

## Research Article

# R2AU-Net: Attention Recurrent Residual Convolutional Neural Network for Multimodal Medical Image Segmentation

Qiang Zuo , Songyu Chen , and Zhifang Wang 

*Electronic Engineering College, Heilongjiang University, Harbin 150000, China*

Correspondence should be addressed to Zhifang Wang; wangzhifang@hlju.edu.cn

Received 31 December 2020; Accepted 26 May 2021; Published 10 June 2021

Academic Editor: Liguozhang

Copyright © 2021 Qiang Zuo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, semantic segmentation method based on deep learning provides advanced performance in medical image segmentation. As one of the typical segmentation networks, U-Net is successfully applied to multimodal medical image segmentation. A recurrent residual convolutional neural network with attention gate connection (R2AU-Net) based on U-Net is proposed in this paper. It enhances the capability of integrating contextual information by replacing basic convolutional units in U-Net by recurrent residual convolutional units. Furthermore, R2AU-Net adopts attention gates instead of the original skip connection. In this paper, the experiments are performed on three multimodal datasets: ISIC 2018, DRIVE, and public dataset used in LUNA and the Kaggle Data Science Bowl 2017. Experimental results show that R2AU-Net achieves much better performance than other improved U-Net algorithms for multimodal medical image segmentation.

## 1. Introduction

Medical image plays a key role in medical treatment. Computer-aided diagnosis (CAD) is designed to provide doctors with accurate interpretation of medical images systematically so as to treat the patients better. Manual segmentation not only relies heavily on doctors' own knowledge and clinical experience in recognition accuracy, but also has very low efficiency. Therefore, the application of deep learning in medical image segmentation has aroused widespread concern. Because medical image labeling requires the experts to spend considerable time and effort, it is difficult to acquire thousands of training images in medical image segmentation tasks. Ciresan et al. [1] trained networks in sliding windows to predict class tags for each pixel by providing local areas (patch) around pixels. However, this network must run networks independently for each patch, and there are plenty of redundancies due to overlapping patches. Furthermore, more maximum pool layers are needed for large patches, which will reduce the positioning accuracy. Full convolutional neural network [2] is one of the earliest applied deep neural networks with image segmentation. Without

traditional full connection layer, it uses deconvolution to restore original images at the last layer of network. Ronneberger et al. [3] extended this system and proposed U-Net, which includes coding path and decoding path. Encoder uses output feature map to characterize original image. Through the information output from encoder, decoder restores details and size of the original image. U-Net adds multiple skip connections between the encoder and decoder, which can transfer the features of the shallow network to the deep network. Thus, it can help the decoding path recover the details of the image better. Since then, U-Net becomes a very popular segmentation network and is applied to medical imaging segmentation including cardiac MRI [4], cardiac CT [5], abdominal CT [6] segmentation and pulmonary nodule detection [7], and liver segmentation [8]. However, target organs vary greatly among different patients, so U-Net will rely extremely on multicascaded CNN. Cascade framework will make dense predictions of ROI, which will lead to repetitive extraction of similar low-level features and result in the waste of computational resource and the increase in model parameters. Therefore, the design of an efficient structure of deep CNN is very important.

So far, many improved versions of U-Net have been proposed. Azad et al. [9] proposed BCDU-Net. The most important changes are for feature extraction method and skipping connections. The original U-Net relies on multi-cascaded CNN, which results in the waste of computing resources and the increase of the number of parameters. U-Net is able to splice shallow features and deep features simply by using skip connection. In this paper, an extended version of U-Net is proposed, which uses recurrent residual convolutional neural networks with attention gate connection (R2AU-Net) for medical image segmentation. The contributions of this paper can be summarized as follows: firstly, R2AU-Net uses more attention gates (AGs) to deal with deep features and shallow features. The AGs use the depth feature map in the decoding path as a gating signal to modify the feature map generated in the coding process and suppress feature responses in the irrelevant background area, so as to highlight features that are useful for a specific task [10]. Secondly, R2AU-Net substitutes recurrent residual convolutional unit for the U-Net basic convolutional unit. In the recurrent residual convolutional unit, recurrent connection and residual connection [11] are added to each convolutional layer [12], thus not increasing network parameters. The use of recurrent connection can enhance the ability of integrating context information. Residual connection can help train deeper network [13, 14]. In addition, batch normalization [15] is used to accelerate the convergence speed of the network. R2AU-Net is evaluated on three datasets: retinal vascular segmentation (DRIVE dataset), skin lesion segmentation (ISIC 2018 dataset), and lung nodule segmentation (lung dataset).

## 2. Proposed Method

Multiple deep learning models are usually taken as functional modules to construct the new network. Inspired by U-Net, R2AU-Net is proposed in this paper. The network structure is shown in Figure 1, which takes advantage of four recently developed deep learning models. There are three differences between R2AU-Net and U-Net. Recurrent convolutional block with the residual unit is used in encoding and decoding paths. Secondly, the skipping connections are replaced by AGs to correct low-resolution features through deep features. The third point is that BN [15] is used to increase the stability of the neural network and speed up the convergence speed of the network in the upsampling process. BN can standardize data, obtain smaller regularization, reduce generalization error, and improve network performance [15].

**2.1. Encoding Path.** The encoding path of R2AU-Net contains four steps. Every step contains a recurrent residual convolutional unit, which consists of two  $3 \times 3$  convolution and adds recurrent connections to each convolutional layer to enhance the capability of integrating contextual information of the model. In addition, residual connections are added to develop more efficient and deeper models. Each time a recursive residual convolutional unit is passed, the

number of feature maps is doubled and the size becomes half of the original. The R2AU-Net model applies concatenation on feature mapping from encoding unit to decoding unit. The recurrent convolutional layers (RCL) in R2CL are performed according to the discrete time steps expressed by RCNN [12]. Suppose the  $u_l$  input sample in the  $l^{\text{th}}$  layer of the R2CL block and a pixel located at  $(i, j)$  in an input sample on the  $k^{\text{th}}$  feature map in the RCL.  $M_{ijk}^l(t)$  denotes the output at time step  $t$  and can be expressed as follows:

$$M_{ijk}^l(t) = (w_k^f)^T u_l^{f(i,j)}(t) + (w_k^r)^T u_l^{r(i,j)}(t-1) + b_k. \quad (1)$$

In formula (1),  $u_l^{f(i,j)}(t)$  and  $u_l^{r(i,j)}(t-1)$  represent standard convolutional layers and the input sample of the  $l^{\text{th}}$  RCL, respectively. The standard convolutional layer and the RCL of the  $k^{\text{th}}$  feature maps are, respectively, weighted by  $w_k^f$  and  $w_k^r$ , and  $b_k$  is the bias. The output of RCL is activated by standard ReLU function  $f$  as follows:

$$F(u_l, w_l) = f(M_{ijk}^l(t)) = (0, M_{ijk}^l(t)). \quad (2)$$

The output of the R2CL unit can be calculated as follows:

$$u_{l+1} = u_l + F(u_l, w_l), \quad (3)$$

where  $u_l$  is an input sample of R2CL layer.  $u_{l+1}$  is both the output of downsampling layer in encoding path and the output of upsampling layer in decoding path, respectively. The basic unit of U-Net convolution is shown in Figure 2(a), and the structure of the R2CL block is shown in Figure 2(b).

Formulas (1) and (2) describe the dynamic characteristics of RCL. When RCL is expanded into  $T$  time steps, the feedforward subnetwork with depth of  $T+1$  will be obtained. In this paper, RCL is expanded to two time steps; namely,  $T=2$ . RCL includes a single convolutional layer and two subsequence recurrent convolutional layers. RCL expansion structure is shown in Figure 3.

**2.2. Decoding Path.** Each step of the decoding path performs the upsampling operation of the output from the R2CL unit of the previous layer. With each upsampling operation, the number of feature maps will be halved and the size will be doubled. At the last layer of the decoding path, the size of the feature map is restored to the original size of the input image. The LNR layer in R2CL is replaced by BN layer, so that the input of each layer keeps the same distribution. In the process of training, the distribution of activation in each layer of neural network will lead to the decrease of training speed. Therefore, BN [15] is used to enhance stability of neural networks after sampling at each step. It improves the stability of the neural network by subtracting the batch mean and dividing the inputs according to the batch standard deviation. BN accelerates training speed and promotes the performance of network model.

The output  $u_l^{\text{up}}$  of BN layer is sent to AGs. R2AU-Net uses AGs to readjust the output features of the encoder before splicing the features on each resolution of the encoder with the corresponding features in the decoder. This module generates a gating signal which controls the importance of

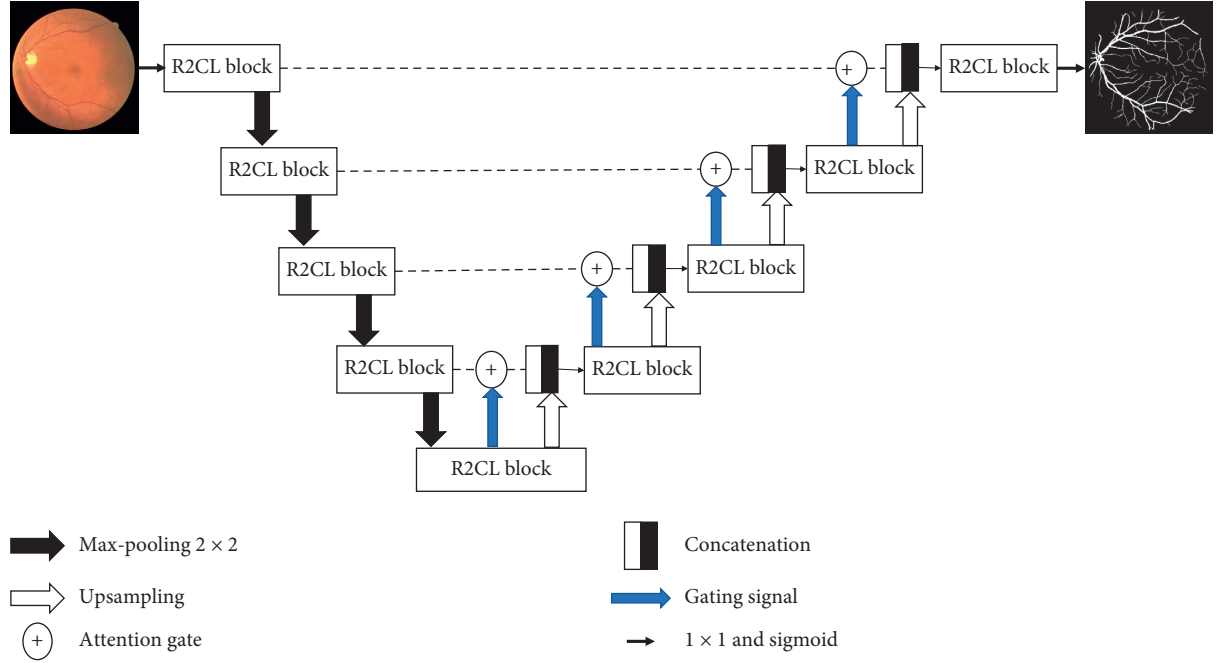


FIGURE 1: Convolutional encoding and decoding unit based on the recurrent residual convolutional layer (R2CL) and R2AU-Net structure with AG connection.

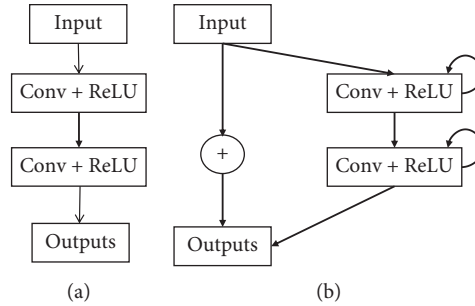


FIGURE 2: Two kinds of convolutional block. (a) The basic unit of the U-Net convolution. (b) R2CL block.

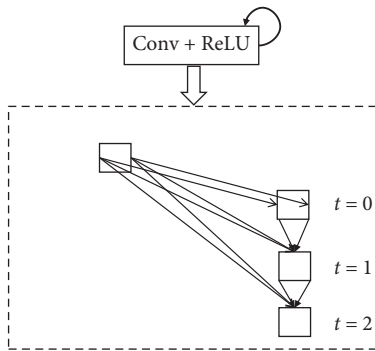


FIGURE 3: Unfolded recurrent convolutional units for  $T=2$ .

features at different locations. AGs gradually suppress feature responses unrelated to background regions without clipping ROI regions between networks.

Figure 4 shows the proposed additive attention map. Attention values are calculated for each pixel  $u_l^{f(i,j)}$ ,

respectively. For the convenience of representation and distinction,  $u_l^{\text{down}}$  and  $\hat{u}_l^{\text{up}}$  are denoted as  $u_l$  and  $g_l$ , respectively. The gating signal  $g_l$  determines the focus area per pixel. In order to obtain higher accuracy, the additive attention [16] is used to obtain the attention coefficient. The additive formula is as follows:

$$Q_L = \Phi(\delta_1(W_u u_l + W_g g_l + b_g)) + b_\Phi, \quad \alpha_l = \delta_2(Q_L), \quad (4)$$

where  $\delta_1$  and  $\delta_2$  denote ReLU and sigmoid activation functions, respectively,  $W_g$  is the weight, and  $b_g$  and  $b_\Phi$  are the bias. Wang et al. [17] used attention based on vector splicing. Linear transformations are calculated using a  $1 \times 1 \times 1$  convolution of tensor channel direction. Grid resampling of attention coefficients is achieved using trilinear interpolation. The update of AG parameter is trained according to backpropagation instead of using sampling-based update method [18].

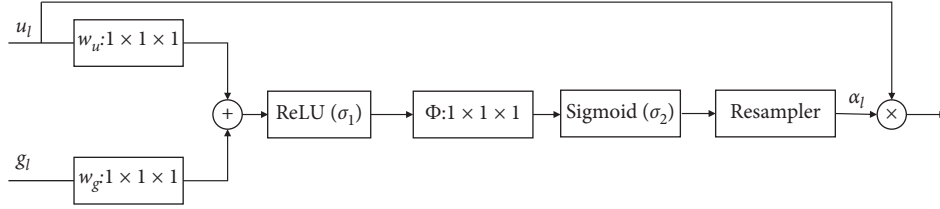


FIGURE 4: Schematic of the proposed additive AG.

Finally, the output of AGs is the multiplication of feature map and attention coefficient by elements, as shown in formula (5).

$$\hat{u}_l^{\text{up}} = u_l \times \alpha_l. \quad (5)$$

Attention coefficients tend to obtain large values in target organ regions and relatively small values in background regions, which can improve the accuracy of image segmentation.

### 3. Experimental Results

The experiments are performed on three datasets: DRIVE, ISIC 2018, and public dataset used in LUNA and the Kaggle Data Science Bowl 2017. The following performance indicators are adopted in this paper: True positive (TP), true negative (TN), false positive (FP), and false negative (FN), including accuracy (AC), sensitivity (SE), specificity (SP), and  $F_1$ -score ( $F_1$ ).

The accuracy rate is used to evaluate the accuracy of pixel classification and obtained by the following formula:

$$AC = \frac{TP + TN}{TP + TN + FP + FN}. \quad (6)$$

Sensitivity represents the proportion of samples that are predicted to be positive in the experimental results. It reflects the situation of positive samples. Sensitivity is calculated by the following formula:

$$SE = \frac{TP}{TP + FP}. \quad (7)$$

The specificity is calculated by the following formula:

$$SP = \frac{TN}{TP + FP}. \quad (8)$$

$F_1$ -score is used to measure the accuracy of binary classification model. It considers both the precision and recall of the classification model. It can be regarded as a harmonic average of model precision and recall.  $F_1$ -score is calculated by the following formula:

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (9)$$

In addition, receiver operating characteristics (ROC) curve and precision recall (PR) curve are used to compare the performance of each network more intuitively. The values of area under curve (AUC) of both ROC curve and PR curve of each network are calculated in this paper.

**3.1. Skin Lesion Segmentation.** The ISIC dataset is published by the International Skin Imaging Collaboration (ISIC), which contains 2594 dermoscopy images of common skin pigmentation lesions. All the images have been annotated by the recognized skin cancer specialists. These annotations include dermoscopic features, which are used to identify the type of skin lesions of the known global and local morphological elements in the image. During the experiment, 1815 images are used for training, 259 images are used for verification, and 520 images are used as testing sets. The size of each image in ISIC is  $700 \times 900$ . Firstly, the input image is preprocessed into  $256 \times 256$ . Training images include original images and ground truth images labeled by professional physicians. Figure 5 shows the segmentation results of ISIC. The first column is the original images, the second column is the ground truth images, and the third column shows the segmentation result of R2AU-Net. The first line of the following skin lesion segmentation figures shows that R2AU-Net can accurately segment the dark skin lesions and will not be affected by the hair around the lesions. For the less obvious light-colored lesions in the second line, R2AU-Net can also segment the lesions well. It can be found that R2AU-Net can accurately segment the image of skin lesions, which is almost identical to the ground truth image. Table 1 shows the comparison of segmentation result among R2AU-Net and other improved versions of U-Net through  $F_1$ -score, sensitivity, specificity, accuracy, and AUC value. R2AU-Net performs well in various indicators. For the dichotomy experiment, ROC curve and PR curve can intuitively compare the performance of each classifier. The AUC values of the ROC curve and PR curve are shown in Figures 6 and 7, respectively. The ROC curve tends to the upper left and the PR curve tends to the upper right, which shows the great performance of the segmentation model.

**3.2. Retinal Vascular Segmentation.** Images of DRIVE dataset are obtained from the diabetic retinopathy screening program in the Netherlands. Screening groups included 400 subjects aged between 18 and 24 who were diabetic. 40 color retina images are randomly selected. The doctor can diagnose, screen, treat, and evaluate a variety of cardiovascular and ophthalmic diseases, such as diabetes, hypertension, arteriosclerosis, and choroidal neovascularization, through the blood vessel segmentation from retinopathy images and the signs of retina blood vessel morphological properties. In the experiment, 20 samples are used for training and 20 samples are used for testing. The size of the original image is  $565 \times 584$ . Obviously, the number of samples is not enough to train a deep neural network model. Therefore, this paper randomly divided the input 20 training images into 190000



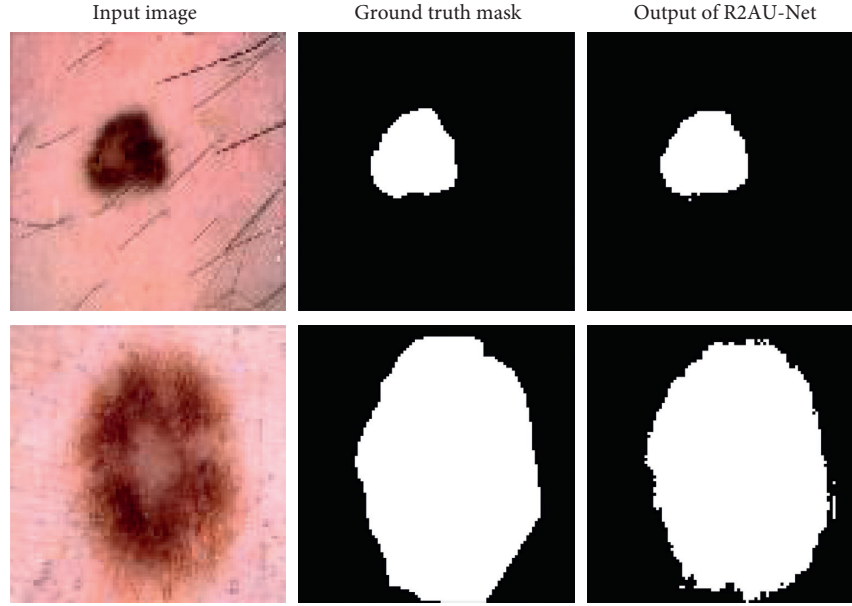


FIGURE 5: Segmentation results of skin lesions.

TABLE 1: Performance comparison of R2AU-Net and other networks on the ISIC dataset.

Methods	F1-score	Sensitivity	Specificity	Accuracy	AUC
SegNet [19]	0.7092	0.8056	0.8147	0.8121	0.8101
U-Net [3]	0.6954	0.7160	0.8636	0.8216	0.7898
Attention U-Net [20]	0.7541	0.7627	0.8966	0.8585	0.8297
RU-Net [21]	0.679	0.792	0.928	0.880	0.8374
R2U-Net [21]	0.691	0.726	0.971	0.904	0.8565
Attention Res U-Net	0.8545	0.8363	0.9519	0.9190	0.8941
R2AU-Net	0.8660	0.8214	0.9699	0.9277	0.8957

patches for training. Among them, 171000 patches are used for training set, and 19000 patches are used for testing set. The data size of the input network is  $64 \times 64$ . The segmentation result of the input image is shown in Figure 8. The first image is the original color image, the second image is the ground truth mask, and the third image is the segmentation result of the R2AU-Net output; most of the blood vessels at the end can still be segmented. Table 2 shows the results of comparative experiments on the DRIVE dataset, including F1-score, sensitivity, specificity, accuracy, and AUC value. From the experimental results, the performance of R2AU-Net is better than the traditional methods and the original U-Net. Figures 9 and 10 show the AUC value of both ROC curve and PR curve of each network.

**3.3. Lung Segmentation.** Public datasets used in LUNA and the Kaggle Data Science Bowl 2017 are provided by the National Cancer Research Center of the United States. This dataset consists of 2D and 3D images. The original size of lung CT images is  $512 \times 512$ , the number of which is 267. 134 images are used for training, 54 images are used for verification, and 79 images are used for testing set. Figure 11

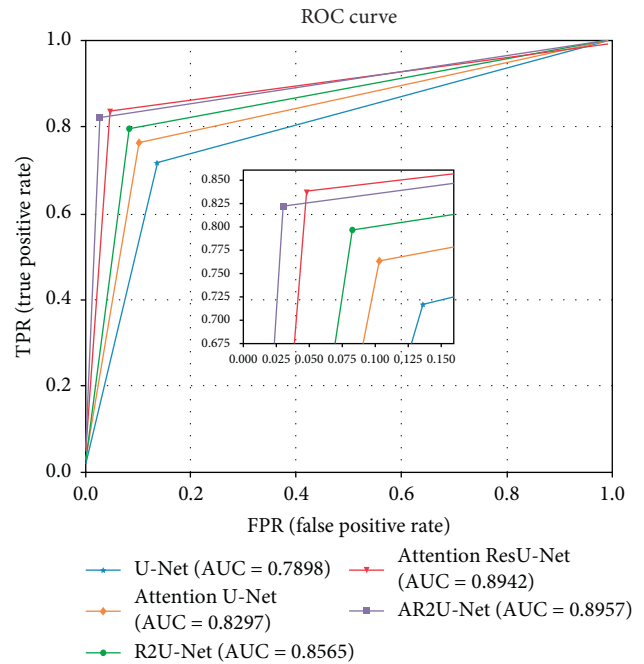


FIGURE 6: ROC curves of skin lesion segmentation of five models.

shows the segmentation results of R2AU-Net on the lung dataset. The first column is the input image, the second column is the ground truth mask, and the third column is the lung segmentation image of R2AU-Net. The third image of the first row shows that very small CT image of lung region is able to be segmented. The segmentation results of R2AU-Net are basically the same as the ground truth image. Table 3 shows the performance comparison of R2AU-Net and other improved versions of U-Net. Figures 12 and 13 show the AUC value of both ROC curve and PR curve of each network.

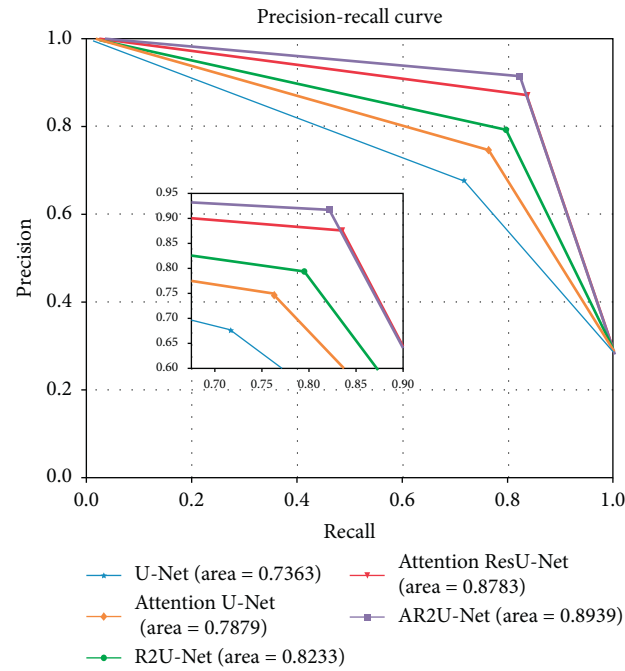


FIGURE 7: PR curves of skin lesion segmentation of five models.

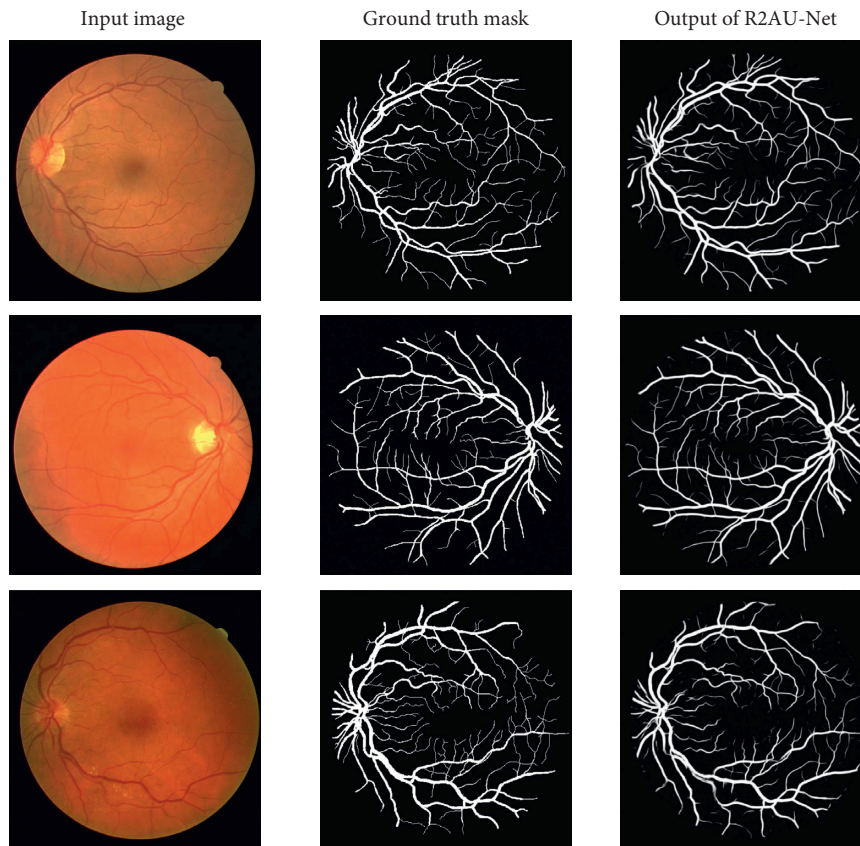


FIGURE 8: Retinal blood vessel segmentation results.

TABLE 2: Performance comparison of R2AU-Net and other networks on the DRIVE dataset.

Methods	F1-score	Sensitivity	Specificity	Accuracy	AUC
Hybrid features [22]	—	0.7252	0.9798	0.9474	0.9648
Trainable COSFIRE filters [23]	—	0.7655	0.9704	0.9442	0.9614
Three-stage filtering [24]	—	0.7250	0.9830	0.9520	0.9620
Deep model [25]	—	0.7763	0.9768	0.9495	0.9720
Cross-modality [26]	—	0.7569	0.9816	0.9527	0.9738
SegNet [19]	0.7992	0.7419	0.9833	0.9526	0.9752
U-Net [3]	0.8155	0.7908	0.9783	0.9544	0.9775
Attention U-Net [20]	0.8003	0.7272	0.9868	0.9538	0.9771
RU-Net [21]	0.8180	0.7999	0.9772	0.9547	0.9773
R2U-Net [21]	0.8187	0.7980	0.9779	0.9550	0.9775
R2AU-Net	0.8213	0.8036	0.9777	0.9555	0.9790

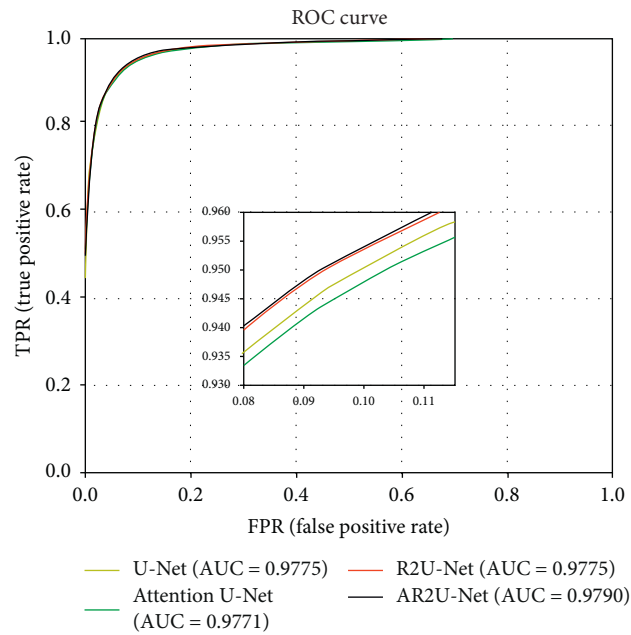


FIGURE 9: ROC curve of retinal blood vessel segmentation of four models.

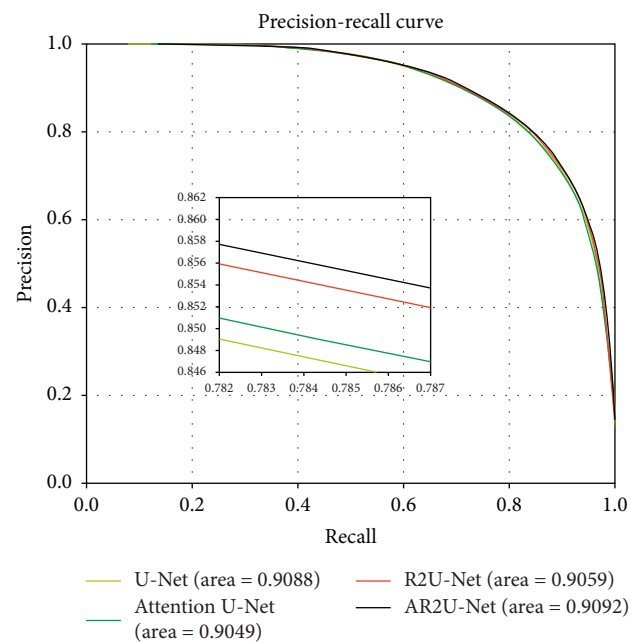


FIGURE 10: PR curve of retinal blood vessel segmentation of four models.

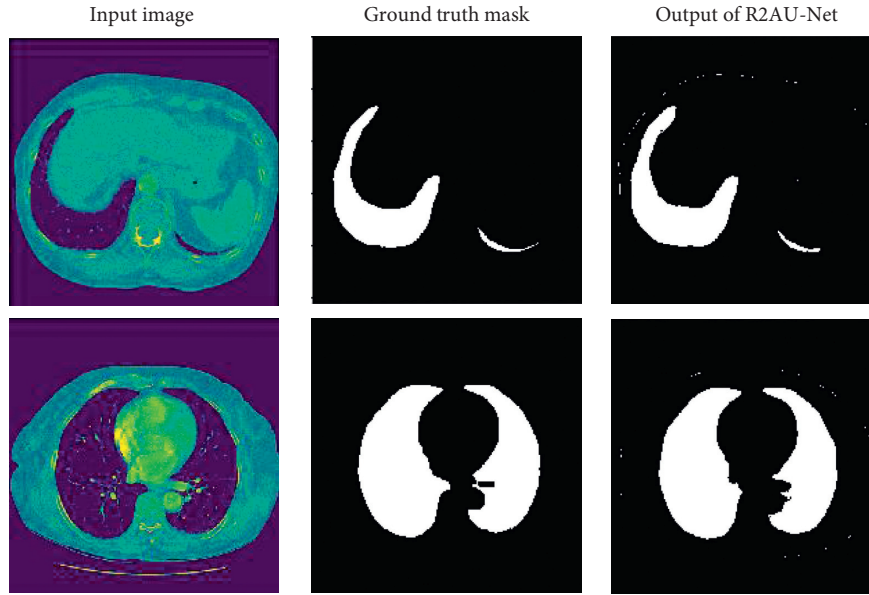


FIGURE 11: Results of the lung region segmented by R2AU-Net.

TABLE 3: Performance comparison of R2AU-Net and other networks on the lung dataset.

Methods	F1-score	Sensitivity	Specificity	Accuracy	AUC
SegNet [19]	0.9515	0.9240	0.9957	0.9820	0.9598
U-Net [3]	0.9714	0.9536	0.9977	0.9892	0.9756
Attention U-Net [20]	0.9780	0.9911	0.9915	0.9915	0.9914
RU-Net [21]	0.9638	0.9734	0.9866	0.9836	0.9800
R2U-Net [21]	0.9809	0.9842	0.9946	0.9926	0.9894
Attention Res U-Net	0.9811	0.9862	0.9942	0.9927	0.9902
R2AU-Net	0.9868	0.9874	0.9968	0.9950	0.9921

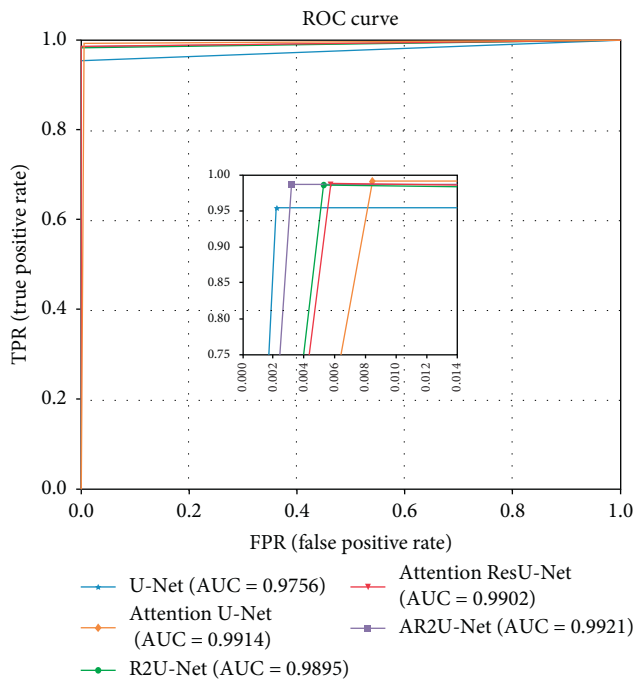


FIGURE 12: ROC curves of the lung region segmented by five models.

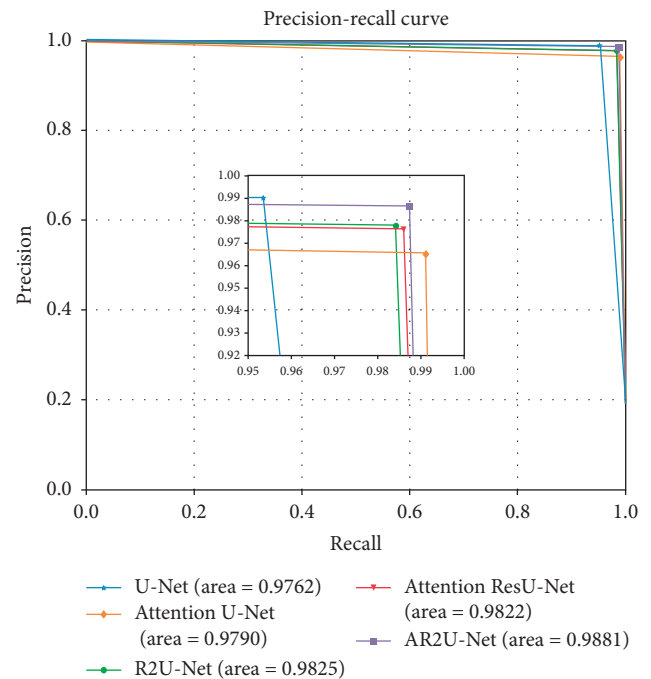


FIGURE 13: PR curves of the lung region segmented by five models.



## 4. Conclusion

In this paper, R2AU-Net is proposed for medical image segmentation. The recurrent residual convolutional block is used to enhance the ability of capturing context information, and AGs are added in the skip connections. Attention gates use deep features of decoding path as gating signal to modify shallow features and suppress feature response of background area, so that the network can obtain more accurate segmentation results. Moreover, BN is used to accelerate the convergence speed and stability of the network in the upsampling process. The experimental results of three datasets show that R2AU-Net has good performance in medical image segmentation.

## Data Availability

The following are the links to the datasets used in this article: skin lesion segmentation (<https://challenge2018.isic-archive.com/>), retinal vascular segmentation (<https://drive.google.com/file/d/17wVfELqgwb4Q02GD247jYjq6lwB0l6/view>), and lung nodule segmentation (<https://www.kaggle.com/kmader/finding-lungs-in-ct-data/data>).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] D. C. Cirean, A. Giusti, L. M. Gambardella et al., "Deep neural networks segment neuronal membranes in electron microscopy images," *Advances in Neural Information Processing Systems*, vol. 25, pp. 2852–2860, 2012.
- [2] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2015.
- [3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Munich, Germany, October 2015.
- [4] M. Khened, V. A. Kollerathu, and G. Krishnamurthi, "Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of Classifiers," *Medical Image Analysis*, vol. 51, pp. 21–45, 2019.
- [5] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Multi-label whole heart segmentation using cnns and anatomical label configurations," in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWH Challenges*, pp. 190–198, Springer, Berlin, Germany, 2018.
- [6] H. R. Roth, H. Oda, Y. Hayashi et al., "Hierarchical 3D fully convolutional networks for multi-organ segmentation," 2017, <https://arxiv.org/abs/1704.06382>.
- [7] F. Liao, M. Liang, Z. Li et al., "Evaluate the malignancy of pulmonary nodules using the 3D deep leaky noisy-or network," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 30, no. 11, pp. 3484–3495, 2017.
- [8] Z. Liu, Y.-Q. Song, V. S. Sheng et al., "Liver CT sequence segmentation based with improved U-net and graph Cut," *Expert Systems with Applications*, vol. 126, pp. 54–63, 2019.
- [9] R. Azad, M. Asadi-Aghbolaghi, M. Fathy et al., "Bi-directional ConvLSTM U-Net with densely connected convolutions," in *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition*, pp. 406–415, Seoul, Republic of Korea, October 2019.
- [10] N. Abraham and N. M. Khan, "A novel focal tversky loss function with improved attention U-Net for lesion segmentation," in *Proceedings of the IEEE 16th International Symposium on Biomedical Imaging (ISBI)*, pp. 683–687, Venice, Italy, April 2019.
- [11] K. He, X. Zhang, S. Ren et al., "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [12] M. Liang and X. Hu, "Recurrent convolutional neural network for object recognition," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3367–3375, Boston, MA, USA, June 2015.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *Computer Vision—ECCV 2016*, Springer International Publishing, Berlin, Germany, pp. 630–645, 2016.
- [14] N. Ibtehaz and M. S. Rahman, "MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Networks*, vol. 121, pp. 74–87, 2020.
- [15] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, pp. 448–456, Lille, France, July 2015.
- [16] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, <https://arxiv.org/abs/1409.0473>.
- [17] X. Wang, R. Girshick, A. Gupta et al., "Non-local neural networks," 2017, <https://arxiv.org/abs/1711.07971>.
- [18] V. Mnih, N. Heess, A. Graves et al., "Recurrent models of visual attention," *Advances in Neural Information Processing Systems*, pp. 2204–2212, MIT Press, Cambridge, MA, USA, 2014.
- [19] 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.
- [20] O. Oktay, J. Schlemper, L. L. Folgoc et al., "Attention U-Net: learning where to look for the pancreas," 2018, <https://arxiv.org/abs/1804.03999>.
- [21] M. Z. Alom, M. Hasan, C. Yakopcic et al., "Recurrent residual convolutional neural network based on u-net (R2U-net) for medical image segmentation," 2018, <https://arxiv.org/abs/1802.06955>.
- [22] E. Cheng, L. Du, Y. Wu, Y. J. Zhu, V. Megalooikonomou, and H. Ling, "Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features," *Machine Vision and Applications*, vol. 25, no. 7, pp. 1779–1792, 2014.
- [23] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable COSFIRE filters for vessel delineation with application to retinal images," *Medical Image Analysis*, vol. 19, no. 1, pp. 46–57, 2015.
- [24] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Blood vessel segmentation of fundus images by major vessel extraction and subimage classification," *IEEE Journal of*

- Biomedical and Health Informatics*, vol. 19, no. 3, pp. 1118–1128, 2015.
- [25] P. Liskowski and K. Krawiec, “Segmenting retinal blood vessels With\_newline deep neural networks,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2369–2380, 2016.
- [26] Q. Li, B. Feng, L. P. Xie et al., “A Cross-modality learning approach for vessel segmentation in retinal images,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 109–118, 2015.