IET Image Processing

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

MDR-Net: Multiscale dense residual networks for liver image segmentation

Lijie Xie [ID] | **Fubao Zhu** | **Ni Yao**

School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan, China

Correspondence
Ni Yao, School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan 450001, China.
Email: yaoni@zzuli.edu.cn

Abstract

Liver image segmentation is an attractive topic in the diagnosis and surgical planning of liver diseases. Although deep learning methods have significantly advanced liver segmentation, existing frameworks fail to clearly determine liver boundaries, especially in medical images where various organs have similar grey levels. In this paper, the authors design a multi-scale dense residual network (MDR-Net) for liver segmentation, which consists of two blocks: a liver segmentation network and an edge-aware network. In the segmentation network, the authors introduce a multi-scale residual pooling module combining channel attention (CA) mechanism and depth-wise separable convolution to accommodate liver scale variation. Furthermore, the authors employ an edge-aware loss network to refine edge information and enhance feature representation, which is beneficial to guide the network to iterate towards the ground truth. The authors' method achieves the best visualization results in qualitative evaluation. In addition, the authors' method achieves 96.189% on 3D-IRCADb and 96.889% on the CHAOS dataset in quantitative evaluation with respect to the dice index.

1 | INTRODUCTION

The liver is an important organ of human metabolism, and it is extremely easy to cause liver diseases such as hepatitis B, liver cirrhosis, and tumours [1]. During the growth of liver tumours, patients rarely experience physical discomfort, and liver lesions are difficult to be detected at an early stage. New medical imaging techniques have shown great advancement in computer-aided detection of lesions in the body very early [2]. Therefore, how to completely obtain the liver region from complex abdominal CT images has a guiding role for doctors to diagnose diseases and follow up targeted treatment.

Numerous studies have used different methods to accurately extract the liver from abdominal CT images, liver segmentation as a prerequisite has very important research and clinical significance. At present, the existing liver segmentation methods are summed up as traditional methods and deep learning-based methods. Traditional segmentation methods include threshold method [3], the region growing method [4], and the active contour model [5]. The training model segmentation method based on deep learning mainly uses the analysis of liver image data; it establishes a training model to achieve the optimal solution through repeated training. The more mainstream segmentation network models are CNN [6], FCN [7], and U-net [8]. On the basis of these three networks, a variety of different network models have evolved, for example ResUNet [9], FCN-VGG16 [10], R2U-net [11], and U-Net++ [12]. The current deep learning method has been used as the main trend method in the field of liver segmentation (see Figure 1). Although the accuracy of these methods is constantly being improved, they have not achieved satisfactory results. The liver segmentation faces many challenges, such as (1) the difference between the grey value of the liver region and the pixel value between adjacent tissues and organs is not obvious; (2) the liver region is not fixed, and its size varies from person to person; and (3) the position of the liver varies greatly between adjacent slices and cannot be fully processed.

To address the above problems, we explore an efficient and effective method for liver segmentation. First, we exploit the channel information of liver images for region matching to capture the dependences between different channels,

| ResUNet [9] | FCN-VGG16 [10] | R2U-net[11] | U-Net++[12] | DDA-net[29] |
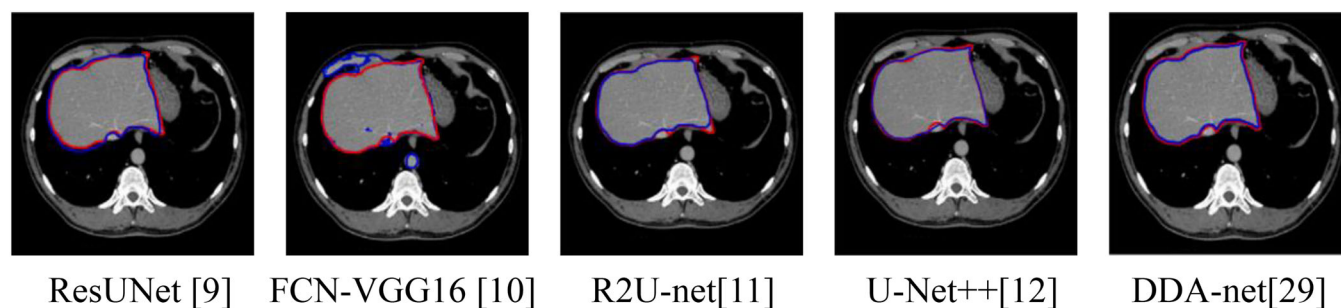
FIGURE 1 Liver segmentation results for different models.

and then propose layer-normalized integrated channel attention (CA) blocks to enhance feature representation. To combine the feature information of diverse scales, a multi-scale dense residual block is designed to obtain the liver information in the image. Finally, we utilize a boundary-aware loss network to reduce the loss of edge information. Our method can automatically segment the liver region and effectively solve the over-segmentation and under-segmentation of the liver.

The main contributions of this paper are as follows:

1. We propose a novel channel representation learning method combining CA mechanism and layer normalization scaling factor to enhance feature information, which can capture the dependences and valuable features from different channels.
2. A multi-scale dense residual module is proposed, which can automatically adapt to changes in liver scale and location.
3. We design an edge-aware loss network to ensure the similarity of edge features between predicted images and ground truth, which effectively solves over- and under-segmentation.
4. Qualitative and quantitative method evaluations are performed on multiple datasets; our method exhibits superiority compared to state-of-the-art methods.

The main work of this paper is as follows: Section 1 mainly summarizes the research status of the field of liver segmentation and the methods and contributions of this paper; Section 2 describes the related work; the proposed method is introduced in Section 3; Section 4 is our experimental part, including the experimental setting, qualitative and quantitative analysis; Section 5 draws the conclusion.

2 | RELATED WORK

The main purpose of liver image segmentation is to completely obtain the liver region from the abdominal CT image, and maintain a clear target edge, which is conducive to the subsequent 3D reconstruction. In this section, we classify the existing liver segmentation methods into model-driven and data-driven frameworks.

2.1 | Model-driven-based method

In the early stages, liver segmentation methods mainly rely on prior assumptions as model-driven algorithms, and they usually reinforce the prior knowledge of target and background regions. Such methods require human–computer interaction and set prior values based on experience to adapt to different methods and results. Siri et al. [13] determined multiple thresholds for liver segmentation from histogram slope profiles. Arica et al. [14] selected a seed point in each selected region for region growth, and then merged the liver regions in combination with region binding to obtain the complete liver region. Reddy et al. [15] split the input image into multiple subdomains, and defined seed points for each grid for regional growth, but over-segmentation and under-segmentation occurred. Liao et al. [16] combined principal component analysis and graph cutting to eliminate complex backgrounds for liver segmentation. However, such methods require human intervention (such as setting seed points, initial contour models, etc.).

2.2 | Data-driven-based method

The neural network introduces spatial information, which solves the problem of unbalanced pixel distribution and noise in the image, and is very suitable for segmenting abdominal CT images. Xie et al. [17] designed a context-aware network with dual-stream pyramid for medical image segmentation. Budak et al. [18] trained dual-branch encoder–decoder networks (EDCNN) for liver and cancer segmentation. Sarmadi et al. [19] adopted CNN to automatically segment liver tumour regions from CT images. Liu et al. [20] designed a two-stage encoder–decoder segmentation network based on Unet and graph cut energy functions for liver segmentation. Chen et al. [21] designed a channel u-net network to extract the spatial information between feature maps and pixels, optimized the spatial channel mechanism, and segmented the liver region. They established a non-linear mapping from the input to the output and utilized robust data contrast for liver segmentation. In a data-driven-based method, the residual network replaces the classical convolution operation, which can boost the accuracy by increasing the depth of the network. In the field of liver segmentation, a large number of studies operate using
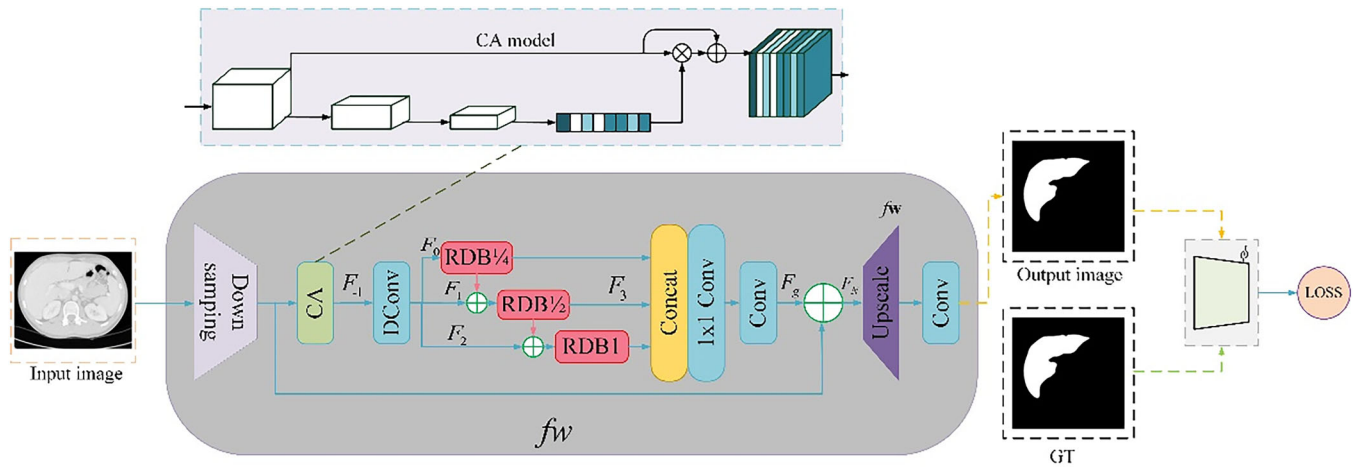
FIGURE 2 Overview of the proposed MDR-Net. MDR-Net, multi-scale dense residual network.

residual networks as the underlying network model. Xie et al. [22] proposed the multi-scale context integrated network for liver CT image segmentation. Bi et al. [9] inferred liver region boundaries by a cascade operation of hybrid fusion, and added residual blocks to obtain boundary parts of healthy and diseased livers. The residual network-based method can achieve fast convergence, and the segmentation accuracy will also be improved.

3 | METHOD

In this section, the proposed MDR-Net network is described in detail. The overall framework is shown in Figure 2.

We first downsample the input CT images, and then combine the CA mechanism and depthwise separable convolution to extract features at different scales (1/4, 1/2, and 1) for the liver region. To extract feature information at different scales, we use dense residual blocks to enhance and fuse the features. Then, the up-sampling is adopted to augment the resolution of the image. Finally, the prediction results are fed into the loss network together with the GT to improve the segmentation accuracy.

3.1 | Network architecture

As shown in Figure 2, our network architecture consists of two modules: a liver segmentation network f_W and an edge-aware loss network ϕ. A liver segmentation network is constructed through several residual blocks, and the input abdominal CT image S is converted into an output liver image O, which is characterized by the mapping $f_W : S \rightarrow O$. Furthermore, skip connections [23] are used to help ease the training of liver segmentation networks. The subsequent loss network takes the obtained liver regions and GTs as input and outputs their feature maps. In particular, our framework employs
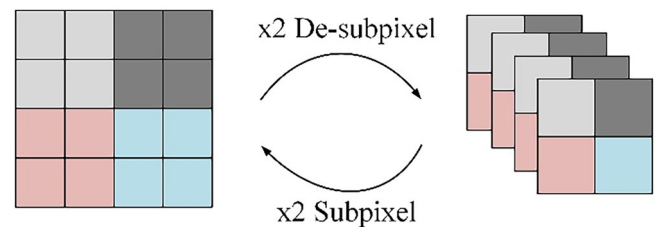
FIGURE 3 Subpixel upsampling and inverse subpixel downsampling.

a shallow feed-forward network instead of the pretrained VGG network.

1. Liver segmentation network: In the liver region feature extraction task, although downsampling is used to speed up the forward propagation of the network, it may also lead to loss of information, especially in the early phase of training. To balance the speed of forward propagation and extract complete feature information, we choose de-depthwise separable convolution [24] instead of traditional convolution operations. In addition, de-subpixeling is adopted instead of the usual pooling operation to restore the resolution, and strided convolution is adopted for downsampling. De-subpixels can reduce the spatial dimension without losing feature information. It rearranges the spatial information into channel information, and finally generates a complete feature map. The specific operation process is shown in Figure 3. Here, we concatenate sub-pixel upsampling with downsampling and produce an identical change.

$$S(V(X)) = X \qquad (1)$$

where S and V denote sub-pixel upsampling and downsampling operations, respectively.
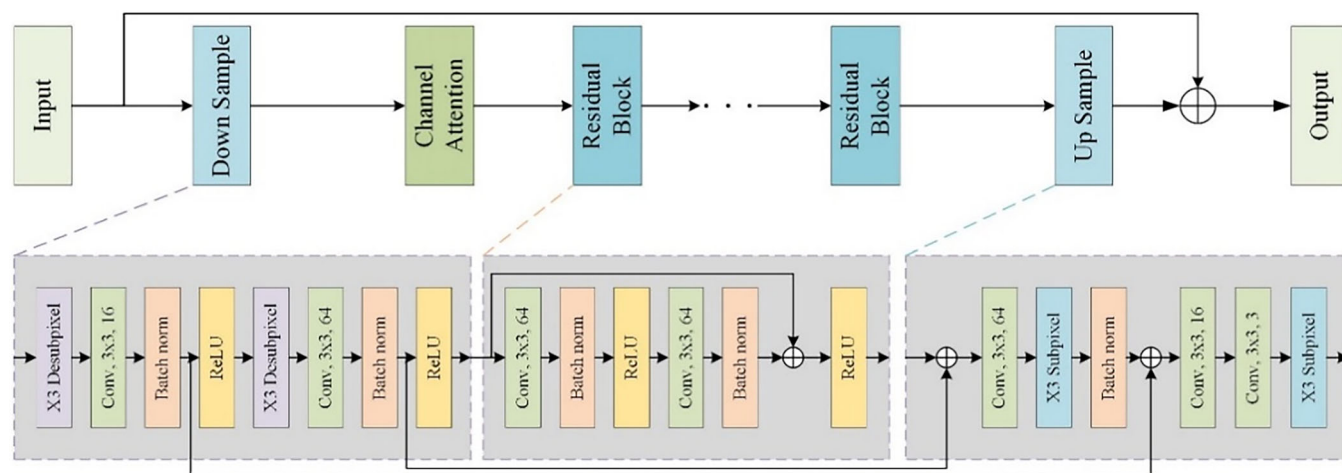
FIGURE 4 Detailed architecture of the liver segmentation network.

2. Edge-aware loss network: The changes of CT values in the liver region and the differences in CT values in adjacent regions were not obvious. However, the feature reconstruction loss detected by the loss network is large, and the accurate edge information is mainly obtained from the segmentation network. Therefore, even an untrained network with specific weights can produce useful features for liver segmentation tasks through some feature learning architecture. This is beneficial to utilize the edge-aware loss network to guide the training of the segmentation network.

In liver CT images, there are smooth regions and high variance regions. The smooth area is relatively smooth, and the difference between the pixel values will be very small, while the fluctuation between the pixel values of the high variance area is very large, and there will be many such high variance areas in the liver due to the lesion. Therefore, in order to obtain high variance regions and capture their texture feature information, edge-aware loss networks play a crucial role. Since the gradients are relatively large in this region, a large edge-aware loss may cause the segmentation network in training to be penalized, which in turn fails to preserve the texture details of the source image.

3.2 | Network details

We take the U-Net network as the basic framework and introduce de-sub-pixel blocks for down-sampling feature extraction. In order to speed up the network convergence, we introduce residual blocks of various scales, which obtain rich features of the liver region by deepening the number of network layers. The network details are shown in Figure 4.

In the downsampling operation, we first use three times the proportion of de-subpixel blocks to operate, and then perform two consecutive 3 × 3 convolutions, batch normalization operations, and ReLU operations. Finally, the obtained

feature information is sent to the residual block as input. Upsampling operation is performed to obtain the original resolution after successive residual block processing. The feature map information obtained by down-sampling and the features obtained by the residual block are jointly sent as input to the up-sampling, which includes three convolution operations and two 3 times inverse sub-pixel block and batch normalization operations.

The CA mechanism can make our network focus on the valuable feature and reduce the overfitting state of the network. CA is essentially a feature-weighted representation that learns the interrelationships between the features. We use CA to extract the global features of the liver part, as shown in Figure 5.

CA can compute the corresponding attention map through down-sampled branches, which is able to aggregate contextual information of the liver in the branch network. The mask is obtained through the sigmoid function, and the normalized information of the obtained attention map is subjected to global average pooling. The original images are multiplied and added to obtain liver feature information.
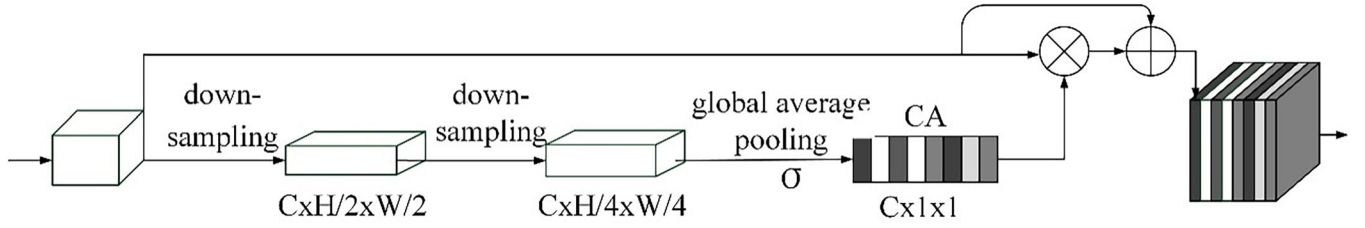
3.3 | Multiscale dense residual networks

Inspired by the classical residual network, we propose a multi-scale dense residual network (MDR-Net) as our residual module, which can take full advantage of features at different scales by extracting the liver region, and has good generalization.

The MDR-Net is a chain process from downsampling to upsampling, which fuses feature information of different scales to obtain liver regions. We downsample the features to different scales and perform feature fusion. The specific network form is as follows:

The dense block preserves the feedforward information and integrates dense features at local and global levels in the liver region. If there are too many network convolution blocks in Figure 6a, it will lead to parameter explosion. Based on this, we introduce a channel compression mechanism into the network,

FIGURE 5 Details of the channel attention module.

FIGURE 6 Network structure: (a) dense block; (b) dense residual network; and (c) our proposed multi-scale dense residual network block.
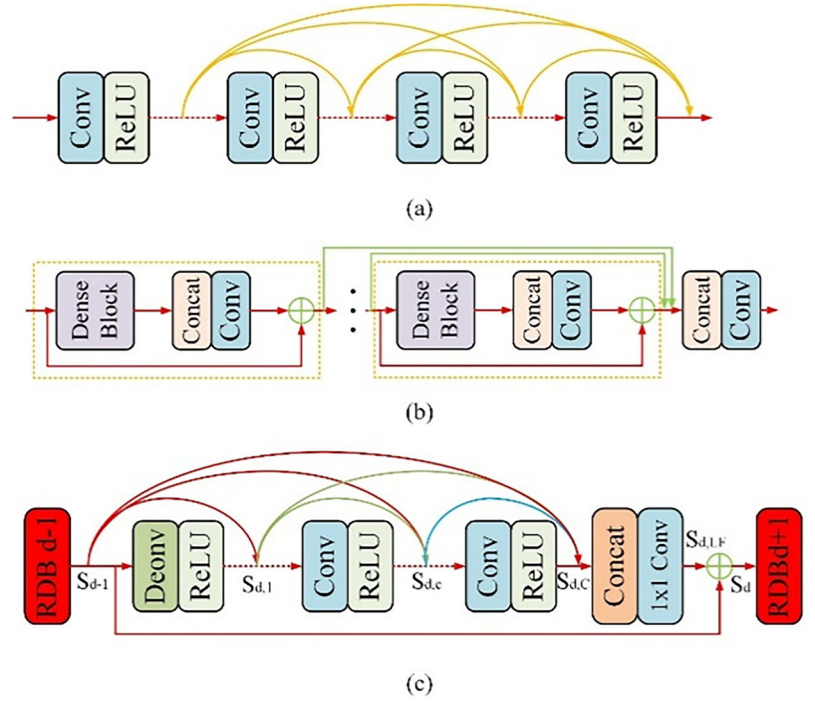
add a 1×1 convolution after each dense block for channel compression, and add skip connections to make its features more obvious. Finally, we downsample to different scale features (1/4, 1/2, and 1), corresponding to our three RDB block information in Figure 2. It can not only aggregate more scale information to address the phenomenon of over- and under-segmentation.

In Figure 6c, we assume that S_{d-1} represents the input of our dth block RDB, and the corresponding S_d is the output. To extract features of different scales, we first restore the image to its original resolution using transposed convolution, so the output of the cth layer can be expressed as

$$S_{d,c} = \sigma(\omega_{d,c}[S_{d-1}, S_{d,1}, \dots, S_{d,c-1}]) \quad (2)$$

where σ is the ReLU activation function, which $\omega_{d,c}$ is the specific weight value of the cth convolutional layer, $[S_{d-1}, S_{d,1}, \dots, S_{d,c-1}]$ representing different maps of the convolution layer.

Then, local feature fusion is used to fuse the feature map of RDB d-1 with all the feature maps of the current RDB. Here, we use the concatenation method and introduce a $1 \times$

1 convolution, which is specifically expressed as follows:

$$S_{d,LF} = S_{LFF}^d([S_{d-1}, F_{d,1}, \dots F_{d,c}, \dots F_{d,C}]) \quad (3)$$

where S_{LFF}^d represents the 1×1 convolution layer in the dth RDB. Finally, the output of the obtained dth RDB block is

$$S_d = S_{d-1} + S_{d,LF} \quad (4)$$

After the feature extraction of dense residual block, the feature map information processed by all dense blocks is fused through 1×1 convolution, and the original resolution of the image is restored by upsampling to obtain the final liver region.

3.4 | Edge-aware network

Boundary losses are applied to segmentation tasks with highly unbalanced samples. Such a loss is a measure of distance on the spatial contour, not a region. In order to make the weight update faster, we use the cross entropy loss function. If the training

error is very large, then the adjustment range of each parameter will also increase, so that the training can quickly converge. In addition, it can avoid gradient dispersion. Hence, we adopt the cross-entropy loss as the loss of our network. The specific expression is as follows:

$$L_c = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \tag{5}$$

where x represents the samples, and n denotes the total number of samples. We choose edge-aware loss as the second auxiliary feature representation to obtain sharp-edged images. It guarantees the similarity between the predicted image and the target image in edge features, which can be specifically represented by the Euclidean distance of multiple feature maps of the neural network layer:

$$Le = \frac{1}{N} ||\phi(y) - \phi(\hat{y})||_F^2 \tag{6}$$

Among them, ϕ represents a network with fixed model parameters. y denotes the GT, while $\hat{y}$ the image of the liver region we generated. Finally our overall loss is

$$L_{total} = L_c + L_e \tag{7}$$

The total loss is to assess the difference between the predicted output and the ground truth, global loss minimization makes the prediction result as close as possible to the actual situation.

4 | EXPERIMENT AND ANALYSIS

4.1 | Datasets

We select a public dataset 3D-IRCADb and CHAOS dataset for qualitative and quantitative analyses.

The 3D-IRCADb dataset contains abdominal CT images of 20 patients with liver disease, 10 males and 10 females. A total of 2823 liver slices were included, and about 75% of them have liver disease. To obtain the basic characteristics of the liver, we used a total of 2782 CT slices for 17 persons as training images, and three persons with a total of 152 slices as test images.

The data set CHAOS contains a total of 6407 liver slices from 40 patients, of which 20 groups of data are labelled and the other 20 groups are unlabelled. Therefore, we select a total of 1872 labelled data for training and 408 for testing. We preprocess the original image with grey-scale mapping and contrast-enhanced histogram modification operations to eliminate the irrelevant tissue in the liver part. The specific details are shown in Table 1. Finally, we use the principle of random allocation to scramble the training set and the test set, and conduct multiple cross-experiment verifications to prevent overfitting in the experiment. The input image resolution is 512×512.

TABLE 1 Display of data augmentation parameters.

| No	Data augmentation	Parameters
1	Horizontal_flip	True
2	Shear	0.12
3	Zoom	0.20
4	Width shift range	0.12
5	Height shift range	0.12
6	Rotation	0-30o

4.2 | Experimental detail

In the testing phase, the proposed MDR-Net model is implemented in Python 3.7 and Tensorflow 1.8, running on a system with NVIDIA GeForce RTX 2060 12GB. The Adam optimizer is used for optimization. The initial learning rate is set to 1e-4. A total of 100 epochs were trained.

4.3 | Qualitative and quantitative analyses

We conduct a comprehensive evaluation on the experimental results. We use deep learning-based and residual network-based comparisons with the proposed methods, in which deep learning methods include FCN-VGG16 [10], Channel-UNet [25], M-SegNet [26], DDANet [27], LadderNet [28] a total of five, and the residual network methods ResUNet [9], R2U-Net [11], U-Net++ [12], CE-Net [28] a total of four kinds. We perform qualitative and quantitative analyses on datasets 3D-IRCADb and CHAOS to verify the superiority and generalization of our method. Figures 7 and 8 show the segmentation results at different scales (large, medium, and small) on two datasets. We can find that the liver segmentation results of different scales segmented by the same segmentation method are quite different from the real labels.

In Figure 7, we present the comparison results in the dataset 3DIRCADb. As can be seen from the figure, all the methods can lock the location of the liver region in the abdominal CT image. We show the test results of the small-scale liver in the first row. Due to the inconspicuousness of the small-scale liver and the uneven distribution of the liver, it is difficult to fully identify the liver region. It can be seen that the M-SegNet, FCN-VGG16, ResUNet, and LadderNet methods cannot accurately determine the exact location of the small-scale liver in the abdominal CT images, which may lead to incorrect segmentation of other tissues. At the same time, although the CE-Net and Channel-UNet methods can identify the liver results, the segmentation results are not complete enough. Although our results have errors in some areas, they are close to the ground truth. In the middle-scale liver segmentation results (second row), the segmentation results of DenseNet, channel-UNet, R2U-Net, ResUNet, and U-Net++ models show over-segmentation and under-segmentation. M-SegNet and the proposed method achieve better segmentation results. In the large-scale liver seg-
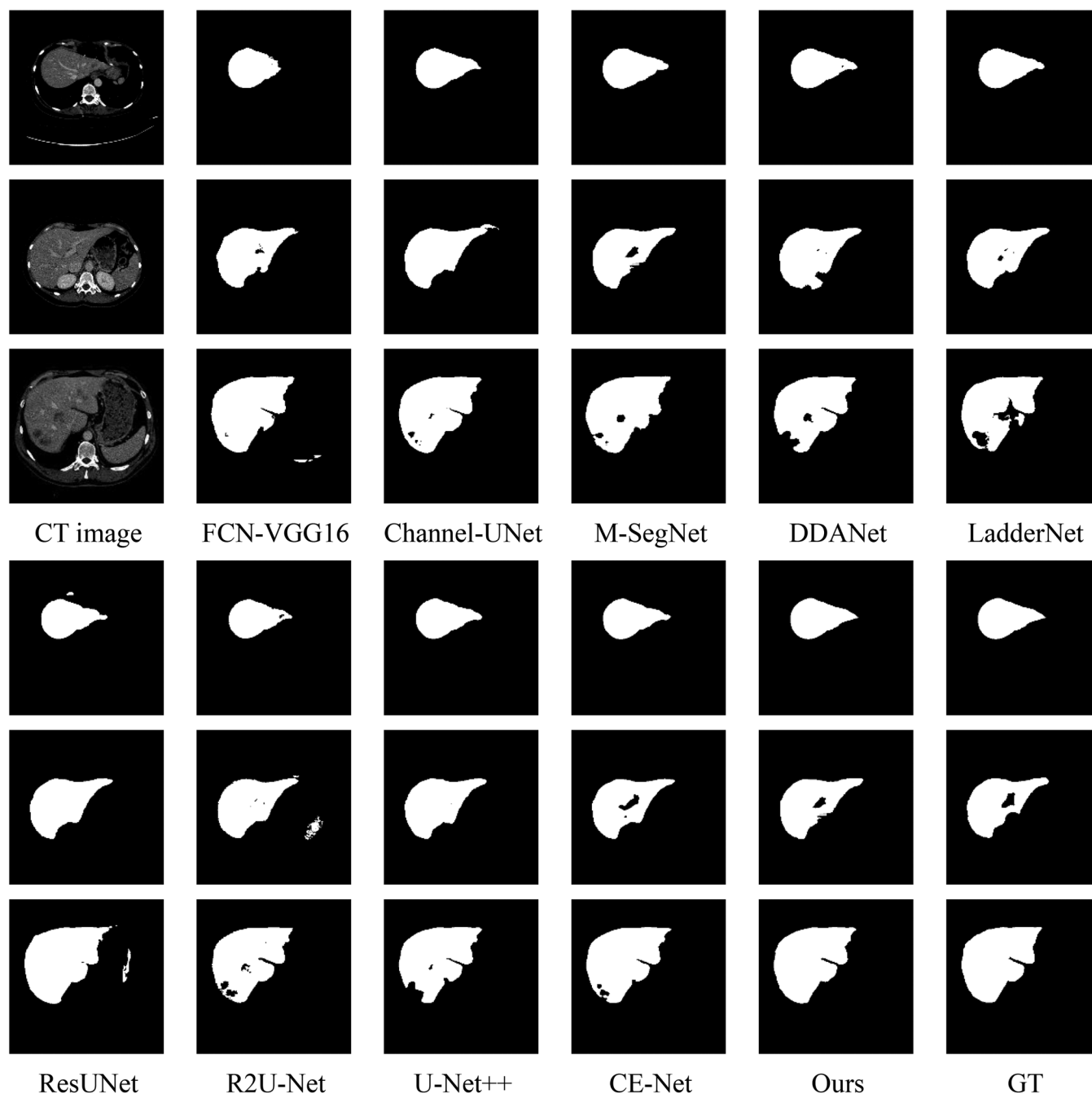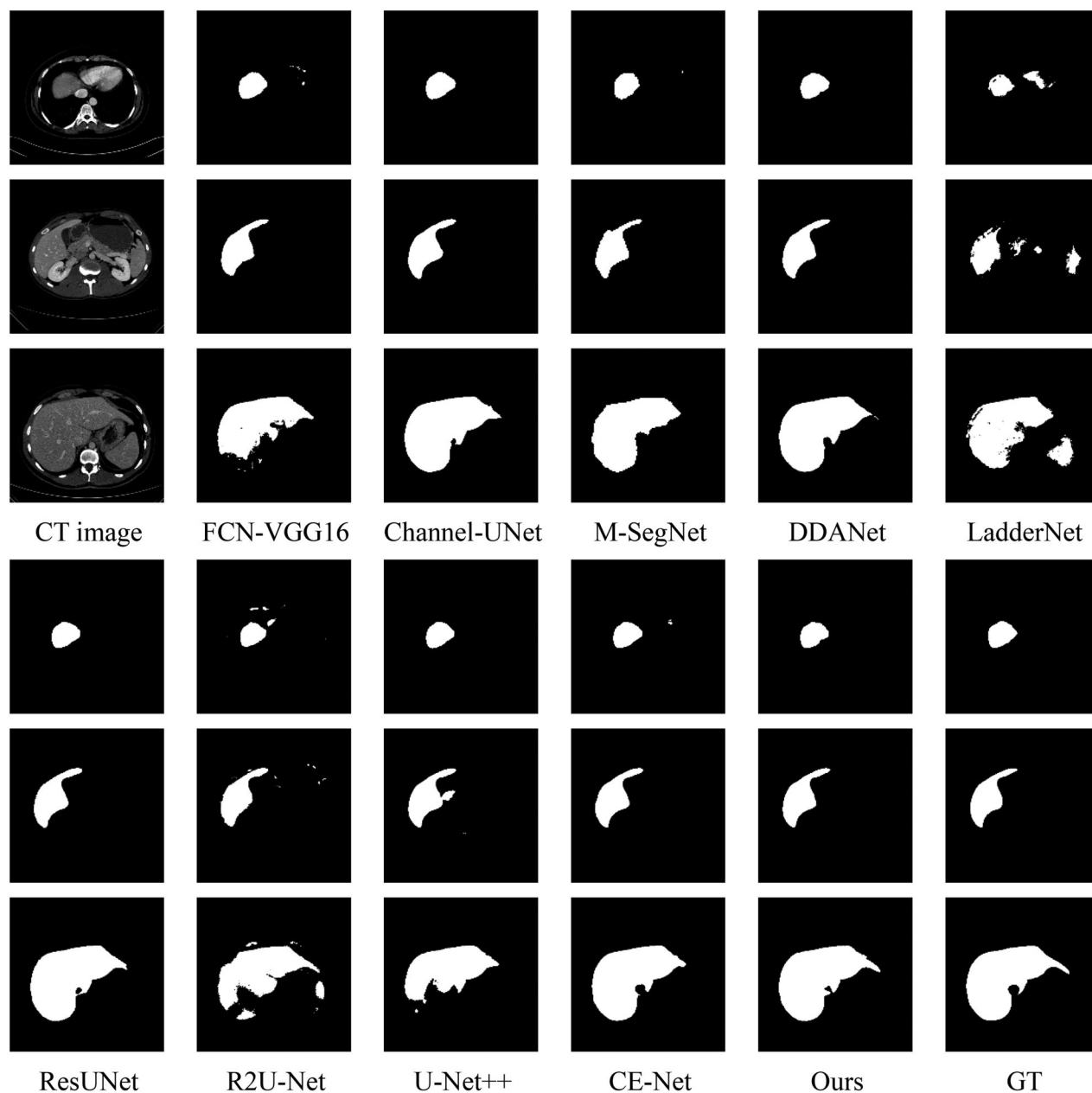
FIGURE 7 Comparison results in the dataset 3DIRCADb. From the first row to the third column: small, medium, and large-scale liver images, respectively.

mentation results (third row), all methods can complete the segmentation task well, but the DDANet, ResUNet, and U-Net++ models will have blurred boundaries. All in all, on the 3DIRCADb dataset, our method can achieve good results for large, medium, and small-scale livers, which proves that our method is effective.

To express the generalization ability of our method, we used six image segmentation evaluation indicators for quantitative analysis, Dice coefficient (DICE), volume overlap error (VOE), relative volume error (RVD), Hausdorff Distance (HD), average perpendicular distance (APD), and Intersection-Over-

Union (IOU). We test it on the CHAOS dataset and also obtain satisfactory results. In Figure 8, we present the comparative results in the dataset CHAOS. In the segmentation results of small-scale livers (the first row), Channel-UNet and ResUNet obtain good segmentation results, FCN-VGG16 and U-Net++ cannot accurately obtain liver boundaries, and R2U-Net and LadderNet produce severe over-segmentation. In particular, LadderNet cannot accurately find the specific location of the liver. In the segmentation results of the mesoscale liver (the second row), CE-Net, M-SegNet, DDANet, ResUNet, and U-Net++ did not accurately find the boundary of the liver and

FIGURE 8 Comparison results in the dataset CHAOS, where the first row to the third column are small, medium, and large-scale liver images, respectively.

the segmentation results were incomplete, while channel-UNet, R2U-Net, and FCN-VGG16 methods will mistake the non-liver region as the final liver result. In the large-scale liver segmentation results (third row), the R2U-Net and LadderNet methods segment the redundant liver parts, DenseNet, DDANet, and U-Net++ cannot completely segment the liver, and ResUNet and M-SegNet cannot accurately determine the boundary of the liver. Our method can achieve good results with the compared methods on all three scales, especially in mesoscale and small scale, outperforming other methods. It is worth noting that the LadderNet generalization ability is not good; it shows good results on the 3DIRCADb dataset, but it performs poorly on the CHAOS dataset. However, our method can achieve good

results in both data sets, which further demonstrates that our method has good robustness and generalization.

In addition to qualitative analysis, we also sorted out the six index parameters of the comparison methods on the two datasets, and the specific results are shown in Tables 2 and 3.

From Table 2, it can be seen that our method achieves the largest advantage among the six metrics, in which DICE, VOE, IOU, HD, and APD metrics get the best results, and RVD gets the second best result. In particular, our performance on DICE is 1% higher than the second CE-Net model, reaching 96.189%. It can be seen that our method can obtain the specific location of the liver and perform accurate segmentation on a large number of CT images. As far as the VOE metric, our method obtains

**TABLE 2** Average scores on dataset 3D-IRCADb, the red font is the best result under each indicator.

	Methods									
Indicator	FCN-VGG16 [10]	Channel-UNet [25]	M-SegNet [26]	DDANet [27]	LadderNet [28]	ResUNet [9]	R2U-Net [11]	U-Net++ [12]	CE-Net [28]	Ours
DICE (%)	92.357	93.138	91.426	90.751	94.821	94.196	90.813	92.956	95.156	96.189
VOE (%)	7.324	7.137	8.935	7.618	6.289	4.238	13.012	9.143	3.156	3.125
RVD (%)	−3.911	−4.825	−5.61	0.441	−4.815	−1.923	−4.163	−5.716	−2.164	−0.721
IOU (%)	88.436	89.352	85.954	85.228	89.218	90.319	85.041	89.015	91.315	91.921
HD (mm)	92.987	48.681	70.318	60.398	38.71	45.921	90.153	59.531	36.214	16.963
APD (mm)	13.426	3.368	8.019	7.314	3.186	4.012	19.150	4.183	2.935	2.465

APD, average perpendicular distance; DICE, dice coefficient; HD, Hausdorff distance; IOU, intersection-over-union; RVD, relative volume error; VOE, volume overlap error.

TABLE 3 Average scores on the dataset CHAOS, the red font is the best result under each indicator.

	Methods									
Indicator	FCN-VGG16 [10]	Channel-UNet [25]	M-SegNet [26]	DDANet [27]	LadderNet [28]	ResUNet [9]	R2U-Net [11]	U-Net++ [12]	CE-Net [28]	Ours
DICE (%)	93.864	95.107	94.957	94.208	79.935	96.015	87.215	91.039	95.931	96.889
VOE (%)	6.954	6.821	7.126	9.015	21.076	5.439	16.279	17.120	3.936	3.831
RVD (%)	−2.377	7.105	5.697	−3.591	−9.079	5.379	−4.103	−7.135	3.810	−3.015
IOU (%)	90.324	92.331	90.985	89.228	69.120	91.930	79.124	84.985	93.102	94.321
HD (mm)	66.981	55.981	51.934	32.398	180.214	22.691	268.156	55.127	29.103	14.225
APD (mm)	12.634	10.349	7.125	4.325	36.724	2.154	29.102	6.102	2.910	1.930

APD, average perpendicular distance; DICE, dice coefficient; HD, Hausdorff distance; IOU, intersection-over-union; RVD, relative volume error; VOE, volume overlap error.

Original image   Predict result   Ground truth   Original image   Predict result   Ground truth
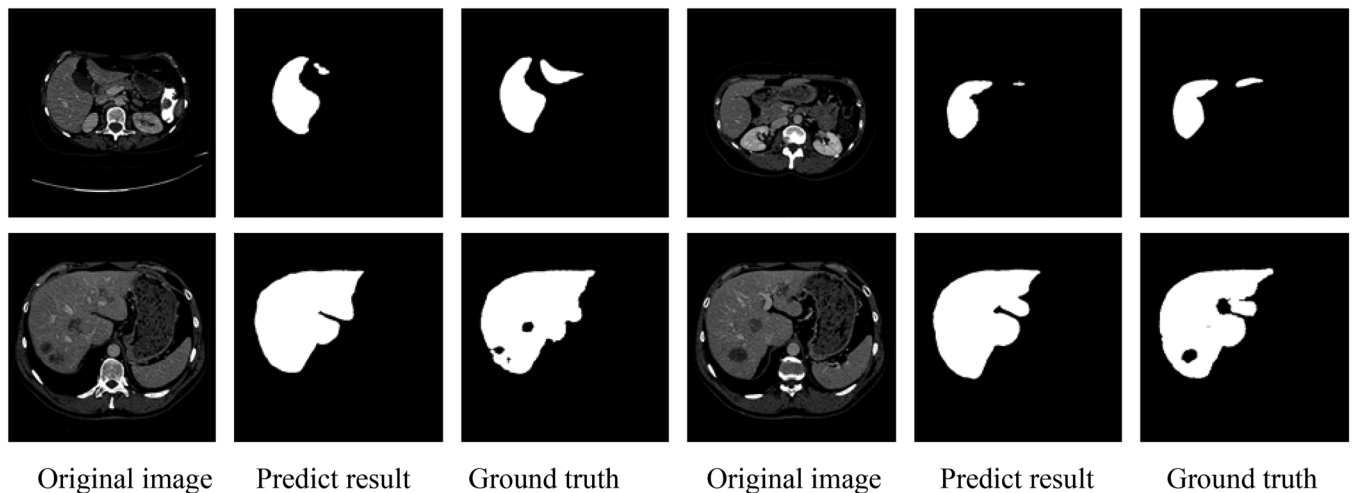
FIGURE 9 Visual comparison of some worst predictions.

the lowest score, which may be caused by the excessive segmentation of other tissues and organs. As far as the VOE metric, our method obtains the lowest score, which may be caused by the excessive segmentation of other tissues and organs.

As can be seen from Table 3, our method also achieves great advantages in the six indicators on the dataset CHAOS, indicating that our method can also have good generalization ability on different datasets. To sum up, our method can obtain the best advantages in terms of DICE, VOE, IOU, HD, and APD, and also obtains the second result in RVD. The effectiveness of our method is proved further.

In addition, we also show some worst predictions, as shown in Figure 9. We can observe that when multiple segmented target regions are far away, the network cannot segment multiple
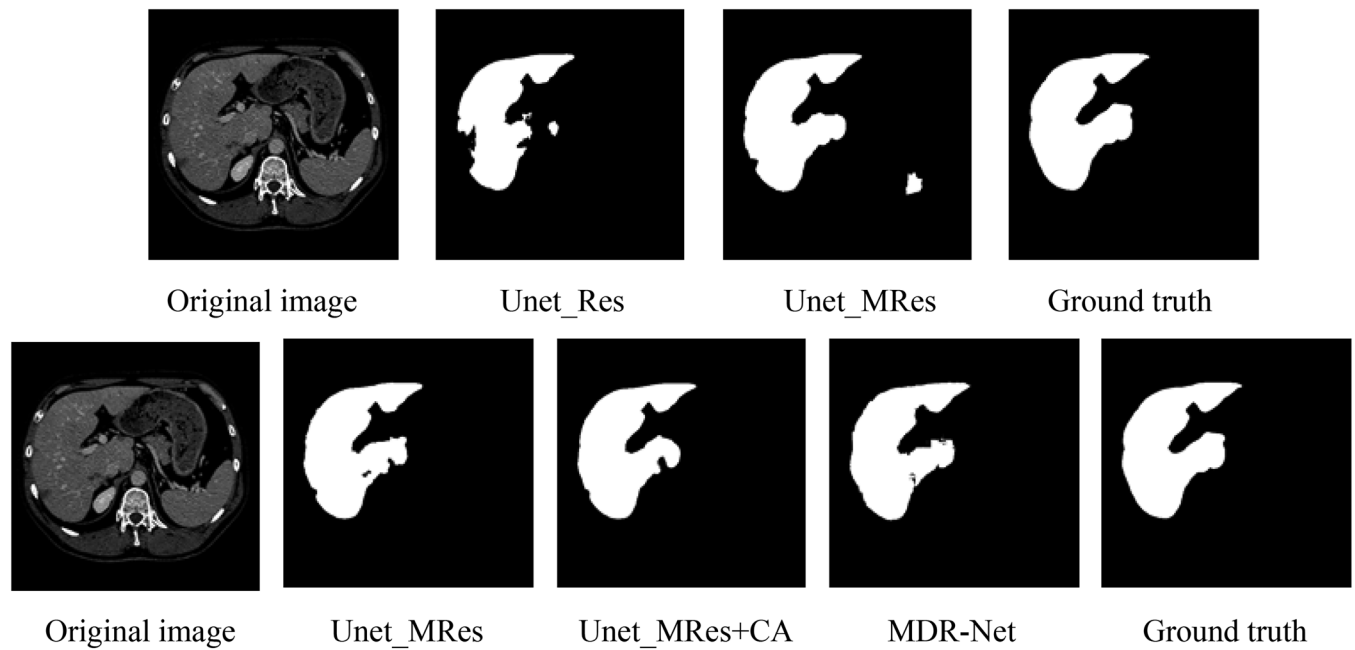
FIGURE 10 Visual comparison of different modules.

TABLE 4 The average time and the amount of parameters included in the processing of images by different network model methods.

Method	Speed/s	Parameters/M
FCN-VGG16 [10]	0.056	18.347
Channel-UNet [25]	0.209	48.317
M-SegNet [26]	0.091	28.931
DDANet [27]	0.159	33.129
LadderNet [28]	0.026	1.687
ResUNet [9]	0.016	21.098
R2U-Net [11]	0.368	100.648
U-Net++ [12]	0.103	9.349
CE-Net [28]	0.031	39.167
MDR-Net(our)	0.057	4.297

We have taken the image resolution of 512×512 as an example for processing.

TABLE 5 Analysis of the performance of each component on the dataset 3DIRCADb.

Methods	DICE (%)	VOE (%)	RVD (%)	IOU (%)	HD (mm)	APD (mm)
1. U-Net	92.167	10.132	−6.934	86.934	57.345	4.013
2. U-Net_Res	93.137	7.354	−5.217	90.310	63.214	6.031
3. U-Net_MRes	94.425	6.349	−5.103	90.339	50.864	4.399
4. U-Net_MRes+CA	95.169	6.137	−4.123	91.479	46.137	2.854
5. MDR-Net (ours)	96.189	3.125	−0.721	91.921	16.963	2.465

APD, average perpendicular distance; DICE, dice coefficient; HD, Hausdorff distance; IOU, intersection-over-union; RVD, relative volume error; VOE, volume overlap error.

target regions completely. For the liver image with lesion area, it will lead to the wrong segmentation of the lesion area.

4.4 | Number of parameters and inference time

In addition, we also show the average processing time of each slice and the parameters of the overall network. The specific parameter performance is shown in Table 4. It can be found from the table that ResUNet achieves the best performance in processing time, and the LadderNet model has the fewest parameters and is a lightweight network. However, R2U-Net is not ideal in terms of parameter performance and speed, and it needs further improvement. Compared to lightweight networks (ResUNet, LadderNet, and CE-Net), the processing speed of our proposed method is not too fast because we use a residual network model and use block stacking of different scales. But in terms of network parameters, our parameters are much smaller than those of large networks, and we get the highest score in performance. The processing speed of our proposed method is not too fast because we use a residual network model and use block stacking of different scales.

4.5 | Ablation experiment

In this section, we mainly test the basic components of our network, including residual models at different scales, an edge-aware loss network, and a CA block. We conduct ablation experiments on the 3DIRCADb dataset, and the obtained parameters are shown in Table 5. In addition, we also show

the effectiveness of different modules for visual comparison, as shown in Figure 10.

1. We first compare Model 2 (U-Net_Res) and Module 3 (U-Net_MRes) to demonstrate the effectiveness of the proposed multi-scale residual network. It can be found that the introduction of the residual block DICE index in the basic network model has improved by more than 1%. Comparing modules 2 and 3, the DICE score obtained by our proposed multi-scale residual network improves by another 1.3% on the basis of the first one.

2. We introduce a CA mechanism in the base network which is effective by comparing Model 3 (U-Net_MRes) and Model 4 (U-Net_MRes+CA). Since the CA mechanism enhances the salient features of pixels in the liver region, the performance enhancement is more significant. Although the performance improvement of DICE is not much, there is a comprehensive improvement in the APD indicator. It shows that the object contour is closer to the real liver, which is also the result of our enlarged liver region features.

3. Finally, we compare the performance of our edge-aware loss network, which is beneficial to obtain the edge information of the liver and obtain the complete information of the liver slice. From the indicators, it can be seen that HD has been greatly improved after using the edge-aware loss network.

In summary, each of our components is critical to our MDR-Net model, and the performance gains outweigh simply adding the performance gains of individual modules.

5 | CONCLUSION

We propose a MDR-Net for liver segmentation. It can learn liver changes at different scales and lock liver boundaries. Trained and tested on 3DIRCADb and CHAOS datasets, our approach achieves competitive performance with clear contour and texture details. Despite the good performance, our method suffers from some limitations. For example, for 3D CT scan data, using 2D network structure for segmentation will lead to a certain loss of spatial information, and then affect the segmentation accuracy. Therefore, we will focus on the 3D-based convolutional neural network algorithm for liver image segmentation in the future work.

AUTHOR CONTRIBUTIONS
Lijie Xie: Investigation; Resources; Software. Fubao Zhu: Investigation; Supervision; Validation. Ni Yao: Conceptualization; Supervision; Visualization

CONFLICT OF INTEREST STATEMENT
The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

DATA AVAILABILITY STATEMENT
Data openly available in a public repository.

ORCID
Lijie Xie https://orcid.org/0000-0001-7284-6847

REFERENCES

1. Kulathilake, K.A.S.H., Abdullah, N.A., Sabri, A.Q.M., et al.: A review on deep learning approaches for low-dose computed tomography restoration. Complex Intell. Syst. 1–33 (2021)
2. Xie, X., Pan, X., Zhang, W., et al.: A context hierarchical integrated network for medical image segmentation. Comput. Electr. Eng. 101, 108029 (2022)
3. Pan, X., Lin, H., Han, C., et al.: Computerized tumor-infiltrating lymphocytes density score predicts survival of patients with resectable lung adenocarcinoma. iScience 25(12), 105605 (2022)
4. Shovkun, A.D., Shevchun, A.F., Shovkun, D.V., et al.: Using optical zone melting for growing single crystals of superconductors. J. Surf. Invest. 16(1), 118–121 (2022)
5. Yu, H., He, F., Pan, Y.: A novel region-based active contour model via local patch similarity measure for image segmentation. Multimed. Tools Appl. 77(18), 24097–24119 (2018)
6. Huo, Y., Terry, J.G., Wang, J., et al.: Fully automatic liver attenuation estimation combing CNN segmentation and morphological operations. Med. Phys. 46(8), 3508–3519 (2019)
7. Guo, X., Schwartz, L.H., Zhao, B.: Automatic liver segmentation by integrating fully convolutional networks into active contour models. Med. Phys. 46(10), 4455–4469 (2019)
8. Seo, H., Huang, C., Bassenne, M., Xiao, R., Xing, L.: Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images. IEEE Trans. Med. Imag. 39(5), 1316–1325 (2020)
9. Bi, L., Kim, J., Kumar, A., Feng, D.: Automatic liver lesion detection using cascaded deep residual networks. arXiv:arXiv:1704.02703 (2017)
10. Ben-Cohen, A., Klang, E., Raskin, S.P., Soffer, S., Ben-Haim, S., Konen, E., et al.: Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection. Eng. Appl. Artif. Intell. 78(2), 186–194 (2018)
11. Alom, M., Yakopcic, C., Hasan, M., Taha, T.M., Asari, V.K.: Recurrent residual U-Net for medical image segmentation. J. Med. Imaging (Bellingham). 6(1), 014006 (2019)
12. Zhou, Z., Siddiquee, M., Tajbakhsh, N., Liang, J.: UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans. Med. Imaging 39(6), 1856–1867 (2020)
13. Siri Sangeeta, K., Pramod Kumar, S., Latte Mrityunjaya, V.: Threshold-based new segmentation model to separate the liver from CT scan images. IETE J. Res. 68(6), 4468–4475 (2020)
14. Arica, S., Avar, T.S., Erbay, G.: A plain segmentation Algorithm Utilizing Region growing technique for automatic partitioning of computed tomography liver images. In: 2018 Medical Technologies National Congress, pp. 1–4 (2018)
15. Fu, S., Wang, C., Luo, J., Liu, X.: Segmentation method of brain tumor MR image based on improved U-Net model. J. China West Norm. Univ. 42, 202–208 (2021)
16. Liao, M., Zhao, Y.Q., Wang, W., Zeng, Y.Z., Yang, Q., Shih, F.Y., et al.: Efficient liver segmentation in CT images based on graph cuts and bottleneck detection. Phys. Med. 32, 1383 (2016)
17. Xie, X., Zhang, W., Pan, X., et al.: CANet: Context aware network with dual-stream pyramid for medical image segmentation. Biomed. Signal Process. Control 81, 104437 (2023)
18. Budak, Ü., Guo, Y., Tanyildizi, E., Saengur, A.: Cascaded deep convolutional encoder-decoder neural networks for efficient liver tumor segmentation. Med. Hypotheses 134, 109431 (2019)

19. Sarmadi, H., Entezami, A., Salar, M., et al.: Bridge health monitoring in environmental variability by new clustering and threshold estimation methods. J. Civ. Struct. Health Monit. 11(3), 629–644 (2021)

20. Liu, Z., Song, Y.Q., Sheng, V.S., Wang, L., Yuan, D.: Liver CT sequence segmentation based with improved U-Net and graph cut. Expert Syst. Appl. 126, 54–63 (2019)

21. Chen, Y., Wang, K., Liao, X., Qian, Y., Wang, Q.: Channel-unet: A spatial channel-wise convolutional neural network for liver and tumors segmentation. Front Genet. 10, 1110 (2019)

22. Xie, X., Pan, X., Shao, F., et al.: Mci-net: Multi-scale context integrated network for liver CT image segmentation. Comput. Electr. Eng. 101, 108085 (2022)

23. Gu, Z., Cheng, J., Fu, H., et al. Ce-net: Context encoder network for 2D medical image segmentation. IEEE Trans. Med. Imaging 38(10), 2281–2292 (2019)

24. Xie, X., Zhang, W., Wang, H., et al.: Dynamic adaptive residual network for liver CT image segmentation. Comput. Electr. Eng. 91, 107024 (2021)

25. Cheng, D., Gong, Y., Zhou, S., Wang, J., Zhang, N.: Person re-identification by multi-channel parts-based CNN with improved triplet loss function. In: Computer Vision & Pattern Recognition. IEEE, New York, pp. 1335–1344 (2016)

26. Yamanakkanavar, N., Lee, B.: A novel M-SegNet with global attention CNN architecture for automatic segmentation of brain MRI. Comput. Biol. Med. 136, 104761 (2021)

27. Rajamani, K.T., Siebert, H., Heinrich, M.P.: Dynamic deformable attention network (DDANet) for COVID-19 lesion semantic segmentation. J. Biomed. Inf. 119, 103816 (2021)

28. LadderNet JTZ J: Multi-path networks based on U-Net for medical image segmentation. arXiv preprint arXiv: 1810.07810, (2018)