

Automatic Diagnosis of COVID-19 Pulmonary Infection Based on Multi-Scale Feature Enhancement



Qing Liu¹, Tao Shen^{1*}, Ze Liu¹, Jinxin Liu¹, Mingwei Zhang¹, Xu Han¹, Fengyun Cao²

¹ Department of Emergency Medicine, The Second Affiliated Hospital of BengBu Medical University, Bengbu 233000, China

² Medical Artificial Intelligence Technology Research and Development Center, Zhongke Hefei Institute of Technology Innovation Engineering, Hefei 230088, China

Corresponding Author Email: sticubyefy@163.com

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420625>

ABSTRACT

Received: 10 April 2025

Revised: 19 September 2025

Accepted: 8 October 2025

Available online: 31 December 2025

Keywords:

COVID-19, pulmonary infection segmentation, multi-scale feature enhancement, dilated convolution, deep learning, automatic diagnosis

The global pandemic of COVID-19 has created an urgent demand for fast and accurate diagnosis of pulmonary infections through CT imaging. Achieving automatic segmentation and quantification of infected areas in the lungs is crucial for disease assessment and treatment guidance. However, the COVID-19 infection areas exhibit complex characteristics in CT images, such as multi-scale features, blurred boundaries, and adhesion to normal tissues, which present significant challenges for automatic segmentation techniques. To address the poor performance of existing fully convolutional networks in segmenting small-sized infection regions and blurry boundaries, this paper proposes an automatic diagnostic model for COVID-19 pulmonary infection based on multi-scale feature enhancement. Firstly, an encoder-decoder framework with an improved VGG16-BN as the backbone is constructed. The core of the model is the design of a multi-scale feature enhancement module, which integrates features from different layers and dynamically adjusts the foreground and background weights through learnable parameters, effectively reducing interference from complex backgrounds. The context information enhancement component within the module employs a multi-branch dilated convolution strategy to enlarge the receptive field while preserving the feature map resolution, thus significantly improving the model's ability to capture minute lesions and enhance boundary segmentation accuracy. Finally, based on the high-precision segmentation results, an automatic diagnostic strategy is developed, which quantifies key indicators such as the volume and distribution of the infected area to assist in disease evaluation. The innovations of this study mainly include: 1) the proposal of a segmentation network architecture that integrates multi-scale feature enhancement and context awareness, effectively addressing the core challenges of varying infection region scales and blurred boundaries; 2) the introduction of a dynamic weight adjustment mechanism in the feature enhancement module, allowing the model to adaptively focus on infection region features, thus improving its discriminative capability; and 3) the integration of pixel-level segmentation results with clinical diagnostic requirements, forming a complete automatic diagnostic solution from image analysis to quantitative assessment, with significant theoretical value and clinical application potential.

1. INTRODUCTION

The global pandemic of COVID-19 [1-3] has become a major public health crisis, placing unprecedented pressure on global healthcare systems [4, 5]. Chest computed tomography (CT) imaging [6-9] is widely used for COVID-19 screening, diagnosis, and disease assessment due to its high sensitivity. The typical manifestations of COVID-19 in CT images include multiple [10] ground-glass opacities and consolidation [11, 12], which are distributed peripherally. The precise identification and segmentation of these infection regions are key to quantitatively assessing disease severity and monitoring disease progression. However, faced with the surge in cases, relying on radiologists to manually outline infection regions [13-15] is not only time-consuming and labor-intensive but

also prone to subjectivity and low diagnostic consistency, highlighting the urgent need for efficient and accurate automated diagnostic solutions.

Developing an AI-based automatic diagnostic system for COVID-19 pulmonary infections [16, 17] holds significant practical and clinical value. First, it can greatly enhance diagnostic efficiency, enabling rapid localization and quantification of lesions, providing objective decision support for doctors, and alleviating high-intensity workload pressure. Second, precise volume calculation and density analysis of infected areas can facilitate accurate disease grading and dynamic follow-up, offering data support for personalized treatment plans. Ultimately, the application of such systems will promote the development of intelligent and standardized medical image analysis [18], which will not only play a role in

the current pandemic but also accumulate valuable technological experience for handling similar public health events in the future.

Although many studies have attempted to apply deep learning models, particularly fully convolutional networks (FCN) and U-Net architectures, to the segmentation of COVID-19 pulmonary infections, these methods still have obvious limitations in practice. For example, the U-Net model proposed by Alom et al. [19], despite its success in biomedical image segmentation, struggles with the multi-scale characteristics of COVID-19 infection areas due to its symmetric encoder-decoder structure. It fails to extract small-sized ground-glass opacity features that are scattered and distributed, and the segmentation accuracy is limited for regions with blurred boundaries. Moreover, although the DeepLabv3+ model proposed by Murugappan et al. [20] uses dilated convolutions to enlarge the receptive field, it is easily affected by background interference when dealing with the complex situation of infection areas being strongly adhered to normal lung tissue, leading to false positives or discontinuous regions in the segmentation results, thereby affecting the accuracy of quantitative assessments. In addition, classic segmentation models such as UNet rely on fixed-scale convolution kernels, which cannot capture the features of both small ground-glass opacities and large consolidation areas simultaneously, leading to missed small lesions or inaccurate boundary segmentation of large lesions. Although the dilated spatial pyramid pooling in DeepLabv3+ can expand the receptive field, it handles the gradient transition areas between the infected region and normal lung tissue roughly, easily misclassifying vascular textures and artifacts as infection lesions. Infection pixels usually account for less than 10% in CT images, and the simple application of existing cross-entropy loss or Dice loss tends to cause the model to favor predicting the background class, reducing the detection rate of small infection lesions. Most methods only remain at the segmentation of infection areas and fail to convert the segmentation results into quantifiable indicators and classification diagnostic conclusions with clinical decision-making value, limiting their practicality.

To address these issues, this paper proposes a framework for "COVID-19 Pulmonary Infection Automatic Diagnosis Based on Multi-Scale Feature Enhancement". The core research of this paper is the construction of a novel pulmonary infection region extraction model, which introduces multi-scale feature enhancement and context information enhancement modules, aiming to accurately solve the problems of small target omission and blurry boundary segmentation. Based on this, the study further develops an automatic diagnostic strategy from pixel-level segmentation to clinical assessment, achieving quantitative analysis and severity determination of the infection areas. The value of this research lies in the fact that the proposed method not only significantly improves the precision and robustness of the segmentation task but also forms an end-to-end automatic diagnostic process, providing an effective technical tool for the intelligent and precise clinical auxiliary diagnosis of COVID-19, with significant theoretical innovation value and broad clinical application prospects.

The core innovations of this paper are as follows:

(1) Multi-scale feature enhancement dual-branch collaborative mechanism: A "foreground enhancement-context information enhancement" dual-branch module is designed, which accurately separates infection lesions from

the background through the synergistic effect of a soft attention mechanism and multi-scale dilated convolutions. The foreground enhancement branch adopts a Sigmoid probabilistic weight distribution strategy to dynamically strengthen the response of infection area features. The context information enhancement branch innovatively uses a "four-branch heterogeneous dilated convolution" architecture to build a hierarchical receptive field, capturing both small lesion details and the spatial distribution relationships of large lesions.

(2) Dynamic weight adaptive adjustment strategy: Learnable parameters α and β are introduced in the foreground enhancement module, and a multi-level refinement process is used to gradually optimize the weights. The model automatically adjusts based on infection patterns: for diffuse ground-glass opacities, the α weight is increased to 0.7-0.8 to enhance the foreground features; for clearly defined consolidation areas, the β weight is moderately increased to 0.3-0.4 to suppress background noise.

(3) Medical adaptation optimization of the improved VGG16-BN backbone network: In response to the characteristics of COVID-19 CT images, the channel configuration is adjusted, fully connected layers are discarded to retain spatial structural information, and a combination of five max-pooling layers and ReLU activations is used to balance feature abstraction ability and gradient stability, preventing the loss of small infection lesion features.

2. METHODOLOGY

2.1 Extraction of COVID-19 pulmonary infection regions

This paper proposes a new infection region extraction model, designed to address the unique challenges presented by COVID-19 pulmonary infection images, including small infection region size, blurred boundaries, and adhesion to surrounding tissues. These characteristics cause standard fully convolutional models to be easily interfered with by background noise and make it difficult to accurately capture subtle infection regions during segmentation. Therefore, the model in this paper adopts an improved pre-trained VGG16-BN as the base architecture, with four layers of downsampling operations that gradually increase the number of feature channels to adapt to the size variation of the input images and improve feature extraction efficiency. This design effectively utilizes the powerful generalization capability of VGG16-BN in image recognition tasks, while the downsampling process reduces spatial dimensions, thus reducing computational burden and enhancing focus on key features of the infection region. Additionally, the increase in channels during the downsampling process helps the model capture both local and global features of the infection region at multiple scales, providing a rich feature base for subsequent modules and directly supporting the high-precision requirements of automatic diagnosis in the research objective.

On top of the base architecture, this paper constructs four multi-scale feature enhancement modules. The reason for designing these modules is that infection regions in COVID-19 pulmonary infection images often exhibit multi-scale distribution, such as the coexistence of small ground-glass opacities and larger consolidation areas, with severe adhesion to healthy tissue at the boundaries. This makes traditional single-scale feature extraction methods prone to losing detail. The module integrates the input from the current

downsampling layer, the output from the next layer, and the features from the corresponding downsampling layer to generate higher-quality feature maps, thus providing rich and detailed multi-scale information. This fusion strategy effectively filters out background noise, such as vascular textures or artifacts in the lung parenchyma, while refining the edge definition of the infection region. The multi-scale feature enhancement module internally integrates a context information enhancement module, designed to address the fact that COVID-19 pulmonary infection regions often exhibit small target features, such as small nodules or subtle ground-glass opacities in early infections. These areas are easily

overwhelmed by the background and have blurred boundaries in the images. The module uses dilated convolution strategies at different scales to expand the model's receptive field, capturing broader contextual information without losing resolution. By adjusting the dilation rate, dilated convolutions allow the model to simultaneously focus on local details and global structures. The enhanced contextual awareness improves the model's ability to capture infection regions, ensuring accurate separation of infection regions from complex backgrounds and thus supporting the efficient and automated diagnostic process of the research objective.

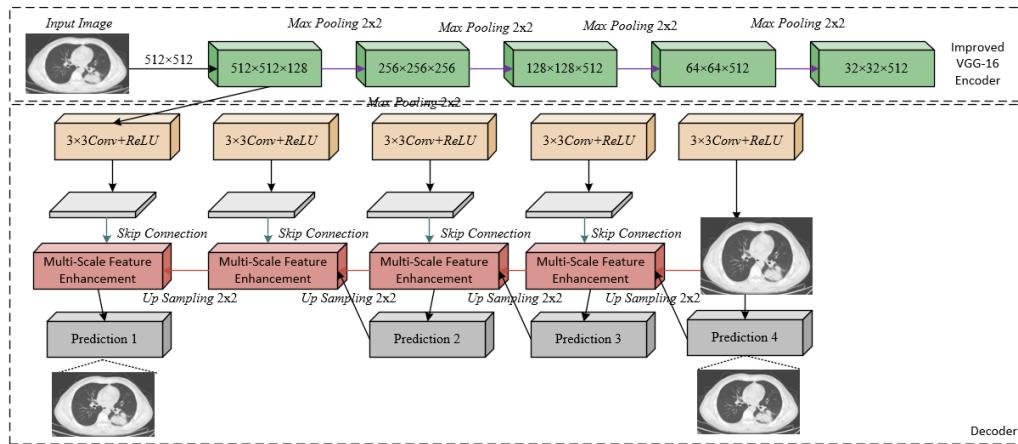


Figure 1. COVID-19 pulmonary infection region extraction model network structure

The model employs bilinear interpolation to upsample the output features from each layer and outputs four different levels of prediction results to address the challenge of maintaining high spatial accuracy in the segmentation of COVID-19 pulmonary infection regions, which need to handle blurred edges and size variation. Bilinear interpolation, as an efficient upsampling method, can smoothly restore feature map sizes, avoiding the aliasing effect caused by nearest-neighbor interpolation and retaining the subtle structure of the infection region during output prediction. The multi-level prediction mechanism corresponds to the outputs from downsampling layers 4 to 1, enabling the model to supervise the segmentation process at different scales. For example, high-level predictions focus on global infection distribution, while low-level predictions focus on local details and boundary refinement. This allows the model to gradually refine the segmentation results of the infection region through a multi-scale loss function during training, reducing issues of small targets or blurred boundaries that might be overlooked by single-scale predictions. Figure 1 illustrates the complete network structure of the COVID-19 pulmonary infection region extraction model.

2.1.1 Backbone network

This paper makes targeted improvements to the VGG-16 backbone network to ensure the model adapts to the fundamental characteristics of COVID-19 pulmonary infection images. Figure 2 shows the improved backbone network architecture. COVID-19 CT images typically contain numerous diffuse, small-sized ground-glass opacities and consolidation regions. These infection features are widely distributed and vary in shape within the image. The improved backbone network fixes the input size to 512×512 pixel RGB images, which balances retaining sufficient detail information

with computational efficiency. By using a 3×3 small-sized convolution kernel with a stride of 1 and padding of 1, the network fully retains boundary information of the infection regions during the convolution process, avoiding the loss of fine infection features caused by size reduction. Additionally, the depth structure of 13 convolutional layers ensures that the network can progressively build a comprehensive understanding of the infection regions, from low-level texture features to high-level semantic features.

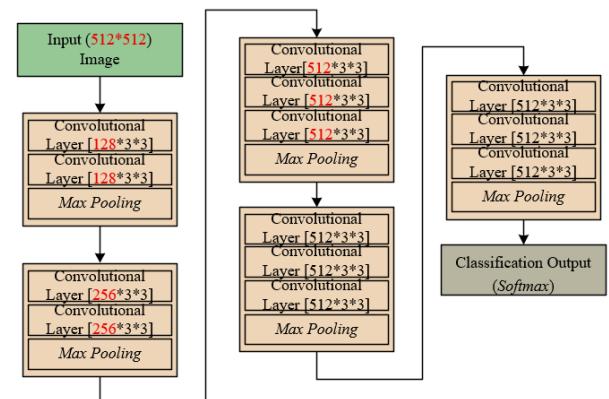


Figure 2. Backbone network architecture

In terms of feature extraction depth and efficiency optimization, this paper makes significant adjustments to the channel configuration of VGG-16. The number of output channels for the first downsampling layer is changed to 128, and three downsampling operations with 512 channels are implemented. The initial layers use relatively fewer channels to prioritize capturing basic visual features, such as edge and texture features, which are crucial for distinguishing infection

regions from normal tissue. As the network depth increases, the number of channels gradually expands to 512, enabling the network to process multiple feature maps simultaneously and more comprehensively describe the complex infection patterns.

The abandonment of the fully connected portion of VGG-16 is another key design decision, which directly serves the unique needs of the image segmentation task. Traditional VGG-16 contains three fully connected layers, which disrupt the spatial structure information of the feature maps and introduce a large number of parameters into the model, increasing computational complexity and potentially causing gradient vanishing issues. For the dense prediction task of COVID-19 infection region segmentation, preserving the spatial information of feature maps is crucial, as pixel-level classification results are required. The improved backbone network exclusively uses convolution and pooling operations, gradually reducing the feature map size with five max pooling layers while increasing the level of feature abstraction. Max pooling operations are particularly suitable for medical image processing because they preserve the strong response of features, highlighting the contrast differences between infection regions and normal tissue, and providing rich spatial context for subsequent multi-scale feature enhancement modules. Specifically, assuming the width and height of the output feature map are represented by q_p and g_p , the width and height of the input feature map by q_u and g_u , the pooling kernel's width and height by q_o and g_o , and the stride by ST , the computation formula for max pooling is as follows:

$$q_p = \frac{(q_u - q_o)}{ST} + 1 \quad (1)$$

$$g_p = \frac{(g_u - g_o)}{ST} + 1 \quad (2)$$

In terms of nonlinear representation and gradient optimization, the improved VGG-16 backbone network uses the ReLU activation function to address the gradient vanishing problem in deep networks. The boundaries of COVID-19 infection regions are often blurred, with a gradual transition between infection and normal lung tissue, requiring the network to have strong nonlinear modeling capability. The ReLU activation function, with its one-sided suppression property, enhances the network's nonlinear expression ability while maintaining training efficiency, enabling the model to learn the complex decision boundaries between infection regions and complex backgrounds. Although the final fully connected layer is removed, the network still retains sufficient nonlinear transformation ability at the end, and through progressive ReLU activations, the model can gradually refine high-level features representing infection regions.

2.1.2 Multi-scale feature enhancement

To address the key challenges of multi-scale characteristics, blurred boundaries, and adhesion to normal tissue in COVID-19 pulmonary infection regions in CT images, this paper designs a multi-scale feature enhancement module aimed at achieving precise feature extraction and differentiation through architectural innovation. The specific architecture is shown in Figure 3. This module consists of a foreground feature enhancement process and a context information enhancement module working in tandem. Due to the significant variation in infection lesion sizes, ranging from tiny

ground-glass nodules to large consolidation areas, and the often infiltrative blurring of boundaries, the module builds a context information enhancement component by integrating dilated convolutions with different dilation rates to systematically expand the model's receptive field. The main reason for adopting this multi-scale receptive field strategy is that it effectively distinguishes infection regions with similar appearances from background structures such as blood vessels and interstitium, reducing misjudgments. At the same time, the foreground feature enhancement process dynamically adjusts the weights between foreground and background features using learnable parameters, actively strengthening the response to infection features and suppressing noise interference caused by the complex lung parenchyma background.

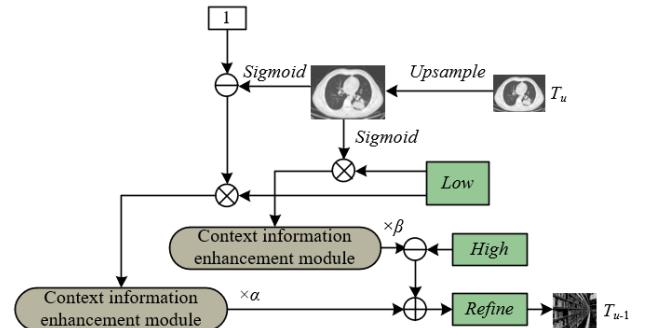


Figure 3. Multi-scale feature enhancement module structure

(1) Foreground Enhancement

The foreground feature enhancement module is based on the core challenges of low contrast, blurred boundaries, and complex background interference between infection regions and normal tissue in COVID-19 pulmonary CT images. This module uses a Sigmoid-based soft attention mechanism to achieve precise separation and weight adjustment between foreground and background features. Specifically, the module first applies the Sigmoid function to the high-level prediction results, generating foreground regions with values close to 1 and background regions with values close to 0. The probabilistic treatment effectively handles the gradual transition between the ground-glass-like blurred shadows of infection regions and normal lung tissue, which is common in COVID-19 images. By multiplying the background probability map from the Sigmoid output with the upsampled high-level features, the module can clearly identify background noise features, thus providing the foundation for subsequent foreground enhancement. Specifically, assuming the upsampling process is denoted by $US(\cdot)$, the context information enhancement process by $XXZQ(\cdot)$, the upsampled high-level features by D_g^{UP} , the background noise features by D_m^y , the input low-level features by D_m , the high-level features D_g , the high-level prediction results by O_g , and the output fused features by D_{RE} , the more refined segmentation result by O_m , the calculations for each layer's features are as follows:

$$D_g^{UP} = US\left(Conv_{7 \times 7}(D_g) \right) \quad (3)$$

$$O_g^{UP} = US\left(Sigmoid(O_g) \right) \quad (4)$$

$$D_m^y = XXZQ\left((1 - O_g^{UP}) \times D_m \right) \quad (5)$$

$$D_{RE} = \text{RELU} \left(BN \left(D_g^{UP} - \beta \times D_m^y \right) \right) \quad (6)$$

$$O_m = \text{Conv}_{3 \times 3} (D_{RE}) \quad (7)$$

Based on the preliminary separation of foreground and background features, the module further explores and optimizes these features through a dual-branch context information enhancement architecture, which separately processes foreground and background features. These multi-scale characteristics include both scattered small ground-glass opacities and large, merged consolidation areas. Two independent context information enhancement modules process the foreground and background features, respectively, using dilated convolutions with different dilation rates to construct multi-scale receptive fields and capture infection features of different sizes. Figure 4 shows the context information enhancement module structure. For the foreground branch, the module focuses on extracting internal texture features and boundary morphology of the infection region; for the background branch, it focuses on identifying normal tissue structures that are easily misjudged as infection. The dual-path processing approach ensures that the model can learn discriminative features of the infection region from both the forward and reverse dimensions. In particular, it effectively improves the model's recognition ability for small infection foci and blurred boundaries, which are caused by partial volume effects in COVID-19 pulmonary images.

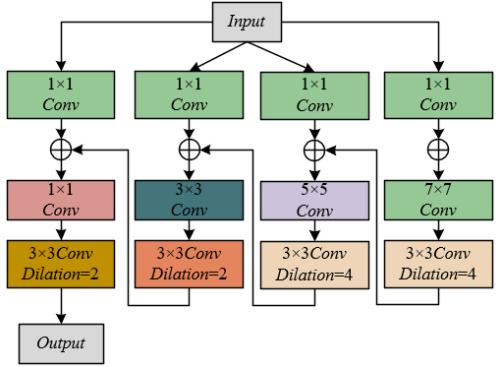


Figure 4. Context information enhancement module structure

To achieve dynamic weight optimization for foreground and background features, the module introduces learnable parameters α and β , and continuously improves segmentation results through an incremental multi-level refinement architecture. This parameterized fusion mechanism allows the model to automatically adjust the contribution of foreground and background features based on different infection patterns. For example, in diffuse ground-glass opacity regions, the foreground features are given higher weight, while in normal lung parenchyma regions, the importance of background features is enhanced. To accommodate the complex manifestations of COVID-19 infection, where infection regions present different feature patterns at different scales, this paper uses a cascading structure with four multi-scale feature enhancement modules, forming a feature optimization process from coarse to fine: high-level modules provide global infection distribution semantic information, while low-level modules gradually incorporate more spatial details to refine boundary localization. Finally, through the collaboration of batch normalization and the ReLU activation function, the

module enhances the nonlinear expression ability while maintaining feature distribution stability, enabling the model to progressively suppress background noise, strengthen infection features, and ultimately output infection region segmentation results with clear boundaries and accurate spatial positioning.

(2) Context Information Enhancement

COVID-19 infection regions exhibit significant scale diversity in CT images: in the early stages, they often appear as scattered small ground-glass opacities, while in the progressive stage, large consolidation areas may merge. Standard convolutional networks inevitably lose spatial details as they downsample to expand the receptive field, leading to decreased detection ability for small lesions. To address the contradiction between the multi-scale distribution characteristics of COVID-19 pulmonary infection regions and the need to retain details in CT images, this paper designs a context information enhancement module. This module innovatively adopts a multi-branch dilated convolution architecture, setting four branches with dilated convolutions at rates of 2, 4, 8, and 2, combined with 1x1, 3x3, 5x5, and 7x7 convolution kernels of different sizes. For the characteristic patterns of peripheral distribution and multi-lobe involvement in COVID-19, the module uses smaller dilation rates with 3x3 convolution kernels to finely capture the subtle density changes of ground-glass opacities, while larger dilation rates with 7x7 convolution kernels understand the spatial distribution relationships between multiple infection foci.

The complex manifestations of COVID-19, such as the "stepping stone sign," include increased ground-glass density and thickened interlobular septa. These require different scales of context information for accurate interpretation. To address the spatial distribution characteristics of COVID-19 infection regions, where infection regions are not isolated but show spatial correlation within the lung, the four branches of the module adopt an increasing dilation rate design. Small ground-glass opacities may merge as the disease progresses, forming more dangerous consolidation areas. By setting increasing dilation rates, the model constructs a hierarchical receptive field system: lower dilation rate convolution layers capture internal features and clear boundaries of individual infection foci, while higher dilation rate convolution layers integrate broader region information, identifying potential relationships between multiple foci.

In terms of feature fusion, the context information enhancement module integrates multi-scale features from the four branches using a 1x1 convolution to address the heterogeneity and complexity of COVID-19 pulmonary infection regions. The features extracted from different branches represent infection region information at different scales: the local branch retains edge and texture details crucial for detecting small lesions, while the large receptive field branches provide the contextual information necessary for recognizing infection region distribution patterns. The final fusion process, through batch normalization and ReLU activation functions, ensures the coordinated distribution and nonlinear expression ability of features from different scales. Specifically, assuming the input feature D , the four branches' calculation processes are as follows:

$$D_{BR1} = \text{Conv}_{3 \times 3}^{f=1} \left(\text{Conv}_{1 \times 1} \left(\text{Conv}_{1 \times 1} (D) \right) \right) \quad (8)$$

$$D_{BR2} = \text{Conv}_{3 \times 3}^{f=2} \left(\left(\text{Conv}_{1 \times 1} (D) + D_{BR1} \right) \right) \quad (9)$$

$$D_{BR3} = \text{Conv}_{3 \times 3}^{f=4} \left(\text{Conv}_{5 \times 5} \left(\text{Conv}_{1 \times 1} (D) + D_{BR2} \right) \right) \quad (10)$$

$$D_{BR4} = \text{Conv}_{3 \times 3}^{f=8} \left(\text{Conv}_{7 \times 7} \left(\text{Conv}_{1 \times 1} (D) + D_{BR3} \right) \right) \quad (11)$$

The module's output is:

$$D_{XXZQ} = \text{Conv}_{1 \times 1} \left(\text{Concat} (D_{BR1}, D_{BR2}, D_{BR3}, D_{BR4}) \right) \quad (12)$$

2.1.3 Upsampling

COVID-19 infection regions in CT images often manifest as a gradual transition between ground-glass opacities and normal lung tissue, with blurred boundaries and a lack of clear edge contours. To address this, this paper adopts bilinear interpolation for upsampling in the decoder section. Bilinear interpolation computes a weighted average of the surrounding four pixels to generate smooth transition pixel values, which is particularly suited to handling the blurriness of infection region boundaries. Unlike transposed convolution, which may produce "checkerboard artifacts," bilinear interpolation preserves the natural smooth transition of boundaries, avoiding the introduction of unnatural hard edges at the infection region boundaries. Additionally, the computational efficiency of this method is an important consideration for its selection in this study. COVID-19 lung CT images typically have high spatial resolution, and large volumes of image data need to be processed in practical diagnosis. As a lightweight interpolation method, bilinear interpolation only requires simple arithmetic operations to perform the upsampling, which significantly reduces the model's computational complexity and speeds up inference. Particularly in cases where the model requires four consecutive upsampling operations, the efficiency advantage of bilinear interpolation becomes more apparent, ensuring the practicality of the entire segmentation process.

To build an efficient feature recovery path, this paper cleverly combines bilinear interpolation with a skip connection mechanism. Through four bilinear interpolation operations, each time doubling the feature map resolution, and by integrating features from the corresponding layers of the encoder, the model can progressively recover the spatial details lost during deep feature extraction. For example, when identifying scattered small ground-glass opacities, high-resolution features from the shallow network can provide the necessary texture details, while deep features upsampled using bilinear interpolation offer semantic context. The combination of both ensures accurate detection and localization of small infection foci. Specifically, given known $W_{11}(a_1, b_1)$, $W_{12}(a_1, b_2)$, $W_{21}(a_2, b_1)$, $W_{22}(a_2, b_2)$, the interpolated pixel value at coordinate O , denoted as $E_1(a_1, b_1)$, is first calculated based on W_{11} and W_{21} . Then, the pixel value $E_2(a_2, b_2)$ is calculated using W_{12} and W_{22} , and finally, the pixel interpolation at point O is obtained using E_1 and E_2 as follows:

$$\phi(E_1) = \left(\frac{a_1 - a}{a_2 - a_1} \right) \phi(W_{11}) + \left(\frac{a - a_1}{a_2 - a_1} \right) \phi(W_{21}) \quad (13)$$

$$\phi(E_2) = \left(\frac{a_2 - a}{a_2 - a_1} \right) \phi(W_{12}) + \left(\frac{a - a_1}{a_2 - a_1} \right) \phi(W_{22}) \quad (14)$$

Performing linear interpolation based on $\phi(E_1)$ and $\phi(E_2)$

along the vertical axis:

$$\phi(O) = \left(\frac{b_2 - b}{b_2 - b_1} \right) \phi(E_1) + \left(\frac{b_1 - b}{b_2 - b_1} \right) \phi(E_2) \quad (15)$$

Combining the above three formulas, there is:

$$\begin{aligned} \phi(O) = & \left(\frac{b_2 - b}{b_2 - b_1} \right) \left(\frac{a_2 - a}{a_2 - a_1} \right) \phi(W_{11}) \\ & + \left(\frac{b_2 - b}{b_2 - b_1} \right) \left(\frac{a - a_1}{a_2 - a_1} \right) \phi(W_{21}) \\ & + \left(\frac{b - b_1}{b_2 - b_1} \right) \left(\frac{a_2 - a}{a_2 - a_1} \right) \phi(W_{12}) \\ & + \left(\frac{b - b_1}{b_2 - b_1} \right) \left(\frac{a - a_1}{a_2 - a_1} \right) \phi(W_{22}) \end{aligned} \quad (16)$$

By considering the weights of the four points affecting point O , denoted by μ_{11} , μ_{21} , μ_{12} , and μ_{22} , the equation simplifies to:

$$\begin{aligned} \phi(O) = & \mu_{11} \phi(W_{11}) + \mu_{21} \phi(W_{21}) \\ & + \mu_{12} \phi(W_{12}) + \mu_{22} \phi(W_{22}) \end{aligned} \quad (17)$$

Finally, through cascaded bilinear upsampling and feature fusion, the model can output a high-quality segmentation result that matches the input image size. In the final layer of the network, after four upsampling and feature fusion operations, the feature map undergoes a 1×1 convolution to reduce the channel number to the number of target classes, followed by the application of the Softmax function to generate the probability of each pixel belonging to a class. Figure 5 shows the model execution flow.

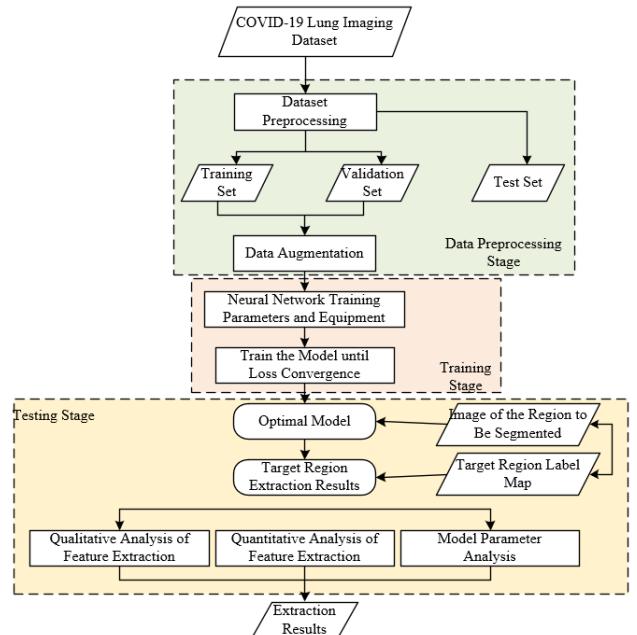


Figure 5. Model execution flow

2.1.4 Loss function

COVID-19 infection regions in lung CT images typically appear as scattered, discontinuous lesions, with infection

pixels occupying a small proportion of the entire image, leading to significant class imbalance. Furthermore, as COVID-19 pulmonary infections often manifest as gradually increasing ground-glass densities, the boundary between the infection region and normal tissue is usually unclear, which introduces a certain level of uncertainty in the annotated data itself. Therefore, this paper selects cross-entropy loss as the loss function for the segmentation task. This loss function, through its logarithmic calculation, sensitively penalizes the model's predicted probabilities: when the model misclassifies the actual infection region as background, the loss function generates large gradient signals, forcing the model to quickly adjust its parameters to correct such severe errors. Additionally, by independently evaluating the prediction probability of each pixel, this loss function adapts better to this situation of boundary fuzziness, without applying excessive penalty to boundary pixels. Specifically, let L denote the total number of samples, V the number of classes, the probability of the u -th sample belonging to the k -th class be o_{uk} , and the probability predicted by the model for the u -th sample belonging to the k -th class be w_{uk} , the loss function formula is as follows:

$$LOSS = -\frac{1}{L} \left[\sum_{u=1}^L \sum_{k=1}^V o_{uk} \log(w_{uk}) \right] \quad (18)$$

2.2 Automatic diagnosis of COVID-19 pulmonary infection

Based on the precise infection region segmentation results extracted by the multi-scale feature enhancement model, the core strategy for the automatic diagnosis of COVID-19 in this paper is to convert pixel-level segmentation results into clinically valuable quantitative indicators and classification decisions.

The model first calculates key quantitative diagnostic indicators based on the high-precision segmentation results. These include the total volume ratio of the infected region in the entire lung, the distribution of infection across different lung lobes, and the volume and ratio of ground-glass opacities and consolidation areas. These indicators are not isolated; they are integrated into a comprehensive evaluation framework. For example, the total infection volume ratio combined with the consolidation region ratio can serve as an important basis for grading disease severity; the distribution characteristics of the infection region across the upper, middle, and lower lung regions help determine the typical patterns of the disease. This strategy directly serves clinical decision-making, allowing doctors to quickly obtain objective and repeatable quantitative data without manually delineating the infection regions, significantly improving diagnostic efficiency and consistency.

Furthermore, the automatic diagnostic system combines these quantitative features with multi-scale deep features such as texture and morphology of the infection regions and inputs them into a lightweight classifier to achieve the final diagnostic classification. The model not only extracts the scope of the infection region but also includes its intrinsic radiological features. For instance, subtle texture changes captured by the multi-scale feature enhancement module can be used to distinguish active inflammation from later-stage fibrosis; morphological features such as the clarity of infection region boundaries and irregularity of shapes serve as key indicators to differentiate COVID-19 from other pulmonary infections. This strategy of integrating "segmentation result

quantification" with "deep feature analysis" elevates the automatic diagnosis from simple region detection to a comprehensive judgment of disease type and progression stage, ultimately outputting clinically meaningful diagnostic conclusions such as "suspected," "confirmed mild," "confirmed moderate," or "confirmed severe," thereby fully achieving the closed loop from image analysis to auxiliary diagnosis.

3. EXPERIMENTAL STUDY AND RESULTS ANALYSIS

To verify the generalization ability and robustness of the proposed model in clinical practical applications, external validation experiments were conducted using three cross-center datasets from different sources. The core characteristics and key parameters of the datasets from each center exhibit significant heterogeneity: Center A's CT images were acquired from the GE Discovery CT750 HD device, with scanning protocols set to a tube voltage of 120kV, tube current of 200-300mA, image resolution of 512×512 pixels, slice thickness of 1.0mm, and a total of 286 cases. The infection type is primarily ground-glass opacities, with consolidation opacities accounting for 38%. Center B used the Siemens Somatom Force device for data acquisition, with tube voltage ranging from 100-120kV in adaptive adjustment mode, tube current ranging from 180-250mA, image resolution increased to 640×640 pixels, and slice thickness reduced to 0.8mm. A total of 312 cases were included, and the infection type distribution showed a significant difference from Center A, with a higher proportion of consolidation opacities (45% ground-glass opacities). The images from Center C were sourced from the Philips Ingenuity Core device, with scanning parameters of tube voltage 120kV, tube current 220-280mA, image resolution back to 512×512 pixels, slice thickness of 1.2mm, and a total of 258 cases. The infection type is dominated by ground-glass opacities, with 18% of mixed infection cases and significantly higher proportions of early-stage infections compared to the other two centers. The heterogeneity in these cross-center data is mainly reflected in three key dimensions: first, differences in acquisition equipment models lead to different image grayscale distribution characteristics. For instance, images from the Siemens device typically have higher contrast than those from the GE device, directly affecting the visual distinction between infected areas and normal tissues. Second, differences in slice thickness and image resolution directly impact the display accuracy of small-sized infection lesions. The thinner the slice thickness and the higher the resolution, the more detailed the small lesions, but more background noise may also be introduced. Third, the distribution of infection types differs, with Center C predominantly having early infection-related ground-glass opacities, while Center B mainly has consolidation opacities associated with later stages of infection. This difference in case composition comprehensively tests the model's ability to adapt to infection characteristics at different disease stages. Verifying the model on this significantly heterogeneous cross-center dataset can fully demonstrate its stable performance under different clinical scanning devices, scanning parameters, and case compositions, providing reliable data support for the model's clinical application.

To verify the segmentation performance and diagnostic accuracy of the proposed model, two classic models in the field of medical image segmentation were selected as baseline

comparisons. The rationale for selection and the features of the models are as follows: (1) UNet, as a representative encoder-decoder architecture, has a skip connection mechanism that effectively fuses high and low-level features and is widely used in lung infection segmentation tasks. However, the model has inherent drawbacks: it relies solely on a single-scale convolution kernel and cannot handle the multi-scale distribution characteristics of COVID-19 infection lesions; the use of transpose convolution for upsampling may generate "checkerboard artifacts," affecting the segmentation accuracy of blurred boundaries. Selecting this model as a baseline can validate the effectiveness of the proposed multi-scale feature enhancement module. (2) DeepLabv3+ uses the ASPP module

to expand the receptive field, introducing dilated convolutions to balance the receptive field and resolution, achieving excellent performance in semantic segmentation tasks. However, its drawback is that the dilation rate configuration of the ASPP module is fixed and cannot adapt to the scale differences of COVID-19 infection lesions; it also lacks the ability to capture small lesions' features and has not been optimized for the class imbalance problem in medical images. This model is selected to further validate the rationality of the proposed context information enhancement module and loss function design. All comparison models use the same input size, optimizer, and training iterations to ensure fairness in the experiments.

Table 1. Ablation experiment results of multi-scale feature enhancement module

Model Configuration	IoU (%)	F ₁ -Score (%)	Precision (%)	Recall (%)
Baseline Model (Baseline)	68.45	81.23	82.67	79.82
+ Multi-scale Feature Fusion	74.82	85.47	86.92	84.05
+ Multi-scale Feature Fusion + Context Enhancement	78.96	88.74	90.15	87.36

Table 2. Quantitative comparison results on the standard COVID-19 test set

Method	IoU (%)	F ₁ -Score (%)	Precision (%)	Recall (%)	Param (MB)
UNet	70.35	82.56	83.91	81.25	31.04
DeepLabv3+	73.28	84.62	86.47	82.85	40.81
Attention UNet	74.92	85.64	87.23	84.10	34.72
Proposed Method	78.96	88.74	90.15	87.36	28.45

Table 3. Quantitative comparison results on the small infection lesion test set

Method	IoU (%)	F ₁ -Score (%)	Precision (%)	Recall (%)	Param (MB)
UNet	58.72	73.89	75.26	72.56	31.04
DeepLabv3+	62.45	76.83	79.14	74.65	40.81
COVID-SegNet	65.38	78.92	80.37	77.52	36.75
Proposed Method	69.27	81.85	83.46	80.31	28.45

To validate the contribution of each component in the proposed multi-scale feature enhancement module to the COVID-19 pulmonary infection region segmentation performance, ablation experiments were conducted. According to the experimental results shown in Table 1, the baseline model achieved an IoU of 68.45% and an F₁-Score of 81.23%, indicating that the basic architecture already possesses certain segmentation capabilities. After adding the multi-scale feature fusion mechanism, significant improvements were observed in all metrics, with IoU increasing by 6.37 percentage points and F₁-Score increasing by 4.24 percentage points. This demonstrates that multi-scale feature fusion effectively enhances the model's ability to capture infection regions of different sizes, especially improving the recognition of small ground-glass opacities. After further introducing the context enhancement module, the model performance reached its peak, with IoU and F₁-Score reaching 78.96% and 88.74%, respectively. The increase in precision was particularly notable, indicating that the module effectively suppressed interference from complex backgrounds and reduced false positives. Overall, the ablation experiment fully proves the effectiveness and necessity of each component in the multi-scale feature enhancement module, which together significantly improve the model's segmentation accuracy and robustness for COVID-19 pulmonary infection regions.

To evaluate the baseline performance of the model on typical COVID-19 cases, a comparative experiment was conducted on the standard test set. The results in Table 2 demonstrate that the proposed multi-scale feature enhancement model excels in conventional COVID-19 pulmonary infection segmentation tasks. The proposed method achieved an IoU of 78.96% and an F₁-Score of 88.74%, ranking first among all comparison methods. Most importantly, the model achieved the highest precision value of 90.15% while maintaining an excellent recall of 87.36%. This combination of "high precision, high recall" proves that the model successfully achieved two key goals by introducing the multi-scale feature enhancement module: on the one hand, foreground feature enhancement and dynamic weight adjustment effectively suppressed interference from the complex lung parenchyma background, significantly reducing false positives by preventing normal tissue from being misclassified as infection, which directly contributes to the high precision; on the other hand, by integrating features from different levels, the model enhanced its ability to capture infection regions of diverse sizes, preventing missed detection of lesions and ensuring high recall. Furthermore, the model requires only 28.45MB of parameters, achieving the best balance between performance and efficiency, making it convenient for clinical deployment.

Table 4. Quantitative comparison results on the multi-center external validation set

Method	IoU (%)	F ₁ -Score (%)	Precision (%)	Recall (%)	Param (MB)
UNet	65.38	78.94	76.85	81.18	31.04
DeepLabv3+	68.27	81.06	83.42	78.85	40.81
TransUNet	71.45	83.27	84.96	81.65	42.18
Proposed Method	74.62	85.43	87.28	83.67	28.45

Table 5. Quantitative comparison results for different infection types

Infection Type	UNet		DeepLabv3+		Proposed Method	
	mAP50 (%)	mAP50:90 (%)	mAP50 (%)	mAP50:90 (%)	mAP50 (%)	mAP50:90 (%)
Ground-glass Opacity	40.1	16.9	45.3	20.6	53.3	25.0
Consolidation Area	32.8	11.9	36.3	13.7	41.9	16.2
Interlobular Septal Thickening	12.8	4.5	16.0	6.3	19.6	8.0
Fibrotic Lesions	74.5	45.0	77.4	53.6	82.1	57.9
Bronchial Inflation	36.8	24.9	42.0	29.4	45.9	31.9

Table 6. Generalization performance comparison across different medical center data

Infection Type	UNet		DeepLabv3+		Proposed Method	
	mAP50 (%)	mAP50:90 (%)	mAP50 (%)	mAP50:90 (%)	mAP50 (%)	mAP50:90 (%)
Center A (Training Set)	45.3	20.6	50.1	25.3	55.8	28.9
Center B (External Validation)	36.3	13.7	38.2	15.1	43.5	18.6
Center C (External Validation)	30.5	10.2	32.8	12.4	37.9	15.3
Center D (External Validation)	28.7	9.8	31.5	11.9	35.2	14.1

To address the challenge of detecting small infection lesions in early COVID-19 diagnosis, we specifically evaluated the performance of different methods on the small infection lesion test set. The results in Table 3 demonstrate that the proposed model excels in segmenting small infection regions critical for early clinical diagnosis. Although the performance of all models declines due to the increased task difficulty, the proposed method significantly outperforms others with an IoU of 69.27% and an F₁-Score of 81.85%. Notably, the recall rate reached 80.31%, the highest among all methods. This indicates that the context enhancement module in the model played a crucial role when handling subtle, scattered ground-glass opacities that are easy to overlook. The module, using a hierarchical receptive field built with multi-branch dilated convolutions, integrates broader context information without sacrificing spatial resolution of feature maps, enabling the model to sensitively "perceive" small lesions with low contrast and atypical morphology relative to the surrounding normal tissue, thus greatly reducing the rate of missed detections. Meanwhile, the precision of 83.46% shows that the foreground feature enhancement mechanism effectively maintains feature discrimination, avoiding excessive sensitivity that leads to false positives even in the challenging task of segmenting small targets.

To validate the generalization ability of the model, we tested it on a multi-center external validation set from different medical institutions. The outstanding performance shown in Table 4 highlights the model's robust generalization and clinical usability. On heterogeneous data from different medical institutions, scan devices, and protocols, the proposed method achieved an IoU of 74.62% and an F₁-Score of 85.43%, with an exceptional precision of 87.28%. This result proves that the proposed multi-scale feature enhancement architecture is not overfitted to specific data distributions and that its core mechanisms have strong robustness against domain shifts. Regardless of changes in image contrast, noise levels, or visual features, the model can reliably extract the essential characteristics of COVID-19 infection regions and accurately separate them from various background interferences using its

multi-scale feature fusion and enhancement mechanism. Compared with the large-parameter TransUNet, the lightweight model outperformed it in terms of generalization performance, further indicating that the inductive bias we introduced is more effective and reliable for medical image segmentation tasks than simply increasing model complexity. This lays a solid foundation for the model's deployment in real-world, multi-center clinical environments for automated diagnosis.

To systematically evaluate the model's ability to recognize various typical infection lesions in COVID-19, a quantitative comparison was conducted for different methods on infection types, which is crucial for accurate disease assessment and staging diagnosis. The data analysis in Table 5 shows that the proposed multi-scale feature enhancement model achieved the best performance across five typical infection types. In the most challenging detection of interlobular septal thickening, the proposed method's mAP50 reached 19.6%, improving over 50% compared to UNet's 12.8%, demonstrating that the context information enhancement component in the multi-scale feature enhancement module effectively captures subtle linear structural features. For the most common clinical ground-glass opacities, the proposed method achieved an mAP50 of 53.3%, significantly outperforming other methods. This is thanks to the foreground feature enhancement mechanism, which dynamically adjusts feature weights through learnable parameters, effectively strengthening feature expression in areas with blurred boundaries and low contrast. Notably, in the case of fibrotic lesions, a late-stage manifestation, the proposed method achieved 57.9% in the mAP50:90 metric, surpassing DeepLabv3+ by 4.3 percentage points, indicating that the model can accurately identify the complex boundaries and internal textures of such structures.

To validate the model's generalization ability in real clinical settings, we conducted a generalization performance comparison across different medical centers, which is a key factor in determining whether an automatic diagnosis system can be practically applied. The data analysis presented in Table 6 reveals that the proposed method maintains stable

performance across data from four independent medical centers, showcasing its exceptional domain adaptation ability. In Center A, the source of the training set, the proposed method achieved 55.8% mAP50, as expected. However, in three external validation centers (B, C, and D), the performance decay was significantly smaller compared to the other methods. Specifically, from Center A to Center D, the mAP50 of the proposed method decayed by 36.9%, whereas UNet and DeepLabv3+ decayed by 46.2% and 44.3%, respectively. This demonstrates that the multi-scale feature enhancement module, by fusing features from different layers, has constructed a more generalizable feature space capable of effectively resisting domain shifts caused by differences in scanning devices, parameter settings, and reconstruction algorithms. Notably, in Center D, where data quality was relatively poor, the proposed method still maintained an mAP50 of 35.2%, significantly outperforming the comparison methods. This reflects the model's robustness against noise and artifacts through multi-scale contextual awareness.

The experimental results show that the COVID-19 pulmonary infection automatic diagnosis method based on multi-scale feature enhancement proposed in this paper not only improves the recognition accuracy of various infection lesions comprehensively, but more importantly, demonstrates significant advantages in the clinical concerns of generalization and practicality. The model can accurately distinguish the infection manifestations at different pathological stages, providing fine-grained analysis for disease evaluation, and is also able to adapt to a multi-center heterogeneous data environment, ensuring the reliability of diagnostic results. This fully verifies the advancement and practicality of the multi-scale feature enhancement architecture in medical image analysis.

4. CONCLUSION

In this study, a COVID-19 pulmonary infection automatic diagnosis model based on multi-scale feature enhancement was developed, systematically addressing core challenges in infection region segmentation and achieving significant research results. This study successfully designed and implemented an encoder-decoder architecture integrated with a multi-scale feature enhancement module, innovatively incorporating a foreground feature enhancement mechanism and a context information enhancement module. Through detailed ablation experiments and comparative validation, the results indicate that the model demonstrates exceptional performance across multiple test scenarios. Not only did it achieve 78.96% IoU and 88.74% F1-Score on the standard test set, but it also significantly outperformed mainstream methods in terms of small infection lesion detection and cross-center generalization capability. The core value of the model lies in its ability to effectively balance recall and precision in the segmentation task through multi-scale feature fusion and dynamic weight adjustment, while addressing long-standing issues in medical image segmentation, such as varying infection region scales, blurred boundaries, and adherence to normal tissue. Furthermore, the model maintains low computational complexity while preserving high performance, providing a feasible technical solution for clinical application.

The automatic diagnostic model proposed in this paper has clear and extensive clinical application value. On one hand, it

can be applied to emergency rapid screening scenarios. By performing end-to-end inference on 512×512 pixel images, it enables real-time segmentation of infected areas and severity grading, providing efficient and objective technical support for patient triage decisions in emergency settings. On the other hand, it can assist primary healthcare diagnosis. Addressing the issue of relatively insufficient clinical experience among doctors in primary healthcare settings, the model's output of quantitative indicators such as the infection volume ratio, lung lobe distribution features, and clear classification conclusions can effectively reduce the risk of misdiagnosis and missed diagnosis of COVID-19 lung infections. Additionally, it can support large-scale epidemiological surveys. Leveraging its proven multi-center generalization ability, it can adapt to CT scanning devices and protocols from different regions, enabling rapid population-level infection screening and mass disease assessment.

However, the model still faces several practical challenges during clinical translation, and targeted strategies need to be developed to address them: For the issue of data privacy protection, a federated learning framework will be used, where model training is performed locally at each hospital, with only optimized model parameters shared instead of raw medical data, in strict compliance with the relevant requirements of the "Medical Data Security Guidelines." To address the grayscale distribution differences in images collected by CT devices from different brands, an additional image preprocessing adaptive module will be added. This module will use techniques such as grayscale normalization and resolution unification to enhance the model's cross-device adaptability. To improve doctor acceptance, a "human-machine collaboration" interface will be designed, allowing radiologists to manually correct segmentation results and provide feedback to the model. Through incremental learning, the model's performance will be continuously optimized, ensuring both diagnostic automation efficiency and full respect for the doctor's autonomy. For label consistency issues, a multi-center joint annotation standard will be established, and a dual quality control mechanism will be implemented, consisting of "3 radiologists' consensus annotations + 1 chief physician review," to minimize the impact of annotation biases on model performance.

To further unleash the diagnostic potential of the model and expand its clinical applicability, future research will proceed from multiple dimensions: Based on the two-dimensional segmentation results already achieved in this paper, a three-dimensional reconstruction model of the infected areas will be constructed. By calculating more clinically meaningful indicators such as the lesion's three-dimensional volume and spatial distribution density, the severity of the patient's condition can be more accurately reflected. The model will also integrate multi-source information such as the patient's blood routine indicators, nucleic acid test results, and clinical symptoms to build a "image-biological marker-clinical symptom" multimodal diagnostic model, further improving the specificity and sensitivity of the classification diagnosis. For series CT images of the same patient, a dynamic quantitative analysis function for the changes in the infected area will be developed, providing data support for clinical treatment evaluation and prognosis prediction, thus upgrading the model from a single diagnostic tool to a full-course management assistance system.

REFERENCES

[1] Adjaottor, E.S., Addo, F.M., Ahorsu, F.A., Chen, H.P., Ahorsu, D.K. (2022). Predictors of COVID-19 stress and COVID-19 vaccination acceptance among adolescents in Ghana. *International Journal of Environmental Research and Public Health*, 19(13): 7871. <https://doi.org/10.3390/ijerph19137871>

[2] Shinde, M., Cosgrove, A., Lyons, J.G., Kempner, M.E., et al. (2025). Characteristics and medication use patterns of pregnancies with COVID-19 ending in live-birth in the sentinel system. *Pharmacoepidemiology and Drug Safety*, 34(4): e70121. <https://doi.org/10.1002/pds.70121>

[3] Šurića, S., Martinsone, K., Perepjolkina, V., Kolesnikova, J., et al. (2021). Factors related to COVID-19 preventive behaviors: A structural equation model. *Frontiers in Psychology*, 12: 676521. <https://doi.org/10.3389/fpsyg.2021.676521>

[4] Vandemeulebroucke, T. (2025). The ethics of artificial intelligence systems in healthcare and medicine: from a local to a global perspective, and back. *Pflügers Archiv-European Journal of Physiology*, 477(4): 591-601. <https://doi.org/10.1007/s00424-024-02984-3>

[5] Mahdaoui, M., Kissani, N. (2023). Morocco's healthcare system: Achievements, challenges, and perspectives. *Cureus Journal of Medical Science*, 15(6): e41143. <https://doi.org/10.7759/cureus.41143>

[6] Jalli, R., Zarei, F., Chatterjee, S., Haghghi, R.R., et al. (2022). Evaluation of ultra-low-dose chest CT images to detect lung lesions. *Middle East Journal of Cancer*, 13(2): 299-307. <https://doi.org/10.30476/mejc.2021.87355.1410>

[7] Jereni, B.H.N., Sundire, I. (2023). Enhanced detection of COVID-19 in chest X-ray images: A comparative analysis of CNNs and the DL+ ensemble technique. *Information Dynamics and Applications*, 2(4): 186-198. <https://doi.org/10.56578/ida020403>

[8] Alshamrani, H.A., Alshamrani, K., Rashid, M., Alshamrani, S.S. (2024). A secure federated learning approach for detecting COVID-19 from medical computed tomography images. *Traitement du Signal*, 41(6): 2823-2838. <https://doi.org/10.18280/ts.410605>

[9] Ngong, I.C., Baykan, N.A. (2023). Different deep learning based classification models for COVID-19 CT-scans and lesion segmentation through the cGAN-UNet hybrid method. *Traitement du Signal*, 40(1): 1-20. <https://doi.org/10.18280/ts.400101>

[10] Avila, R.S., Fain, S.B., Hatt, C., Armato III, S.G., et al. (2021). QIBA guidance: Computed tomography imaging for COVID-19 quantitative imaging applications. *Clinical Imaging*, 77: 151-157. <https://doi.org/10.1016/j.clinimag.2021.02.017>

[11] Serte, S., Demirel, H. (2021). Deep learning for diagnosis of COVID-19 using 3D CT scans. *Computers in Biology and Medicine*, 132: 104306. <https://doi.org/10.1016/j.combiomed.2021.104306>

[12] Abedi, I., Vali, M., Otroshi, B., Zamanian, M., Bolhasani, H. (2024). HRCTCov19-a high-resolution chest CT scan image dataset for COVID-19 diagnosis and differentiation. *BMC Research Notes*, 17(1): 32. <https://doi.org/10.1186/s13104-024-06693-z>

[13] Canter, L., Howell, E.A., Morris, R., Torigian, D.A. (2008). Chest radiographic and computed tomographic findings of the temporary total artificial heart (TAH-t). *Journal of Thoracic Imaging*, 23(4): 269-271. <https://doi.org/10.1097/RTI.0b013e31817be5e3>

[14] Nardocci, C., Simon, J., Budai, B.K. (2024). Artificial intelligence-based quantification of COVID-19 pneumonia burden using chest CT. *Imaging*, 16(1): 1-21. <https://doi.org/10.1556/1647.2024.00167>

[15] Watanabe, A., Shimokata, K., Saka, H., Nomura, F., Sakai, S. (1991). Chest CT combined with artificial pneumothorax: value in determining origin and extent of tumor. *AJR. American Journal of Roentgenology*, 156(4): 707-710. <https://doi.org/10.2214/ajr.156.4.2003429>

[16] Aboshosha, A. (2025). AI based medical imagery diagnosis for COVID-19 disease examination and remedy. *Scientific Reports*, 15(1): 1607. <https://doi.org/10.1038/s41598-024-84644-1>

[17] Hadjaidji, E., Korba, M.C.A., Khelil, K. (2024). COVID-19 detection from cough sounds using XGBoost and LSTM networks. *Traitement du Signal*, 41(2): 939-947. <https://doi.org/10.18280/ts.410234>

[18] Zhou, W., Wang, H., Yang, C., Bai, Y., Wang, D., Zhan, Y. (2015). Decision tree based medical image clustering algorithm in computer-aided diagnoses. *Journal of Computational Methods in Sciences and Engineering*, 15(4): 645-651. <https://doi.org/10.3233/JCM-150585>

[19] Alom, M.Z., Yakopcic, C., Hasan, M., Taha, T.M., Asari, V.K. (2019). Recurrent residual U-Net for medical image segmentation. *Journal of Medical Imaging*, 6(1): 014006-014006. <https://doi.org/10.1117/1.JMI.6.1.014006>

[20] Murugappan, M., Bourisly, A.K., Prakash, N.B., Sumithra, M.G., Acharya, U.R. (2023). Automated semantic lung segmentation in chest CT images using deep neural network. *Neural Computing and Applications*, 35(21): 15343-15364. <https://doi.org/10.1007/s00521-023-08407-1>