side} adopt class-balanced cross-entropy loss function for balancing the loss between positive and negative classes. α is a class-balancing weight, $\alpha = |Y_{edge}|/|Y|$, $|Y_{edge}|$ expresses the size of the edge ground truth (GT) label set. $Y(p) \in \{0, 1\}$, $\hat{Y}_D(p|X, W_D) \in [0, 1]$ and $\hat{Y}_{D\_s}(p|X, W_D) \in [0, 1]$ are the GT, the predicted final edge probability and the edge probability from Side5 output at pixel $p \in \{1, 2 \cdots, |X|\}$, respectively. W_D stands for the backbone's parameters.

3.4 Multi-scale residual dual attention (MSRDA) module

To strengthen semantic information in captured features, we design a novel parallel dual attention module as shown in Fig. 2, including channel-wise and spatial-wise attention sub-modules.

In the channel attention sub-module, the input feature F is processed by max pooling and average pooling operations to aggregate each feature map's spatial information, respectively. Then a fast and shared *1D* convolution with kernel size 3 rather than MLP [29] follows to capture non-linear dependencies across all channels, due that the dimensionality reduction in MLP may have a negative impact on the final accuracy performance [30]. In the end, the channel-wise attention feature F_c is expressed as:

$$
F_c = F \otimes \sigma(1Dconv(AvgPool(F)) \oplus 1Dconv(MaxPool(F))), \tag{2}
$$

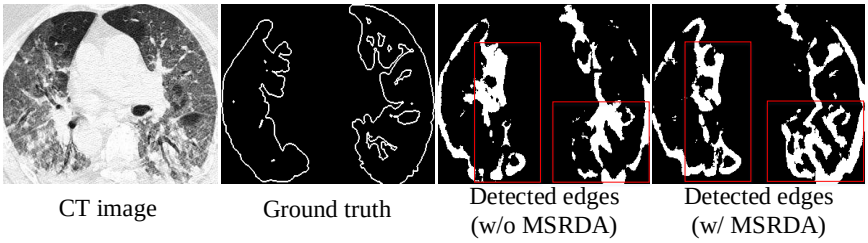

Fig. 3. Visualization of the edge detection with (w/) and without (w/o) the MSRDA module.

where $\oplus$ denotes the element-wise summation; σ is the sigmoid function; $\otimes$ stands for the element-wise multiplication. In the spatial attention sub-module, max pooling and average pooling operations are performed along the channel axis, respectively. Two produced feature maps are concatenated and forwarded to a standard convolution with kernel size 7, generating the spatial attention map. The spatial-wise attention feature F_s is computed as:

$$F_s = F \otimes \sigma(Conv7 \times 7([AvgPool(F); MaxPool(F)])). \tag{3}$$

To make full use of more features and enrich context information, we propose the residual structure to combine original features with attention features. The residual dual attention feature $F^{'}$ is represented as:

$$F^{'} = F_c \oplus F_s \oplus F, \tag{4}$$

where $\oplus$ is the element-wise summation operation. Then the residual dual attention mechanism is used for multiple basic descriptors generated by residual blocks to enhance feature representations, forming the multi-scale residual dual attention (MSRDA) module as shown in Fig. 1. Visualized results in Fig. 3 clarify that the MSRDA module helps detect more accurate and complex edges, which is beneficial to COVID-19 edge detection.

3.5 Canny operator module

Canny operator [28] is considered as one of the most classical algorithms for image edge detection. It is simple and its specific steps are as follows:

Step1: Use Gaussian filter to smooth the image.

Step2: Calculate the amplitude and direction of the image gradient after filtering.

Step3: Perform the non-maximum suppression for gradient amplitude to obtain thinner edges.

Step4: Select two thresholds T1 and T2 and connect edges. A pixel can be regarded as a strong edge point, weak edge point, or non-edge point when its

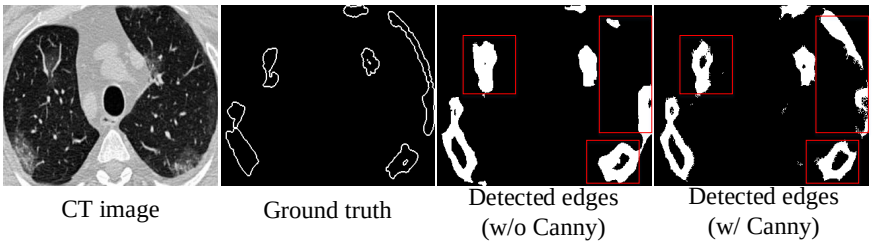

Fig. 4. Visualization of the edge detection with (w/) and without (w/o) the Canny module.

gradient is beyond T1, between T1 and T2, or lower than T2. When strong edge points appear in the 8 neighborhoods around the weak edge point, the weak edge point is changed into a strong edge point to augment the strong edge set.

The Canny feature F_{Canny} of the input image X is obtained through Canny operator, and then combined with weighted deep learning feature F_w. The final fusion feature F_{f_fuse} is represented as:

$$F_{f\_fuse} = Conv1 \times 1(F_{Canny} \oplus F_w), \tag{5}$$

where $\oplus$ is the element-wise summation operation. The visualization in Fig. 4 demonstrates that our Canny module helps detect thinner, clearer and sharper edges, contributing to COVID-19 infection edge detection.

3.6 Global loss function

Based on the final fusion features F_{f_fuse}, our final prediction $\hat{Y}$ is computed as:

$$\hat{Y} = \sigma(F_{f\_fuse}), \tag{6}$$

where σ is the sigmoid function. Our COVID Edge-Net adopts the same loss function as the backbone (Sec. 3.3), considering two losses. However, $\hat{Y}_D(p|X, W_D) \in [0, 1]$ in backbone's L_{fuse} is replaced by $\hat{Y}(p|X, W) \in [0, 1]$, which is our final predicted edge probability at pixel p. $\hat{Y}_{D\_s}(p|X, W_D) \in [0, 1]$ in backbone's L_{side} is changed into $\hat{Y}_s(p|X, W) \in [0, 1]$ that is our Side5 output's edge probability in Fig. 1, where Canny features are incorporated.

4 Experiments and discussions

4.1 Experimental settings

The experimental dataset and augmentations. Our experiments rely on two COVID-19 CT datasets [20]: COVID-19 CT segmentation dataset and Segmentation dataset nr.2 (13th April), which are publicly available. COVID-19

CT segmentation dataset [20] contains 100 CT images from different COVID-19 patients and is collected by the Italian Society of Medical and Interventional Radiology. We randomly select 50 CT images as training samples and the remaining 50 images for testing, whose GT edges are generated via their segmentation labels. The larger Segmentation dataset nr.2 (13th April) released later consists of 829 slices (373 infected slices and 456 non-infected slices) extracted from 9 CT volumes of real COVID-19 patients, where 373 infected CT images are used as our experimental data for COVID-19 lung lesion edge detection. Among them, 186 randomly selected slices are regarded as the training set and the rest as the test set, whose GT edges are also produced based on segmentation marks given by radiologists. These experimental datasets suffer from a small sample size, thus, we augment training samples in each dataset by resizing each CT image with scaling factors {0.5, 0.75, 1, 1.25, 1.5} referring to [12] and employing random mirroring and cropping during the training process. In the test phase, each CT slice in COVID-19 CT segmentation dataset is resized to 512 $\times$ 512, and Segmentation dataset nr.2 (13th April)'s each CT image is cropped into 576 $\times$ 576.

Training settings. Our model is based on ResNet18 [21] pre-trained on ImageNet [22]. K in Sec. 3.3 is set as one, and weighting factors in loss function are set as $w_1 = w_2 = 1$. Two thresholds T1 and T2 we used in the Canny operator are set to 200 and 100, respectively. We use SGD optimization, the learning rate is initialized to 0.05 and the "poly" policy is used for its decay. The crop size, batch size, training epoch, momentum, weight decay, and seed are set as 352×352, 16, 50, 0.9, 1e-4, 1, respectively. Our experiments depend on PyTorch and one NVIDIA 2080 Ti GPU. For fair comparisons, all edge detection experiments use the same training settings.

Evaluation Metrics. Following [12], Maximum F-measure (MF) at optimal dataset scale (ODS) with different matching distance tolerances is used as a common metric for edge detection performance, where F-measure is the harmonic average of precision and recall as shown in Eq. (7) and ODS means each image in the dataset uses the same threshold to evaluate image edges and achieve the entire dataset's maximum F-measure.

$$F - measure = \frac{2 \cdot precision \cdot recall}{precision + recall}. \tag{7}$$

We also consider two cases of MF (ODS) metric under one matching distance tolerance: with morphological thinning (w/ MT) and without morphological thinning (w/o MT).

4.2 Comparison with state-of-the-arts

We compare the performance of our COVID Edge-Net model with state-of-the-art edge detection methods [28, 14, 12] on COVID-19 CT segmentation dataset and Segmentation dataset nr.2 (13th April), and evaluate MF(ODS) with different matching distance tolerances that are set as 0.02, 0.06, 0.10, respectively.

Table 1. Performance comparison of edge detection on COVID-19 CT segmentation dataset

Edge detection methods	MF(0.02)		MF(0.06)		MF(0.10)	
	w/ MT	w/o MT	w/ MT	w/o MT	w/ MT	w/o MT
Canny[28]	32.20%	45.15%	32.55%	46.77%	32.55%	46.77%
CASENet[14]	60.38%	42.87%	84.37%	56.71%	90.73%	66.78%
DFF[12]	**84.14**%	54.53%	90.83%	72.02%	93.45%	72.89%
Ours	83.58%	**63.80**%	**94.70**%	**85.30**%	**95.94**%	**91.57**%

Table 2. Performance comparison of edge detection on Segmentation dataset nr.2 (13th April)

Edge detection methods	MF(0.02)		MF(0.06)		MF(0.10)	
	w/ MT	w/o MT	w/ MT	w/o MT	w/ MT	w/o MT
Canny[28]	3.46%	5.76%	3.46%	5.76%	3.46%	5.76%
CASENet[14]	79.67%	74.74%	89.92%	**98.96**%	89.92%	**98.96**%
DFF[12]	75.38%	64.31%	90.77%	98.62%	90.77%	98.62%
Ours	**92.54**%	**78.29**%	**94.03**%	96.05%	**94.03**%	96.05%

On the COVID-19 CT segmentation dataset, Table 1 shows that ours exceeds the DFF method in almost all cases, and is completely superior to Canny and CASENet methods. When the matching distance tolerance is 0.10, ours achieves state-of-the-art results that are 95.94% MF(ODS) with MT and 91.57% MF(ODS) without MT. Specifically, our method surpasses DFF with 2.49% and 18.68% under 0.10 matching distance tolerance, respectively. As the matching distance tolerance is stricter, our approach is 3.87% and 13.28% better than the DFF under 0.06 matching distance tolerance, respectively. Under the strictest matching distance tolerance (0.02), our approach is 9.27% higher than DFF when ignoring MT, 23.20% and 20.93% better than CASENet with and without MT, 51.38% and 18.65% higher than Canny with and without MT. We also present performance comparison results of COVID-19 lung lesion edge detection on the larger Segmentation dataset nr.2 (13th April) in Table 2. It is evident that our proposed approach is far better than Canny method in all cases. Meanwhile, when the matching distance tolerance is the strictest 0.02, ours is obviously superior to other competing approaches. Specifically, ours achieves promising results that are 92.54% MF(ODS) with MT and 78.29% MF(ODS) without MT, which are 17.16% and 13.98% better than DFF considering MT and ignoring MT, 12.87% and 3.55% higher than CASENet with and without MT, respectively. Under 0.06 and 0.10 matching distance tolerances, our proposed approach has 94.03% MF(ODS) with MT, which is beyond DFF 3.26% and 4.11% higher than CASENet method; however, DFF and CASENet have slightly better MF(ODS) leaving the MT out of consideration. In the case of higher matching distance tolerance, COVID-19 lung lesion edges produced by CASENet and DFF are more accurate without MT; our proposed method is able to get more promising results

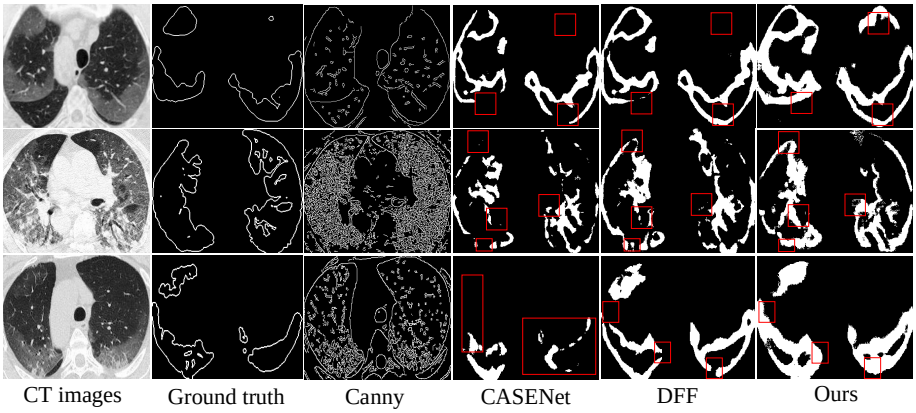

Fig. 5. Qualitative comparison of edge detection on COVID-19 CT segmentation dataset.

after MT. In general, our proposed method outperforms other state-of-the-arts in most cases. In addition, qualitative comparisons are also provided by visualizing some edge detection results in Fig. 5 and Fig. 6. On two datasets, we observe that Canny detects excessive useless edges because of the low contrast between infected regions and non-infected regions. Compared with other deep learning methods: CASENet and DFF, results in red boxes demonstrate that our COVID Edge-Net method has the capability of predicting more accurate, more continuous, sharper and clearer object edges, having higher edge detection performance by taking full advantage of multi-scale semantic information and fused features.

4.3 Ablation study

In this subsection, we conduct ablation experiments to validate the performance of each newly proposed module in our COVID Edge-Net, taking COVID-19 CT segmentation dataset as an example, results of which are displayed in Table 3. DFF via ResNet18 is our baseline model. The comparison of Baseline (DFF) and Baseline+Canny clearly shows that Canny operator boosts performance without MT. Thus, the Canny operator module has the ability of refining and enhancing edges in COVID-19 CT images. Baseline (DFF) and Baseline+MSRDA demonstrate that the MSRDA module increases the baseline performance with MT. MSRDA strategy enables a model to identify much more edges that cannot be detected by the baseline. With these two complementary modules, the performance of our model is significantly enhanced no matter whether MT is considered or not considered. Furthermore, our MSRDA module is compared with multi-scale residual convolutional block attention module (MSRCBAM), where convolutional block attention module (CBAM) is proposed in [29]. It can be shown that Baseline+MSRDA is superior to Base-

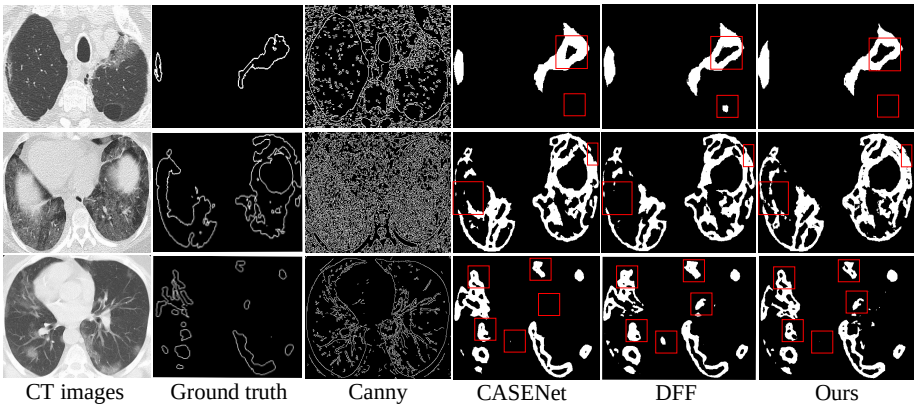

Fig. 6. Qualitative comparison of edge detection on Segmentation dataset nr.2 (13th April).

Table 3. Ablation study of our COVID Edge-Net on COVID-19 CT segmentation dataset

Edge detection methods	MF(0.02)		MF(0.06)		MF(0.10)	
	w/ MT	w/o MT	w/ MT	w/o MT	w/ MT	w/o MT
Baseline (DFF)	84.14%	54.53%	90.83%	72.02%	93.45%	72.89%
Baseline+Canny	73.41%	55.67%	90.36%	74.31%	93.64%	83.11%
Baseline+MSRCBAM	72.67%	44.59%	87.45%	63.59%	92.89%	78.12%
Baseline+MSRDA	**84.85**%	44.58%	94.70%	60.61%	95.71%	66.49%
Baseline+MSRCBAM+Canny	79.57%	60.57%	93.39%	80.16%	93.52%	85.56%
Ours	83.58%	**63.80**%	**94.70**%	**85.30**%	**95.94**%	**91.57**%

line+MSRCBAM with MT, and the whole method we proposed completely outperforms Baseline+MSRCBAM+Canny. Above all demonstrate the effectiveness of the integration of the two modules (i.e., MSRDA and Canny modules) into the Baseline (DFF).

4.4 Additional experiments

We conduct COVID-19 segmentation experiments to further verify the performance of our algorithm on two COVID-19 segmentation datasets. The one-channel edge feature map from each edge detection model (e.g., our model, Canny, DFF) is replicated into a 64-channel map and then replaces Inf-Net's 64-channel edge feature [13]. We introduce GT edges via same operation as well. To make fair comparisons, we use the same training parameters and evaluated metrics as Inf-Net, and comparisons are shown in Table 4 and Table 5. The segmentation effect with the best edge features (GT edges) is far beyond original Inf-Net in Table 4 and Table 5, which reflects that better edges benefit segmentation. Furthermore, on the COVID-19 CT segmentation dataset, we observe

Table 4. Performance comparison of segmentation on COVID-19 CT segmentation dataset

Segmentation methods	Dice	Sen.	Spec.	S_α	$E_\emptyset^{mean}$	MAE
U-Net [23]*	0.439	0.534	0.858	0.622	0.625	0.186
Dense-UNet [26]*	0.515	0.594	0.840	0.655	0.662	0.184
Attention-UNet [25]*	0.583	0.637	0.921	0.744	0.739	0.112
U-Net++ [24]*	0.581	0.672	0.902	0.722	0.720	0.120
Gated-UNet [27]*	0.623	0.658	0.926	0.725	0.814	0.102
Inf-Net [13]*	0.682	0.692	0.943	0.781	0.838	0.082
Inf-Net(with Canny's edges)	**0.722**	**0.823**	0.920	0.773	0.865	0.083
Inf-Net(with DFF's edges)	0.713	0.732	0.946	**0.798**	0.870	0.075
Inf-Net(with our edges)	0.718	0.736	**0.948**	**0.798**	**0.872**	**0.074**
Inf-Net(with GT edges)	**0.780**	**0.821**	**0.952**	**0.861**	**0.888**	**0.059**

Note: Dice: Dice similarity coefficient, Sen.: Sensitivity, Spec.: Specificity, S_α: Structure Measure, $E_\emptyset^{mean}$: Enhance-alignment Measure, MAE: Mean Absolute Error.
*: All experiment data here refer to [13].

Table 5. Performance comparison of segmentation on Segmentation dataset nr.2 (13th April)

Segmentation methods	Dice	Sen.	Spec.	S_α	$E_\emptyset^{mean}$	MAE
Inf-Net [13]	0.802	0.831	0.961	0.861	0.938	0.020
Inf-Net(with Canny's edges)	0.808	**0.855**	0.964	0.861	**0.945**	0.019
Inf-Net(with DFF's edges)	0.812	0.849	0.963	0.867	0.942	**0.018**
Inf-Net(with our edges)	**0.814**	0.850	**0.967**	**0.868**	0.943	**0.018**
Inf-Net(with GT edges)	**0.894**	**0.904**	**0.988**	**0.930**	**0.965**	**0.012**

Note: Dice: Dice similarity coefficient, Sen.: Sensitivity, Spec.: Specificity, S_α: Structure Measure, $E_\emptyset^{mean}$: Enhance-alignment Measure, MAE: Mean Absolute Error.

that Inf-Net(with our edges) outperforms other excellent segmentation models [13] (e.g., U-Net-based models [23–27] and Inf-Net) for all metrics and exceeds Canny and DFF methods in segmentation for most metrics from Table 4. Similarly, Table 5 describes that Inf-Net(with our edges) absolutely defeats Inf-Net and outperforms Canny and DFF approaches in segmentation under most evaluation metrics on the Segmentation dataset nr.2 (13th April). Obviously, COVID Edge-Net is able to extract more accurate and richer edges than Inf-Net's edge extraction submodule, further enhancing segmentation performance.

5 Conclusions

In this paper, we present the first COVID Edge-Net for automatic COVID-19 lung lesion edge detection. On the one hand, an effective MSRDA module is designed and combined with edge detection backbone to extract more distinguishable deep learning features by capturing richer contextual relationships from CT

scans. On the other hand, our network leverages Canny features to further enrich edge information by multi-feature fusion. Our proposed method achieves state-of-the-art COVID-19 edge detection performance compared to other competing approaches, and significantly benefits segmentation performance. It has the potential to be developed as a clinical tool for COVID-19 CT images analysis. Code related to this paper is available at: https://github.com/Elephant-123/COVID-Edge-Net.

References

1. Zhu, N., Zhang, D., Wang, W., et al.: A novel coronavirus from patients with pneumonia in china, 2019. New England Journal of Medicine (2020)
2. Wang, C., Horby, P.W., Hayden, F.G., Gao, G.F.: A novel coronavirus outbreak of global health concern. The lancet **395**(10223), 470–473 (2020)
3. Oudkerk, M., Büller, H.R., Kuijpers, D., et al.: Diagnosis, prevention, and treatment of thromboembolic complications in covid-19: report of the national institute for public health of the netherlands. Radiology **297**(1), E216–E222 (2020)
4. Coronavirus covid-19 global cases by the center for systems science and engineering at johns hopkins university, https://coronavirus.jhu.edu/map.html. Accessed 24 November 2020
5. Liang, T., et al.: Handbook of covid-19 prevention and treatment. The First Affiliated Hospital, Zhejiang University School of Medicine. Compiled According to Clinical Experience **68** (2020)
6. Shan, F., Gao, Y., Wang, J., et al.: Lung infection quantification of covid-19 in ct images with deep learning. arXiv preprint arXiv:2003.04655 (2020)
7. Fang, Y., Zhang, H., Xie, J., et al.: Sensitivity of chest ct for covid-19: comparison to rt-pcr. Radiology p. 200432 (2020)
8. Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., Xia, L.: Correlation of chest ct and rt-pcr testing for coronavirus disease 2019 (covid-19) in china: a report of 1014 cases. Radiology **296**(2), E32–E40 (2020)
9. Ng, M.Y., Lee, E.Y., Yang, J., et al.: Imaging profile of the covid-19 infection: radiologic findings and literature review. Radiology: Cardiothoracic Imaging **2**(1), e200034 (2020)
10. Kang, H., Xia, L., Yan, F., et al.: Diagnosis of coronavirus disease 2019 (covid-19) with structured latent multi-view representation learning. IEEE transactions on medical imaging (2020)
11. Wang, J., Bao, Y., Wen, Y., et al.: Prior-attention residual learning for more discriminative covid-19 screening in ct images. IEEE Transactions on Medical Imaging (2020)
12. Hu, Y., Chen, Y., Li, X., Feng, J.: Dynamic feature fusion for semantic edge detection. arXiv preprint arXiv:1902.09104 (2019)
13. Fan, D.P., Zhou, T., Ji, G.P., et al.: Inf-net: Automatic covid-19 lung infection segmentation from ct images. IEEE Transactions on Medical Imaging (2020)
14. Yu, Z., Feng, C., Liu, M.Y., Ramalingam, S.: Casenet: Deep category-aware semantic edge detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5964–5973 (2017). https://doi.org/10.1109/CVPR.2017.191
15. Qiu, Y., Liu, Y., Xu, J.: Miniseg: An extremely minimum network for efficient covid-19 segmentation. arXiv preprint arXiv:2004.09750 (2020)

16. Wang, Y., Zhang, Y., Liu, Y., et al.: Does non-covid19 lung lesion help? investigating transferability in covid-19 ct image segmentation. arXiv preprint arXiv:2006.13877 (2020)
17. Müller, D., Rey, I.S., Kramer, F.: Automated chest ct image segmentation of covid-19 lung infection based on 3d u-net. arXiv preprint arXiv:2007.04774 (2020)
18. Zhou, T., Canu, S., Ruan, S.: An automatic covid-19 ct segmentation network using spatial and channel attention mechanism. arXiv preprint arXiv:2004.06673 (2020)
19. Chen, X., Yao, L., Zhang, Y.: Residual attention u-net for automated multi-class segmentation of covid-19 chest ct images. arXiv preprint arXiv:2004.05645 (2020)
20. Covid-19 ct segmentation dataset, https://medicalsegmentation. com/covid19/
21. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90
22. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009). https://doi.org/10.1109/CVPR.2009.5206848
23. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015). https://doi.org/10.1007/978-3-662-54345-0_3
24. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 3–11. Springer (2018). https://doi.org/10.1007/978-3-030-00889-5_1
25. Oktay, O., Schlemper, J., Folgoc, L.L., et al.: Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 (2018)
26. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A.: H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes. IEEE transactions on medical imaging **37**(12), 2663–2674 (2018)
27. Schlemper, J., Oktay, O., Schaap, M., et al.: Attention gated networks: Learning to leverage salient regions in medical images. Medical image analysis **53**, 197–207 (2019)
28. Canny, J.: A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence (6), 679–698 (1986)
29. Woo, S., Park, J., Lee, J.Y., So Kweon, I.: Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV). pp. 3–19 (2018). https://doi.org/10.1007/978-3-030-01234-2_1
30. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q.: Eca-net: Efficient channel attention for deep convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11534–11542 (2020).