IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 15 September 2023, accepted 9 November 2023, date of publication 20 November 2023,
date of current version 8 December 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3334621*

RESEARCH ARTICLE

# EU-Net: Efficient Dense Skip-Connected Autoencoder for Medical Image Segmentation

**LIZHUANG LIU[1], JIACUN QIU[1], KE WANG[2], QIAO ZHAN[3], JIAXI JIANG[1], ZHENQI HAN[1], JIANXIN QIU[2], TIAN WU[4], JINGHANG XU[4], AND ZHENG ZENG[4]**

[1]Shanghai Advanced Research Institute, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Pudong, Shanghai 201210, China
[2]Radiology Department, Peking University First Hospital, Xicheng, Beijing 100034, China
[3]Department of Infectious Diseases, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu 210029, China
[4]Department of Infectious Diseases, Peking University First Hospital, Xicheng, Beijing 100034, China

Corresponding author: Zhenqi Han (hanzq@sari.ac.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61972007; in part by the Peking University Medicine Fund of Fostering Young Scholars' Scientific and Technological Innovation under Grant BMU2021PYB021; and in part by the Science and Technology Service Network Plan of Dongwan Chinese Academy of Sciences (STS) under Grant 20211600200102.

**ABSTRACT** This paper introduces EU-Net, an efficient and enhanced U-Net-like architecture designed for medical image segmentation. It comprises a lightweight encoder and decoder connected through dense skip-connections. To further improve the robustness of the EU-Net, chain EU-Net is proposed. Chain EU-Net is based on a streamlined architecture that uses multiple EU-Net to build light weight deep neural networks by dense skip-connection. Compared to traditional segmentation algorithms such as U-Net and its variants, our neural network structure possesses both lightweight and stability simultaneously. EU-Net and chain EU-Net are evaluated on three typical medical image segmentation tasks: GLAS (Gland segmentation) dataset, RITE (Retinal Images Vessel Tree Extraction) dataset and LiTS (Liver Tumor Segmentation Challenge) dataset. In addition, we used PUFH (Peking University First Hospital) dataset. Experimental results show that the proposed methods achieve state-of-the-art performance with very few parameters.

**INDEX TERMS** Medical image segmentation, deep learning, dense skip-connection, efficiency.

## I. INTRODUCTION

Medical image segmentation is a significant work, which provides the shape contour and range of organ regions, to help doctors make better clinical decisions in many areas, e.g., vessel detection, gland disease segmentation and tumour diagnosis.

Different to common RGB image, medical images usually suffer from high noise, weak contrast and blurred boundaries, making it difficult to extract discriminative features. Early methods often based on edge detection [1], [2], activate contours [3], [4] and graph theory [5], [6], etc. Recently, convolutional neural networks (CNN) appeal more research attraction due to its excellent discriminating feature extraction, such as U-Net [7], Res-UNet [8], UNet ++ [9] and so on. Among those CNN methods, U-Net and its variants achieved outstanding results by applying the encoder-decoder architecture and skip-connection structure in medical image segmentation. CNN's methods are successful for two main reasons. The first is that multiple encoder-decoder structures are adopted to extract different levels of features by down sampling. Specifically, the initial layers of encoder in U-Net capture low level features, such as edges and small anatomical structures, while deeper layers extract semantic features. The second is that U-Net and its variants use skip connection to fusion different level of features among adjacent layers or even all layers. The structure of skip-connection plays an important role in preventing the loss of semantic information and gradient information.

Medical image segmentation still faces great challenges in practical application. First, U-Nets and its variants are susceptible to noise, especially in some complex contexts. Second, the skip-connection causes the large semantic gap between low-feature map and high-feature map, leading information loss. Third, U-Net and its variants have millions of parameters, which makes them hard to embed into

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan-Li Sun.

micro hardware. Striking a balance between performance and lightweight is an important challenge. In addition, acquiring medical image data typically requires a significant amount of time and financial resources, along with specialized knowledge for accurately annotating structures within the images. Therefore, many medical image datasets are relatively small, posing challenges for training deep learning models. In medical images, different structures may appear at varying frequencies. For example, in tumor segmentation, the frequency of positive samples (tumor regions) is often much lower than that of negative samples (non-tumor regions). This leads to class imbalance issues that require special attention and handling. In medical applications, there is often a need to deploy models on embedded devices such as mobile devices or medical instruments.

In this paper, in order to overcome the above problems, we propose EU-Net, a lightweight dense skip-connected similar U-Net architecture for medical semantic segmentation, which explicitly exploits the inception and dense skip-connected encoder-decoder structures to reduce semantic loss between low features and high features. To future enhance discriminating feature extraction, we cascade multiple EU-Net for segmentation which is called chain EU-Net. To verify the effectiveness of our approach, we conduct extensive experiments on GLAS datasets and RITE datasets respectively, whose results show that our methods achieve state-of-the-art performance with few parameters. In conclusion, the main contributions of this paper are as follows:

1) An efficient and dense skip-connected autoencoder called EU-Net is proposed for medical image segmentation, which embraces efficient to extract features.
2) We propose a chain EU-Net to enhance the robustness, which consists of multiple lightweight dense skip-connected autoencoder.
3) We conduct extensive experiments to empirically analyses the proposed methods. The results on four dataset show the effectiveness of the methods.
4) We collect a new dataset for liver occupying lesions based on clinical data from Peking University First Hospital.

The rest of the paper is organized as follows. Section II reviews the related works on deep learning for medical semantic segmentation. Section III describes the basic ideal of proposed EU-Net and chain EU-Net in this study. The detailed results are presented and discussed in Section IV and Section V. We conducted ablation experiments in Section VI. Section VII studied the combination of EU-Net and attention. Finally, we conclude our work in Section VIII.

This study was approved by the Ethics Committees of Peking University First Hospital.

## II. RELATED WORK

The main goal of medical image segmentation is to obtain the pixel level prediction of human tissue. Medical image segmentation is a very challenging task because of the high noise, low contrast and blurred boundary of medical image. In this section, we briefly demonstrated methods based deep learning for medical image segmentation and structures related to our model.

Long et al. [10] proposed fully convolutional networks (FCN) to view semantic segmentation tasks as per-pixel classification. U-Net [7], derived from FCN, is a groundbreaking structure as it uses skip connections to concatenate encoders and decoders to make low semantic features to be learned in deeper layers. At the same time, the $2 \times 2$ pooling or $2 \times 2$ interpolate makes U-Net learn different levels of features. After U-Net came out, the researchers make great effort to modify U-Net-based structure for better performance. Inspired by the ResNet [11], Res-UNet [8] replaces each submodule of U-Net with a residual structure. UNet++ [9], proposed by Zhou et al., consists of an encoder and decoder that are connected through a series of nested dense convolutional blocks. Taking advantage of full-scale skip connections and deep supervisions, Huang et al. proposed UNet 3+ [12]. UXNet [13] proposes a novel architecture, which makes model have the ability to search for the best feature aggregation strategy; DeepLabV3 [14] uses Res-Net as encoders and uses depth-wise separable convolution to decoder modules; V-Net [15] improves U-Net to be able to recognize 3D medical scans; KiU-Net [16], [17] developed a structure that was good at detecting small edges with fewer parameters. Although those structures have excellent performance, it is difficult for them to maintain the generalization ability and robustness of the model, while further reduce the number of parameters. Inception [18] constructs are an important way to improve network performance by fuse multiple scales feature information. In inception, firstly, the number of channels is reduced by $1 \times 1$ convolution to aggregate the information, and then feature extraction and pooling of different scales are carried out to obtain information of multiple scales. Finally, the features are superimposed and output.

## III. METHODS

Fig.1 shows a general overview of the chain EU-Net. The method is composed of four lightweight EU-Net blocks by dense skip connection. All EU-Net blocks have same structure which is a lightweight and dense skip-connected encoder-decoder. In this section, we start by describing the proposed EU-Net block. Then, the method of chain EU-Net is presented.

### A. EU-Net: AN EFFICIENT U-Net-LIKE STRUCTURES

As illustrated in Fig.2, the EU-Net is an encoder-decoder model that is similar with U-net. Compared to U-net, the EU-Net has three aspects of transformation. First, inspired by inception, five-parallel inceptions were applied to extract 2,3,4,5 and 6 channel feature maps respectively from a medical image. Then concatenate them to a 20-channel

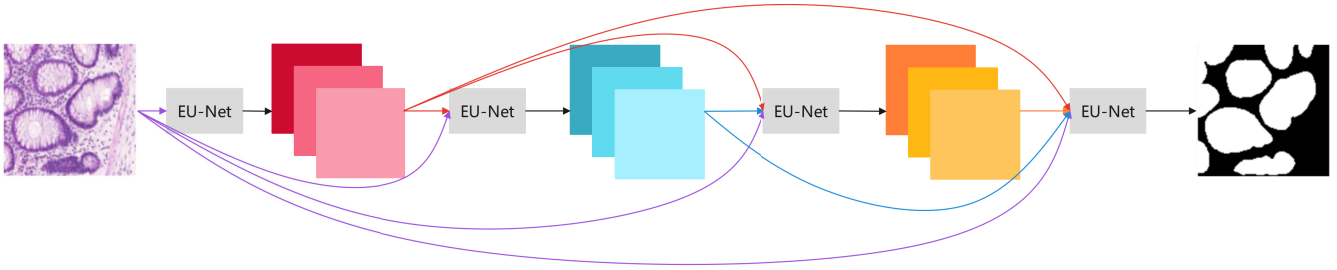


**FIGURE 1.** The overview of the chain EU-Net.

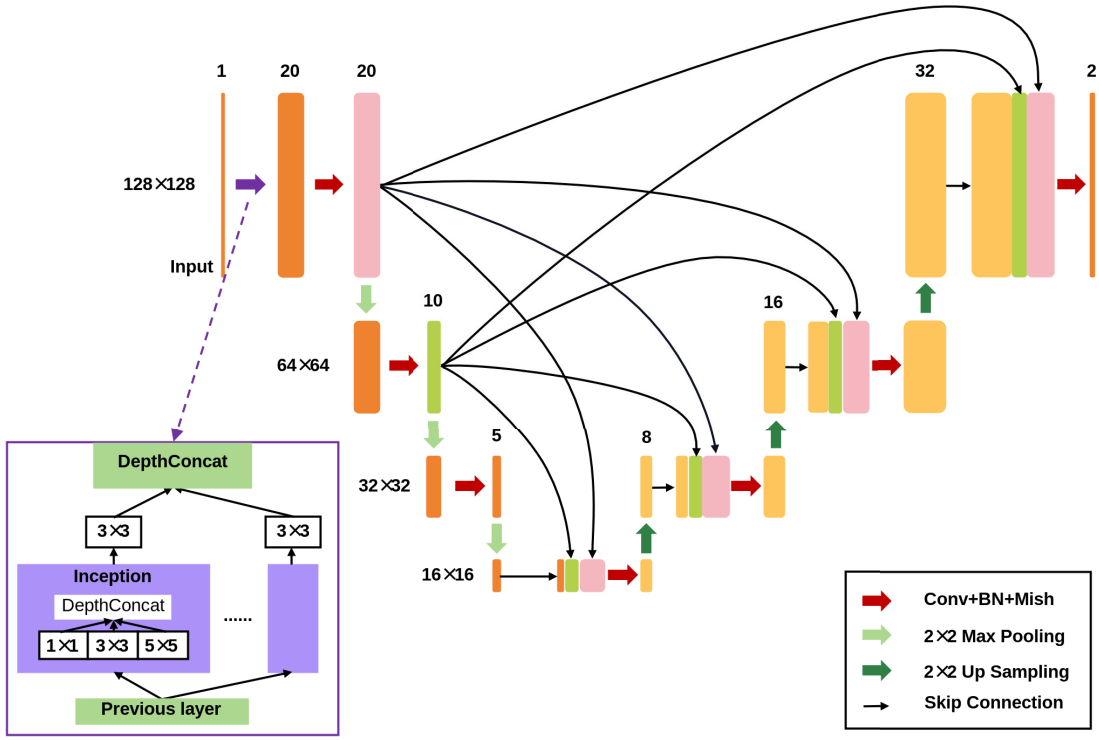


**FIGURE 2.** Lightweight and dense skip-connected EU-Net.

feature map. The 20-channel feature map contains abundant low-level features. All five-parallel inception blocks output 2,3,4,5 and 6 feature maps respectively by a convolution operation.

Second, a lightweight and dense skip-connected encoder-decoder structure was cascaded after concatenating the 20-channel feature map. The encoder structure consists of three encoder layers and corresponding max-pooling operators. Each encoder operator includes convolution, batch normalization [15] and mish activation function [16].

The size of convolutional filter is $1 \times 1$ in first encoder layer and $3 \times 3$ in other encoder layers. Mish activation function is a novel smooth and non-monotonic neural activation function which can be defined as (1), where $x$ represents the input to the activation function.

$$f(x) = x \cdot tanh(ln(1 + e^x)) \tag{1}$$

The spatial dimension of the feature maps is reduced via a max-pooling operator. On decoder structure, it consists of three decoder layers and corresponding upsampling operators. Each decoder includes concatenation, convolution, batch normalization and mish activation. The concatenation means that the last layer feature maps are concatenated channel-wisely with feature maps of similar size from first and second encoder layer by global pooling.

Third, in order to reduce the large semantic gap between low-feature map and high-feature map, at the end of encoder-decoder, a decoder layer was attached and transferred the concatenated feature maps to 2-channel feature maps.

### B. CHAIN EU-NET

In order to extract more discriminating features from medical images, we concatenate multiple lightweight and dense

skip-connected EU-Net to the deeper network by dense skip connection, as shown in Fig.1. The deeper network named chain EU-Net whose shapes just like a chain. Specifically, we regard a single EU-Net as a block and concatenate the origin image with all outputs of the block as a new input for the next block. By repeating this operation, the network becomes deeper and acquires better discriminating feature extraction ability. Since each autoencoder block contains extremely small number of parameters, the total number of parameters of the network still remain relatively small.

### C. LOSS

Each lightweight EU-Net get an output and has a binary cross-entropy loss. In our chain EU-Net, the final loss is the sum of all lightweight EU-Net loss. So, the final loss for our chain EU-Net can be defined as (2):

$$Loss = \sum_{i=1}^{N} l_i \tag{2}$$

$l_i$ is a the binary-cross entropy loss which represent difference between the prediction of $i$-th lightweight EU-Net and the ground truth. N means that the chain EU-Net have N lightweight EU-Net. The loss of the lightweight EU-Net shows as follows:

$$\begin{aligned} l = -\frac{1}{wh} \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} & (p(x,y) \log(\hat{p}((x,y))) \\ & + (1 - p(x,y)) \log(1 - \hat{p}(x,y)) \end{aligned} \tag{3}$$

where $w$ and $h$ are the dimensions of image, $p(x, y)$ corresponds to the image and $\hat{p}(x, y)$ represents the output prediction at specific pixel location $(x, y)$.

## IV. EXPERIMENTS

In this section, to evaluate the effectiveness of the proposed lightweight EU-Net and chain EU-Net, several experiments were conducted on the GLAS, RITE, LiTS and PUFH datasets.

### A. DATASET

Gland Segmentation (GLAS) dataset [19] contains 165 images of different sizes derived from 16 H&E-stained histological sections of stage T3 or T4 colorectal adenocarcinoma. Because each section comes from different patients and is processed on different occasions, the data exhibits high level of inter-subject variability. The pictures of GLAS are divided into 80 train images, 5 validation images, and 80 test images.

Retinal Images vessel Tree Extraction (RITE) dataset [20] contains 40 images for retinal nerve segmentation. Retinal nerve segmentation has great effects on diagnosis and treatment of various ophthalmic diseases. RITE dataset was split into 18 for training, 2 for validating and 20 for testing. To fit the model, we converted the label images of the RITE dataset from grayscale images to binary images, converting pixels with grayscale values greater than 32 to 1, and pixels with grayscale values less than 32 to 0.

Liver Tumor Segmentation(LiTS) datasets contains 131 contrast-enhanced 3D abdominal training CT scans. The tumor segmentation in LiTS is very challenging because of tumor's heterogeneous and diffuse shape. In LiTS, the ground truth segmentation provides three different labels: liver, tumor and background. In LiTS, liver segmentation task consider background as negative class and others as positive class. Tumor segmentation task consider tumor as positive class and others as negative class. We split these CT scans at 8:1:1 ratio into train set, validation set and test set respectively. In the same time, we remove irrelevant useless details which don't contain tumor.

Peking University First Hospital(PUFH) datasets contain 200 3D abdominal CT scans. In PUFH, the ground truth segmentation provides three different labels: liver, liver space-occupying lesions and background. We split these CT scans at 8:1:1 ratio into train set, validation set and test set respectively and remove irrelevant useless details.

### B. EVALUATION METRICS

The Dice [21] and Jaccard [22] coefficient was used as the evaluation metric for each segmentation result, which is expressed as follows:

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

$$Jaccard = \frac{TP}{TP + FP + FN}$$

where TP represents the true positives; FP represents the false positives; FN represents the false negatives.

### C. IMPLEMENTATION DETAILS

The proposed method is implemented by pytorch-1.10 on ubuntu 18.04 System with NVIDIA RTX 8000 GPUs. To compare fairly with other methods, we pre-process the images by resizing them to 128 × 128 pixels. In training phase, the lightweight EU-Net is initialized with random metric. We use the Adam optimizer to train the network with parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The learning rate is 0.01. The weight decay is 0.04. and the batch size is set as 8. We use grid search to optimize hyperparameters. Training stops when the network is converged.

## V. DISCUSSION WITH FOUR DATASETS

We compare the results of the proposed method with that of previous outstanding deep learning-based methods on the five medical image segmentation tasks. Tabel.1 shows the results of ligthweight EU-Net and Chain EU-Net on five image segmentation tasks. It can be clearly seen that our lightweight Chain EU-Net architectures achieve best performance with fewer parameters than other networks. At the same time,

**TABLE 1. Comparison of our methods and other state-of-the-art methods on the four datasets mentioned above. Our results and other best results are highlighted on bold. Where LiTS-T means tumour segmentation task on LiTS dataset, and LiTS-L means liver segmentation task on LiTS dataset.**

| Dataset<br>Methods | GLAS<br>Dice | RITE<br>Dice | LiTS-T<br>Dice | LiTS-L<br>Dice | PUFH<br>Dice | Parameters | FLOPs |
|---|---|---|---|---|---|---|---|
| U-Net [7] | **0.8777** | 0.7694 | 0.8733 | 0.9686 | 0.8745 | 31M | 12.6B |
| Resnet34-UNet [8] | 0.8497 | 0.4456 | 0.8453 | 0.9592 | 0.8837 | 23M | 1.9B |
| UNet++ [9] | 0.8671 | 0.7661 | **0.8795** | 0.9655 | 0.8880 | 9M | 8.6B |
| Attention U-Net [23] | 0.8733 | **0.7707** | 0.8589 | **0.9714** | 0.8859 | 34M | 16.6B |
| SegNet [19] | 0.8746 | 0.7000 | 0.8614 | 0.9632 | 0.8295 | 29M | 457.2B |
| R2UNet [24] | 0.7717 | 0.7005 | 0.6877 | 0.9247 | 0.8031 | 39M | 38.1B |
| FCN8s [10] | 0.8482 | 0.4456 | 0.6515 | 0.9564 | 0.8598 | 18M | 6.3B |
| Mobile-UNet [25] | 0.8411 | 0.5375 | 0.8668 | 0.9636 | 0.8791 | 15M | 10.4B |
| ELU-Net [26] | 0.8654 | 0.7338 | 0.8357 | 0.9616 | **0.8897** | 0.80M | 1.2B |
| KiU-Net [17] | 0.8601 | 0.7576 | 0.8740 | 0.9705 | 0.8607 | 0.29M | 69.6B |
| EU-Net(ours) | 0.8547 | 0.7651 | 0.8432 | 0.9687 | 0.9096 | 0.03M | 0.17B |
| Chain EU-Net(ours) | **0.8806** | **0.7741** | **0.8882** | **0.9724** | **0.9208** | 0.12M | 0.68B |

EU-Net performs better than a considerable number of other networks with the minimal parameters.

On the GLAS dataset, Chain EU-Net performs the best among all networks, while U-Net excels among methods other than ours. As Tabel.1 shows, Chain EU-Net uses only 0.39% of the parameters of U-Net's but still improves performance slightly, and EU-Net, with only 0.1% of the parameters, loses only 2.6% of its performance. Specifically, compared to EU-Net, Chain EU-Net improves the Dice and Jaccard accuracy on GLAS from 0.8547 to 0.8806, achieving 3.03% increase respectively.

Fig.3 shows the medical segmentation images of U-Net, KiU-Net, our EU-Net and chain EU-Net. Among the methods other than ours, KiU-Net has the fewest parameters, while U-Net exhibits the best performance. It can be seen that our results are more distinguishable than other networks. Especially for the prediction of red block, our results are more consistent with the original image, and there are fewer areas of wrong prediction.

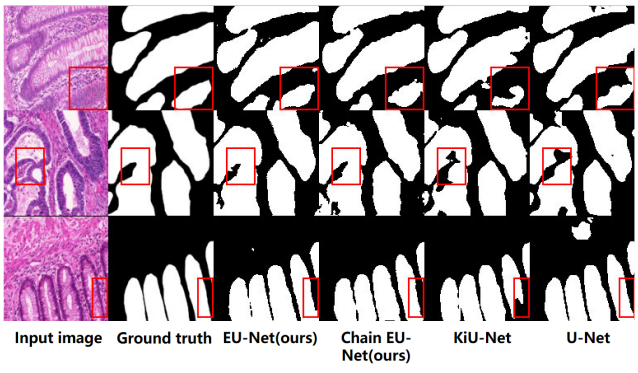


**FIGURE 3. Comparison of qualitative results between SegNet, KiU-Net and our method in GLAS dataset.**

The RITE dataset exhibits highly complex textures, making it exceptionally suitable for testing the network's generalization capability. As Tabel.1shows, the Attention U-Net achieves the best performance among methods other than ours on RITE dataset. Chain EU-Net, with only 0.35% of the parameters of Attention U-Net, achieves superior results compared to Attention U-Net. EU-Net achieves a 99.27% performance with only 0.09% of the parameters of Attention U-Net.

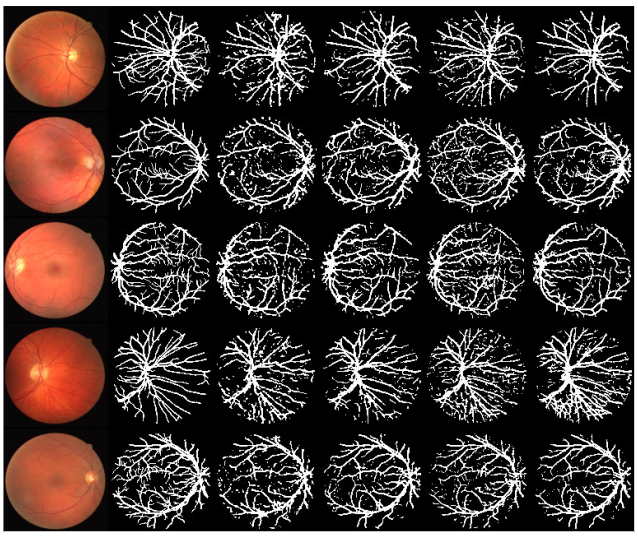


**FIGURE 4. Comparison of qualitative results between Attention U-Net, KiU-Net and our method in RITE dataset.**

Fig.4 shows qualitative results between Attention U-Net, KiU-Net and our method in RITE dataset. For the RITE dataset, our results have better continuity and consistency in the task of vascular segmentation. In addition, the prediction results of our EU-Net and chain EU-Net have less breakpoints and noise points.

The tasks of liver segmentation, tumor segmentation, and liver space-occupying segmentation serve as invaluable aids to medical professionals, enhancing their diagnostic capabilities and delivering substantial practical utility. The experimental results show that Chain EU-Net also has excellent performance in these tasks. Excluding Chain EU-Net, U-Net++, Attention U-Net, and ELU-Net excel in the

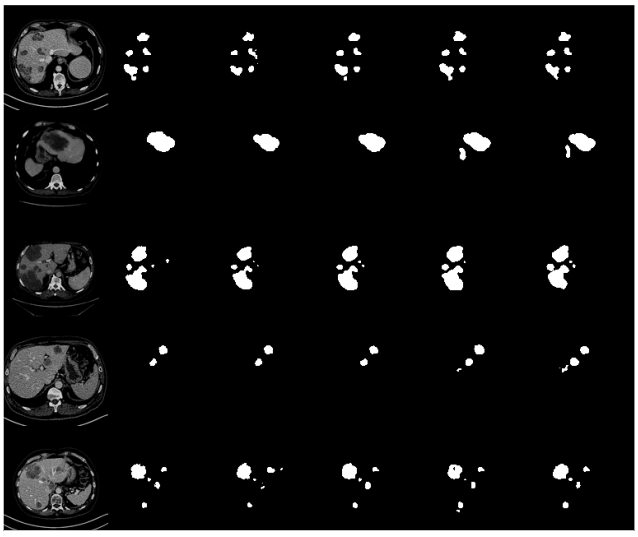


**FIGURE 5.** The tumor segmentation result comparison with other methods on LiTS dataset.

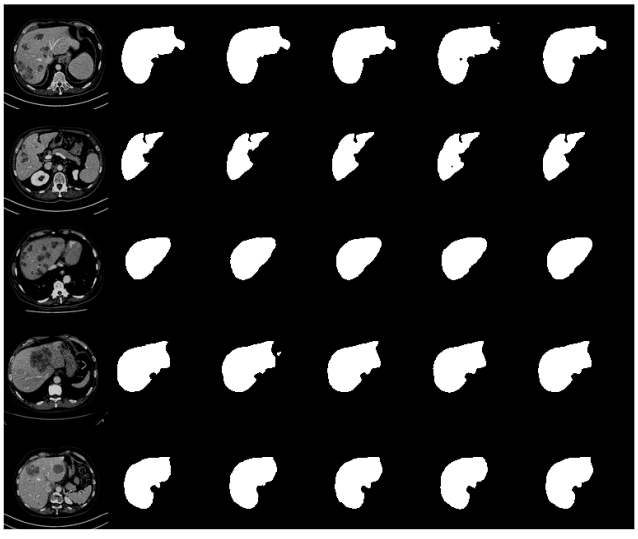


**FIGURE 6.** The liver segmention result comparsion with other methods on LiTS dataset.

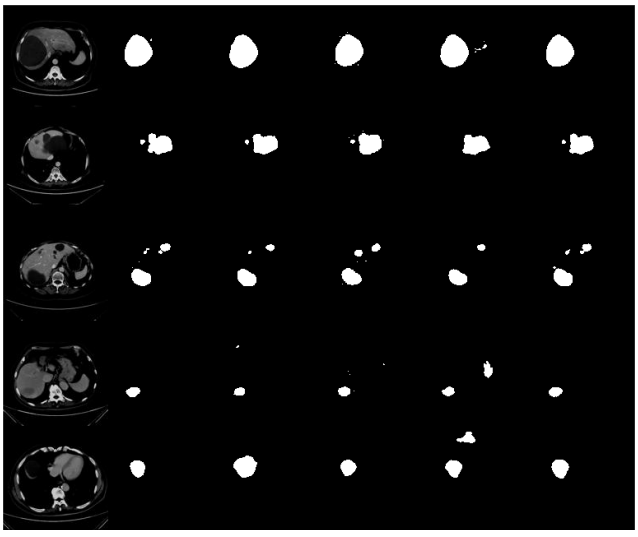


**FIGURE 7.** The liver space-occupying lesions segmentation result comparsion with other methods on PUFH dataset.

LiTS tumor segmentation task, LiTS liver segmentation task, and liver space-occupying segmentation task, respectively. Chain EU-Net achieves superior final results with parameter counts of 1.33%, 0.35%, and 15%, respectively. It's worth mentioning that, on the PUFH dataset, EU-Net outperforms any other method except for Chain EU-Net even with significantly fewer parameters.

Fig.5 and Fig.6 shows the tumour results and liver segmentation results comparsion with other methods on LiTS dataset. Fig.7 shows the results of liver sapce-occupying lesion segmentation on PUFH dataset. It is abundantly clear that our methodology showcases outstanding performance across a diverse array of segmentation tasks, encompassing the precise delineation of small tumors, the intricate segmentation of complex tumor structures, the precise mapping of liver space-occupying lesions, and the comprehensive segmentation of liver.

## VI. ABLATION STUDY

We conduct an ablation study to analyze the effectiveness of different innovation points in the proposed method (EU-Net).

**TABLE 2.** Comparison of Chain EU-Net with different depth. The number of depth means that Chain EU-Net is consisted of n EU-Nets.

| Depth | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Parameters | | 0.03M | 0.06M | 0.09M | 0.12M | 0.15M |
| GLAS | Dice | 0.855 | 0.872 | 0.876 | **0.881** | 0.874 |
| | Jaccard | 0.746 | 0.773 | 0.779 | **0.787** | 0.770 |
| RITE | Dice | 0.765 | 0.764 | 0.767 | **0.774** | 0.766 |
| | Jaccard | 0.620 | 0.619 | 0.621 | **0.632** | 0.620 |

### A. DEPTH

The Chain EU-Net is composed of multiple lightweight dense-connected EU-Nets blocks. To figure out how the performance varies while we change the depth of our chain EU-Net, we did global depth experiment. Specifically, we kept the number of filters of each lightweight dense skip-connected EU-Net block unchanged. Then we increased the depth of our chain EU-Net by densely skip connecting from one EU-Net block to two blocks, three blocks, four blocks, five blocks respectively, which means we added or deleted EU-Net blocks from chain EU-Net.

Table.2 shows the result of our chain EU-Nets that have different numbers of lightweight EU-Net blocks by dense skip-connection on GLAS and RITE dataset. The number in depth line indicates that the chain network is composed of n EU-Net blocks by dense skip-connection. We can see that the Chain EU-Net has the best performance when it contains four EU-Net blocks.

**TABLE 3. Performance of two methods on GLAS and RITE dataset. The Convolution method in table means replacing the inception block in EU-Net with a regular convolutional layer.**

| Method | | Inception(EU-Net) | Convolution |
|---|---|---|---|
| Parameters | | 29961 | 30561 |
| GLAS | Dice | **0.8547** | 0.8516 |
| | Jaccard | **0.7462** | 0.7416 |
| RITE | Dice | **0.7651** | 0.7427 |
| | Jaccard | **0.6195** | 0.5907 |

### B. INCEPTION BLOCKS

We use an inception block to reduce parameter calculation and alleviate the gradient vanishing issue. To validate the effectiveness of the inception block, we replaced the inception block in EU-Net with regular convolutional layers and conducted experiments on the GLAS and RITE dataset. We can clearly observe the effectiveness of the inception block in the Table.3.

### C. DENSE CONNECTIONS

To study the effectiveness of different dense connectivity in EU-Net, we removed some of the skip connections in EU-Net, resulting in four different structures as shown in the Fig.8. The four structures in Fig.8 are carefully chosen. The connectivity method in Fig.8-a is similar to U-Net, where layers with the same size are concatenated from the encoder to the decoder. The connectivity method in Fig.8-b involves concatenating layers with rich original image information into each layer of the decoder. The connectivity method in Fig.8-c can be combined with the connectivity method in Fig.8-b to form a complete EU-Net. Fig.8-d does not contain any skip connections.

For these experiments, we use GLAS and RITE dataset. The experimental results are shown in Table.4. It can be clearly observed that the performance of the complete EU-Net is the best.

**TABLE 4. Performance of EU-Net and other dense connections base on EU-Net.**

| Datasets | GLAS | | RITE | |
|---|---|---|---|---|
| Methods | Dice | Jaccard | Dice | Jaccard |
| EU-Net | **0.8547** | **0.7462** | **0.7651** | **0.6195** |
| Fig.8-a | 0.8278 | 0.7062 | 0.7261 | 0.5699 |
| Fig.8-b | 0.8374 | 0.7206 | 0.7373 | 0.5839 |
| Fig.8-c | 0.8342 | 0.7156 | 0.7140 | 0.5552 |
| Fig.8-d | 0.8230 | 0.6992 | 0.4571 | 0.2963 |

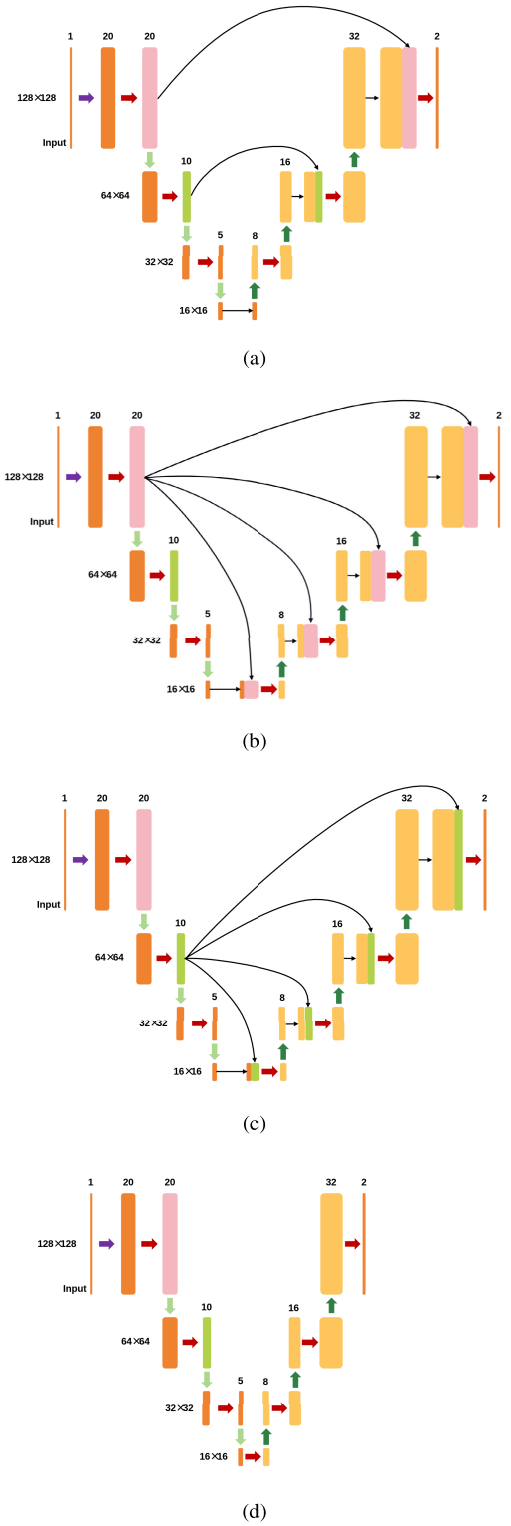


**FIGURE 8. The structures of four different dense connection.**

## VII. COMBINING WITH ATTENTION

Inspired by Attention U-Net [23], we add attention blocks at the positions shows in Fig.9. As Tabel.5 shows, Attention EU-Net achieve a significant improvement.

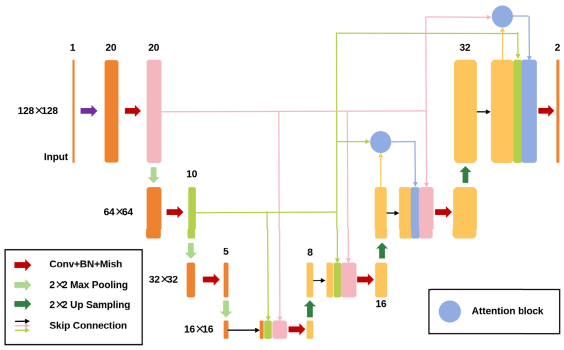


**FIGURE 9.** The structure of attention EU-Net.

**TABLE 5.** Performance of EU-Net and Attention EU-Net on four datasets.

| Method | | EU-Net | Attention EU-Net |
|---|---|---|---|
| Parameters | | 29961 | 30873 |
| GLAS | Dice | 0.8547 | **0.8613** |
| | Jaccard | 0.7462 | **0.7563** |
| RITE | Dice | 0.7651 | **0.7750** |
| | Jaccard | 0.6195 | **0.6326** |
| LiTS-tumour | Dice | 0.8432 | **0.8826** |
| | Jaccard | 0.7288 | **0.7899** |
| LiTS-liver | Dice | 0.9419 | **0.9573** |
| | Jaccard | 0.8901 | **0.9181** |
| PUFH | Dice | 0.9096 | **0.9170** |
| | Jaccard | 0.8341 | **0.8398** |

## VIII. CONCLUSION

This paper shows the power of lightweight dense skip EU-Net structure for medical segmentation area. Compared to U-Net and its variants, EU-Net addresses the limitations they face in terms of stability and efficiency, making it particularly well-suited for deployment on embedded devices. Experimental results on four datasets demonstrate that EU-Net achieves robust performance with fewer parameters. To further characterize the performance and robustness of the network, we propose Chain EU-Net which is consisted of multiple lightweight dense EU-Nets by dense skip-connection. The Chain EU-Net achieves better performance and robustness four datasets, with only few parameters increase. Furthermore, we adjust the depth of Chain EU-Net to analyze the performance and robustness, study the effectiveness of inception blocks, skip connections, and attention EU-Net. In the future, we plan to explore suitable attention mechanisms to integrate with the chain EU-Net, aiming to achieve superior and more stable performance.

## DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

## REFERENCES

[1] Z. Yu-Qian, G. Wei-Hua, C. Zhen-Cheng, T. Jing-Tian, and L. Ling-Yun, "Medical images edge detection based on mathematical morphology," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 6492–6495.

[2] M. Gudmundsson, E. A. El-Kwae, and M. R. Kabuka, "Edge detection in medical images using a genetic algorithm," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 469–474, Jun. 1998.

[3] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, Jan. 1988.

[4] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[5] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proc. 8th IEEE Int. Conf. Comput. Vision. ICCV*, 2001, pp. 105–112.

[6] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut' interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph. (TOG)*, vol. 23, no. 3, pp. 309–314, 2004.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Munich, Germany: Springer, 2015, pp. 234–241.

[8] X. Xiao, S. Lian, Z. Luo, and S. Li, "Weighted res-UNet for high-quality retina vessel segmentation," in *Proc. 9th Int. Conf. Inf. Technol. Med. Educ. (ITME)*, Oct. 2018, pp. 327–331.

[9] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Granada, Spain: Springer, 2018, pp. 3–11.

[10] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[12] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet3+: A full-scale connected UNet for medical image segmentation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 1055–1059.

[13] Y. Ji, R. Zhang, Z. Li, J. Ren, S. Zhang, and P. Luo, "UXNet: Searching multi-level feature aggregation for 3D medical image segmentation," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*. Lima, Peru: Springer, 2020, pp. 346–356.

[14] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*.

[15] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.

[16] J. M. J. Valanarasu, V. A. Sindagi, I. Hacihaliloglu, and V. M. Patel, "KiU-Net: Towards accurate segmentation of biomedical images using over-complete representations," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*. Lima, Peru: Springer, 2020, pp. 363–373.

[17] J. M. J. Valanarasu, V. A. Sindagi, I. Hacihaliloglu, and V. M. Patel, "KiU-Net: Overcomplete convolutional architectures for biomedical image and volumetric segmentation," *IEEE Trans. Med. Imag.*, vol. 41, no. 4, pp. 965–976, Apr. 2022.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[19] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[20] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.

[21] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945.

[22] P. Jaccard, "The distribution of the flora in the Alpine zone. 1," *New Phytologist*, vol. 11, no. 2, pp. 37–50, Feb. 1912.

[23] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.

[24] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," *J. Med. Imag.*, vol. 6, no. 1, Mar. 2019, Art. no. 014006.
[25] J. Jing, Z. Wang, M. Rätsch, and H. Zhang, "Mobile-UNet: An efficient convolutional neural network for fabric defect detection," *Textile Res. J.*, vol. 92, nos. 1–2, pp. 30–42, Jan. 2022.
[26] Y. Deng, Y. Hou, J. Yan, and D. Zeng, "ELU-Net: An efficient and lightweight U-Net for medical image segmentation," *IEEE Access*, vol. 10, pp. 35932–35941, 2022.

**LIZHUANG LIU** received the Ph.D. degree from Xi'an Jiaotong University, China, in 2003. He is currently a Researcher with the Shanghai Advanced Research Institute, Chinese Academy of Sciences. His current research interests include computer vision, image processing, and deep learning.

**JIACUN QIU** received the B.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2022. He is currently pursuing the M.S. degree in electronic information with the University of Chinese Academy of Sciences, China. His current research interests include deep learning, medical image segmentation, and computer vision.

**KE WANG** received the B.M. and M.D. degrees from Peking University Beijing, China, in 2012 and 2015, respectively. She is currently an Attending Radiologist with the Peking University First Hospital, Beijing. Her current research interests include abdominal imaging, CT and MRI technology, and artificial intelligence applications in medicine.

**QIAO ZHAN** received the B.M. degree from Nanjing Medical University, Jiangsu, China, in 2017, and the joint M.D. and Ph.D. degrees from Peking University, Beijing, China, in 2022. She is currently an Attending Physician with The First Affiliated Hospital of Nanjing Medical University, Nanjing, China. Her current research interests include the theory of infectious and hepatic diseases, artificial intelligence application in medicine, and viral immunology.

**JIAXI JIANG** received the master's degree in statistics from George Washington University, in 2020, and the Ph.D. degree in optical engineering with the School of Optoelectronics, Beijing Institute of Technology, in 2022. He was a Research Assistant with the Shanghai Advanced Research Institute, Chinese Academy of Sciences, Beihang University, and the Beijing Institute of Technology, from 2020 to 2022. His current research interests include the virtual simulation of medical treatment, medical image segmentation, and medical-surgical robots. He has won awards in various simulation competitions and entrepreneurship competitions.

**ZHENQI HAN** received the master's degree from the Computer Science Department, Shanghai University. He is currently a Research Assistant with the Shanghai Advanced Research Institute, Chinese Academy of Sciences. He has published more than 20 research articles and patents. His current research interests include computer vision and 3D deep learning and their applications in real scenarios.

**JIANXIN QIU** received the M.D. degree. He is currently the Secretary of the Community Party Branch and the Deputy Medical Director of the Radiology Department, Peking University First Hospital (presiding over the work), the Supervisor of postgraduate students, and the Deputy Director of the Imaging Department, Peking University Medical Department. His current research interests include the MRI quantitative study of ischemic cardiomyopathy, MRI multi-parameter evaluation of pulmonary nodules, and prospective study of radiofrequency ablation of adrenal adenoma.

**TIAN WU** received the B.M. degree from Jilin University, Jilin, China, in 2017, and the M.D. degree from Peking University, Beijing, China in 2020. She is currently pursuing the Ph.D. degree with the Department of Infectious Diseases, Peking University First Hospital. Her current research interests include infectious diseases, liver diseases, and virology.

**JINGHANG XU** received the M.D. degree from Peking University, Beijing, China, in 2008. She is currently the Deputy Director of the Department of Infectious Diseases, the Chief Physician, and an Associated Professor with the Peking University First Hospital, Beijing. Her current research interests include infectious diseases, liver diseases, genetics, and virology.

**ZHENG ZENG** received the B.S. degree from the Guangxi Medical University, Guangxi, China, in 1986, the M.S. degree from the First Medical School of Peking University, in 1992, and the joint M.D. and Ph.D. degree from the Peking University Health Science Center, Beijing, China, in 2000. He is currently a Professor (the Chief Physician) with the Peking University First Hospital, Beijing. His current research interests include bioinformatics, population genetics, virology, and radiomics.