Pak. j. life soc. Sci. (2024), 22(2): 15314-15321 E-ISSN: 2221-7630;P-ISSN: 1727-4915

Pakistan Journal of Life and Social Sciences

www.pjlss.edu.pk

https://doi.org/10.57239/PJLSS-2024-22.2.001108

RESEARCH ARTICLE

Application of Multi Scale Fusion U-Net in Medical Image Segmentation

Dai Lina[1*], Md Gapar Md Johar[2], Mohammed Hazim Alkawaz[3]

[1] School of Graduate Studies, Management and Science University,40100 Shah Alam, Selangor, Malaysia

[1] School of Information Technology and Engineering, Guangzhou College of Commerce, Guangzhou, China

[2] Software Engineering and Digital Innovation Center, Management and Science University,40100 Shah Alam, Selangor, Malaysia

[3] Department of Computer Science, College of Education for Pure Science, University of Mosul, Mosul, Nineveh, Iraq

ARTICLE INFO

Received: Sep 17, 2024

Accepted: Nov 22, 2024

Keywords

Application

Multi Scale Fusion

U-Net

Medical Image

Segmentation

***Corresponding Author:**

dailn526@163.com

ABSTRACT

O Multiscale U-Net, as an important model for deep learning in the field of medical image segmentation, has been widely studied and applied in recent years. In this paper, we systematically review the algorithmic variants of multiscale U-Net and the specific applications of different algorithms in image segmentation of different epidemics. However, multiscale U-Net still faces many challenges in practical applications. The range and quality of the training data restrict the model's capacity to generalize. The lack of high-quality and accurately labeled medical image datasets becomes a major bottleneck. In this paper, we also look forward to the future development direction of multiscale U-Net in the field of medical image segmentation. We propose that combining migration learning and semi-supervised learning approaches can effectively alleviate the problem of insufficient data; model compression and optimization can reduce the demand for computational resources; and fusion of multimodal data and development of interpretable models can help to improve the clinical applicability of models. Finally, we emphasize the importance of multidisciplinary collaboration and open data sharing to advance the field.

1. INTRODUCTION

Segmenting medical images is an important challenge in computer-aided diagnosis, preoperative planning and treatment monitoring, and its goal is to accurately segment specific anatomical structures or lesion regions to provide effective support for medical diagnosis and treatment. In clinical diagnosis and treatment, medical picture segmentation is important, especially in tasks like organ segmentation, tumor detection, and other tasks with a broad variety of applications [1, 2]. By accurately segmenting complex anatomical structures and lesion regions, medical image segmentation technology can assist doctors in identifying lesions more accurately and provide an important reference for personalized treatment plans. However, due to the anatomical variations of different patients, noise, and the diversity of lesion area characteristics, medical image segmentation always faces many challenges [3].

Deep learning techniques, specially convolutional neural networks (CNNs), have produced impressive results in image segmentation tasks lately. The encoder-decoder architecture-based U-Net model has been widely used in medical image segmentation tasks because of its accuracy in feature extraction and resilience to small-sample data [4]. The encoder-decoder design of the U-Net model, which was suggested by Ronneberger et al., has been widely used in the field of segmenting medical images, demonstrating its powerful characterization of complex image features [5]. However, the original U-Net still has some limitations in the segmentation of complex medical images, especially when dealing with multi-scale features and fine structures, it is difficult to realize the balance between global information and local details.

To address these problems, the researchers developed a multiscale U-Net model that incorporates a multiscale feature fusion technique to improve segmentation accuracy. The multiscale U-Net model combines multiscale features in the decoding stage, which can effectively improve the model's ability to discriminate different scale structures, and has achieved excellent performance in medical image segmentation of brain tumors, heart lesions and other diseases [6-8]. For example, Zhao et al. obtained global contextual information by aggregating large-span features between different layers via a multiscale subtraction module (MMSM). Within the same layer, multi-scale feature fusion from pixel to region is realized by extracting feature self-differentiation information through an improved multi-scale subtraction module, which achieves good results on polyp, breast and retina segmentation datasets [9]. Zhang et al. proposed three different multiscale dense connections for encoders, decoders, and connections between them in a U-shaped architecture. They suggested a multi-scale dense connection U-Net for biological image segmentation based on the three dense connections [10]. These variants further extend the application scope of the U-Net model, especially when dealing with multi-scale information of complex medical images, showing significant effect enhancement.

E provide a thorough analysis of the uses and advancements of multiscale U-Net models in medical picture segmentation in this research. This paper first introduces several common multiscale U-Net variants and their structural features, and then analyzes the performance of these methods in the field of medical image segmentation, focusing on the application effect of multiscale U-Net in the segmentation of different lesions and anatomical structures. The future development direction of multi-scale U-Net in the field of medical image segmentation is also summarized. The contents covered in the review will provide insights and references for future research on multiscale U-Net in medical image segmentation.

2. U-Net basic architecture

Ronneberger et al. first suggested the deep learning model U-Net in 2015 [5]. This model works particularly well with tiny amounts of data and is intended for medical image segmentation applications. The proposed U-Net initially validated the model's performance on a cell segmentation task of electron microscope images, demonstrating the power of U-Net when dealing with small datasets.

The architecture of U-Net is a symmetric U-shape that includes contracting (downsampling) paths and expanding (upsampling) paths. This design allows the network to capture contextual information while maintaining edge information, which is ideal for processing multi-scale spatial information.One of the core features of U-Net is its ability to efficiently utilize a limited number of training samples to make accurate pixel-level predictions. Additionally, U-Net uses skip connections to directly merge low-level and high-level information, improving segmentation accuracy and recovering the image's fine structure. Figure 1 depicts the U-Net architecture [5].

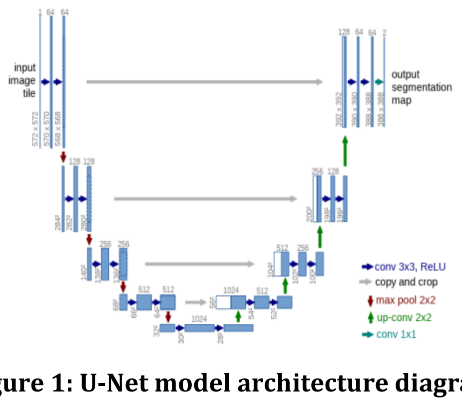

Figure 1: U-Net model architecture diagram

3. MULTI SCALE U-NET

A type of enhanced model based on the U-Net architecture, multiscale U-Net was developed specifically to enhance medical picture segmentation performance, especially for dividing intricate medical images with various scale features. The encoder and decoder, the two symmetrical branches that make up the typical U-Net topology, use skip connections to carry features straight from the encoding stage to the decoding stage, thus preserving the spatial information of the image. However, medical images usually contain fine-grained and coarse-grained information, such as in brain MRI, where tumor areas may be relatively small but the surrounding normal tissues or anatomical structures are large and widely distributed. It is difficult for standard U-Net to effectively handle these structures with different sizes in the same model. Multiple scale It can be partially resolved by U-Net, which incorporates data from multiple scales into the model by implementing multi-scale feature extraction and fusion techniques.

3.1 U-Net++

A more sophisticated feature fusion approach is used in U-Net++, an enhanced version of the U-Net architecture, to enhance medical image segmentation performance. It was first suggested by Zhou and colleagues [11]. U-Net++ introduces sub-networks and nested jump links to optimize feature delivery and fusion. A sequence of nested dense convolutional blocks connects the encoder and decoder that make up U-Net++. U-Net++'s primary goal is to close the semantic gap between the encoder and decoder feature mappings prior to fusion. The U-Net++ architecture is depicted in Figure 2.4 [11].

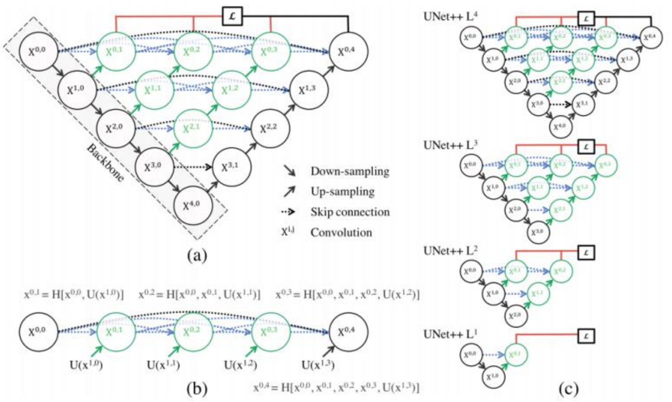

Figure 2: U-Net++architecture diagram

Zhou et al. conducted segmentation experiments on datasets of real cell nuclei, colon polyps, liver, and lung nodules. According to the experimental results, U-Net++ outperformed U-Net and Wid U-Net++ in a number of medical image segmentation tasks without the use of deep supervision (DS) thanks to its enhanced feature fusion and deep supervision techniques [11]. It is especially suitable for those medical image processing scenarios that require high segmentation accuracy and rich feature hierarchy, such as tumor segmentation and organ localization. In addition, the flexible structure and pruning capability of U-Net++ makes it possible to adjust to the needs of specific tasks in order to achieve the best balance of performance and efficiency.

3.2 U-Net3+

Although U-Net++ uses dense skip connections, Huang et al.'s 2020 proposal, U-Net 3+, suggests full-scale skip connections because it contends that U-Net++ does not fully utilize multi-scaled information [12]. Deep supervision learns hierarchical representations from full-scale aggregated feature maps, whereas full-scale skip connections incorporate high-level semantics and low-scale information from feature maps at various scales. In addition to increasing computational efficiency by lowering network parameters, the suggested U-Net 3+ generates more precise location-aware and boundary-enhanced segmentation maps. The method's experimental results, which were confirmed on the liver and spleen, demonstrate that U-Net 3+ works better than other cutting-edge algorithms. The segmentation results emphasize the organs and provide logical borders. A comparison of the U-

Net, U-Net++, and U-Net3+ network architectures is displayed in Figure 3 below [12]. Large-scale feature maps from the decoder and small, same-scale feature maps from the encoder are present in each decoder layer of U-Net 3+, which fully captures coarse-grained semantics and fine-grained details. The suggested U-Net 3+ generates a side output from each decoder step that is supervised by the real labels, as opposed to the full-resolution feature maps created in U-Net++ for depth supervision.

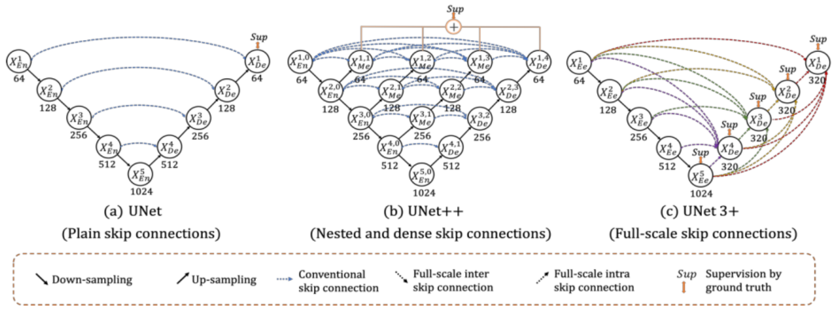

Figure 3: Comparison of U-Net, U-Net++, and U-Net3+ structures

3.3 Variants of multi-scale U-Net combined with other mechanisms

Researchers have also combined multiscale U-Net with other methods to design many classical medical image segmentation models. In order to improve the performance of medical picture segmentation, Cai et al. created MA-Unet, an enhanced U-Net that combines an attentional mechanism with a multiscale mechanism [13]. Experimental results on the lung segmentation dataset and the esophagus dataset used by MA-Unet show that MA-Unet still achieves higher segmentation accuracies than Attention U-Net at relatively low parameter counts, finding a healthy balance between preserving performance and cutting down on computational resource usage [13].

For medical picture segmentation, Chen et al. proposed a TransAttUnet, which combines a Transformer with a Multi-scale Skip Connection [14]. TransAttUnet enhances the representation of multi-scale contextual information by combining two attention mechanisms, Transformer Self Attention and Global Spatial Attention, to capture feature dependencies over large distances and construct global spatial linkages. Multi-scale jump connections are also used in the model to aggregate the up-sampled features in the decoder block to further improve the feature representation and segmentation performance.TransAttUnet exhibits consistent performance improvement on several different medical image datasets, demonstrating its applicability and versatility in various medical image segmentation tasks.

Heidari et al.'s HiFormer creates two multiscale feature representations by utilizing a CNN encoder and the Swin Transformer module [15]. By incorporating a Double-Level Fusion module into the encoder-decoder structure's jump connections, local and global features acquired from CNNs and Transformers can be meticulously fused. The results are superior to other methods for abdominal CT segmentation, dermatologic image segmentation, and myeloma cell segmentation.

4. APPLICATION OF VARIOUS VARIANTS OF MULTI-SCALE U-NET IN MEDICAL IMAGE SEGMENTATION

4.1 Multi scale U-Net for cardiac medical image segmentation

Wang et al. proposed the MMNet model, which uses inflated convolution to extract multi-scale features in the encoding stage to capture semantic information in different perceptual domains [16]. The full-size jump connection structure was reconstructed in the decoding stage, so that the feature information obtained at different levels could be fully utilized in contextual semantic fusion to improve the segmentation accuracy. MMNet achieves high-precision segmentation results on several cardiac datasets. The Dice coefficients were 93.9% and 96.2% for end-systole and end-diastole on the ACDC 2017 dataset, and reached 98.0% on the MICCAI 2009 dataset.

Zhang et al. proposed a model called MSF-TransUNet [17]. The model combines Convolutional Neural Network (CNN) and Transformer architectures and uses a Multiscale Fusion Module (MSFM) to enhance the feature representation (cardiac multiscale). The encoder captures complex cardiac structures through multi-scale convolution and channel attention mechanisms. The MSFM is introduced in the jump connection for fusing multiscale features from different encoder layers to provide detailed segmentation information. On the CAMUS dataset, MSF-TransUNet achieves an average Dice coefficient of 0.9243 and Hausdorff distance (HD) of 9.1902, which significantly outperforms other benchmark models, such as U-Net and TransUNet.

Singh et al. suggested a model named MADRU-Net [18]. For left atrial segmentation in cardiac MRI images, our model combines a deep residual U-Net with a multiscale attention mechanism [18]. The model can more accurately represent the intricate aspects of the heart structure because to the multiscale attention mechanism, which extracts features at several sizes and improves feature transmission through deep residual connections. By using the multiscale attention mechanism and deep residual connections, the model is able to effectively improve the accuracy of feature transfer and thus enhance the capture of cardiac structures. When processing complicated structures in cardiac MRI images, MADRU-Net performs better than other cutting-edge U-Net-based segmentation techniques.

4.2 Multi scale U-Net for medical image segmentation of breast cancer

An enhanced version of Fully Convolutional DenseNet, ASPP-FC-DenseNet, was proposed by Hai et al.[19]and is based on merging multi-scale feature extraction for automatic breast tumor segmentation. Multi-scale image characteristics with varying sample rates are extracted via null-space pyramid pooling, thus enhancing the segmentation accuracy for tumors of different sizes and shapes. Digital molybdenum mammography images gathered from Henan Provincial People's Hospital's radiology department served as the dataset for this research. The experimental findings demonstrate that ASPP-FC-DenseNet performs better in a number of evaluation measures, including Dice index, intersection and concurrency ratio (IoU), and pixel accuracy, when compared to the original FC-DenseNet and other baseline models (such as U-Net, PSPNet, Deeplab v3+, etc.). However, due to the introduction of the ASPP module and multiple null convolutional layers, the overall computational complexity and memory consumption are still high, despite keeping the number of parameters relatively stable when expanding the sensory field. In the original FC-DenseNet, more downsampling operations result in the features of small tumors being easily lost in the downsampling process. Therefore, this is improved in the article by reducing the number of downsampling operations, but this may also affect the model's adaptability and processing capability at different scales.

MDF-Net (Multi-scale Dynamic Fusion Network) is a network that Qi et al. proposed for ultrasound breast tumor segmentation [20]. To increase the model's flexibility in response to patient variations in lesion size and shape, MDF-Net presents a multi-scale dynamic fusion mechanism (MDFM) for fusing the original segmentation data at both coarse and fine scales.Two publicly accessible datasets of breast ultrasound images were used to assess MDF-Net, namely Breast The Breast Ultrasound Lesions Dataset and the Breast Ultrasound Dataset are the two publicly accessible breast ultrasound picture datasets that MDF-Net uses for evaluation.MDF-Net effectively fuses coarse-to-fine-scale data by implementing a multi-scale dynamic fusion method, which improves tumor feature capture and flexibility to changes in tumor size and shape. Although MDF-Net improves resistance to noise through deep supervision and multiscale fusion, there are still some challenges when dealing with particularly low-contrast ultrasound images.

4.3 Multi scale U-Net for segmentation of various other lesions

The Triple Unet with Multi Scale Input Features and Dense Skip Connection (TMD Unet) model was suggested by Tran et al. [21]. This model is made up of three parallel sub-U-Net models that form a triple structure. Instead than using normal convolution, we employed dilated convolution and incorporated dense skip connections and multi-scale input characteristics. Tran et al. conducted medical image segmentation experiments on seven datasets including colon polyp dataset, cell nucleus dataset, skin lesion dataset, and spleen, demonstrating the effectiveness of their proposed multi-scale U-Net model in segmenting medical images of various diseases. It also reflects the generalization ability of the model. However, due to the use of three parallel sub U-Net models and dense skip connections in TMD-Unet, the number of model parameters increases, resulting in higher computational complexity and memory usage.

Ahmed et al. proposed a multi-scale U-Net variant called DoubleU NetPlus. This model adopts a dual stacked U-Net structure and combines a multi-scale residual feature fusion networ [22]. EfficientNetB7 is used as the feature encoder. Integrated multi-core residual convolution module and attention based hollow space pyramid pooling module. And improved the selective modeling of medical picture aspects by introducing the Triple Attention Gate and Hybrid Triple Attention Module. Six publicly available benchmark datasets were utilized to assess the DoubleU-NetPlus model's performance, including DRIVE (diabetes retinal image), LUNA (lung nodule analysis CT image), CVCclinicDB (colonoscopy image), 2018 DSB (nuclear microscopic image), and ISBI 2012 (Drosophila brain microscopic image). And achieved good experimental results. However, due to the use of a dual U-Net stacked structure, an EfficientNetB7 encoder, and multiple attention modules in the model, the computational complexity and parameter count of the model are relatively high, resulting in a longer training time. Below is a comparative analysis of various multi-scale U-Net variant algorithms, as shown in Table 1.

Table 1: Comparative analysis of multi-scale U-Net variant algorithms

Models	Model Type	Datasets Used	Advantages	Disadvantages
MSF-TransUNet	Multi-Scale Fusion TransUNet (CNN + Transformer with Multi-Scale Fusion Module)	CAMUS, ACDC	Effectively improved the accuracy of segmentation, solved the problem of class imbalance, and enhanced the segmentation performance of complex structures.	High computational complexity, Requires high-performance hardware
MADRU-Net	MADRU-Net (Multiscale Attention-Based Deep Residual U-Net)	ACDC 2017, 2018 ASC, LAScarQS 2022	Effectively improve the accuracy of feature transmission, thereby enhancing the capture of cardiac structures.	High computational complexity, Dependent on extensive data augmentation
MMNet	MMNet (Multi-scale Multi-skip Connection Network)	MICCAI 2017 ACDC, MICCAI 2009, MICCAI 2018	Effectively extract heart edge information and maintain the segmentation accuracy of the model for complex structures. Effectively prevents overfitting of the model, especially when the dataset is small	High computational complexity, Issues with small sample sizes and blurry edges
ASPP-FC-DenseNet	ASPP-FC-DenseNet with Multi-Scale Context Feature Extraction	Digital mammography images from a certain hospital	ASPP for Multi-Scale Feature Extraction, Better Boundary Clarity, High Segmentation Precision	High Computational Complexity, Loss of Small Lesion Information
MDF-Net	Multi-Scale Dynamic Fusion Network	Breast Ultrasound Dataset	Dynamic Fusion of Coarse and Fine Scale Features, Enhanced Adaptability to Patient Variability	High Computation and Training Resource Demand
TMD-Unet	Triple U-Net with Multi-Scale Input Features and Dense Skip Connection	ISBI 2012, MICCAI 2015, DSB 2018 etc.	Multi-Scale Input Features, Enhanced Low-Level Feature Utilization, High Generalization	High Computational Complexity, Long Training Time
DoubleU-NetPlus	Double U-Net with Multi-Kernel Residual Convolution and Attention Modules	DRIVE, LUNA, BUSI, CVCclinicDB, DSB 2018, ISBI 2012	Efficient Net Encoder, Multi-Kernel Residual Convolution, Triple Attention Mechanism	Increased Parameter Size, High Computation Requirement

5. CHALLENGES AND FUTURE DEVELOPMENT DIRECTIONS

With the increasing application of multi-scale U-Net and its variants in medical image segmentation, although they have achieved significant results in multiple application scenarios, there are still some urgent challenges that need to be addressed. The following are the challenges and development directions that multi-scale U-Net and its related models may face in the future.

5.1 Computational complexity and hardware requirements

By adding multi-scale feature fusion, attention mechanisms, and other modules, multi-scale U-Net models typically enhance segmentation performance; however, this also significantly increases the model's computational complexity and parameter count. This high complexity puts higher demands on hardware resources and training time, especially requiring more powerful GPUs or TPUs to process high-resolution medical images. In the future, how to effectively reduce computational complexity and memory consumption while maintaining or improving model performance is an important issue for researchers to address. The amount of parameters and computational complexity of models can be decreased in the future by examining methods such knowledge distillation, quantization, pruning, and model compression. Designing lightweight models suitable for embedded devices and mobile facilitates the real-time application of multi-scale U-Net in clinical environments.

5.2 Generalization ability and small sample learning

Medical images often face problems such as limited data volume, expensive labels, and inconsistent quality, which makes the generalization ability of the model a key challenge. The multi-scale U-Net model is prone to overfitting in some cases, especially when the data volume is insufficient, and its performance may significantly decrease. Therefore, enhancing the model's capacity to generalize across various datasets, imaging equipment, and imaging settings is a pressing issue that has to be resolved. The model's flexibility can be improved by investigating more efficient semi-supervised learning, transfer learning, and data augmentation strategies to make the most of a small amount of labeled data and a big number of unlabeled data.

5.3 Multiscale and multimodal fusion

The diversity of medical images is such that information from a single modality is often insufficient to support accurate lesion area segmentation. Multimodal fusion (e.g., MRI, CT, PET, etc.) can provide richer anatomical and lesion information. However, the application of multi-scale U-Net variants in multimodal data fusion still faces challenges, including how to effectively align and extract features from different modalities, and how to effectively fuse information between different modalities.

5.4 Multidisciplinary collaboration and open platform construction

In order to better promote the development of medical image segmentation technology, it is recommended to strengthen cooperation in fields such as computer science, medical imaging, and clinical medicine, and establish open research platforms and data sharing mechanisms. Through interdisciplinary collaboration, promote resource integration and knowledge sharing, accelerate technological innovation and application implementation.

In summary, the application of multiscale U-Net and its variants in the field of medical image segmentation is promising, but also faces many challenges. Through further research in computational complexity, generalization ability, multimodal fusion, and multidisciplinary collaboration, we believe that these models can better serve the actual medical image analysis tasks and provide more reliable support for clinical diagnosis and treatment.

Acknowledgement: The authors are grateful to the Management and Science University in Malaysia, Guangzhou College of Commerce in China, and the University of Mosul in Iraq for their supports. This work is funded by Guangdong Province Undergraduate Teaching Quality and Teaching Reform Project (No. 2022JXGG01), Guangdong Higher Education Association's "14th Five Year Plan" 2024 Higher Education Research Project (No. 24GYB75), Project of the Teaching Quality Management Professional Committee of Private Universities of Guangdong Higher Education Teaching Management Association for the year 2024 (No. GDZLGL2428), Guangdong Provincial Higher Vocational Education Teaching Quality Project (No.2023JG452), Guangdong Provincial Education

Science Foundation (No. 2023GXJK407), Guangzhou Business School 2022 Campus Quality Engineering Construction Project (No. 2022JXGG68), 2022 Guangzhou Business School Teaching Quality and Teaching Reform Project (No. 2022JXGG30), 2024 Guangzhou Business School Campus level Research Project (No. 2024XJYB024).

REFERENCES

Ma, J., et al., *Segment anything in medical images.* 2024. 15(1): p. 654.

Qureshi, I., et al., *Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends.* 2023. 90: p. 316-352.

Xun, S., et al., *Generative adversarial networks in medical image segmentation: A review.* 2022. 140: p. 105063.

Azad, R., et al., *Medical image segmentation review: The success of u-net.* 2024.

Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. 2015. Springer.

Rehman, M.U., et al., *Bu-net: Brain tumor segmentation using modified u-net architecture.* 2020. 9(12): p. 2203.

Allah, A.M.G., A.M. Sarhan, and N.M.J.E.S.w.A. Elshennawy, *Edge U-Net: Brain tumor segmentation using MRI based on deep U-Net model with boundary information.* 2023. 213: p. 118833.

Diniz, J.O.B., et al., *An automatic approach for heart segmentation in CT scans through image processing techniques and Concat-U-Net.* 2022. 196: p. 116632.

Zhao, X., et al., *M ^{2} SNet: Multi-scale in multi-scale subtraction network for medical image segmentation.* 2023.

Zhang, J., et al., *Mdu-net: Multi-scale densely connected u-net for biomedical image segmentation.* 2023. 11(1): p. 13.

Zhou, Z., et al., *Unet++: Redesigning skip connections to exploit multiscale features in image segmentation.* 2019. 39(6): p. 1856-1867.

Huang, H., et al. *Unet 3+: A full-scale connected unet for medical image segmentation*. in *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. 2020. IEEE.

Cai, Y. and Y. Wang. *Ma-unet: An improved version of unet based on multi-scale and attention mechanism for medical image segmentation*. in *Third international conference on electronics and communication; network and computer technology (ECNCT 2021)*. 2022. SPIE.

Chen, B., et al., *Transattunet: Multi-level attention-guided u-net with transformer for medical image segmentation.* 2023.

Heidari, M., et al. *Hiformer: Hierarchical multi-scale representations using transformers for medical image segmentation*. in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2023.

Wang, Z., et al., *MMNet: A multi-scale deep learning network for the left ventricular segmentation of cardiac MRI images.* 2022. 52(5): p. 5225-5240.

Zhang, N. and J. Li. *MSF-TransUNet: A Multi-Scale Fusion Approach for Precise Cardiac Image Segmentation*. in *Proceedings of the 2024 4th International Conference on Artificial Intelligence, Big Data and Algorithms*. 2024.

Singh, K.R., et al., *MADRU-Net: Multi-Scale Attention-based Cardiac MRI Segmentation using Deep Residual U-Net.* 2023.

Hai, J., et al., *Fully convolutional densenet with multiscale context for automated breast tumor segmentation.* 2019. 2019(1): p. 8415485.

Qi, W., H.-C. Wu, and S.-C.J.I.T.o.I.P. Chan, *Mdf-net: A multi-scale dynamic fusion network for breast tumor segmentation of ultrasound images.* 2023.

Tran, S.-T., et al. *Tmd-unet: Triple-unet with multi-scale input features and dense skip connection for medical image segmentation*. in *Healthcare*. 2021. MDPI.

Ahmed, M.R., et al., *DoubleU-NetPlus: a novel attention and context-guided dual U-Net with multi-scale residual feature fusion network for semantic segmentation of medical images.* 2023. 35(19): p. 14379-14401.