International Journal of Innovative Computing, Information and Control
Volume **22**, Number **2**, April **2026**

ICIC International ©2026 ISSN 1349-4198
pp. **393–411**

LUNG NODULES SEGMENTATION BASED ON A NOVEL BIDIRECTIONAL LSTM NEURAL NETWORK

Tianwen Wang[1], Guangxu Li[2,3] and Tohru Kamiya[1,*]

[1]Department of Mechanical and Control Engineering
Kyushu Institute of Technology
1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka 804-8550, Japan
wang.tianwen172@mail.kyutech.jp; *Corresponding author: kamiya@cntl.kyutech.ac.jp

[2]School of Optometry & Ophthalmology
Tianjin Medical University Eye Hospital
Tianjin Medical University
No. 22, Qixiangtai Road, Heping District, Tianjin 300070, P. R. China

[3]Tianjin Key Laboratory of Optoelectronic Detection Technology and System
Tiangong University
No. 399, Binshui West Road, Xiqing District, Tianjin 300387, P. R. China
liguangxu@tiangong.edu.cn

Received July 2025; revised November 2025

Abstract. *As diversity of characteristics, quantitatively analyzing the pulmonary nodules in CT images remains a big challenge. Existing CNN-based segmentation methods are hindered by redundant characteristics of convolutional operation, and scarcity of annotated data for the peripheral tissues. In this paper, we proposed a novel network for automatic pulmonary nodule segmentation, named RICSBU-Net, which models the relationship between the encoder and decoder via a bidirectional convolutional LSTM. The main contribution includes the following. Firstly, we replace U-Net's encoder with a pretrained ResNet-18 to overfit and prevent semantic feature loss typically associated with pooling layers. Secondly, an ICS-Attention module is introduced to mitigate performance degradation caused by variations in nodule morphology, location, and density. This module merges fine-grained spatial and semantic information to amplify the weight of important features. Thirdly, the ICS-BiConvLSTM module replaces the skip connection and concatenator, with the ICS-Attention module emphasizing key features to reduce parameters, and BiConvLSTM enhancing the extraction of local features. Various experiments on the LUNA16 dataset demonstrate that RICSBU-Net surpasses other advanced methods in both quantitative assessment and qualitative metrics, achieving a DSC of 95.84%. Comparing to the other up-to-date segmentation methods, RICSBU-Net delivers more refined and consistent boundary delineation, particularly for nodules near adjacent tissues.*

Keywords: Pulmonary nodule segmentation, ICS-Attention, Bidirectional ConvLSTM, Deep learning

1. **Introduction.** Lung cancer is the primary cause of cancer-related fatalities globally, claiming over one million lives each year. According to 2024 cancer statistics, approximately 340 people die from lung cancer per day [1]. Lung cancer presents with subtle pathology and is commonly diagnosed at the metastatic stage, resulting in low survival rates and poor prognosis for patients [2]. Computed Tomography (CT) provides superior clarity in detecting lung abnormalities compared to X-rays [3]. A randomized controlled

---

DOI: 10.24507/ijicic.22.02.393

trial on lung cancer screening indicates that early identification of nodular lesions can lower lung cancer mortality by 25% or more [4]. However, due to the multi-slice nature of CT images, manual segmentation by radiologists is both time-consuming and labor-intensive, and the subjective nature of interpretation potentially affecting the robustness and precision of results, especially when dealing with lung nodules with indistinct boundaries [5]. Therefore, an accurate and automatic computer-assisted lung nodule segmentation technique is essential for the diagnosis of nodules.

Nodules appear as approximately rounded and indistinct lesions with diameters not exceeding 30 mm on CT imaging [6], typically manifesting as small, ill-defined spots. According to their attenuation characteristics on imaging, lung nodules are classified into solid nodules, ground-glass nodules (GGN), and part-solid nodules [7]. Solid nodules will obscure internal blood vessels and bronchi, while ground-glass nodules, although not obscuring internal structures, have indistinct boundaries and appear less clear on imaging. Part-solid nodules exhibit characteristics of both types. The diversity and irregularity in the location, size, and shape of pulmonary nodules lead to some nodules having complex attachments (such as to blood vessels, bronchi, the lung wall, or the pleura), which complicates segmentation (see Figure 1).

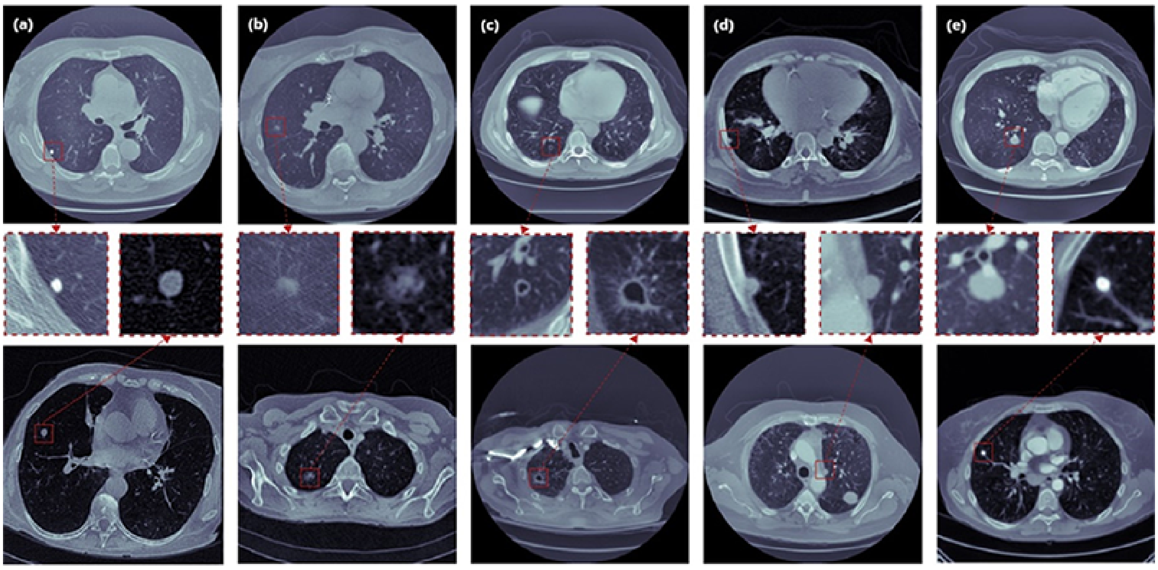


FIGURE 1. Different types of lung nodules: (a) Solid nodules, (b) GGNs, (c) solid nodules with a cavity, (d) and (e) adhesive

Although considerable progress has been made in the field of lung nodule segmentation, existing approaches still exhibit several limitations. Traditional image processing methods rely heavily on manually crafted low-level features, such as region growing, fuzzy clustering, and morphological operations [8, 9, 10, 11]. These approaches often struggle when encountering nodules with indistinct boundaries, complex attachments, or irregular shapes. The fundamental limitation arises from their insufficient feature representation capacity, making them vulnerable to noise interference and variations in nodule heterogeneity [12].

Deep learning has significantly improved medical image segmentation performance through automatic feature learning and has been widely applied in CT image analysis [13]. Its high degree of automation and efficiency aligns better with clinical requirements, further driving the continuous evolution of lung nodule segmentation methods [14]. Early studies relied mainly on classical architectures such as CNN [15], FCN [16], and U-Net

[17]. However, the representational capacity of these basic models remains limited, particularly in multiscale feature fusion and boundary detail modeling, motivating numerous subsequent improvements. Broadly, existing advancements can be categorized into three major directions.

The first category focuses on multiscale feature fusion [18, 19, 14, 20, 21, 22]. Through mechanisms such as dense skip connections, deep supervision, wavelet fusion, or dilated convolutions, these methods enhance robustness across nodules of varying sizes and morphological patterns. However, multi-scale architectures often increase network complexity and rely on static feature concatenation, lacking fine-grained semantic selection capabilities and potentially introducing redundant representations.

The second category emphasizes residual enhancement and lightweight design, represented by architectures such as ResBCDU-Net [23], SAtUNet [24], and DA-Net [25]. These methods improve gradient flow and enhance deep semantic representation, resulting in more effective feature extraction. Nevertheless, their ability to preserve shallow spatial details remains limited, frequently leading to coarse boundary segmentation, particularly in complex scenarios with subtle boundary variations.

The third category incorporates attention mechanisms, including AFN [26], cascaded attention networks [27], and various attention-enhanced U-Net variants [28, 29]. Although attention mechanisms strengthen the network's focus on critical regions, most existing designs primarily model local spatial or channel dependencies and struggle to capture cross-scale structural relationships. In addition, they fail to fully leverage the complementary nature of mask and boundary information, thus facing performance limitations in challenging cases such as juxta-pleural nodules and ground-glass nodules [30].

To address these challenges, the proposed method introduces more targeted architectural innovations, as Figure 2 illuminated. First, the ICS-Attention mechanism is designed to selectively integrate shallow spatial details with high-level semantic representations, overcoming the common issue in U-Net where shallow features are suppressed by dominant deep features. Furthermore, ICS-Attention is coupled with a BiConvLSTM module [31] to replace the static skip-connections used in traditional U-Net architectures. This combination enables dynamic and bidirectionally dependent feature fusion, allowing the network to capture temporal-like structural correlations across scales and enhancing its ability to delineate complex nodule boundaries. In addition, the encoder is replaced by a pre-trained ResNet-18 [32], which strengthens deep semantic feature extraction and ensures stable gradient propagation. Finally, the use of advanced preprocessing techniques to generate enhanced three-channel medical images enriches the structural information of the

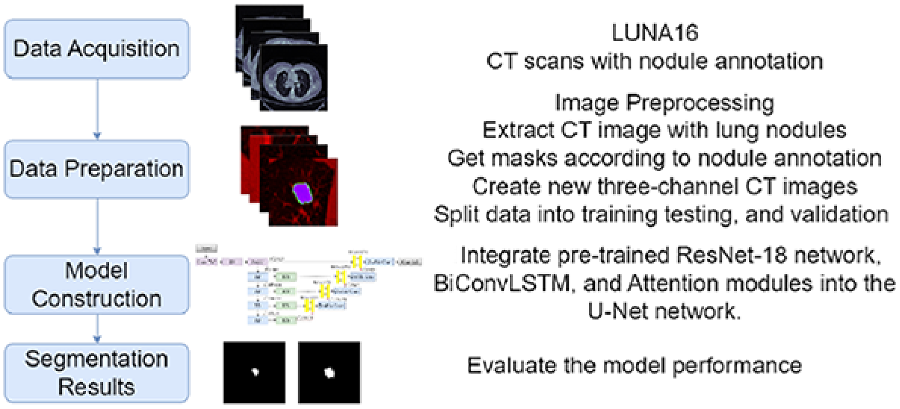


FIGURE 2. Deep learning-based lung nodule segmentation process

input, enabling the model to better handle nodules with fuzzy boundaries and subtle intensity variations. These improvements collectively provide the proposed method with clear advantages in structural information modeling, boundary refinement, and false-positive suppression, making it more effective for accurate and robust lung nodule segmentation.

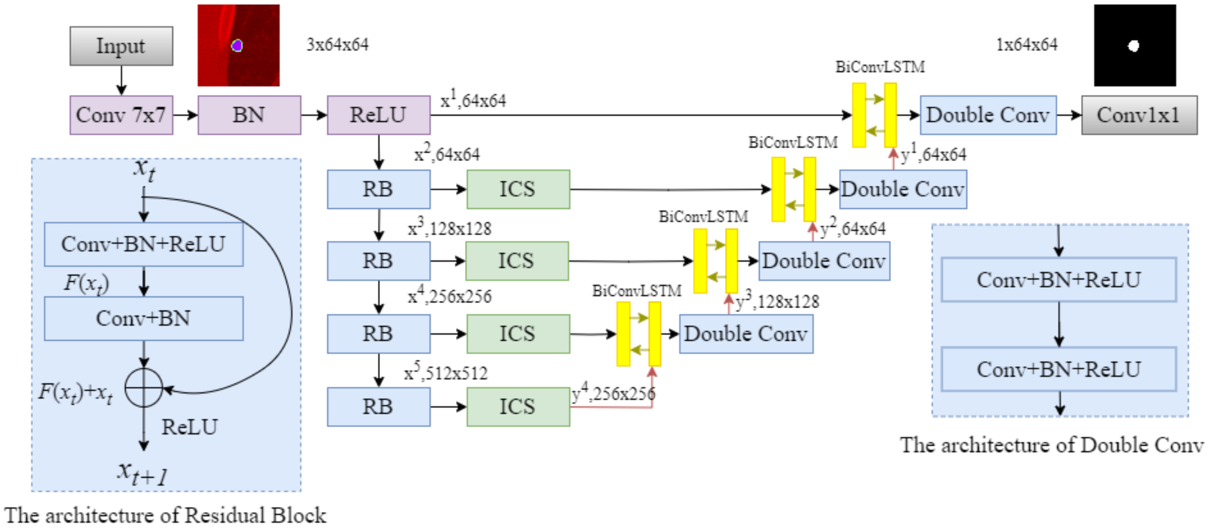
The main innovations presented in this paper are as follows.

- A full-process RICSBU-Net network for pulmonary nodule segmentation is proposed.
- The ICS-Attention mechanism is formulated to integrate detailed spatial and high-level semantic information, and is coupled with the BiConvLSTM module to substitute the skip connections and concatenation components in the U-Net framework.
- Advanced preprocessing techniques are employed to improve raw image quality, producing new three-channel medical images.

The structure of this paper is organized as follows: Section 2 provides a detailed description of the proposed model, Section 3 outlines the experimental setup and presents an in-depth discussion of the experimental results, and Section 4 concludes the paper and discusses future research directions.

2. Method.

2.1. **Framework overview.** Skip connections in the U-Net architecture integrate detailed encoder features with reconstructed decoder features, aiding in recovering fine details lost during spatial compression. However, the irreversible information loss induced by max pooling constrains the network's ability to capture the intricate surface structures and complex topological features of pulmonary nodules. This loss of critical details undermines the efficacy of basic skip connections in integrating multi-scale features, diminishing the model's capacity for comprehensive multi-level semantic representation and compromising its robustness in accurately segmenting pulmonary nodules. To overcome the aforementioned issues, the proposed RICSBU-Net comprises three primary components: a pre-trained ResNet-18 encoder, the ICS-BiConvLSTM module, and a decoder. The detailed structure is presented in Figure 3.



The architecture of Residual Block

The architecture of Double Conv

FIGURE 3. Architecture of RICSBU-Net. These blocks represent the hierarchical structure, while the arrows indicate the inter-layer information flow. The purple arrows signify the upsample operation, while the black arrows indicate copying. The green block represents t the attention module proposed in this paper (Improved Channel Spatial Attention).

The network receives a three-channel image as input, which will be explained in Section 2.2. The ResNet-18 encoder initiates with an initial convolution layer that effectively captures the preliminary features of the input image, establishing a solid foundation for the post-feature processing. The image is then processed through a series of residual modules. The first residual unit preserves the image size while capturing basic shallow features, whereas the second, third, and fourth residual modules progressively extract complex high-level features, enhancing the model's ability to capture detailed structures in the image. The structure of the residual modules, depicted in the lower-left corner of Figure 3, comprises two convolutional layers followed by batch normalization (BN) and ReLU activation operations [33]. The residual block can be expressed in a general form (1).

$$\begin{aligned} y_t &= h(x_t) + F(x_t, W_i) \\ x_{t+1} &= f(y_t) \end{aligned} \tag{1}$$

where, x_t and x_{t+1} represent the input and output vectors of the t-th unit; $h(x_t)$ is the identity mapping; the function $F(x_t, W_i)$ denotes the residual mapping that needs to be learned. $f(y_t)$ is an activation function (ReLU).

This formulation plays a crucial role in preventing the loss of shallow features in deep convolutional networks. By introducing the identity mapping $h(x_t)$, the residual block enables shallow features to be directly propagated to deeper layers rather than being overwhelmed by increasingly abstract high-level representations. This direct feature bypass not only preserves essential edge and texture information but also alleviates gradient vanishing, ensuring that shallow layers receive sufficient gradient flow during backpropagation. Consequently, the network achieves more effective fusion of shallow structural details and deep semantic information, which is particularly important for accurately delineating lung nodule boundaries and refining fine-grained edge structures.

ResBlock modules not only streamline the training process but also facilitate deeper network architectures. The pre-trained ResNet-18 replaces the original encoder, significantly accelerating training, effectively preventing the loss of shallow features caused by pooling layers and mitigating the vanishing gradient problem that arises with increasing network depth. Additionally, the ICS-BiConvLSTM module replaces traditional skip connections and concatenation mechanisms, balancing the extraction of key features and local information. This enables efficient multi-scale information fusion, effectively reducing the semantic disparity between the encoder and decoder.

2.2. **Improved spatial-channel attention module.** The attention mechanism represents a sophisticated technology designed to capture long-range feature interactions and enhance representational capacity [34]. This implies that attention mechanisms are capable of encoding the spatial dependencies between features while also assessing the relative importance of each feature based on its assigned weight. Hu et al. [35] developed the SE-Net architecture, which employs "Squeeze-and-Excitation" (SE) blocks to model inter-channel dependencies and adaptively adjust the responses of channel-specific features, thereby enhancing channel relationships. Although SE-Net, as the first mechanism to effectively learn channel attention, demonstrated outstanding performance, it overlooked spatial attention, which is equally important for determining "where" to focus [36]. To address this, BAM [37] and CBAM [38] modules, which combine spatial and channel attention, were introduced based on an efficient architecture, optimizing channel attention through the integration of max-pooled features. A normalization-based attention module (NAM) [39] was introduced to suppress insignificant weights. To balance performance and complexity, the Efficient Channel Attention (ECA) module [40] uses minimal parameters while significantly enhancing performance. Attention mechanisms can be integrated with any feed-forward convolutional neural network. They overcome several limitations in traditional neural networks, including performance degradation with longer input, inefficient computation due to improper input ordering, and inadequate feature extraction and enhancement [41].

In neural networks, enhancing critical features while suppressing irrelevant information, together with integrating multi-level representations, can substantially improve both the performance and computational efficiency of deep learning models. Although attention mechanisms have been widely incorporated into various architectures, the pursuit of higher accuracy often comes at the cost of increased computational complexity. To address this issue, we propose an improved channel – spatial attention module, ICS-Attention (Figure 4), which boosts feature representation capability without increasing computational overhead. The module is lightweight and effectively aggregates both channel and spatial information.

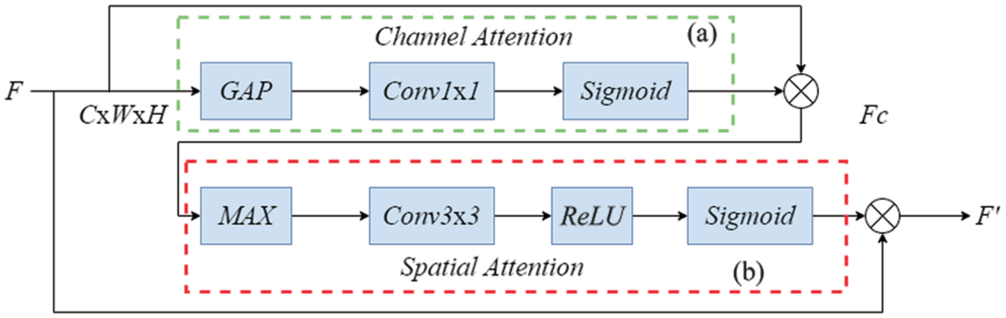


Figure 4. ICS-Attention block: (a) is the channel attention module, and (b) is the spatial attention module

The Channel Attention (CA) mechanism, shown in Figure 4(a), enhances the original SE-Net design by replacing the fully connected (FC) layers with a 1D convolution. Unlike FC layers, which treat the channel descriptor as a global vector and connect each channel to all others, the 1D convolution introduces an explicit local receptive field along the channel dimension. This enables the module to capture local inter-channel interactions and correlations among adjacent channels, while dramatically reducing parameters through weight sharing. Thus, the 1D convolution not only improves the modeling of structured channel dependencies but also maintains high computational efficiency.

In addition to the improved channel attention, ICS-Attention incorporates a spatial attention module (Figure 4(b)) to further refine feature representation. Max-pooling is applied along the channel dimension to emphasizing strong spatial responses, followed by a dilated 3×3 convolution, which effectively enlarges the receptive field and captures broader contextual information without increasing computational cost. A ReLU activation and a Sigmoid function are then employed to produce the final spatial attention map. By sequentially combining channel attention (“what is important”) with spatial attention (“where it is important”), ICS-Attention provides a progressive and complementary enhancement of feature discrimination.

Given $F \in R^{C \times H \times W}$ as the model input, where H, W, and C correspond to the height, width, and channels. To summarize spatial features, global average pooling is applied across each channel, producing the average pooled feature vector Z_G. Then, 1D convolution is used to generate the channel weights, reducing parameter overhead while maintaining efficiency and accuracy through the optimal capture of local inter-channel interactions. The computation of channel attention is as follows:

$$Z_G = GAP(F) = \frac{1}{WH}\sum_{i=1}^{W}\sum_{j=0}^{H} F(i,j) \tag{2}$$
$$M_C(F) = \sigma(f_{1\times 1}(Z_G))$$

where, σ represents the Sigmoid function, and the operation denotes 1D convolution. $M_C(F)$ captures the significance of each feature across the channels. By applying $M_C(F)$ elementwise to the input feature F, we obtain the channel-wise attention-modulated feature F_C.

$$F_C = M_C(F) \bigodot F \tag{3}$$

Precise segmentation of lung nodule boundaries necessitates the extraction of high-resolution spatial features. To tackle this, a spatial attention module is employed to extract high-resolution spatial features, as depicted in Figure 4(b). In contrast to channel attention, which highlights the relevance of specific features, spatial attention focuses on pinpointing spatial regions with crucial information thereby complementing the channel attention mechanism [42]. To assess spatial attention, we begin by performing max-pooling operations along the channel dimension of F_C to produce a compact feature descriptor Z_M. In this formula, we assume the pooling window has a size of $k \times k$, and the pooling operation has a stride of s. It has been demonstrated that performing pooling operations along the channel axis effectively emphasizes the informative regions [43]. After that, one 3×3 dilated convolution is applied to utilizing contextual information effectively. Finally, the attention map $M_S(F_C)$ is generated using the ReLU activation followed by the Sigmoid function. In summary, spatial attention is expressed as Formula (4).

$$Z_M = \max 0 \le m < k,\ 0 \le n < kF_C(s\cdot i + m, s\cdot j + n) \tag{4}$$
$$M_S(F_C) = \delta(\mathtt{ReLU}(f_{3\times 3}(Z_M)))$$

where, the matrix $M_S(F_C)$ captures the significance of the features at each position (i,j). Subsequently, the final feature F' is then derived by multiplying F with $M_S(F_C)$.

$$F' = M_S(F_C) \bigodot F \tag{5}$$

In summary, channel and spatial attention is complementary, focusing on "what" and "where", respectively. Integrating the proposed ICS-Attention module into the model effectively emphasizes lung nodule features while attenuating surrounding features.

2.3. **ICS-BiConvLSTM fusion module.** The incorporation of the ICS-BiConvLSTM module significantly bridges the semantic disparity between the encoder and decoder. The ICS-Attention module enhances feature selection by emphasizing key features, while BiConvLSTM improves the capture of local features [44], achieving a balance between global and local information. As shown in Figure 5, the BiConvLSTM is fed with multi-level feature maps that merge the high-resolution features F' refined by the ICS-Attention module and the upsampled decoder features d_t. The BiConvLSTM module enhances decision-making by analyzing the bidirectional dependencies of input data. Therefore, using BiConvLSTM to combine the two aforementioned features not only captures local details but also strengthens feature representation across both temporal and spatial dimensions.

In particular, the bidirectional ConvLSTM treats the encoder outputs and the upsampled decoder features as a sequential input, where the forward ConvLSTM propagates high-level semantic context from the encoder to guide decoder reconstruction, while the backward ConvLSTM feeds refined high-resolution spatial details back to regularize semantic representations. Through the gated state transitions, each hidden state integrates

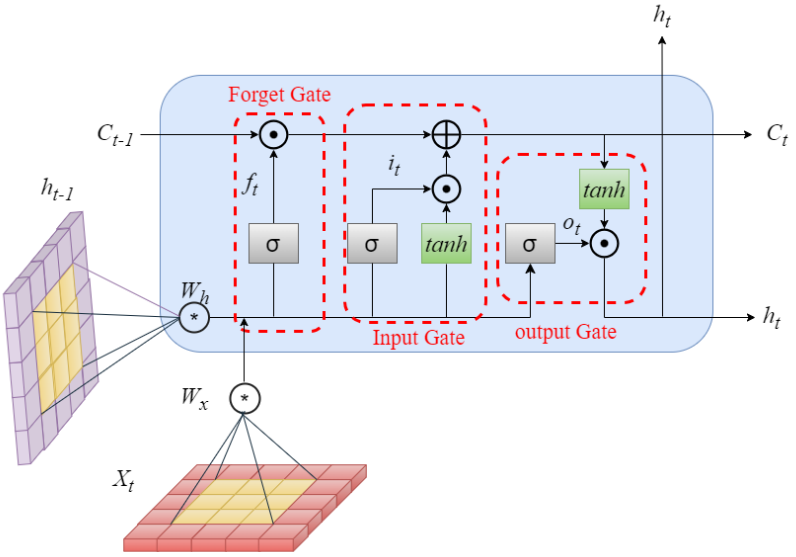


FIGURE 5. The structure of ConvLSTM unit

information from both directions, enabling the model to align high-level semantics with fine-grained spatial structures. This bidirectional information propagation effectively narrows the semantic gap between the encoder and decoder, ensuring that the fused features are both semantically consistent and spatially precise.

The BiConvLSTM model comprises a memory unit (C_t) and three gates (input gate (i_t), forget gate (f_t), and output gate (o_t)), including a forward ConvLSTM and a backward ConvLSTM [45]. The ConvLSTM module performs state transitions through convolution operations, effectively preserving spatial information. The ConvLSTM architecture, as illustrated in Figure 5, can be mathematically expressed as follows:

$$i_t = \sigma\left(W_{xi} * X_t + W_{hi} * h_{t-1} + W_{ci} \bigodot C_{t-1} + b_i\right) \tag{6}$$

$$f_t = \sigma\left(W_{xf} * X_t + W_{hf} * h_{t-1} + W_{cf} \bigodot C_{t-1} + b_f\right) \tag{7}$$

$$C_t = f_t \bigodot C_{t-1} + i_t \bigodot \tanh\left(W_{xc} * X_t + W_{hc} * h_{t-1} + b_c\right) \tag{8}$$

$$o_t = \sigma\left(W_{xo} * X_t + W_{ho} * h_{t-1} + W_{co} \bigodot C_{t-1} + b_o\right) \tag{9}$$

$$h_t = o_t * \tanh(C_t) \tag{10}$$

where, $*$ and $\bigodot$ represent convolution operation and Hadamard product, respectively. X_t denotes the input tensor, h_t denotes the hidden state tensor, C_t denotes the memory cell tensor, W_{x*} and W_{h*} are the learnable weight matrices for each gate. b_i, b_f, b_c and b_o are four offset terms.

Therefore, the output of the BiConvLSTM for the two sets of single-directional state variables (forward and backward states) can be represented as

$$y_t = \tanh\left(W_y \overrightarrow{h} * \overrightarrow{h_t} + W_y \overleftarrow{h} * \overleftarrow{h_t} + b\right) \tag{11}$$

where, $\overrightarrow{h}$ and $\overleftarrow{h}$ denote the forward and backward internal state tensors, respectively. b is the offset term, and y_t is the resulting tensor of the BiConvLSTM. The Tanh function facilitates the smoothing of forward and backward state outputs via nonlinear transformation, enabling the model to effectively capture intricate spatiotemporal dependencies and achieve a balanced representation of both local and global features.

2.4. **Image auxiliary process.** To enhance the network's focus on key features, we redefined the three channels of the CT images to highlight the essential information required by the network, ensuring compatibility with the masks and improving training efficiency. The process of image transformation is outlined as follows.

Image Cropping. The original lung CT images are sized at 512 × 512 pixels, with the pulmonary nodules occupying only a limited region of the image. This results in a significant imbalance between the number of pixels representing the nodule class and those representing the background class during segmentation, which adversely affects the training performance of neural networks [46]. Moreover, using the entire image as input can slow down the training speed. To address this, the study uses lung nodule position annotations from XML files provided by radiologists to perform central cropping on the original CT images, resizing them to 64 × 64 pixels, as illustrated in Figure 6. During the cropping process, the cropping rule strictly follows the principle of centering on the lung nodule. If the size is not allowed, spatial padding is applied to ensuring that the cropped image's size remains consistent with the center of the lung nodule, thereby ensuring the complete preservation of the lung nodule region. The left panel displays the original 512 × 512 CT scan, while the right side shows the zoomed-in 64 × 64 CT scan focused on the pulmonary nodule. Output images are incorporated into the first channel of the new three-channel image, as shown in Figure 7(a). Similarly, the lung nodule masks undergo the same operation, cropped to 64 × 64 pixels and placed in the third channel, as shown in Figure 7(c). Establishing a direct relationship between the input images and ground truth helps emphasize key features, thereby improving segmentation accuracy.

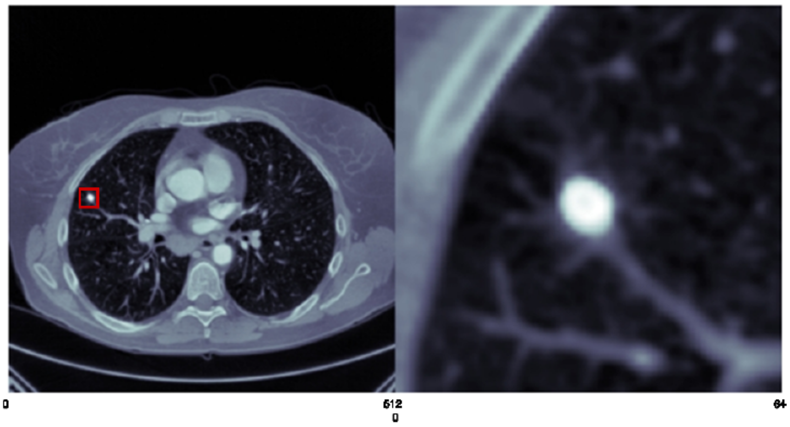


FIGURE 6. CT image cropping

Edge Detection. Edges are important features in images, and edge detection can enhance the details in medical images, making the boundaries between organs and the background clearer, while also improving image quality and reducing noise. The Canny operator [47] achieves robust edge extraction through a multi-step procedure, which involves Gaussian smoothing, gradient computation, non-maximum suppression, double thresholding, and edge linking, effectively handling variations in image intensity and lighting conditions. Therefore, this study employs the Canny operator for precise delineation of lung nodule edges, as shown in Figure 7(b), which is incorporated into the second channel of the new three-channel image.

To summarize, our novel three-channel image comprises the original CT image, a lung nodule edge map, and a lung nodule mask. The resulting three-channel image is colorful as depicted in Figure 7(d). This newly generated three-channel image reduces noise while enhancing the clarity of the key details of the pulmonary nodules.

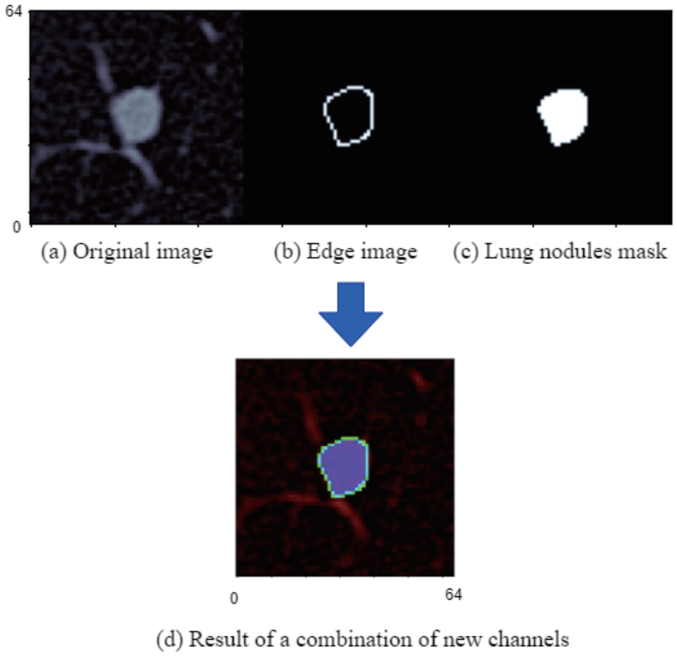


FIGURE 7. New three-channel CT image

3. **Experimental Results.** This study utilizes the Lung Nodule Analysis 2016 (LUNA16) [48] dataset, a subset of the LIDC-IDRI dataset [49]. The LIDC-IDRI dataset comprises 1018 cases, with each CT scan annotated by four radiologists, and the annotation information is stored in XML files. However, the dataset is heterogeneous, containing not only a wide variety of CT images with different doses but also numerous academic terms. This heterogeneity makes it less than ideal for use in the field of pulmonary nodule segmentation. To make the data more accessible, the organizers of LUNA16 removed data with scan spacing greater than 3mm and excluded discontinuous CT scan data based on the original LIDC-IDRI dataset. The data are saved in Meta Image (.mdh) format, rather than the standard DICOM format, and each CT image has a resolution of 512×512 with three channels.

The ground truth of the CT scans is obtained from the annotation information stored in the files. Initially, derive the gold standard masks by utilizing the position and radius information of lung nodules annotated in the .mdh files. Secondly, perform a dilation operation on gold standard masks and original images. Finally, cluster the pixels and obtain the mean of the cluster centers as the threshold for binary lung nodule masks. In Figure 8, the first image represents the original CT scan, the middle image shows the gold standard label, and the right one displays the lung nodule mask.

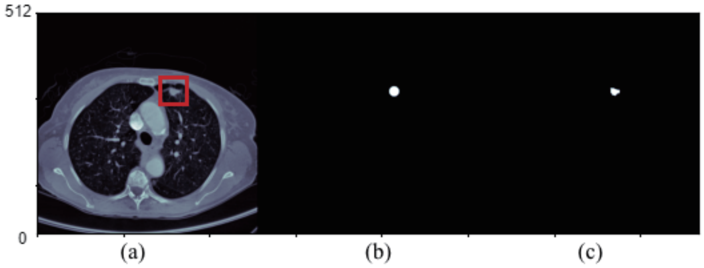


FIGURE 8. (a) Original CT image; (b) gold standard label; (c) lung nodule mask

After extracting the ground truth, we optimized the input raw images through preprocessing to boost the deep learning network's training performance.

The proposed RICSBU-Net was evaluated on the LUNA16 dataset, comprising 1,186 annotated lung nodules across 888 CT scans. The ground truth (GT) was generated as described previously. The evaluation process was carried out using the Torch framework, with the network trained for 200 epochs and a batch size of 16. Additionally, assess the performance of the proposed ICS-Attention module on the CIFAR-10 dataset, consisting of 6,000 images across ten categories, each with a resolution of 32 × 32 pixels. Finally, 5,000 images were used for training and 1,000 for testing.

A compound loss function is employed to monitor the progress of model in minimizing prediction errors throughout the optimization process. BCEDiceLoss merges binary cross-entropy loss with Dice loss with weights set to $[0.5, 1]$. It is defined as

$$BCEDice = 0.5 * BCELoss + DiceLoss \tag{12}$$

$$BCELoss = \frac{-1}{N} + \sum_{i=1}^{N} y_i * \log(P(y_i)) + (1 - y_i) * \log(1 - P(y_i)) \tag{13}$$

where y_i represents the predicted class label, and $P(y_i)$ represents the predicted probability. The BCE loss calculates the disparity between the predicted probability $P(y_i)$ and the true class y_i for each pixel in the segmentation mask. If the predicted probability closely matches the true label, the BCE loss for that pixel is minimized; conversely, a larger discrepancy leads to a higher loss, indicating greater prediction error. During training, the BCE loss is backpropagated to iteratively refine the model's parameters, reducing the pixel-wise loss in the segmentation mask, which is critical for accurately delineating small and complex lesion boundaries, quantifying the overlap between the predicted segmentation and the GT.

3.1. Evaluation metrics.

The following evaluation metrics were employed to assess the segmentation performance of lung nodules: Dice, Sensitivity (SEN), Specificity (SVD), and Intersection over Union (IoU). SEN measures the model's capacity to accurately detect positive cases, SVD reflects the model's effectiveness in identifying negatives, and IoU evaluates the precision and effectiveness of the model in segmentation tasks.

$$DiceLoss = 1 - Dice \tag{14}$$

$$Dice = \frac{2 * TP}{2 * TP + FP + FN} \tag{15}$$

$$SEN = \frac{TN}{TN + FP} \tag{16}$$

$$SVD = \frac{2 * TP}{2 * TP + FP + FN} \tag{17}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{18}$$

Here, *TP* represents true positives, *FP* represents false positives, *FN* represents false negatives, and *TN* refers to true negatives in both the actual and segmented masks of lung nodules.

3.2. Results.

3.2.1. *ICS-Attention evaluation.*

Two independent experiments were conducted on the CLAIR-F10 dataset: the first evaluated the performance of ICS-Attention relative to SE-Net, and the second assessed the performance differences between ICS-Attention and

CBAM. As shown in Table 1, ICS-Attention demonstrated superior accuracy on ResNet-18 compared to SE-Net. A comparative analysis between ICS-Attention and CBAM, both lightweight and easily integrable modules, was conducted under two scenarios: in Scenario 1, the attention module is incorporated into every block, while in Scenario 2, it is applied only before the classifier. Figure 9 and Figure 10 depict the training curves for CBAM and ICS-Attention under these respective scenarios. The findings indicate that deploying ICS-Attention solely before the classifier results in a more rapid reduction in the loss function and lower loss values per epoch. The model was trained for 15 epochs using a batch size of 4. Table 2 further highlights that incorporating the attention module in every block yielded suboptimal results, whereas ICS-Attention significantly outperformed CBAM in accuracy, demonstrating superior performance.

Table 1. Performance of SE-Net and ICS-Attention

Method	ResNet	Accuracy (%)
SE-Net	ResNet	92.4
ICS-Attention	ResNet	92.8

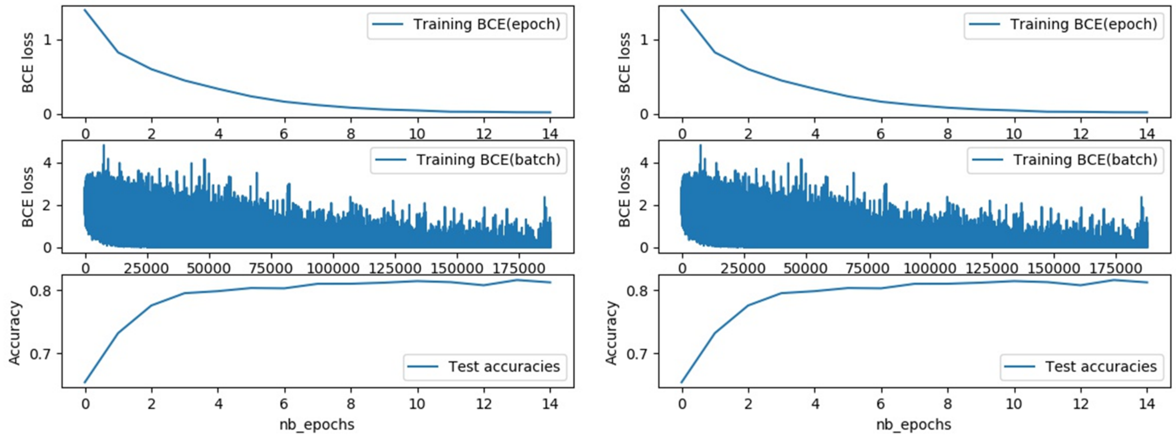


Figure 9. Training curve of attention module added in every block (The right graph is CBAM and the left is ICS-Attention)

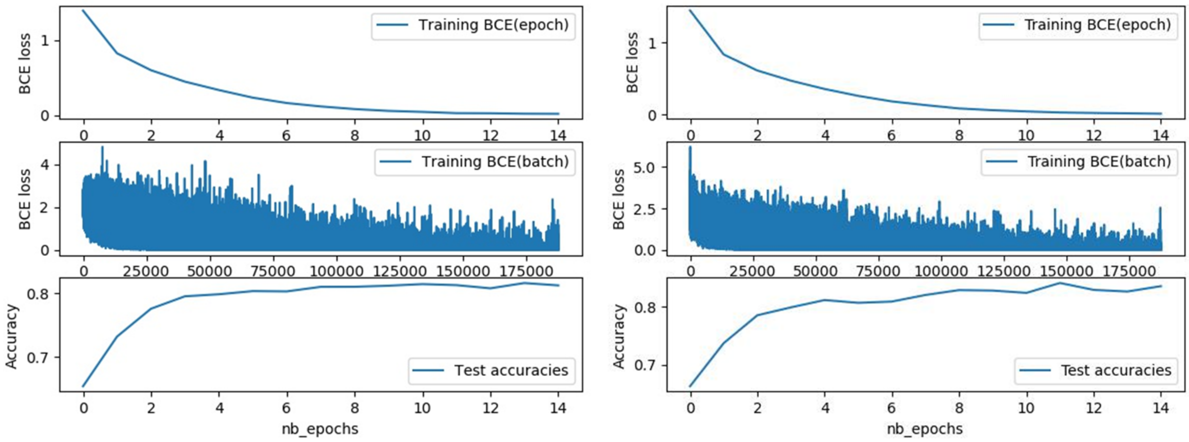


Figure 10. Training curve of attention module adds only before the classifier (The right graph is CBAM and the left is ICS-Attention)

TABLE 2. Comparison of CBAM and ICS-Attention module

Model	Test accuracy (%)
ResNet-18	84.36
ResNet-18 + CBAM Block	82.06
ResNet-18 + CBAM Classifier	83.35
ResNet-18 + ICS-A Block	83.25
ResNet-18 + ICS-A Classifier	85.09

TABLE 3. Performance of SE-Net and ICS-Attention

	Average loss	Dice	SEN	SVD	IoU
Original CT	0.0985	0.9105	0.9251	0.9965	0.8407
Three-Channel CT	0.0419	0.9584	0.9841	0.9983	0.9217

3.2.2. *Lung nodule segmentation.* The proposed RICSBU-Net model's performance is benchmarked against other networks using several metrics, including the Dice, average loss, sensitivity (SEN), specificity (SVD), and IoU. Before conducting a comparative performance analysis with other models, we first assess the performance of RICSBU-Net on two different input sets: the original CT images without preprocessing and the newly generated three-channel images, as shown in Table 3. The results indicate that the RICSBU-Net model achieves substantially better performance when using the enhanced three-channel CT images as input.

To ensure fairness in all comparative experiments, all models – including U-Net++, ResUNet, and the proposed RICSBU-Net – are trained under the same settings. Specifically, the batch size is set to 16, the initial learning rate is 1×10^{-3}, and the training process runs for 200 epochs. The SGD optimizer is used, and the learning rate is adjusted using a cosine annealing scheduling strategy to achieve a smoother and more stable optimization process. In addition, an early stopping mechanism is applied, where training is terminated if the validation performance does not improve for 7 consecutive epochs. These unified configurations ensure that the performance comparison across different models is fair, reliable, and reproducible.

The Dice coefficient shows a marked improvement of 0.0479, while the IoU demonstrates a substantial increase of 0.081. These enhanced images establish a direct correlation between input and ground truth images, effectively eliminating extraneous information from the original images. This refinement significantly boosts the model's capability to identify and learn essential features of Pulmonary nodules with greater accuracy and efficiency. In Figure 11, the qualitative comparison of the original CT image and the new three-channel CT images for lung nodule segmentation using the RICSBU-Net module shows that the proposed three-channel images yield significantly more accurate segmentation results than the original CT images. The figure illustrates the segmentation outcomes for several representative nodules, revealing that the predicted results from the new three-channel CT images achieve highly accurate segmentation. These images exhibit superior handling of fine details and edges, closely aligning with the true labels. To validate the effectiveness and robustness of our method, a detailed comparative evaluation was conducted against several existing models. In this process, consistent evaluation metrics were applied, and the same dataset was used across different medical image segmentation models (Note: the RICSBU-Net model uses the novel three-channel CT images proposed in this paper as input). Multiple performance metrics, including average loss, DICE coefficient, sensitivity

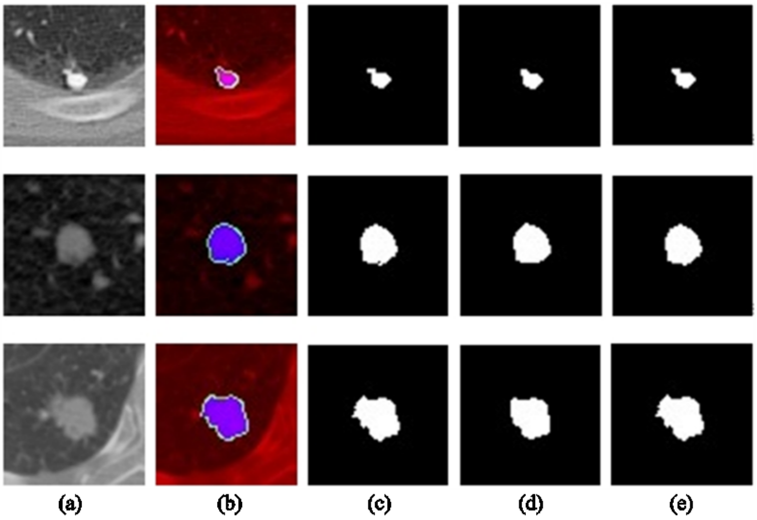


Figure 11. Qualitative comparison of pulmonary nodule segmentation in different CT scans using the RICSBU-Net module: (a) Original CT image, (b) new three-channel CT image, (c) mask, (d) predicted image of original CT image, and (e) predicted image of new three-channel CT image

Table 4. Performance of SE-Net and ICS-Attention

Model	Average loss	Dice	SEN	SVD	IoU
U-Net	0.2250	0.7941	0.9302	0.9812	0.6643
ResUNet	0.1755	0.8326	0.9745	0.9887	0.7172
SegNet	0.1446	0.8782	0.8661	0.9961	0.7931
BCDU-Net	0.1463	0.8821	0.8979	0.9962	0.8000
Res34UNet++	0.1432	0.8838	0.8752	0.9966	0.8050
RICSBU-Net	0.0419	0.9584	0.9841	0.9983	0.9217

(SEN), specificity (SVD), and IoU, were employed during the evaluation. This comparative analysis aims to highlight the exceptional performance of our method, as shown in Table 4. The results clearly demonstrate that our model has made significant progress in lung nodule segmentation, outperforming other methods across all five evaluation metrics.

It is noteworthy that the SVD value of RICSBU-Net (99.83%) is close to 1, which might raise concerns about potential biases, especially in cases with high background pixel ratios, possibly leading to inflated specificity. To mitigate the impact of background pixels on the results, we implemented targeted measures in the data processing stage, such as cropping and region focusing techniques, ensuring that the model primarily concentrates on the nodule regions and minimizing interference from background pixels. Although the SVD value is close to 1, this does not indicate overfitting or bias in the model. A comprehensive comparison (as shown in Table 4) was also performed using other evaluation metrics (such as Dice coefficient, IoU, and sensitivity), and the results show that RICSBU-Net outperforms other existing methods in lung nodule segmentation, thereby validating its high efficiency in nodule detection. Therefore, the high SVD value is a result of the model's effective prediction in the background areas and does not affect its performance in nodule identification.

The visualization in Figure 12 presents segmentation results from five different networks. The first column showcases original lung nodule CT images used as input to the models. In the second column, ground truth lung nodule contours annotated by medical professionals

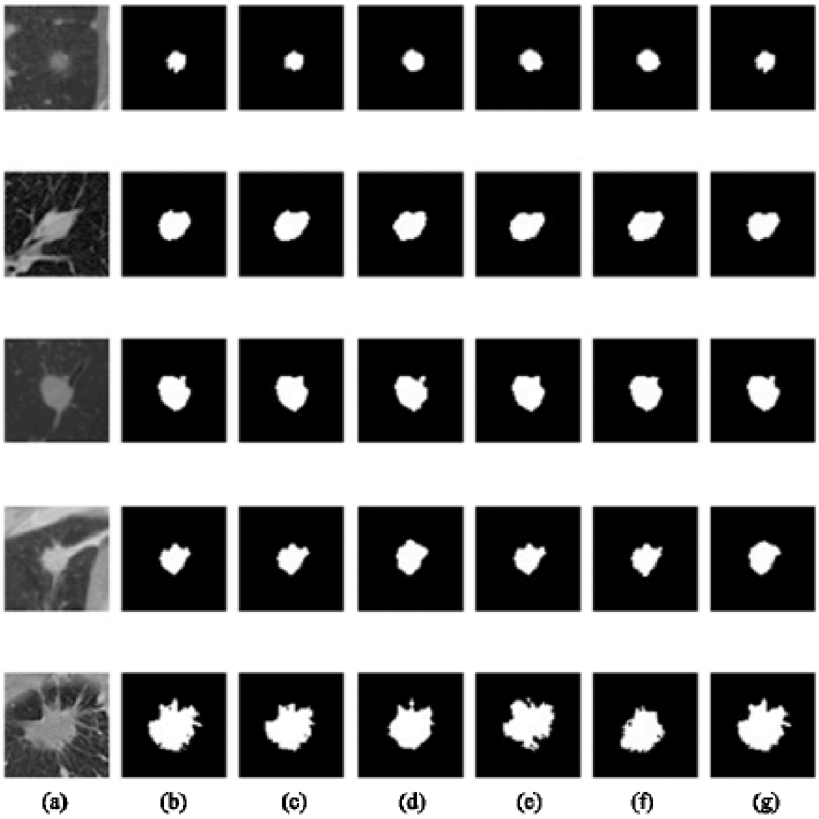


FIGURE 12. Qualitative assessment of different models for pulmonary nodule segmentation: (a) Original CT image, (b) true label, (c) U-Net, (d) SegNet, (e) ResUNet, (f) BCDU-Net, and (g) RICSBU-Net

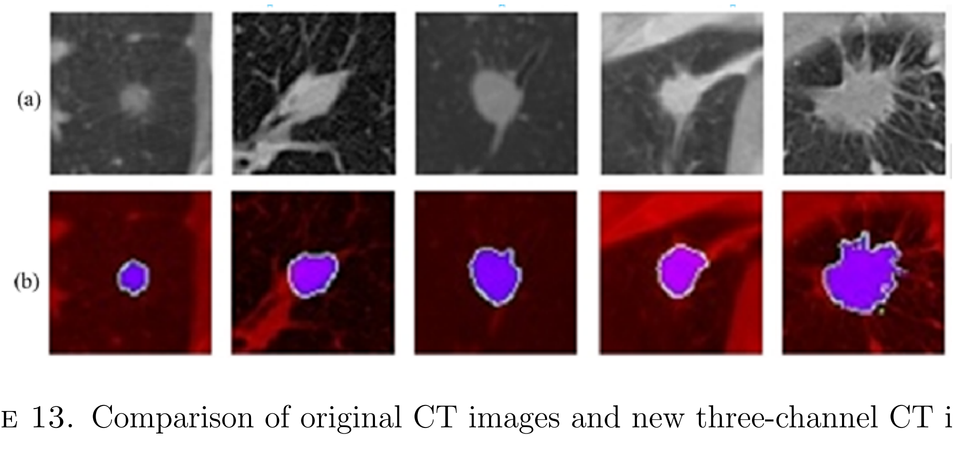


FIGURE 13. Comparison of original CT images and new three-channel CT images

are displayed. The final column exhibits the lung nodule segmentation outcomes achieved using the RICSBU-Net algorithm proposed in this study.

Compared to the RICSBU-Net network, other networks exhibit a reduced ability to extract edge features of pulmonary nodules, resulting in issues such as missed segmentations and excessive cutting. Especially in scenarios involving smaller lung nodules, the algorithm proposed in this study demonstrates superior capability in accurately delineating the edges of the lung nodules. Figure 13 presents the new three-channel CT images corresponding to the original CT images. It is evident that the processed CT images, compared to the unprocessed ones, reduce the interference of irrelevant features and noise, making the edges of the lung nodules more distinct.

In summary, the RICSBU-Net model demonstrates high precision and reliability across diverse lung nodule segmentation scenarios, substantiated by both quantitative and visual analyses, establishing its superiority over other leading-edge methods.

4. Conclusion and Future Work.

4.1. Conclusion. This research presents an innovative and efficient 2D Pulmonary nodule segmentation network, RICSBU-Net, built upon the U-Net architecture. The outlined strategy comprises three key steps. First, we preprocess the original CT images using center-cropping techniques to resize the images to 64×64 pixels, thereby balancing the pixel ratio between the segmentation target and the background. This process generates new three-channel images associated with the mask, effectively removing irrelevant features and noise while highlighting key characteristics of the lung nodules, improving the network's proficiency in detecting small nodules. Second, in response to the heterogeneity of Pulmonary nodules, we present a streamlined ICS-Attention module. This module leverages both meaningful semantic information and fine-grained spatial features, and, when integrated with the BConvLSTM module, replaces the traditional skip connections and concatenation operations in U-Net. This integration not only highlights key features and reduces redundant parameters but also effectively preserves surface-level features such as the edge details of the lung nodules, thus further balancing the extraction of local and global features and improving segmentation accuracy. Finally, the network utilizes a pre-trained ResNet-18 encoder, which prevents the loss of shallow features typically caused by pooling layers in traditional U-Net architectures, while also deepening the network structure to avoid vanishing gradients. The pre-trained encoder also significantly accelerates the training process and improves network efficiency. Comprehensive experiments on the LUNA16 dataset demonstrate the superior performance of RICSBU-Net across multiple segmentation evaluation metrics, particularly excelling in boundary precision and segmentation continuity for small nodules, surpassing existing methods in these areas.

4.2. Future work. In the future, we plan to further enhance the performance and applicability of RICSBU-Net in lung nodule segmentation in the following key areas. First, although RICSBU-Net performs excellently in 2D segmentation, we aim to extend it to 3D CT images to better handle volumetric data and improve segmentation accuracy in 3D space, particularly for small or peripheral nodules. Second, lung nodule segmentation can be enhanced by integrating multimodal data, such as PET and CT images, to capture more comprehensive features of the nodules. This fusion can help overcome some limitations posed by single-modality CT scans. Finally, we will focus on testing the model's generalization ability across different datasets, particularly those with variations in imaging protocols and quality, to ensure robustness in real-world applications. By addressing these issues, we aim to improve the robustness, scalability, and adaptability of the model, further advancing the automation of lung nodule detection and cancer diagnosis.

Acknowledgment. This work is partially supported by the China Scholarship Council. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] R. L. Siegel, A. N. Giaquinto and A. Jemal, Cancer statistics, 2024, *CA: A Cancer Journal for Clinicians*, vol.74, no.1, 2024.

[2] L. Zhi, W. Jiang, S. Zhang and T. Zhou, Deep neural network pulmonary nodule segmentation methods for CT images: Literature review and experimental comparisons, *Computers in Biology and Medicine*, 107321, 2023.

[3] S. Kadry, E. Herrera-Viedma, R. G. Crespo, S. Krishnamoorthy and V. Rajinikanth, Automatic detection of lung nodule in CT scan slices using CNN segmentation schemes: A study, *Procedia Computer Science*, vol.218, pp.2786-2794, 2023.

[4] W. Hendrix, M. Rutten, N. Hendrix, B. van Ginneken, C. Schaefer-Prokop, E. T. Scholten, M. Prokop and C. Jacobs, Trends in the incidence of pulmonary nodules in chest computed tomography: 10-year results from two Dutch hospitals, *European Radiology*, vol.33, no.11, pp.8279-8288, 2023.

[5] C.-F. J. Kuo, C.-C. Huang, J.-J. Siao, C.-W. Hsieh, V. Q. Huy, K.-H. Ko and H.-H. Hsu, Automatic lung nodule detection system using image processing techniques in computed tomography, *Biomedical Signal Processing and Control*, vol.56, 101659, 2020.

[6] T. Zhou, Y. Dong, H. Lu, X. Zheng, S. Qiu and S. Hou, *APU-Net: An Attention Mechanism Parallel U-Net for Lung Tumor Segmentation*, BioMed Research International, 5303651, 2022.

[7] K. Loverdos, A. Fotiadis, C. Kontogianni, M. Iliopoulou and M. Gaga, Lung nodules: A comprehensive review on current approach and management, *Annals of Thoracic Medicine*, vol.14, no.4, pp.226-238, 2019.

[8] S. P. Sahu, P. Agrawal, N. D. Londhe et al., A new hybrid approach using fuzzy clustering and morphological operations for lung segmentation in thoracic CT images, *Biomedical and Pharmacology Journal*, vol.10, no.4, pp.1949-1961, 2017.

[9] J. Dehmeshki, H. Amin, M. Valdivieso and X. Ye, Segmentation of pulmonary nodules in thoracic CT scans: A region growing approach, *IEEE Transactions on Medical Imaging*, vol.27, no.4, pp.467-480, 2008.

[10] Y. R. Baby and V. K. R. Sumathy, Kernel-based Bayesian clustering of computed tomography images for lung nodule segmentation, *IET Image Processing*, vol.14, no.5, pp.890-900, 2020.

[11] X. Ye, M. Siddique, A. Douiri, G. Beddoe and G. Slabaugh, Shape-based CT lung nodule segmentation using five-dimensional mean shift clustering and MEM with shape information, *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp.482-485, 2009.

[12] S. Wang, A. Jiang, X. Li, Y. Qiu, M. Li and F. Li, DPBET: A dual-path lung nodules segmentation model based on boundary enhancement and hybrid transformer, *Computers in Biology and Medicine*, vol.151, 106330, 2022.

[13] A. Bruzadin, M. Boaventura, M. Colnago, R. G. Negri and W. Casaca, Learning label diffusion maps for semi-automatic segmentation of lung CT images with COVID-19, *Neurocomputing*, vol.522, pp.24-38, 2023.

[14] Y. Cai, Z. Liu, Y. Zhang and Z. Yang, MDFN: A multi-level dynamic fusion network with self-calibrated edge enhancement for lung nodule segmentation, *Biomedical Signal Processing and Control*, vol.87, 105507, 2024.

[15] S. Wang, M. Zhou, O. Gevaert, Z. Tang, D. Dong, Z. Liu and T. Jie, A multi-view deep convolutional neural networks for lung nodule segmentation, *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp.1752-1755, 2017.

[16] J. Long, E. Shelhamer and T. Darrell, Fully convolutional networks for semantic segmentation, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.3431-3440, 2015.

[17] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: The 18th International Conference*, Munich, Germany, pp.234-241, 2015.

[18] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh and J. Liang, UNet++: A nested U-Net architecture for medical image segmentation, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: The 4th International Workshop, DLMIA 2018, and the 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018*, Granada, Spain, pp.3-11, 2018.

[19] H. Huang, R. F. Lv, J. L. Tao et al., CT image lung nodule segmentation algorithm based on improved U-Net++, *Acta Photonica Sinica*, vol.50, no.2, pp.73-83, 2021.

[20] S. A. Agnes, A. A. Solomon and K. Karthick, Wavelet U-Net++ for accurate lung nodule segmentation in CT scans: Improving early detection and diagnosis of lung cancer, *Biomedical Signal Processing and Control*, vol.87, 105509, 2024.

[21] P. Dutande, U. Baid and S. Talbar, LNCDS: A 2D-3D cascaded CNN approach for lung nodule classification, detection and segmentation, *Biomedical Signal Processing and Control*, vol.67, 102527, 2021.

[22] Y. Zhang and X. Ye, Pulmonary nodule image segmentation algorithm based on WU-Net network, *Application Research of Computers*, vol.3, 2022.

[23] Y. Jalali, M. Fateh, M. Rezvani, V. Abolghasemi and M. H. Anisi, ResBCDU-Net: A deep learning framework for lung CT image segmentation, *Sensors*, vol.21, no.1, 268, 2021.

[24] S. Selvadass, P. M. Bruntha, K. M. Sagayam and H. Günerhan, SAtUNet: Series atrous convolution enhanced U-Net for lung nodule segmentation, *International Journal of Imaging Systems and Technology*, vol.34, no.1, e22964, 2024.

[25] M. Maqsood, S. Yasmin, I. Mehmood, M. Bukhari and M. Kim, An efficient DA-Net architecture for lung nodule segmentation, *Mathematics*, vol.9, 1457, 2021.

[26] G. Hou, J. Qin, X. Xiang, Y. Tan and N. N. Xiong, AF-Net: A medical image segmentation network based on attention mechanism and feature fusion, *Computers, Materials & Continua*, vol.69, no.2, pp.1877-1891, 2021.

[27] Z. Zhou, F. Gou, Y. Tan and J. Wu, A cascaded multi-stage framework for automatic detection and segmentation of pulmonary nodules in developing countries, *IEEE Journal of Biomedical and Health Informatics*, vol.26, no.11, pp.5619-5630, 2022.

[28] Y.-S. Huang, P.-R. Chou, H.-M. Chen, Y.-C. Chang and R.-F. Chang, One-stage pulmonary nodule detection using 3-D DCNN with feature fusion and attention mechanism in CT image, *Computer Methods and Programs in Biomedicine*, vol.220, 106786, 2022.

[29] T. Wang, G. Li, C. Tang and T. Kamiya, Res BCDU-SCB Net: A lung CT image segmentation based on SC-BLSTM, *Proc. of the 2023 8th International Conference on Biomedical Signal and Image Processing*, pp.42-46, 2023.

[30] J. Song, C. Yang, L. Fan, K. Wang, F. Yang, S. Liu and J. Tian, Lung lesion extraction using a toboggan based growing automatic segmentation approach, *IEEE Transactions on Medical Imaging*, vol.35, no.1, pp.337-353, 2015.

[31] D. Bhattacharyya, N. T. Rao, E. S. N. Joshua and Y. Hu, A bi-directional deep learning architecture for lung nodule semantic segmentation, *The Visual Computer*, vol.39, no.11, pp.5245-5261, 2023.

[32] K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for image recognition, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.770-778, 2016.

[33] K. He, X. Zhang, S. Ren and J. Sun, Identity mappings in deep residual networks, *Computer Vision – ECCV 2016: The 14th European Conference*, Amsterdam, The Netherlands, pp.630-645, 2016.

[34] F. Gao, L. Su and X. Niu, Research on Mongolian-to-Chinese neural machine translation based on transformer, *Computer Applications and Software*, vol.37, no.2, pp.141-146, 2020.

[35] J. Hu, L. Shen and G. Sun, Squeeze-and-excitation networks, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.7132-7141, 2018.

[36] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu and T.-S. Chua, SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.5659-5667, 2017.

[37] J. Park, BAM: Bottleneck attention module, *arXiv Preprint*, arXiv: 1807.06514, 2018.

[38] S. Woo, J. Park, J.-Y. Lee and I. S. Kweon, CBAM: Convolutional block attention module, *Proc. of the European Conference on Computer Vision (ECCV)*, pp.3-19, 2018.

[39] Y. Liu, Z. Shao, Y. Teng and N. Hoffmann, NAM: Normalization-based attention module, *arXiv Preprint*, arXiv: 2111.12419, 2021.

[40] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo and Q. Hu, ECA-Net: Efficient channel attention for deep convolutional neural networks, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.11534-11542, 2020.

[41] S. Hao, D.-H. Lee and D. Zhao, Sequence to sequence learning with attention mechanism for short-term passenger flow prediction in large-scale metro system, *Transportation Research Part C: Emerging Technologies*, vol.107, pp.287-300, 2019.

[42] S. Zagoruyko and N. Komodakis, Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer, *arXiv Preprint*, arXiv: 1612.03928, 2016.

[43] V. Badrinarayanan, A. Kendall and R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.39, no.12, pp.2481-2495, 2017.

[44] S. Mahmud, N. Ibtehaz, A. Khandakar, M. S. Rahman, A. J. R. Gonzales, T. Rahman, M. S. Hossain, M. S. A. Hossain, M. A. A. Faisal, F. F. Abir et al., NABNet: A nested attention-guided BiConvLSTM network for a robust prediction of blood pressure components from reconstructed arterial blood pressure waveforms using PPG and ECG signals, *Biomedical Signal Processing and Control*, vol.79, 104247, 2023.

[45] X. Shi, Z. Chen, H. Wang, D. Yeung, W. Wong and W. Woo, Convolutional LSTM network: A machine learning approach for precipitation nowcasting, *Advances in Neural Information Processing Systems*, vol.28, 2015.

[46] S. Zhong, X. Guo and Y. Zheng, Improving lung nodule segmentation methods based on the U-Net network, *Computer Engineering and Applications*, vol.56, no.17, pp.203-209, 2020.

[47] J. Li, L. Zhang and W. Zheng, Improved faster R-CNN and adaptive Canny algorithm for defect detection using eddy current thermography, *AIP Advances*, vol.14, no.2, 2024.

[48] A. A. A. Setio, A. Traverso, T. de Bel, M. S. N. Berens, C. van den Bogaard, P. Cerello, H. Chen, Q. Dou, M. E. Fantacci, B. Geurts et al., Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge, *Medical Image Analysis*, vol.42, pp.1-13, 2017.

[49] R. J. Suji, S. S. Bhadouria, J. Dhar and W. W. Godfrey, Optical flow methods for lung nodule segmentation on LIDC-IDRI images, *Journal of Digital Imaging*, vol.33, no.5, pp.1306-1324, 2020.

Author Biography

Tianwen Wang received her master's degree from the School of Electronics and Information Engineering, Tiangong University, China, in 2019 and 2022, respectively.

Ms. Wang is currently pursuing her Doctor degree in the Department of Mechanical and Control Engineering at Kyushu Institute of Technology, Japan. Her research interests focus on the medical application of image analysis. She is currently working on automatic segmentation of CT image.

Guangxu Li received the Ph.D. degree from the Kyushu Institute of Technology, Japan, in 2013.

Dr. Li is currently an Assistant Professor with the School of Optometry & Ophthalmology, Tianjin Medical University Eye Hospital, China, he joined Tiangong University and the Tianjin Key Laboratory of Optoelectronic Detection Technology and System since 2014. From 2020 to 2021, he was a Visiting Scholar with the Department of Bioengineering, University of Washington, America. His research interests include computer aided diagnosis, optometry and computational optical imaging. After joining academia, he became a member of the Council of the Tianjin Stereology Society.

Tohru Kamiya received his B.A. degree in Electrical Engineering from Kyushu Institute of Technology, Japan, in 1994, the master's and Ph.D. degrees from Kyushu Institute of Technology, in 1996 and 2001, respectively.

Dr. Kamiya is a Professor in the Department of Mechanical and Control Engineering at Kyushu Institute of Technology. His research interests focus on the medical application of image analysis. He is currently working on automatic segmentation of multi-organ of abdominal CT image, and temporal subtraction of thoracic MDCT image sets.