Architectures and Applications of U-net in Medical Image Segmentation: A Review

Jundi Wang*, Lei Han, and Dongsheng Ran

Beijing University of Posts and Telecommunications, Beijing, China
jundiwang@bupt.edu.cn, hanl@bupt.edu.cn, rands@bupt.edu.cn
*corresponding author

Abstract—Recently, with the increasing application of deep learning in the medical field, convolutional neural networks, represented by U-Net, has been widely applied in medical image segmentation. The improved U-shaped network structure based on U-Net has gradually become a hot topic in medical image segmentation research. This article summarizes the improvement work related to U-Net from three perspectives: modifying skip connections, adding or replacing blocks and concatenating multiple neural networks. Then, taking the segmentation of retina, lungs, brain, abdomen, and other organs as examples, the characteristics and difficulties of various organ segmentation were introduced. Finally, a summary and outlook were made.

Keywords-component; U-Net; skip connections; semantic segmentation; deep learning

1. Introduction

The main task of medical image segmentation is to segment regions of interest from medical images, such as organs, lesion areas, tumors, etc [1][2][3][4]. In the past, doctors had to manually segment regions of interest from medical images. However, manual image segmentation is time-consuming [5][6][7]. In addition, analysis of medical images relies heavily on the experience of doctor. For the medical underdevelopment regions, insufficient experience of doctors may lead to misjudgment. A well-trained neural network can analyze medical images in batches, which is significantly faster than manual analyzing. Moreover, neural networks can learn experience from the annotations of a large number of excellent doctors, thus can provide diagnostic recommendations for doctors. Therefore, segmentation medical images using neural networks can improve the efficiency of image segmentation and fully utilize the experience of doctors [8][9][10].

Traditional image segmentation techniques include region-based segmentation methods and boundary-based segmentation methods [11][12]. The region-based segmentation methods rely on the spatial local features of images, such as grayscale and texture. The boundary-based segmentation methods mainly use gradient information to determine the boundary of the target. With the development of artificial intelligence, deep learning has achieved great success in the field of image analysis. Neural networks can automatically extract features from images. A large number of parameters and nonlinear structures make neural networks highly capable of fitting. In 2015, Ronneberger et al. proposed U-Net in the ISBI Cell Tracking Challenge and achieved good segmentation results [13]

The semantic structure of target organs in medical image is relatively simple [14][15][16][17]. The method of fusing high-level and low-level features through skip connections in U-Net is suitable for medical image segmentation. Therefore, the U-Net architecture has been widely applied in the field of medical image segmentation [18][19][20]. Researchers have proposed some new architectures based on U-Net. Some have changed or replaced some modules in U-Net [21][22][23]. Some introduce new neural network architectures into U-Net. The rest of this paper is organized as follows. Section 2 introduces the basic structure and several improved structures of U-Net. The distribution of various improved structures between 2018 and 2023 is summarized. Section 3 reviews the application of U-Net and its improved structure in image segmentation of different organs, including retinal vessels, lung, brain, abdomen and other medical image segmentation. The significance and difficulties of different organ segmentation are explained. Finally, conclusions are presented in Section 4.

2. U-Net and Its Variants

2.1. U-Net

2015, Ronneberger et al. [13] proposed U-Net for medical image segmentation and achieved good results. U-Net is a typical Encoder Decoder structure and a symmetric network structure. Because its shape resembles the letter U, the authors named it U-Net. As shown in Figure 1, the U-Net has a contracting path on the left side, an expansive path on the right side and skip connections in the middle. The purple arrows in the figure represent 3x3 convolution and ReLU. The input is a single channel 572x572 image, which is sent to two 3x3 convolution layers, resulting in a 64 channel 568x568 feature map. Three feature maps form a stage, followed by a green arrow for down sampling. The green arrow indicates the 2x2 max pooling operation. Then, after 5 stages of convolution and pooling, a feature map of 1024 channels for 28x28 was obtained. Next, perform the decoding operation by up sampling the low-resolution image through up-convolution.

979-8-3503-0247-9/23/$31.00 ©2023 IEEE
DOI 10.1109/ISSSR58837.2023.00022

The red arrow indicates 2x2 up-convolution. The yellow arrow means crop the last feature map of each stage in the Encoder, and then concatenate it with the first feature map in the corresponding stage in the Decoder. Then, each stage continues to perform two convolution and ReLU operations. After passing through four Decoder stages and 1x1 convolution, two single channel 572x572 feature maps are obtained. These two feature maps are the target and background, respectively.
The shallow layers in U-Net capture the shallow features of images, such as colors, textures, etc. The deep layers have the ability to capture some deep features because of the large receptive field. Skip connections fuse shallow features with deep features, allowing the network to better capture image details.

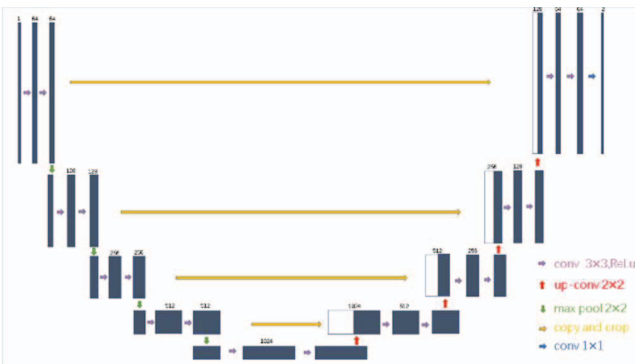

Figure 1. Structure of U-Net

U-Net focus on the segmentation of 2D images, but many medical images such as CT images and MRI images are 3D images. Abdulkadir [24] proposed a 3DU-Net, as shown in Figure 2. Compared to 2D U-Net, the main change of 3DU-Net is to replace the original 2D convolution and pooling with 3D convolution and pooling.

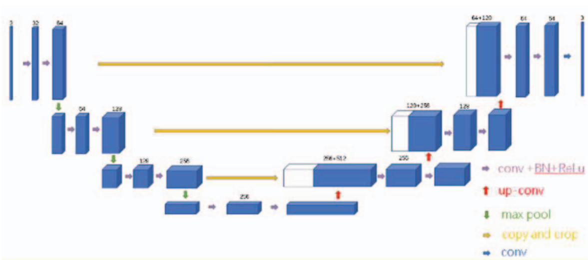

Figure 2. Structure of 3DU-Net

2.2. Variants of U-Net

After the proposal of U-Net, researchers have made improvements to U-Net for specific medical image analysis scenarios. Among them, the main improvement directions include modifying skip connections, adding or replacing modules, introducing new architectures, and concatenating multiple neural networks.

2.2.1. Modifying Skip Connections

Many researchers modify skip connections to improve the performance of U-Net.
Edwin Thomas et al. [43] change the skip connection of U-Net to a new type of connection that combines attention mechanisms, as shown in Figure 3. This new connection can alleviate the problem of large semantic gaps in feature maps caused by the long skip connection structure between the encoder and decoder in U-Net. Liu et al. [66] propose a medical image segmentation model, TransUNet+, which has an excellent performance in small organ segmentation. This model incorporates a score matrix with Transformer to enhance the skip connection. TransUNet+ improves global attention and achieves good results on a variety of image segmentation datasets.

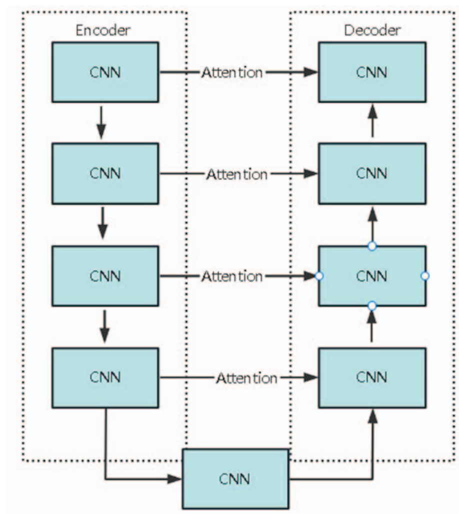

Figure 3. The architecture of Multi-Res-Attention UNet

2.2.2. Adding or Replacing Blocks

Adding blocks refers to adding blocks with specific functions to U-Net to optimize the network. Replacing blocks mainly involves modifying the convolutional layer or sampling layer of blocks in U-Net.
Jin et al. [26] propose DUNet which replaces standard convolutions with deformable convolutions, as shown in Figure 4. Xin et al. [29] propose an improved MR-UNet, as shown in Figure 5. MR-UNet contains two new blocks: Multiconv and Resconv. The one uses different size convolution kernels to extract the characteristics of different thicknesses. The another uses different convolution layers to connect the shallow Semantic information to each layer of the decoding stage, which reduces the semantic diversity between the encoder and decoder.. Sun et al. [35] propose a dual attention 3D-UNet, as shown in Figure 6. The authors replace up sampling in 3D-UNet with Dup sampling to improve the convergence speed. The network integrates spatial attention

blocks and channel attention blocks in the network structure which can improve global correlation.

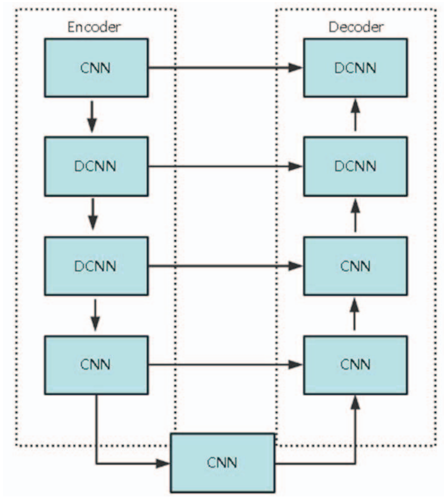

Figure 4 . The architecture of DUNet

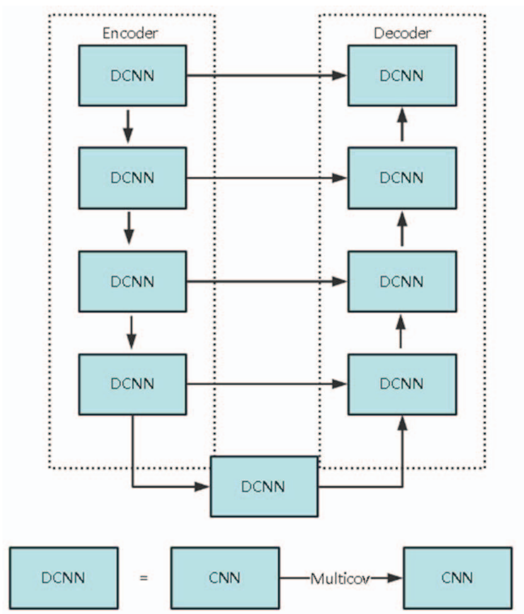

Figure 5 . The architecture of MRUNet

Yin et al. [39] propose a SD-UNet which adds the Squeeze and Attention (SA) block and the Dense ASPP block to U-Net, as shown in Figure 8. In SD-UNet, SA blocks fully utilize global information to better explore the diversity and link between pixels. The dense ASPP block captures multi-scale information of lung infection areas. Guo et al. [47] analyze allocation mechanisms between self-attention and convolution. The authors build a parallel non isomorphic block and named the U-shaped network with this block UNet2022, as shown in Figure 8. This network model has significant advantages over nnUNet.

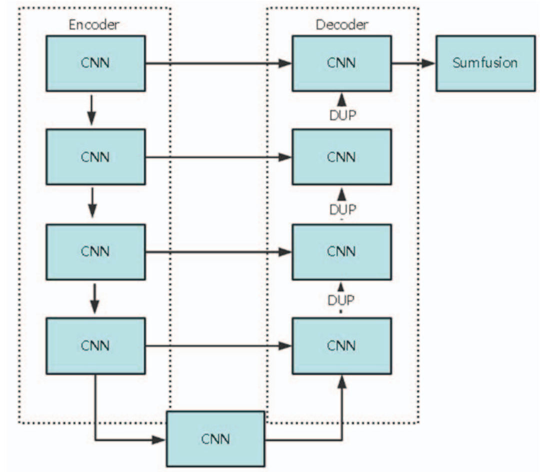

Figure 6 . The architecture of dual-attention 3D-UNet

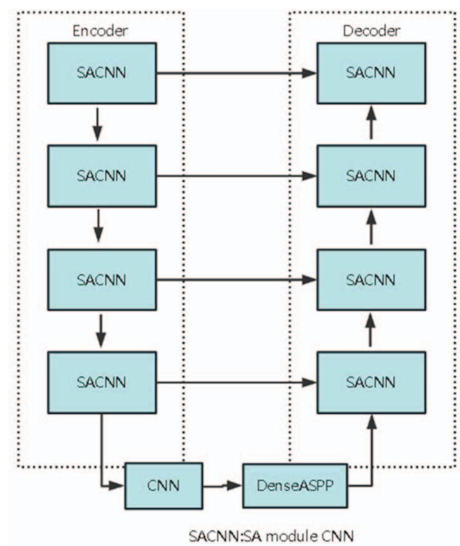

Figure 7 . The architecture of SD-UNet

Mohammed Yusuf Ansari et al. [50] propose Res-PAC-UNet. The network adopts a fixed width residual UNet backbone and pyramid Atrus convolution, achieving model simplification for liver CT segmentation. Xiao et al. [33] develop a 3D-Res2UNet which introduces Res2Net into 3D-UNet, as shown in Figure 9. 3D-Res2UNet has a symmetric network with skip connections and effective learning ability for multi-scale information. This network can provide more precise expression of multi-scale information.

2.2.3. Concatenating Multiple Neural Networks

Introducing a new architecture refers to introducing some new network structures, and the new networks often differ significantly from U-Net.

Yun et al. [31] design a U-Net based network architecture called MTPA-Unet, which adds Transformer into U-net，as shown in Figure 8. MTPA-Unet generates multi resolution images, and inputs them to every stage of encoder to recognize different levels of information.

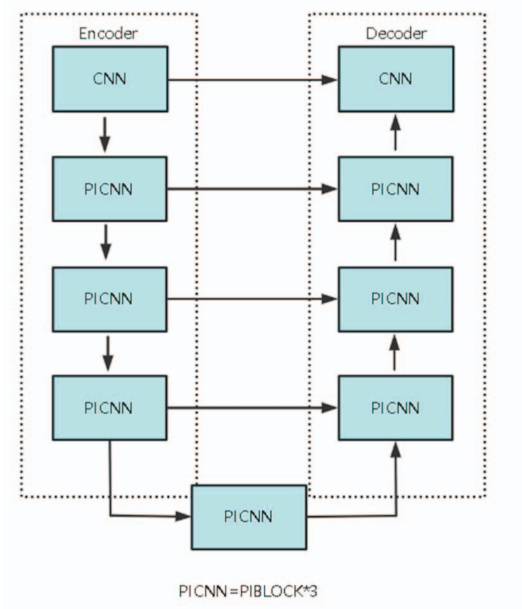

Figure 9 . . The architecture of UNet2022

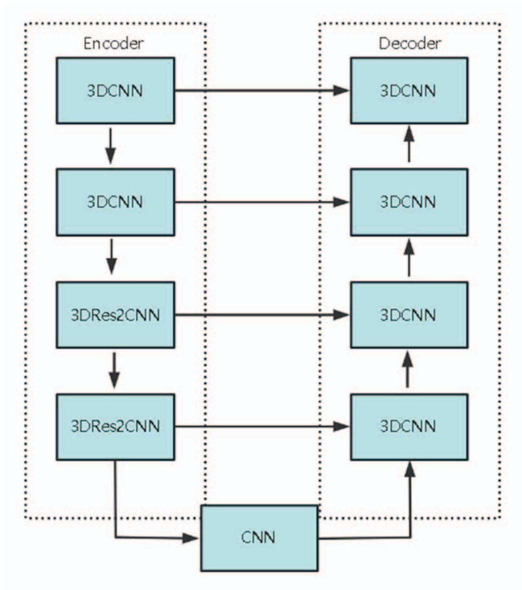

Figure 10 . The architecture of 3D-Res2UNet

Zhuang [25] proposes the LadderNet which adopts multiple pairs of encoder-decoder branches, as shown in Figure 11. Skip connection is made between each pair of adjacent encoders and decoders at each level, which makes the LadderNet have more information flow paths. This not only improves segmentation accuracy, but also reduces the number of parameters by sharing weights with each residual block. Baccouche et al. [56] propose an architecture called Connected-UNets, as shown in Figure 12. Two Unets are connected using modified skip connections, and ASPP is integrated into both standard Unets.

The distribution of U-Net modified in these three ways from 2018 to 2023 is shown in Table 1. As shown in Table 1, adding or replacing blocks is the most popular method in improving U-Net. Many network structures combine one or two of the above methods.

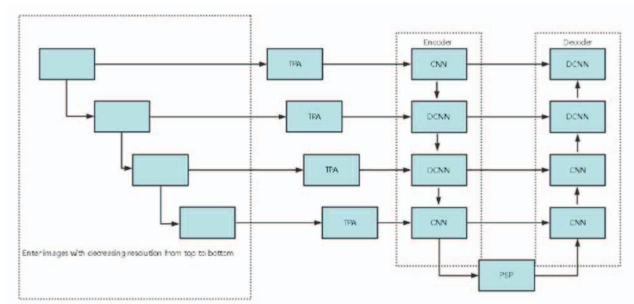

Figure 10 . The architecture of MTPAUNet

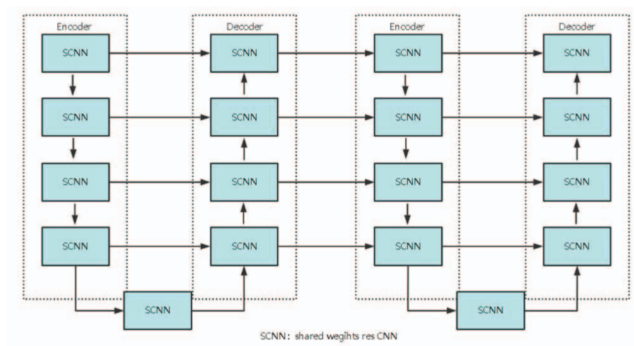

Figure 11 . The architecture of LadderNet

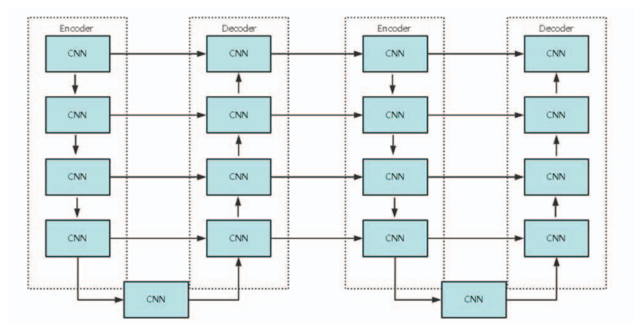

Figure 12 . The architecture of Connected-UNets

3. Application of U-Net in Medical Image Segmentation

In section 2, The U-Net and various modified networks based on U-Net are introduced. In this section, we will specifically describe the application of U-Net in image segmentation of different organs, including retinal vascular segmentation, lung segmentation, brain segmentation, abdominal segmentation, and other medical image segmentation.

Table 1. The distribution of U-Net modified in these three ways from 2018 to 2023.

Year	Modifying skip connections	Adding or replacing blocks	Introducing new architectures
2018	[25]	[49]	[25]
2019	/	[26]	/
2020	/	[36][33]	/
2021	[43][27][38][43][56]	[27][34][37][45]	[56]
2022	[44][48]	[28][29][31][35][39][42][44][47][48][50][55]	[31][54]
2023	/	[46]	/

3.1. Application of U-Net in retinal vascular image segmentation

The bottom of the eyeball is covered with abundant capillaries. Many potential diseases can lead to capillary bleeding, such as retinal vein occlusion and highly myopic retinopathy. By segmenting retinal vascular images, doctors can evaluate the patient's eye condition. The bottom of the eyeball is usually filled with very small capillaries, making the segmentation of retinal capillaries complex. How to utilize global information as much as possible while ensuring high segmentation accuracy has become the key to retinal vascular segmentation. Xin et al. [29] proposed an improved MR-UNet. MR-UNet introduce Multiconv block to improve the accuracy to segment micovascular, and Resconv block to reduce the semantic gap between the encoding structure and decoding structure. In the retinal dataset DRIVE, START, and CHASE_DB1, the accuracy rates of this model are 0.9705, 0.9747, and 0.9778, respectively. The segmentation effect is shown in Figure 13 (a).

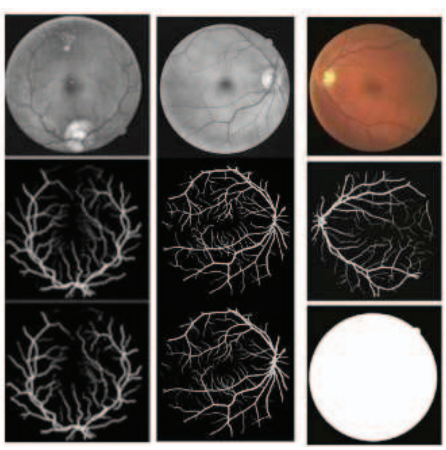
(a) (b) (c)

Figure 13 . The segmentation effect images on the DRIVE dataset: (a) MRUNet, (b) LadderNet, and (c) DUNet.

To improve the segmentation accuracy of U-Net, Zhuang [25] proposes LadderNet. Compared with U-Net, LadderNet has more flow paths of information. This not only improves segmentation accuracy but also reduces the number of parameters. Achieved better performance than R2-UNet on multiple datasets. The segmentation effect is shown in Figure 13 (b). Jin et al. [26] propose DUNet. DUNet can adaptively adjust the receiving field to adapt to the size and shape of blood vessels. DUNet achieve accuracy rates of 0.9566 and 0.9641 on DRIVE and START. The segmentation effect is shown in Figure 13 (c). Du et al. [27] propose the PSP-UNet segmentation algorithm integrating the attention mechanism. The precision of DRIVE and CHASE DB1 datasets are 0.9556 and 0.9590. Experimental results show that the network enhances pixel detection capability of blood vessels, focuses on feature information and minimizes the detection of irrelevant information as much as possible. Gobinath et al. [28] propose the SER-UNet to accurately locate vascular regions and focus on small vascular features. The accuracy on the START and DRIVE datasets are 0.09831 and 0.989. Yun et al. [31] propose the MTPA_Unet. By inputting multi resolution images, the network recognizes different levels of information. The network model get accuracy rates of 0.9718, 0.9762, and 0.9773 on the CHASE DB1, DRIVE, and STARE, respectively.

3.2. Application of U-Net in Lung Image Segmentation

According to the global cancer statistics, cancer is one of the most common cancer. Lung nodules is early indicators of lung cancer, lung nodules accurate detection and image segmentation are of great significance for early diagnosis of lung cancer.

Xiao et al. [33] propose a 3D-Res2UNet. The Dice on the public dataset LUNA16 reached 0.9530, with a recall rate of 0.991, indicating well in the pulmonary nodules image segmentation. When lung cancer is confirmed later, the symptoms of the lung have shifted from nodules to tumors. By reason of the high discrepancies in terms of the appearance of tumors, accurately segmenting lung tumors from images is challenging. Yang et al. [34] proposed a 3DU-Net equipped with ResNet architecture and dual path deep supervision mechanism, which enhances the ability to perceive both global and local information. The model has achieved good results on multiple datasets, with excellent performance in small tumors. In order to balance the integrity of segmenting large tumors with the accuracy of segmenting small tumors, Sun et al. [35] introduce dual attention on 3D-UNet. The MIoU score on the public dataset LIDC-IDRI reached 0.894. Yahyatabar et al. [36] propose a neural network model called Dense-UNet. Dense UNet increases the information flow in the whole network by making dense connections between different layers, so that Dense UNet maintains the robustness of segmentation while having fewer parameters. The model in JSRT and Montgomery datasets are evaluated. The experimental results show that the proposed model achieves a good balance between performance and parameters.

In recent years, COVID-19 has become one of the world's largest health crises. There are some problems need to solve in the segmentation of lung images with COVID-19. Firstly, the degree of pulmonary infection varies greatly, and there is a blurry edge of the infected area. In the meantime, computer tomography (CT) inevitably has the hassle of noise interference. Alex et al. [37] propose the ADID-UNET. To solve the gradient vanishing problem of traditional neural network structures, dense networks are used to replace sampling layers to enhance feature propagation. Replacing standard with extended convolution, effectively increases the receiving domain of neural network, and has a better effect on edge feature extraction. ADID-UNET introduces an attention gate module into the network structure to increase the weight of target area and ignore background information as much as possible. Through the experiment, Dice and F1 scores of ADID-NET are 0.8031 and 0.82. Ma et al. [38] propose the PPM-Unet. They use pyramid pooling modules to replace traditional direct connections. The attention mechanism was used to better solve the problem of blurred edges in the stained area, and the background and target areas are well segmented. Through experiments, the network can accurately segment the infection area of COVID-19. Yin et al. [39] propose SD-UNet by incorporating SA blocks and dense ASPP blocks into UNet networks. The experimental results show that the proposed SD UNet on the lung infection area of COVID-19 are closer to the actual situation.

3.3. Application of U-Net in Brain Image Segmentation

MRI is often used for the diagnosis of brain diseases. There is inevitable random noise in MRI imaging [40]. The goal of brain image recognition is generally to find brain tumor, and there is often data imbalance between brain tumors and background, with tumor areas typically accounting for only 1.5% of the total volume of MRI [41]. Different tumor types often have significant differences in size. Tumors often have irregular and variable boundaries and structures.

Aghalari et al. [42] add two path residual blocks into UNet. The new network utilizes local features and global features. It has fewer parameters compared to U-Net while achieving good results. Thomas et al. [43] propose a multi resolution attention U-Net. They introduce attention mechanism into skip connection to solve the problem of large semantic gaps in feature maps caused by skip connection structure. Huang et al. [44] propose the DO-UNet which integrate attention mechanism and multi-scale feature fusion to achieve the automatic segmentation of brain tumor. To prevent gradient vanishing, residual modules are used to replace convolutional blocks in the U-Net model. Multi-scale feature fusion is added to the skip connections of U-Net to more effectively fuse shallow-level and deep-level features. In addition, during the decoding stage, attention mechanisms are added to avoid information redundancy, in order to increase the weight of effective information and minimize the interference of invalid information on the model.

Xiao et al. [45] propose MVHS Net for brain image segmentation. For lightweight structures, multi view fusion convolution and MVHS blocks are introduced into 3D-UNet. By adding these blocks, the network performance is improved and redundant feature information parameters are optimized. The experimental results obtained from the BraTS 2018 challenge dataset indicate that MVHSNet greatly optimizes computational parameters while maintaining image segmentation accuracy. Konar et al. [46] propose 3D-QNet, which demonstrates its advantages in semantic segmentation on the BRATS 2019 dataset and the LiTS17 dataset.

3.4. Application of U-Net in Abdominal Image Segmentation

The abdomen is a complex space. The images of abdominal organs have a high similarity. Moreover, due to differences in instrument parameters in CT imaging, there are differences in the grayscale values of the imaging, making it difficult to determine the spatial boundaries of adjacent organs. Therefore, abdominal organ segmentation faces enormous challenges.

Guo et al. [47] propose UNet2022 which contain a parallel non isomorphic block. UNet2022 achieves good results in Multi-organ CT segmentation (Synapse) dataset, Automated cardiac diagnosis (ACDC) dataset and Neural structures segmentation dataset. Zhao et al. [48] propose TU-Net which achieves the best accuracy in segmenting 8 abdominal organs on the Synapse dataset. TU-Net is based on the Transformer architecture and uses continuous convolutional layers with small convolutional kernels to extract features. The skip connection is replaced with cross attention skip connections.

In order to achieve precise segmentation of liver regions, Sun et al. [49] propose 3D UNet-C2. The network effectively utilize the three-dimensional information of CT images. A primary model is obtained by selecting clear images and capturing liver regions as samples. Then the network parameters are initialized using this model to enable the network to converge. Finally, on the basis of the original model, three-dimensional conditional random field is used to optimize the liver segmentation boundary. The segmentation results are superior to 3DNet and Vnet models. Ansari et al. [50] propose a Res-PAC-UNet, which provides a low memory footprint method for accurate liver CT segmentation.

3.5. Application of U-Net in Other Medical Image Segmentation

3.5.1 Image Segmentation of Hypopharyngeal Cancer

Hypopharyngeal cancer (HPC) is a rare malignant tumor. Due to the lack of specific clinical manifestations of early stomach cancer are easy to be misdiagnosed as other pharyngeal diseases. HPC total 5 years survival rate is only 30% to 35% [51]. Required for a long time, because of MRI in the diagnosis of stomach cancer tend to have uneven luminance, edge contour image fuzzy, susceptible to noise and other issues, as shown in Figure 14 [53]. These have brought many difficulties to image segmentation of hypopharyngeal cancer using deep learning. Ran et al. [52]

propose a prognosis classification model based on Densenet. Experiment shows that good prediction performance. Zhang et al. [54] propose a hybrid network of bilateral encoder for hypopharyngeal cancer segmentation, called Twist-Net. Twist-Net proposes bilateral transformation (BT) blocks and bilateral aggregation (BG) blocks to fuse deep semantic characteristics graph and shallow semantic features. Designed a M piece of multi-scale information to learn. In order to avoid over fitting dataset is not big enough, a transfer learning method is proposed. The network in a private HPC on the dataset, the highest 82.98% of Dice. And CHASE_DB1 and BraTS2018 both public datasets to verify the results, good results have been achieved.

3.5.2 Heart Image Segmentation

Refers to the inner diameter of the left atrium of the left atrium, volume increases. Common in the heart of the left atrial disease, such as high blood pressure cause damage to the heart. Research has shown that studying and judging the left atrium can be used to predict whether there is the risk of cardiovascular disease. Wong et al. [55] propose GCW-UNet which can accurately segment the left atrium region in MRI images through the newly designed channel weight, and obtain MRI images with different resolutions through Gaussian blur. And on the test dataset, Dice's average similarity coefficient reached 0.9357.

Table 2. Difficulties and coping methods in major image segmentation of the retina, lungs, brain, and abdomen.

	Difficulties	Coping methods
Retinal vascular image segmentation	Retinal blood vessels are abundant, with capillaries intersecting and obstructing. It is difficult to obtain global information while ensuring high segmentation accuracy.	Using multiple pairs of encoder decoder branches and constructing skip connections between each adjacent pair of encoders and decoders [25]. Introducing Transformer to improve segmentation accuracy [31]. Introducing variable convolution [26], multi-scale convolution [29], attention mechanism [27], and residual convolution blocks [29] to obtain global information.
Lung Image Segmentation	The degree of pulmonary infection varies greatly. The image has blurred edges in the infected area. CT images have noise interference.	Introducing attention mechanism [37] [38] [39] and dual path supervision mechanism [34]. Introducing Res2Net to improve the gradient vanishing problem [33]. Introducing dense networks to replace sampling layers [39]. Replacing ordinary convolutions with dilated convolutions [37].
Brain Image Segmentation	There is noise in the MRI image. Long MRI imaging time leads to blurred images. Brain tumor vary greatly in shape, location and size. The tumor boundary and structure are irregular.	Introducing path residual blocks [42]. Introducing feature fusion and attention mechanism in skip connections [43][44]. Introducing attention mechanism in the network to increase the weight of small features [44].
Abdominal Image Segmentation	There is noise in CT imaging. The similarity of abdominal organs is high. It is difficult to determine the spatial boundaries of adjacent organs.	Using cross attention skip connections [48]. Adopting continuous convolutional layers with small convolutional sums. Introducing attention mechanisms [48] [50].

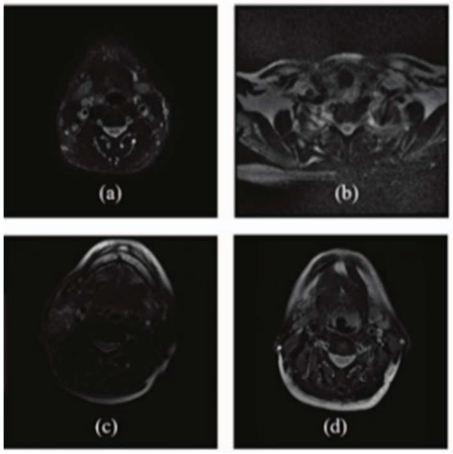

Figure 14 . MRI characteristics of HPC. (a) blurry HPC MRI.(b) noisy HPC MRI (c) uneven brightness HPC MRI. (d) shows a better shot in the HPC MRI.[52]

3.5.3 Breast Image Segmentation

At present, the main diagnostic method of breast cancer is ultrasound screening. Baccouche et al. [56] propose the Connected-UNets. Two U-Nets are connected using a modified skip connection. And ASPP is integrated into two standard U-Nets to emphasize the encoder decoder network architecture to focus on global information. On CBIS-DSM, INbreast, and private datasets, Dice scores of Connected-UNets are 0.8952, 0.9528, and 0.9588, respectively.

In summary, U-Net has been widely used in the field of medical image recognition, playing a certain auxiliary role for doctors to diagnose related diseases. However, different target organs have different requirements. For example, image recognition of the lungs, abdomen, liver and breast is often judged through CT images, while MRI is often used in hypopharynx, brain and other areas. Different imaging instruments have different problems to solve. For example, there is the inevitable noise interference problem in MRI images [57] [58] [59] [60]. What's more, the migration problem caused by different parameter designs of CT imaging equipment leading to different grayscales [61] [62] [63] [64] [65]. In addition, retinal image segmentation is the image processing of fundus color images. Researchers have made various modifications to U-Net to address various issues. For example, adding attention modules, modifying skip connections, and changing convolutional pooling modules. Researchers are also accustomed to transplanting architectures and models in other fields to medical image segmentation. For example, introducing Transformer, an architecture in the field of natural language processing (NLP), into U-Net. This not only makes the model converge faster, but also expands the Receptive field. The difficulties and coping methods of image segmentation in different organs that are specifically introduced are shown in Table 2.

4. CONCLUSION

This article reviews the network structure and basic principles of U-Net, and explains the reasons why most of the research in the field of medical image recognition is based on U-Net network structure. The article summarizes several typical modifications to U-Net, including modifying skip connections, adding or replacing blocks and introducing new architectures. This article also explores and summarizes the relevant applications of U-Net network in retinal image segmentation, lung image segmentation, brain image segmentation, abdominal image segmentation and other medical image segmentation. Finally, the difficulties and coping methods for medical segmentation of different organs are summarized.
Overall, it is foreseeable that U-Net will experience more improvements. And it will continue to promote the development of medical image segmentation and achieve more achievements

REFERENCES

[1] S. Deep, R. Kumar, "Modified Double U-Net Architecture for Medical Image Segmentation," IEEE Transactions on Radiation and Plasma Medical Sciences, Vol.7, pp. 151-162, 2023.

[2] Q. Sun, M. Dai, F. Cai, L. Wei, C. Yang, et al., "UCR-Net: U-shaped context residual network for medical image segmentation," Computers in Biology and Medicine, Vol.151, pp. 106203, 2022.

[3] O. Ali, H. Ali, S. Shah, and A. Shahzad, "Implementation of a Modified U-Net for Medical Image Segmentation on Edge Devices," IEEE Transactions on Circuits and Systems II: Express Briefs, Vol.69, pp. 4593-4597, 2022.

[4] S. Kurtek, J. Su, C. Grimm, M. Vaughan, R. Sowell, and A. Srivastava, "Statistical analysis of manual segmentations of structures in medical images," COMPUTER VISION AND IMAGE UNDERSTANDING, Vol.117, pp. 1036-1050, 2013.

[5] N. Shareef, D. Wang, and R. Yagel, "Segmentation of medical images using LEGION," IEEE Transactions on Medical Imaging, Vol.18, pp. 74-91, 1999.

[6] J. Egger, R. Colen, B. Freisleben, and C. Nimsky, "Manual refinement system for graph-based segmentation results in the medical domain(Article)," Journal of Medical Systems, Vol.36, pp. 2829-2839, 2012.

[7] S. Chen, Z. Gamechi, F. Dubost, G. Tulder, and M. Bruijne, "An end-to-end approach to segmentation in medical images with CNN and posterior-CRF," Medical image analysis, Vol.76, pp. 102311, 2022.

[8] P. Malhotra, S. Gupta, D. Koundal, A. Zaguia, and W. Enbeyle, "Deep Neural Networks for Medical Image Segmentation," Journal of Healthcare Engineering, Vol.2022, pp. 9580991, 2022.

[9] X. Wang, L. Gu, and Z. Wang, "Computer Medical Image Segmentation Based on Neural Network," IEEE Access, Vol.8, pp. 158778-158786, 2020.

[10] M. Huang, Y. Wu, "Semantic segmentation of pancreatic medical images by using convolutional neural network," Biomedical Signal Processing and Control, Vol.73, pp. 103458, 2022.

[11] E. Giora, C. Casco, "Region- and edge-based configurational effects in texture segmentation," Vision Research, Vol.47, pp. 879-886, 2007.

[12] Y. Meng, H. Zhang, Y. Zhao, X. Yang, Y. Qiao, L. MacCormick, et al., "Graph-Based Region and Boundary Aggregation for Biomedical Image Segmentation," IEEE transactions on medical imaging, Vol.41, pp. 690-701, 2022.

[13] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv preprint arXiv:1505.04597. https://doi.org/10.48550/arXiv.1505.04597.

[14] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," Electronic Markets, Vol.31, pp. 685-695, 2021.

[15] R. Diaz, "Machine Learning and Deep Learning," INGENIARE - Revista Chilena de Ingeniería, Vol.29, pp. 180-181, 2021.

[16] E. Akleman, "Deep Learning," Computer, Vol.53, pp. 17, 2020.

[17] L. Cao, "Deep Learning Applications," IEEE Intelligent Systems, Vol.37, pp. 3-5, 2022.

[18] J. Huang, D. Shao, H. Liu, Y. Xiang, L. Ma, S. Yi, et al., "A lightweight segmentation method based on residual U-Net for MR images," Journal of Intelligent & Fuzzy Systems, Vol.42, pp. 5085-5095, 2022.

[19] S. Sangui, T. Iqbal, P. Chandra, S. Ghosh, and A. Ghosh, "3D MRI Segmentation using U-Net Architecture for the detection of Brain Tumor," Procedia Computer Science, Vol.218, pp. 542-553, 2023.

[20] J. Jin, H. Zhu, Y Teng, Y. Ai, C. Xie, X. Jin, "The Accuracy and Radiomics Feature Effects of Multiple U-net-Based Automatic Segmentation Models for Transvaginal Ultrasound Images of Cervical Cancer," Journal of Digital Imaging, Vol.35, pp. 983-992, 2022.

[21] M. Benazzouz, M. Benomar, Lamine, and Y. Moualek, "Modified U-Net for cytological medical image segmentation," International Journal of Imaging Systems & Technology, Vol.32, pp. 1761-1773, 2022.

[22] X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, "U-Net-Based Medical Image Segmentation," Journal of Healthcare Engineering, Vol.2022, pp. 4189781, 2022.

[23] Z. Zhang, C. Wu, S. Coleman, and D. Kerr, "DENSE-INception U-net for medical image segmentation," Computer methods and programs in biomedicine, Vol.192, pp. 105395, 2020.

[24] Ö. Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. arXiv preprint arXiv:1606.06650. https://doi.org/10.48550/arXiv.1606.06650.

[25] J. Zhuang, "LadderNet: Multi-path networks based on U-Net for medical image segmentation," Electrical Engineering and Systems Science, Vol.2, 2018.

[26] Q. Jin, Z. Meng, D. Pham, Q. Chen, L. Wei, and R. Su, "DUNet: A deformable network for retinal vessel segmentation," Knowledge-Based Systems, Vol.178, pp. 149-162, 2019.

[27] X. Du, J. Wang, and W. Sun, "UNet retinal blood vessel segmentation algorithm based on improved pyramid pooling method and attention mechanism," Physics in medicine and biology, Vol.66, pp. 175013, 2021.

[28] C. Gobinath, P. Gopinath, "Design and development of SER-UNet model for glaucoma image analysis," Journal of Electronic Imaging, Vol.31, pp. 041215, 2022.

[29] Y. Xin, L. Li, and L. Tao, "MR-UNet: An UNet model using multi-scale and residual convolutions for retinal vessel segmentation," International Journal of Imaging Systems & Technology, Vol.32, pp. 1588-1603, 2022.

[30] X. Liu, D. Zhang, J. Yao, and J. Tang, "Transformer and convolutional based dual branch network for retinal vessel segmentation in OCTA images," Biomedical Signal Processing and Control, Vol.83, pp. 104604, 2023.

[31] Y. Jiang, J. Liang, T. Cheng, X. Lin, Y. Zhang, and J. Dong, "MTPA_Unet: Multi-Scale Transformer-Position Attention Retinal Vessel Segmentation Network Joint Transformer and CNN," Sensors, Vol.22, pp. 4592, 2022.

[32] R. Siegel, K. Miller, and A. Jemal, "Cancer statistics, 2020," CA: A Cancer Journal for Clinicians, Vol.70, pp. 7-30, 2020.

[33] Z. Xiao, B. Liu, L. Geng, F. Zhang, and Y. Liu,"Segmentation of Lung Nodules Using Improved 3D-UNet Neural Network," Symmetry, Vol.12, pp. 1787, 2020.

[34] J. Yang, B. Wu, L. Li, P. Cao, and O. Zaiane, "MSDS-UNet: A multi-scale deeply supervised 3D U-Net for automatic segmentation of lung tumor in CT," Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society, Vol.92, pp. 101957, 2021.

[35] J. Sun, W. Chen, L. Zhang, and X. Yan, "Research on Lung Tumor Cell Segmentation Method Based on Improved UNet Algorithm," Scientific Programming, Vol.2022, pp. 1-8, 2022.

[36] M. Yahyatabar, P. Jouvet, and F. Cheriet, "Dense-Unet: a light model for lung fields segmentation in Chest X-

Ray images," Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, Vol.2020, pp. 1242-1245, 2020.

[37] A. Raj, H. Zhu, A. Khan, Z. Zhuang, Z. Yang, V.Mahesh, and G. Karthik, "ADID-UNET-a segmentation model for COVID-19 infection from lung CT scans," PeerJ. Computer science, Vol.7, pp. e349, 2021.

[38] Y. Ma, P. Feng, P. He, Y. Ren, X. Guo, and B. Wei, "Segmenting lung lesions of COVID-19 from CT images via pyramid pooling improved Unet," Biomedical physics & engineering express, Vol.7, pp. 045008, 2021.

[39] S. Yin, H. Deng, Z. Xu, Q. Zhu, and J. Cheng, "SD-UNet: A Novel Segmentation Framework for CT Images of Lung Infections," Electronics, Vol.31, pp. 130, 2022.

[40] H. Sung, J. Ferlay, R. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: a cancer journal for clinicians, Vol.71, pp. 209-249, 2021.

[41] P. Wang, A. Chung, "Relax and focus on brain tumor segmentation," Medical image analysis, Vol.75, pp. 102259, 2022.

[42] M. Aghalari, A. Aghagolzadeh, and M. Ezoji, "Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks," Biomedical Signal Processing and Control, Vol.69, pp. 102841, 2021.

[43] E. Thomas, S. Pawan, S. Kumar, Anmol Horo, S. Niyas, S. Vinayagamani, C. Kesavadas, and J. Rajan, "Multi-Res-Attention UNet : A CNN Model for the Segmentation of Focal Cortical Dysplasia Lesions from Magnetic Resonance Images," IEEE journal of biomedical and health informatics, Vol.25, pp. 1724-1734, 2021.

[44] T. Huang, Y. Liu, "Research on the magnetic resonance imaging brain tumor segmentation algorithm based on DO-UNet," International Journal of Imaging Systems and Technology, Vol.33, pp. 143-157, 2023.

[45] Z. Xiao, K. He, J. Liu, and W. Zhang, "Multi-view hierarchical split network for brain tumor segmentation," Biomedical Signal Processing and Control, Vol.69, pp. 102897, 2021.

[46] D. Konar, S. Bhattacharyya, T. Gandhi, B. Panigrahi, and R. Jiang, "3-D Quantum-Inspired Self-Supervised Tensor Network for Volumetric Segmentation of Medical Images," IEEE Transactions on Neural Networks and Learning Systems, pp. 1-14, 2023.

[47] J. Guo, H. Zhou, L. Wang, and Y. Yu, UNet-2022: Exploring Dynamics in Non-isomorphic Architecture. arXiv preprint arXiv:2210.15566. https://doi.org/10.48550/arXiv. 2210.15566.

[48] J. Zhao, D. Wu, and Z. Wang, "TU-Net: U-shaped Structure Based on Transformers for Medical Image Segmentation," Data Science, Vol.1628, pp.376-386, 2022.

[49] M. Sun, J. Xu , W. Ma, and Y. Zhang, "A new fully convolutional network for 3D liver region segmentation on CT images," Chinese Journal of Biomedical Engineering, Vol.37, pp. 385-393, 2018.

[50] M. Ansari, Y. Yang, S. Balakrishnan, J Abinahed, A. Ansari, M. Warfa et al., "A lightweight neural network with multiscale feature enhancement for liver CT segmentation," Scientific reports, Vol.12, pp. 14153, 2022.

[51] J. Newman, T. Connolly, E. llling, M. Kilgore, J. Locher, and W. Carroll, "Survival trends in hypopharyngeal cancer: A population-based review(Article)," Laryngoscope, Vol.125, pp. 624-629, 2015.

[52] D. Ran, L. Han, "A classification model for prognosis based on Densenet," 2022 IEEE 17th Conference on Industrial Electronics and Applications (ICIEA), pp. 517-521, 2022.

[53] S. Zhang, Y. Miao, J. Chen, X. Zhang, L. Han, Z. Huang et al.,"CCS-Net: Cascade Detection Network with the Convolution Kernel Switch Block and Statistics Optimal Anchors Block in Hypopharyngeal Cancer MRI," IEEE journal of biomedical and health informatics, Vol.27, pp. 1-11, 2022.

[54] S. Zhang, Y. Miao, J. Chen, X. Zhang, L. Han, D. Ran et al., "Twist-Net: A multi-modality transfer learning network with the hybrid bilateral encoder for hypopharyngeal cancer segmentation," Computers in Biology and Medicine, Vol.154, pp. 106555, 2023.

[55] K. Wong, A. Zhang, K. Yang, S. Wu, and D. Ghista, "GCW-UNet segmentation of cardiac magnetic resonance images for evaluation of left atrial enlargement," Computer Methods & Programs in Biomedicine, Vol.221, pp. 106915, 2022.

[56] Baccouche, Z. Garcia, C. Olea, and Elmaghraby, "Connected-UNets: a deep learning architecture for breast mass segmentation," NPJ breast cancer, Vol.7, pp. 151, 2021.

[57] J. Wu, S. Yang, F. Gou, Z. Zhou, P. Xie, N. Xu et al., "Intelligent Segmentation Medical Assistance System for MRI Images of Osteosarcoma in Developing Countries," Computational and mathematical methods in medicine, Vol.2022, pp. 7703583, 2022.

[58] S. Talukder, N. Noyes, "Analyzing brain tumor MRI to demonstrate GPU-based medical image segmentation," Journal of Computing Sciences in Colleges, Vol.37, pp. 45-62, 2021.

[59] L. Anand, K. Rane, L. Bewoor, J. Bangare, J. Surve, M. Raghunath et al.,"Development of Machine Learning and Medical Enabled Multimodal for Segmentation and Classification of Brain Tumor Using MRI Images," Computational intelligence and neuroscience, Vol.2022, pp. 7797094, 2022.

[60] N. Feng, X. Geng, and L. Qin, "study on MRI Medical Image Segmentation Technology Based on CNN-CRF Model," IEEE Access, Vol.8, pp. 60505-60514, 2020.

[61] S. Rakesh, S. Mahesh, "Nodule segmentation of lung CT image for medical applications," Global Transitions Proceedings, Vol.2, pp. 80-83, 2021.

[62] Q. He, J. Shao, J. Pu, M. Zhou, M. M, S. Xiang et al., "FRFCM clustering segmentation method for medical CT image feature diagnosis," Journal of Intelligent & Fuzzy Systems, pp. 1-9, 2020.

[63] M. Ma, H. Xia, Y. Tan, H. Li, and S. Song, "HT-Net: hierarchical context-attention transformer network for medical ct image segmentation," Applied Intelligence, Vol.52, pp. 10692-10705, 2022.

[64] L. Liu, L. Wang, D. Xu, H. Zhang, A. Sharma, S. Tiwari et al., "CT Image Segmentation Method of Liver Tumor Based on Artificial Intelligence Enabled Medical Imaging," Mathematical Problems in Engineering, Vol.2021, pp. 1-8, 2021.

[65] M. van Eijnatten, R. van Dijk, J. Dobbe, G. Streekstra, J. Koivisto, and J. Wolff, "CT image segmentation methods for bone used in medical additive manufacturing," Medical Engineering & Physics, Vol.51, pp. 6-16, 2018.

[66] Y. Liu, H. Wang, Z. Chen, K. Huangliang, and H. Zhang, "TransUNet＋: Redesigning the skip connection to enhance features in medical image segmentation," Knowledge-Based Systems, Vol.256, pp. 109859, 2022.