Journal of Electronics and Information Science (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/jeis.2024.090105
ISSN 2371-9524 Vol. 9 Num. 1

Improved Dense Recurrent Residual U-Net for Skin Lesion Segmentation

Yang Yuan*, Dechun Zhao, Zixin Luo

College of Bioinformation, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China
Corresponding author: 1302629079@qq.com

Keywords: Medical image processing, segmentation of skin lesion, U-Net, Squeeze-and-Excitation block, feature adaptation module

Abstract: Accurate segmentation of skin lesion areas is of great significance for computer-aided diagnosis. However, due to the irregular shape, boundary blurring, and noise interference of skin lesion images, accurate segmentation is difficult and has low precision. Therefore, it proposes an improved dense recurrent residual U-Net model. Firstly, This improved network use of dense recurrent residual connections in the Squeeze-and-Excitation convolution block design to alleviate gradient vanishing and provide accurate location information for segmentation; Secondly, the integration of feature adaptive modules between the encoder and decoder to enhance feature fusion between adjacent layers. Finally, a combined Dice and cross-entropy loss function is adopted to mitigate the class imbalance issue in skin lesion image segmentation. The model is evaluated on the public dataset ISIC 2017, achieving Jaccard, Dice, and accuracy scores of 78.86%, 86.92%, and 94.61% respectively. The experimental results demonstrate that the proposed model outperforms other networks in terms of segmentation performance and provides more accurate segmentation results.

1. Introduction

Skin cancer is one of the most prevalent and deadly cancer types, with an estimated 18.1 million cases of non-melanoma skin cancer patients in 2020, of which 9.9 million were fatal cases[1]. Early diagnosis is crucial to improve the survival rate of skin cancer patients. Skin malignancies include squamous cell carcinoma, basal cell carcinoma, malignant melanoma, and others[2]. Although dermatoscopy is the main method for clinical diagnosis of skin cancer, processing a large number of dermatoscopic images consumes a considerable amount of time and is subject to subjective judgments[3,4]. Therefore, it is of great significance to develop computer-aided diagnostic systems to assist doctors in automatically identifying skin lesion areas and providing accurate diagnostic results[5].

Machine learning has been widely used in skin lesion segmentation[6]. However, commonly used methods such as the adaptive global threshold based on color model normalization by Thanh et al[7]. and the automatic threshold determination method based on type-2 fuzzy logic algorithm by Emin et al[8]. require manually setting parameters and have relatively cockamamie steps. To address these

issues, Alexander et al[9]. proposed an iterative random region merging method for macroscopic image segmentation, while Jaisakthi[10] combined the GrabCut and K-means algorithms to achieve skin lesion segmentation. Additionally, Ahn et al[11]. developed a method that utilizes sparse reconstruction and background detection, combined with a Bayesian framework for lesion area and boundary segmentation. However, these methods still have limitations in accuracy when dealing with cases where normal skin and lesion areas are highly similar.

In recent years, with the rapid development of computers, deep learning has attracted widespread attention and application in the field of computer vision. Convolutional neural networks have achieved significant breakthroughs in medical image classification[12], multi-object detection[13], and tissue pathological image segmentation tasks[14]. Especially in the field of medical image segmentation, deep learning has achieved remarkable results. Medical image segmentation is a dense pixel classification problem. Long et al[15]. proposed a fully convolutional network (FCN) for semantic segmentation, while Nasr E et al[16]. combined the FCN model and dense pooling layers for skin lesion image segmentation. However, due to the small receptive field, FCN cannot obtain global information, resulting in insufficient segmentation accuracy. To address this issue, Ronneberger et al[17]. proposed a symmetric structure-based U-Net, which has become one of the most commonly used frameworks in medical image segmentation. However, there is a semantic gap between the encoder and decoder of U-Net, and multiple downsampling operations can lead to information loss.Therefore, Ibteha et al[18]. proposed the MultiResUNet model, which replaces the original U-Net convolution block with multi-level residual blocks to learn multi-scale feature information, and optimizes skip connections to residual paths to reduce semantic gaps between the encoder and decoder. In addition, Rehman et al[19]. preprocessed images and introduced conditional random fields into the segmentation network to improve the accuracy of skin lesion segmentation. Gu et al[20]. used edge-guiding modules in U-Net networks to improve the segmentation accuracy of skin lesion boundaries. Hu et al[21]. proposed AS-Net networks, which combine channel and spatial attention mechanisms to improve the recognition ability of skin lesion segmentation. For 3D medical image segmentation problems, Lin Wei et al[22]. proposed a 3D U-Net model, which achieved excellent results in segmentation outcomes. However, due to the low contrast, blurry boundaries, different sizes and shapes of lesion areas, as well as many interference noises (such as bubbles, blood vessels, light spots, and hairs), these factors can seriously affect the segmentation performance[23]. Therefore, achieving accurate skin lesion segmentation remains a highly challenging task.

This study proposes an improved dense recurrent residual U-Net model (IDR2U-Net) for automatic segmentation of skin lesion areas. Based on the original U-Net network, the model is improved by designing Squeeze-and-Excitation (SE) convolution blocks and using dense recurrent residual connections to address issues such as semantic gap and feature information loss. In addition, a feature adaptive module is used in the skip connections of the encoder and decoder to improve the efficiency of feature fusion and enhance the segmentation accuracy of boundary-blurry images. To address the class imbalance issue in skin lesion images, a combined loss function of joint cross-entropy and Dice coefficient is used. Experimental results on the ISIC 2017 dataset demonstrate the superior performance of the proposed method in terms of segmentation accuracy.

2. Methods

2.1 Network

This article proposes a skin lesion image segmentation algorithm based on the U-Net model, which uses SE blocks, dense recurrent residual convolution (DRR), and a feature adaptation module (FAM). In the feature encoding section, each layer uses an SE convolution block to extract features, which

can re-adjust the weight of channel features and enhance useful channel information. The dense recurrent residual (DRR) connection method is used to further enhance the model's feature extraction ability, increase gradient flow, and reduce semantic differences. Then, a max pooling layer with a stride of 2 is used for downsampling. In the feature decoding section, features from the encoding stage are merged with upsampled features through the FAM, which can adaptively match the distribution of two feature maps in the decoder, improve feature fusion efficiency, and bridge the semantic gap. After that, a SE convolution block with DRR connection is added, and finally, a 1×1 ordinary convolution block is connected to the output image. The specific structure is shown in **Figure 1**.

In this network, the input image size is 256×256 with 3 channels, and the output image size is also 256×256 with 2 channels, corresponding to the two prediction results: skin lesion area and non-skin lesion area. SE-DRR refers to the improved SE convolution block using DRR connection method, FAM refers to the feature adaptation module, Down-sampling refers to the downsampling layer, Up-sampling refers to the upsampling layer, Final convolution refers to a 1×1 ordinary convolution block, and Skip connection refers to the skip connection.

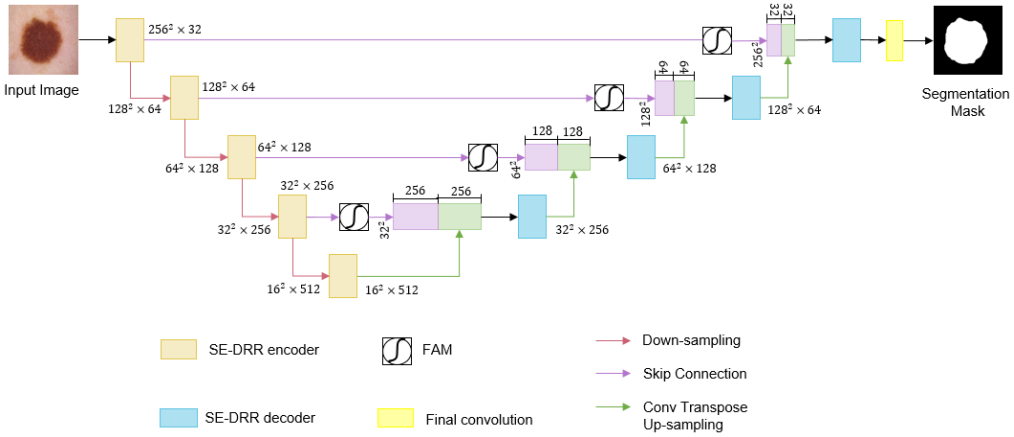

Figure 1: IDR2U-Net network structure diagram.

2.2 SE-DRR module

2.2.1 SE-conv

Channel attention is used to enhance the discrimination ability of feature information and suppress irrelevant background information. By explicitly modeling the inter-channel dependencies, the network can assign more weights to important features, greatly improving the model's performance and accuracy. In this study, we adopt the Squeeze-and-Excitation block (SE) structure[24], which is shown in **Figure 2**. Among them, F_{tr} represents the traditional convolution operation. Let x and u_c be the input ($c_1 \times h \times w$) and output ($c_2 \times h \times w$) of F_{tr}, respectively. The calculation formula is as follows:

$$F_{tr}: \quad u_c = v_c * x = \sum_{s=1}^{c_1} v_c^s * x^s \tag{1}$$

Among them, v_c represents the c convolution kernel, x_s represents the s input covered by the current convolution kernel.

Firstly, the output u_c undergoes global average pooling operation is $F_{sq}(\cdot)$, with the formula as follows:

$$z_c = F_{sq}(u_c) = \frac{1}{w*h}\sum_{i=1}^{w}\sum_{j=1}^{h} u_c(i,j) \tag{2}$$

Then, the feature map is subjected to two stages of full connection operation is $F_{ex}(\cdot)$. The calculation formula is as follows:

$$F_{ex}(z_c, w) = \text{sigmoid}(w * \text{Re}\,\text{lu}(w_1 z_c)) \tag{3}$$

Among them, w_1 and w_2 represent linear layers, s_c represents the feature weights of each channel, and finally undergoes $F_{scale}(\cdot)$ operation, each channel value multiplied by its weight s_c. The final output result is $u_c \times s_c$. The introduction of SE block slightly increases the computational complexity of the entire network, but the segmentation effect is significantly improved.

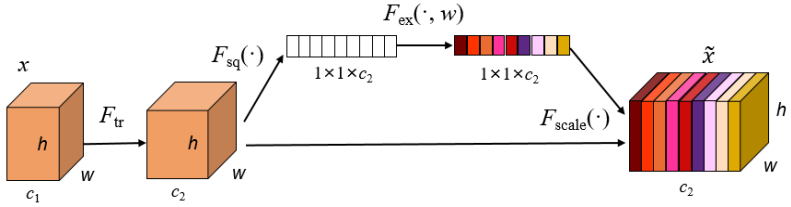

Figure 2: SE module.

2.2.2 DRR connection

The convolutional block structure in Convolutional Neural Networks (CNNs) has a significant impact on medical image segmentation results. To improve feature propagation and further optimize segmentation accuracy, this study proposes an improvement to the original convolutional block and the connections between them. **Figure 3** shows the connection structure diagram of the convolutional block. The original U-Net convolutional block is used for image feature extraction, consisting of two 3×3 convolutions and ReLU activation functions. The improved SE-conv block uses recurrent and residual structure connections to reduce overfitting in the network and also has a regularization effect[25]. Secondly, dense skip connections are added between the recurrent residual convolutional blocks to increase gradient flow and reduce semantic differences. The optimized network has stronger feature extraction capabilities and more stable training results.

The recurrent operation is because Recurrent Neural Networks (RNNs) receive an input at each time step and also consider the output state of the previous time step. Assuming x_l is the input of the neural network at the l^{th} layer, and there is a pixel position (i, j) at the k^{th} feature map of the recurrent layer, then the network output at that position is:

$$y_{ijk}^{l}(t) = \left(w_k^f\right)^{\mathrm{T}} * x_l^{f(i,j)}(t) + \left(w_k^r\right)^{\mathrm{T}} * x_l^{r(i,j)}(t-1) \tag{4}$$

$$F\left(x_l, w_l\right) = \max\left(0, y_{ijk}^{l}(t)\right) \tag{5}$$

In the formula, $F(x_l, w_l)$ represents the output value of the ReLU operation for the l^{th} layer. Then, the output is passed through the residual layer, The resulting output is:

$$x_{l+1} = x_l + F\left(x_l, w_l\right) \tag{6}$$

x_l represents the input to the recurrent layer and is added to the output of the recurrent layer. In this paper, dense connections are adopted between the convolutional layer and the recurrent layer to better implement feature extraction across layers. Therefore, the final output is:

$$y_{ijk}^{l}(t)=\left(w_{k}^{f}\right)^{\mathrm{T}} * H\left(x_{l}^{f(i,j)}\right)(t)+\left(w_{k}^{r}\right)^{\mathrm{T}} * H\left(x_{l}^{r(i,j)}\right)(t-1) \tag{7}$$

In the formula, $H(\cdot)$ is a composite function of the output from the previous convolutional block.

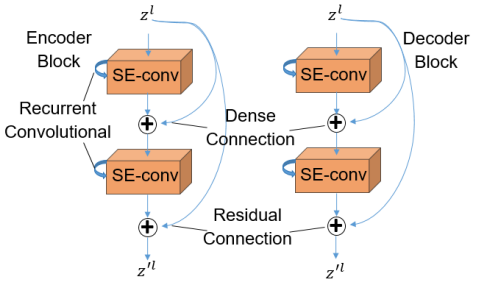

Figure 3: SE-DRR module.

2.3 FAM module

To improve the low efficiency of feature fusion, large noise of feature maps, and spatial information loss in the original U-Net structure, this article proposes a new feature adaptive module to record the feature distribution between the encoder and decoder[26]. The feature adaptive module mainly consists of multiple 3×3 and 1×1 convolution layers, as shown in **Figure 4**. Specifically, this module can adjust the number of channels through 1×1 convolution layers to maintain the dimensionality of feature maps unchanged, and increase the receptive field and learn richer features through 3×3 convolution layers. It can capture the relationship between adjacent pixels and merge multiple convolution kernels for multiple convolutions to achieve nonlinear perception. By combining multiple convolution layers, the feature adaptive module can learn richer, more abstract feature representations, thereby improving the performance of deep neural networks. Compared with the attention gate (AG)[27] module, this module uses relatively few parameters, has lighter weight, and has high storage efficiency.

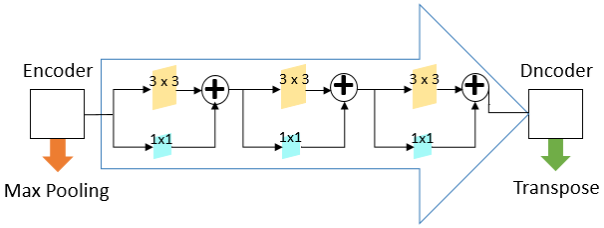

Figure 4: FAM module.

2.4 Joint Cross-entropy, Dice and BCE

Due to the significant variations in lesion size in the skin lesion dataset, it was observed that the convergence speed of the model was influenced by the size of the lesion. To address this issue and reduce overfitting, a combined loss function of cross-entropy, BCE and Dice coefficient was employed for training the segmentation network. The formula for this combined loss function is defined as follows:

$$L_{\mathrm{Dice}}=1-2\times\frac{\sum_{i}P_{i}G_{i}}{\sum_{i}P_{i}+\sum_{i}G_{i}} \tag{8}$$

$$L_{\mathrm{BCE}}=-\sum_{i}\left[\left(1-G_{i}\right)\ln\left(1-P_{i}\right)+G_{i}\ln\left(P_{i}\right)\right] \tag{9}$$

In the formula, G_i represents the true class label of the i pixel, and P_i represents the predicted class label of the i pixel.

The final loss function is composed of the above two loss functions, and its formula is defined as follows:

$$L=\alpha L_{BCE}+(1-\alpha)L_{Dice} \tag{10}$$

The coefficients α and (1-α) represent the weighting factors for the corresponding loss functions. The loss function used in this paper combines the advantages of both, which enhances the accuracy of segmenting skin lesions of different sizes.

3. Experiment

3.1 Datasets

The dataset used in this experiment is the publicly available ISIC 2017, which is a dermatology image dataset that includes single, multiple, and lesion-based skin images from around the world. It contains more than 2000 different types of skin cases with high-quality images and a resolution of 2048×1536 pixels. The ISIC 2017 dataset includes images of different types and severity levels of skin lesions, including 514 basal cell carcinomas, 327 squamous cell carcinomas, 374 melanomas, and 1372 benign lesions. All images have been manually annotated by professional doctors and provide detailed metadata information for algorithm training, testing, and evaluation. The dataset has been divided into training set (2000 images), validation set (150 images), and test set (600 images).

To obtain higher computational efficiency, the size of all images was uniformly adjusted to 256×256 pixels. Preprocessing operations such as image resizing were applied to the input images to normalize the data fed into the model and have it follow a standard normal distribution, which facilitates better training and application of deep learning models.

3.2 Experimental setup and implement details

3.2.1 Training and testing

Experimental hardware environment: The processor is Intel (R) Core (TM) i5-7500 CPU @ 3.40GHz, the memory (RAM) is 16.0 GB, and the GPU is NVIDIA V100. The software environment is Windows 10, Python3.9, CUDA11.2, CUDNN11.2, using the Tensorflow 2.9.1 deep learning framework for overall model construction.

During the training process, we modified the initial learning rate, momentum, weight decay regularization and other parameters of each network to acquire their best performance. The input image size is adjusted to 256x256 pixels. In the experimental stage, Due to the memory constraints of the GPU, we utilize the Adam optimizer for optimization and gradient update. The training batch size is set to 16, and the initial learning rate is set to 0.0002. The momentum parameter is set to the default value of 0.99, and the training iterates for 200 epochs.The preliminary experimental results shown indicate that the best segmentation effect is achieved when α is set to 0.6. When the cross-entropy loss function accounts for 0.6. Therefore, the weight factor for the evaluation indicators in this article is selected as this value.

3.2.2 Evaluation

To evaluate the segmentation results and compare them with other methods, the following evaluation metrics are adopted in this article. Jaccard index (JAC) indicates the overlap rate between

the network's segmented skin lesion area and the true label area; Dice coefficient (DICE) indicates the similarity between the predicted skin lesion area and the true label area; Accuracy (ACC) indicates the overall accuracy of all pixel classifications; Sensitivity (SEN) indicates the recall rate of the skin lesion area; Specificity (SPE) indicates the accurate segmentation degree of non-lesion areas. The relevant metrics are calculated as follows:

$$SEN = \frac{TP}{TP + FN} \tag{11}$$

$$SPE = \frac{TN}{TN + FP} \tag{12}$$

$$JAC = \frac{TP}{TP + FP + FN} \tag{13}$$

$$DICE = \frac{2 \times TP}{2 \times TP + FN + FP} \tag{14}$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{15}$$

In the formula, TP, FP, FN, and TN represent true positive, false positive, false negative, and true negative, respectively.

3.3 Result and analysis

To validate the performance of the proposed IDR2U-Net model for skin lesion segmentation, we conducted comparative experiments with U-Net[28], R2U-Net[29], and DA-Net[30] under the same experimental settings. The loss function curves of these networks on the validation set of the ISIC-2017 dataset are shown in **Figure 5**. Table 1 presents the evaluation results of skin lesion segmentation on the test set using various models.

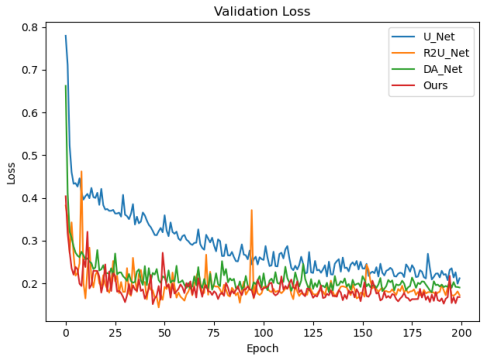

Figure 5: Loss curve of U-Net, R2U-Net, DANet and Ours.

The loss function curve shows that the IDR2U-Net model has the smallest loss, converges fastest, and does not exhibit significant overfitting, indicating good performance for the model.

Table 1: Segmentation results of different networks.

Model	JAC/%	DICE/%	ACC/%	SEN/%	SPE/%
U-Net	68.93	78.98	90.37	83.45	96.43
R2U-Net	74.09	82.47	92.14	84.58	96.34
DA-Net	75.37	84.37	93.06	82.16	98.49
ours	78.86	86.92	94.61	86.75	96.28

As we can see from Table 1, the method proposed in this article achieves the best segmentation performance on skin lesion areas, achieving the maximum values of Jaccard index, Dice coefficient, accuracy, and sensitivity of 78.86%, 86.92%, 94.61%, and 86.75% respectively. In terms of Dice coefficient, our method outperforms U-Net, R2UNet, and DANet by 7.94%, 4.45%, and 2.25% respectively. In terms of accuracy, our method outperforms U-Net, R2UNet, and DANet by 4.24%, 2.47%, and 1.55% respectively. Due to the noise interference and blurred lesion boundaries commonly found in the ISIC 2017 dataset, U-Net and R2UNet suffer from excessive feature information loss, resulting in low segmentation accuracy and phenomena of over-segmentation and under-segmentation in the segmentation results.

Figure 6 demonstrates an example of segmentation results comparing U-Net, R2UNet, DANet, and our proposed method. As shown in the figure, the segmentation results of U-Net and R2UNet are relatively poor with insufficient refinement in the segmentation boundaries. Although DANet has clearer boundaries compared to the previous two methods, its accuracy is still low. By incorporating attention mechanisms, our proposed method can improve the accuracy and boundary segmentation precision of skin lesion areas.

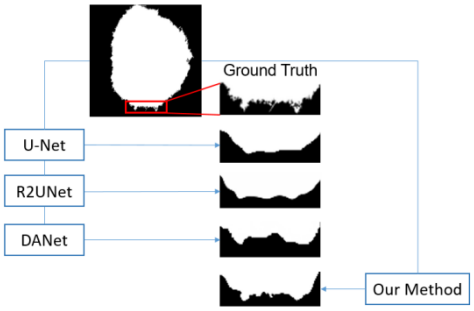

Figure 6: Instance of segmentation.

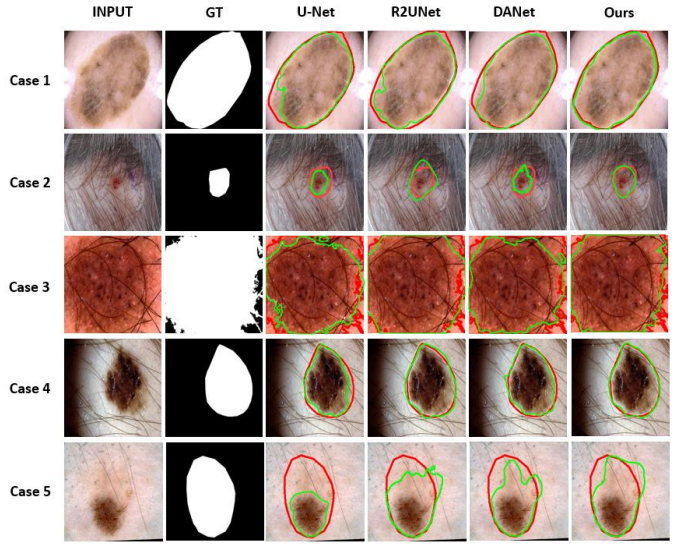

Figure 7: Different networks image segmentation results. Red line: ground truth; Green line: different networks image segmentation results.

The segmentation results of all networks are shown in Figure 7. On one hand, the compared other networks exhibit varying degrees of over-segmentation and under-segmentation phenomena, and have lower accuracy in segmenting skin lesion images with blurry boundaries. As shown in **Figure 7**, in Case 1, 2, 4, and 5, U-Net has obvious under-segmentation, while R2UNet has obvious over-segmentation, and the proposed method has fewer occurrences of these two situations, with more

obvious improvement in segmentation accuracy. Case 2 has the most obvious mis-segmentation of the skin lesion area. On the other hand, due to the low contrast of skin lesion images, it is difficult to accurately segment boundaries, as shown in Case 3 and 4. DANet introduces a dual attention mechanism on the basis of U-Net, which improves the accuracy of boundary segmentation, but the effect is not significant. Especially in Case 3, the previous methods have insufficient refinement of the boundary segmentation. Overall, the proposed method can effectively improve the above two problems, and compared to other models, it has better segmentation results and accuracy, closer to the true labels.

4. Discussion

This section conducted 5 sets of ablation experiments on the ISIC 2017 dataset using the original U-Net network as the baseline model (Model 1). Model 2 added SE convolution blocks to Model 1; Model 3 further incorporated DRR modules based on Model 2; Model 4 introduced FAM modules on top of Model 1. Models 2, 3, and 4 aimed to assess the performance improvement of skin lesion image segmentation with the proposed optimization modules, including SE blocks, DRR connections, and FAM. Model 5 represents the final model used in this paper. The results of the ablation experiments are presented in Table 2.

From Table 2, we can observe that compared to Model 1, Model 2 has improved performance across all three evaluation metrics, indicating that the proposed SE convolution block has a significant impact on skin lesion segmentation accuracy. Model 3, which incorporates the DRR connection on top of Model 2, sees a further substantial improvement in all three metrics, with the JAC and DICE showing the most significant gains of 9% and 3.15% compared to Model 1. This verifies that the SE-DRR module effectively extracts features and enhances useful information for better network training and segmentation results. Model 4 shows a more pronounced improvement in accuracy compared to other models, and analysis of segmentation results indicates that the FAM module effectively addresses boundary blurring in images. This further confirms that the module has a significant impact on feature fusion. Model 5, which integrates all three modules, achieves the best performance across all evaluation metrics, with ACC reaching its maximum value. Compared to the baseline network, it has improved accuracy by 4.24 percentage points, indicating that the proposed network generates more accurate segmentation results that are closer to the true labels and outperforms other models. This verifies the effectiveness of the proposed network modifications for skin lesion segmentation.

Table 2: Ablation experimental results on ISIC 2017 dataset.

Improvement measures				**Evaluation indices (%)**		
U-Net	**Add SE**	**Add DRR Connection**	**Add FAM**	**JAC**	**DICE**	**ACC**
√				68.93	78.98	90.37
√	√			75.81	83.76	92.98
√	√	√		77.93	86.14	93.52
√			√	76.57	85.48	94.05
√	√	√	√	78.86	86.92	94.61

5. Conclusion

This article addresses the issues of low feature fusion efficiency, information loss, and lack of connection between local and global information in skin lesion image segmentation, resulting in low accuracy. An improved dense recursive residual U-Net (IDR2U-Net) for skin lesion image

segmentation is proposed. Through ablation experiments and comparative experiments, it is shown that the proposed network outperforms other networks in performance evaluation metrics and skin lesion segmentation results. The proposed network not only promotes feature fusion between the encoder and decoder, but also enhances the correlation between global and detailed information features, providing a reference for computer-aided diagnosis of skin lesion areas and guidance for other medical image segmentation tasks.

Due to the low contrast and extremely blurred boundaries of skin lesion images, the obtained segmentation results are not ideal, indicating that there is still room for improving the accuracy of skin lesion image segmentation. Preprocessing of skin lesion images, such as enhancing and denoising, can be further explored to prepare for improving the overall performance of segmentation. In future research, more advanced deep learning methods can be applied to skin lesion region segmentation to further optimize the segmentation effect.

References

[1] Sung H, Ferlay J, Siegel R, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: A Cancer Journal for Clinicians, 2021, 71(3): 209-249.

[2] Kasmi R, Mokrani K. Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule [J]. IET Image Processing, 2016, 10 (6): 448-455.

[3] Dai Duwei, Dong Caixia, Xu Songhua, et al. Ms RED: A novel multi-scale residual encoding and decoding network for skin lesion segmentation [J]. Medical Image Analysis, 2022, 75: 102293.

[4] Yin W, Zhou D M, FAN T, et al.Image segmentation of skin lesions based on dense atrous spatial pyramid pooling and attention mechanism[J].Journal of Biomedical Engineering, 2022, 39(06):1108-1116.

[5] Liu Qi, Wang Jingkun, Zuo Mengying, et al. NCRNet: Neighborhood context refinement network for skin lesion segmentation [J]. Computers in Biology and Medicine, 2022, 146: 105545.

[6] Shan Pufang, Fu Chong, Dai Liming, et al. Automatic skin lesion classification using a new densely connected convolutional network with an SF module [J]. Medical & Biological Engineering & Computing, 2022, 60 (8): 2173-2188.

[7] Thanh DNH, Erkan U, Prasath S, et al. A skin lesion segmentation method for dermoscopic images based on adaptive thresholding with normalization of color models [C] //International Conference on Electrical and Electronic Engineering. Istanbul, Turkey: IEEE, 2019: 116-120.

[8] Emin MY, Murat B. Accurate segmentation of dermoscopic images by image thresholding based on type-2 fuzzy logic [J]. IEEE Trans Fuzzy Systems, 2009, 17(4): 976-982.

[9] Alexander W, Jacob S, Paul F. Automatic skin lesion segmentation via iterative stochastic region merging [J]. IEEE Transactions on Information Technology in Biomedicine: A Publication of the IEEE Engineering in Medicine and Biology Society, 2011, 15(6): 929-936.

[10] Hamed G YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms. Comput Meth Prog Bio 2020, 200(2).

[11] Ahn E, Kim J, Bi L, et al. Saliency-Based Lesion Segmentation via Background Detection in Dermoscopic Images [J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(6): 1685-1693.

[12] Yin XH, Wang YC, Li DY. Suvery of medical image segmentation technology based on U-Net structure improvement. Ruan Jian Xue Bao/Journal of Software, 2021, 32(2):519-550.

[13] Xie Fengying, Yang Jiawen, Liu Jie, et al. Skin lesion segmentation using high-resolution convolutional neural network [J]. Computer Methods and Programs in Biomedicine, 2020, 186: 105241.

[14] Zhang Y, Liang F M, Liu J M. Skin Lesion Segmentation Based on Classification Activation Mapping and Visual Field Attention[J].Computer Engineering and Application, 2023, 59(21):187-194.

[15] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.

[16] Nasr-E E, Rafiei S, Mohammad H, et al. Dense pooling layers in fully convolutional network for skin lesion segmentation [J]. Computerized Medical Imaging and Graphics, 2019, 78(C).

[17] Ronneberrger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C] //International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.

[18] Ibtehaz N, Rahman MS. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Networks, 2020, 121: 74-87.

[19] Rehman H, Nida N, Shah SA, et al. Automatic melanoma detection and segmentation in dermoscopy images using deep

RetinaNet and conditional random fields [J]. Multimedia Tools and Applications, 2022, 81(18): 25765-25785.
[20] Gu Rui, Wang Lituan, Zhang Lei. DE-Net: A deep edge network with boundary information for automatic skin lesion segmentation [J]. Neurocomputing, 2022, 468: 71-84.
[21] Kai Hu, Jing Lu, Dongjin Lee. AS-Net: Attention Synergy Network for skin lesion segmentation [J]. Expert Systems with Applications, 2022, 201: 117112.
[22] Ling Wei, Fan Hong, Hu Chenxi, et al. Improved Medical Image Segmentation Model Based on 3D U-Net [J]. Journal of Donghua University, 2022, 39(04): 311-316.
[23] Sarker MK, Rashwan HA, Akram F, et al. SLSNet: Skin lesion segmentation using a lightweight generative adversarial network [J]. Expert Systems with Applications, 2021, 183: 115433.
[24] Ma Shiqiang, Tang Jijun, Guo Fei. Multi-Task Deep Supervision on Attention R2U-Net for Brain Tumor Segmentation [J]. Frontiers in Oncology, 2021, 11: 704850.
[25] Alom Z, Asari VK, Parwani A, et al. Microscopic nuclei classification, segmentation, and detection with improved deep convolutional neural networks (DCNN) [J]. Diagnostic Pathology, 2022, 17(1): 1-17.
[26] Wu Huisi, Chen Shihuai, Chen Guilian, et al. FAT-Net: Feature adaptive transformers for automated skin lesion segmentation [J]. Medical Image Analysis, 2022, 76: 102327.
[27] Tran TT, Pham VT. Fully convolutional neural network with attention gate and fuzzy active contour model for skin lesion segmentation [J]. Multimedia Tools and Applications, 2022, 81(10): 13979-13999.
[28] Alom MZ, Yakopcic C, Taha, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation [DB/OL].http://arxiv.org/abs/1802.06955, 2018-02-20/2018-05-29.
[29] Fu Jun, Liu Jing, Tian Haijie, et al. Dual attention network for scene segmentation [DB/OL]. http://arxiv.org/abs/1809.02983, 2019-04-21/2020-12-18.
[30] Qiu FBZ. An efficient U-shaped network combined with edge attention module and context pyramid fusion for skin lesion segmentation. [J]. Medical & biological engineering & computing, 2022, 60 (7): 1987-2000.