Research Article

A Novel Brain Image Segmentation Method Using an Improved 3D U-Net Model

Zhuqing Yang

Jiangsu Vocational College of Information Technology, Wuxi, Jiangsu 214153, China

Correspondence should be addressed to Zhuqing Yang; yangzq@jsit.edu.cn

Received 26 July 2021; Revised 10 August 2021; Accepted 13 August 2021; Published 19 August 2021

Academic Editor: Chenxi Huang

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

Medical image segmentation (IS) is a research field in image processing. Deep learning methods are used to automatically segment organs, tissues, or tumor regions in medical images, which can assist doctors in diagnosing diseases. Since most IS models based on convolutional neural network (CNN) are two-dimensional models, they are not suitable for three-dimensional medical imaging. On the contrary, the three-dimensional segmentation model has problems such as complex network structure and large amount of calculation. Therefore, this study introduces the self-excited compressed dilated convolution (SECDC) module on the basis of the 3D U-Net network and proposes an improved 3D U-Net network model. In the SECDC module, the calculation amount of the model can be reduced by $1 \times 1 \times 1$ convolution. Combining normal convolution and cavity convolution with an expansion rate of 2 can dig out the multiview features of the image. At the same time, the 3D squeeze-and-excitation (3D-SE) module can realize automatic learning of the importance of each layer. The experimental results on the BraTS2019 dataset show that the Dice coefficient and other indicators obtained by the model used in this paper indicate that the overall tumor can reach 0.87, the tumor core can reach 0.84, and the most difficult to segment enhanced tumor can reach 0.80. From the evaluation indicators, it can be analyzed that the improved 3D U-Net model used can greatly reduce the amount of data while achieving better segmentation results, and the model has better robustness. This model can meet the clinical needs of brain tumor segmentation methods.

1. Introduction

A basic task in medical IS [1, 2] is to extract specific organs and tumors from different types of medical images. Organ and tumor segmentation provides an important basis for cancer diagnosis, surgical planning, and pathological analysis. The segmentation of organs and tumors in the clinic faces two main challenges. The first challenge is the quality of medical images. Due to the diversity of clinical needs, doctors usually choose different imaging examination methods to produce different types of medical images, such as computed tomography (CT) [3], magnetic resonance imaging (MRI) [4], X-ray [5], and ultrasound [6]. Different types of imaging equipment have different principles and will be interfered by various factors during the image acquisition process. There are many reasons for interference, which can be summarized into the following three points.

One is that due to individual differences, the organs or tumors of patients of different body shapes and ages will show different shapes of anatomical structures on medical images. Second, in the image acquisition process, the patient's breathing, blood flow, heart circulation, and other factors will cause the anatomical structure to deform. The overlapping of soft tissues such as organs causes unclear boundaries. The third is the defects of the equipment itself and the interference from the outside. The above factors will reduce the image quality and increase the difficulty of medical IS. The second challenge is the characteristics of the disease itself. Some tumors appear in the early stage of medical images with small volume, small difference in texture from surrounding tissues, and low contrast, which affects the accuracy of segmentation. These small tumors are easily misdiagnosed as calcification or fatification or even ignored, leading to misdiagnosis or missed diagnosis of

cancer. It is an arduous and complicated task to screen out images with tumors and segment them from a large number of medical images. Therefore, there is an urgent clinical need to develop a method that can accurately and automatically segment tumor regions from a large number of medical images to assist doctors in cancer diagnosis.

Traditional medical IS methods include threshold method [7, 8], graph cut method [9, 10], and region growing method [11, 12]. The threshold method mainly calculates the threshold value suitable for the segmentation task according to the manually designed criteria and compares the gray value of each pixel in the image with the threshold value, thereby separating the target from the background. The core idea of the region growing method is to select a seed pixel in different regions and use the texture, gray, gradient, color, and other characteristics as the criterion to measure the similarity between all other pixels in the region and the seed pixel. Pixels that meet the similarity criterion are classified into one category to achieve IS. The core idea of the graph cut method is to take the pixels of the image as the vertices of the graph and establish the graph structure from the predefined vertex connection relationship. By designing a suitable cutting criterion to remove the edges that do not conform to the criterion, several unconnected subgraphs are obtained. These subgraphs constitute the final segmentation result. Traditional IS methods rely on manually set parameters, and setting these parameters requires a lot of medical expertise. With the emergence of bottlenecks in traditional medical segmentation methods, there are more and more medical IS studies based on deep learning methods. The segmentation model based on deep learning does not require manual setting of parameters. Under the guidance of supervision information, features can be automatically learned from the given training samples, which significantly improve the efficiency and accuracy of IS.

In recent years, medical IS methods based on deep learning have made good progress. All fully connected layers of neural networks such as AlexNet [13] and VGG Net [14] are replaced with convolutional layers, and fully convolutional networks (FCNs) [15] are proposed, and this network is applied to the field of IS for the first time. In order to extract multiscale features of an image, convolution or pooling operations are usually used to change the feature size of the image. Since the size of the IS result needs to be consistent with the original image size, an upsampling operation is used to restore the feature map size to the original image size. The part of the network used to extract features is usually called the encoder, and the part of the network that is restored from the features to the original image size is called the decoder. The encoder and decoder form an encoding-decoding structure. Based on the encoding-decoding structure, U-Net [16] introduces skip connections to integrate low-level semantic information with high-level semantic information, which further improves the segmentation performance of the network. The success of U-Net made other networks choose U-Net as the backbone of the model. Iglovikov and Shvets [17] regard VGG Net as the encoder of U-Net and improve the performance of U-Net by pretraining the weights of VGG Net

on ImageNet. Attention U-Net [18] introduces an attention mechanism into the decoder of U-Net [19], which effectively suppresses the influence of areas that are not related to the target in the medical image. Sun et al. [20] introduce the Attention-Up module to improve the symmetric structure of U-Net to an asymmetric structure.

The above methods mainly use 2D models to segment clinical medical images. However, clinical medical images are usually stacked by multiple slices, and adjacent slices of the image are sometimes related to each other. Only the 2D model cannot learn the features between image slices. With the continuous deepening of research, many 3D segmentation models for 3D medical images have begun to emerge. V-Net [21] and 3D U-Net [22] replace two-dimensional convolution with three-dimensional convolution and realize the transformation from the 2D model to the 3D model. Compared with the 2D model, the 3D model encodes the image in three directions and extracts the three-dimensional features of the image. 3D models need to consume a lot of computing and storage resources when calculating, so the image needs to be cropped into image blocks of a certain size and input into the network for calculation. In order to take advantage of the respective advantages of the 2D and 3D medical IS models, H-Dense U-Net [23] combines the 2D segmentation model with the 3D segmentation model. The 2D model is used to extract the intraslice features of the three-dimensional medical image, and the 3D model is used to extract the interslice features, and the segmentation accuracy is improved by fusing the intraslice features and the interslice features.

Although the CNN-based IS model performs well in a variety of segmentation tasks, its performance still has room for improvement. Since CNN-based IS models are mostly two-dimensional models, they are not suitable for three-dimensional medical imaging. Therefore, this article uses 3D U-Net model to segment medical images. The work done in this study is summarized as follows:

- (1) The advantages and disadvantages of various network models in medical IS are compared and analyzed, and the 3D U-Net network is used for brain IS tasks.
- (2) Due to the complex network structure and large amount of calculation in the 3D segmentation model, this paper introduces the SECDC module optimization model. The introduction of the SECDC module can reduce the calculation parameters and effectively reduce the amount of calculation.
- (3) The abovementioned improved 3D U-Net network is applied to the brain image dataset. The IS evaluation index is used to verify the experimental results, and the results show the effectiveness of the model used. Based on the better segmentation performance of the used model, it has certain clinical significance for the diagnosis of diseases.

2. D U-Net Model

2.1. Network Structure. At present, there are two strategies for 3D medical IS tasks. One is to send the 2D slices of the volume into the 2D network for training. The training speed

of this strategy is fast, but the spatial location information is insufficient. The second is to send the volume directly into the 3D network for training. This method has a large amount of parameters, a long training time, and high requirements on hardware conditions, but its segmentation effect is better than that of a 2D network. The use of CNN network can make 2D biomedical IS reach the accuracy close to human manual segmentation. It is because of such successful applications that people begin to use CNN network to segment 3D data. In 2016, Çiçek et al. [24] proposed a 3D U-Net network structure for learning 3D segmentation from sparsely annotated stereo data. The 3D U-Net model structure is shown in Figure 1.

2.2. Network Advantages and Disadvantages. The biggest feature of the 3D U-Net network is the U-shaped codec structure and jump connection so that the shallow features can be well integrated with the high-level abstract features. These features are very effective for medical images with continuous structure, fuzzy boundaries, and simple semantics. Without considering the calculation and memory performance, the 3D U-Net network can combine the information between image layers to ensure the continuity of changes between the image masks of the interlayer. This feature is easier to obtain better results than using 2D slices for training.

The network structure of 3D U-Net [25] realizes the segmentation in two scenarios. One is for semiautomatic segmentation, and the other is for fully automatic segmentation. In semiautomatic segmentation, users mark some slices in the volume to be segmented. The network then learns from these sparse annotations and provides a dense 3D segmentation result. In fully automated segmentation, it is assumed that there is a representative training set with sparse annotations. Trained on this dataset, the network can intensively segment new volumetric images.

The main disadvantage of the 3D U-Net network lies in the memory usage, which makes it impossible to use the entire 3D patch as input. Therefore, it needs to be tailored. Generally, the entire volume is cut into a series of 3D patches of the same size as input. Using patch for training will limit the size of the maximum receptive field that the network can reach, resulting in the loss of certain global information. If the target to be segmented is much larger than the patch block, it is difficult for the network to learn the overall structure of the target.

3. Image Segmentation Based on Improved 3D U-Net Model

3.1. Model Segmentation Process. Generally speaking, the IS process based on the network model is shown in Figure 2. First, preprocess the image to be segmented. Preprocessing mainly includes cutting a series of standardized processing. Second, the training data is input into the model structure, and the IS model is obtained through training. In this study, we use an improved 3D U-Net model for training. Third, when the network training is over, the performance of the

trained network needs to be tested. By inputting the test data, the evaluation index is calculated according to the test result. Multidimensional evaluation of trained network performance based on different evaluation indicators is carried out.

3.2. Improved 3D U-Net Model. This article introduces the SECDC module to optimize the convolution operation in 3D U-Net. This module first uses $1 \times 1 \times 1$ convolution to reduce the dimensionality of the network layer. Then, feature mapping is performed on the input layer with normal convolution and dilated convolution with an expansion rate of 2. Then, learn the importance of different layers through the 3D-SE module, and finally, use $1 \times 1 \times 1$ convolution to upgrade the dimension. The structure of the module is shown in Figure 3.

3.2.1. 3D-SE Module. The Squeeze-and-Excitation (SE) module [26] is mainly used to measure the relationship between channels so that the model can automatically learn the importance of different channel features. Two key operations, Squeeze and Excitation, are included in the SE module. The Squeeze operation aggregates the feature maps obtained by convolution to obtain the feature map with dimension $W \times H$ as the feature descriptor, so as to obtain the information of the global receptive field. The Excitation operation is a self-screening mechanism that uses a sample specialized activation function to evaluate the weights of all channels. The module structure is shown in Figure 4.

The mapping rules of the SE module can be expressed as $F_{tr}: X \rightarrow U, X \in R^{H' \times W' \times C'}, U \in R^{H \times W \times C}$. In the convolutional map, let the convolution kernel be $V = [v_1, v_2, \dots, v_c]$, where v_c represents the c th convolution kernel, and the output $U = [u_1, u_1, \dots, u_c]$ can be expressed as

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s, \quad (1)$$

where $*$ represents convolution and v_c^s represents a 3D convolution kernel. It inputs spatial features on a channel to learn the relationship between feature spaces. However, the convolution results of each channel are summed so that the feature relationship of each channel is merged with the spatial relationship of the convolution kernel. This fusion is not good for the training of the model. The SE module avoids this unprofitable fusion and aims to allow the model to directly obtain the characteristic relationship of each channel.

Compared with the original SE module, the improved 3D SE module can be applied to three-dimensional convolution. It mainly focuses on the importance of different channel features in three-dimensional space features. Figure 5 gives the structure of the 3D SE module.

The output formula of the 3D SE module is as follows:

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s, \quad (2)$$

FIGURE 1: 3D U-Net model structure.

FIGURE 2: Image segmentation process based on the network model.

FIGURE 3: SECDC module structure.

FIGURE 4: SE module structure.

FIGURE 5: 3D SE module.

where the difference between SE and 3D SE modules is that γ_c^s represents a 4D convolution kernel, which can be directly combined with the 3D convolution operation in 3D U-Net.

3.2.2. $1 \times 1 \times 1$ Convolution. Add $1 \times 1 \times 1$ convolution before and after convolution, as shown in Figure 3. $1 \times 1 \times 1$ convolution has two main functions. One is to realize

cross-channel interaction and information integration. The $1 \times 1 \times 1$ convolution keeps the width, height, and depth of the data unchanged and linearly combines multiple features to achieve cross-channel information integration and improve the expressive ability of the network. The second is to reduce and increase the number of channels of the convolution kernel. Since the $3 \times 3 \times 3$ convolution is very time-consuming to perform convolution operations on the convolutional layer of hundreds of filters, the $1 \times 1 \times 1$

convolution is used to reduce and increase the dimensionality before and after the $3 \times 3 \times 3$ convolutional layer. Dimensional operation reduces the number of parameters and shortens time.

3.2.3. Dilated Convolution. Dilated convolution [27, 28] injects holes in the standard convolution map. Compared with normal convolution, the dilated has one more hyperparameter called the dilation rate, which represents the number of intervals between the convolution kernels. The dilatation rate of the convolution kernel for normal convolution operation is 1. The advantage of dilated convolution is that, without pooling loss information, the receptive field is enlarged so that each convolution output contains a larger range of information. The SECD module structure is shown in Figure 3.

4. Experiment

4.1. Segmentation Tasks and Experimental Datasets. The segmentation task is to accurately segment the three tumor regions of brain tumor patients, namely, whole tumor (WT), tumor core (TC), enhanced tumor (ET), and background region. Among them, WT refers to the area containing all tumor types, TC refers to the area containing all gangrene types, and ET refers to the area containing all enhanced tumor types. The dataset used in the experiment is BraTS2019 [29]. The experimental dataset selected MRI images of 248 HGG patients and 68 LGG patients. The MRI image of each patient includes four registered modal images and real segmentation label images. Each modal and real segmentation label has 155 scanned images of different layers with a size of 240×240 .

Before network training, preprocess the MRI image. Cut and standardize the image first. The size of the processed image is 160×160 , and methods such as flipping, rotation, and elastic deformation continue to be used for data enhancement. The test data enhancement here refers to multiple transformations of an image, including folding, rotating, cropping, scaling, and adding random noise. Input it into the model and calculate multiple versions of data and finally get the average output as the final output of the image. Use 5-fold cross validation in the experiment to avoid data bias. Images of 180 HGG patients and 50 LGG patients were used as training samples. Images of 68 HGG patients and 18 LGG patients were used as the test set. The network inputs the data of each patient, including preprocessed MRI images of FLAIR, T1, T1ce, and T2, and real segmentation labels. The network output is a segmentation map of each patient, including 4 types of segmentation labels: WT, TC, ET, and background area.

4.2. Evaluation Index. Four evaluation indicators are used in the experiment to evaluate the segmentation performance of the algorithm. The four evaluation indicators are Dice coefficient, Positive Predictive Value (PPV), Sensitivity, and Hausdorff distance. The calculation formulas of each indicator are as follows:

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

$$\text{PPV} = \frac{TP}{FP + TP},$$

$$\text{Sensitivity} = \frac{TP}{TP + FN},$$

$$\text{Hans}(A, P) = \max \left[\sup_{a \in A} \inf_{p \in P} d(a, p), \sup_{p \in P} \inf_{a \in A} \text{dis}(a, p) \right], \quad (3)$$

where TP and TN represent pixels as True Positive and True Negative and FP and FN represent pixels as False Positive and False Negative. In the calculation formula of Hausdorff distance, sup represents the supremum, inf represents the inferior, A is the manually marked tumor area, a is the point on A , P is the predicted tumor area, p is the point on P , and $\text{dis}(a, p)$ represents the function used to calculate the distance between two points. Dice coefficient, PPV, and Sensitivity are used to evaluate the overlap between the real value and the predicted result. Large values of these three indicators represent good performance of the algorithm. Hausdorff distance is used to calculate the distance between the true value boundary and the predicted area boundary. A small Hausdorff value represents good performance of the algorithm.

4.3. Experimental Setup and Environment. Before starting to train the entire model, the hyperparameters in the model need to be set reasonably. This study uses a five-fold cross-validation method to determine the optimal hyperparameters to improve the effect and performance of the network. In the experiment, all weights are initialized randomly using $N(0, 1)$ Gaussian distribution. According to the graphics card and memory conditions, set batch_size to 1. The initial learning rate is set to 0.001 using Adam optimization algorithm [30]. The Adam optimization algorithm calculates the gradient's first-order moment estimation and second-order moment estimation to adapt the learning rate. The exponential decay rate of the first-order moment estimation is set to 0.88, and the exponential decay rate of the second-order moment estimation is set to 0.97. The comparison algorithms used in the experiment are CNN, U-net, and 3D U-net. The experimental environment is shown in Table 1.

4.4. Experimental Results and Analysis. The segmentation results of the four comparison models on the dataset are shown in Table 2. In order to compare the experimental data more vividly, this study uses a histogram to show the difference between each model in each indicator, as shown in Figure 6.

Table 2 shows the segmentation results of the four models for the three tissues. It can be seen from Table 2 that, in terms of Dice, PPV, and Sensitivity, 3D U-net and U-Net score higher than CNN. This shows that U-Net is more

TABLE 1: Experimental environment.

Configuration	Details
Operating system	Windows10 DircetX12 64
RAM	32G
Processor	Intel core i7-87003
Graphics card	Nvidia GeForce GTx 1080Ti
Programming framework	PyTorch
Architecture platform	CUDA8.0

TABLE 2: Image segmentation results of each comparison model.

Index	Segmented organization	CNN	U-net	3D U-net	Proposed
Dice	WT	0.7992	0.8285	0.8428	0.8787
	TC	0.7844	0.8006	0.8211	0.8456
	ET	0.7020	0.7602	0.7993	0.8004
	Mean	0.7619	0.7964	0.8211	0.8416
PPV	WT	0.8002	0.8473	0.8687	0.8992
	TC	0.7946	0.8331	0.8505	0.8668
	ET	0.7553	0.7717	0.7899	0.8321
	Mean	0.7837	0.8174	0.8364	0.866
Sensitivity	WT	0.8335	0.8586	0.8640	0.8899
	TC	0.8664	0.8903	0.8947	0.9078
	ET	0.7608	0.8192	0.8193	0.8335
	Mean	0.8202	0.856	0.8593	0.8771
Hausdorff	WT	2.9746	2.5762	2.6227	2.5002
	TC	1.8345	1.7481	1.7008	1.5101
	ET	3.2023	2.8350	2.8088	2.6443
	Mean	2.6704	2.3864	2.3774	2.2182

FIGURE 6: Continued.

FIGURE 6: Segmentation comparison of each model on different indicators.

TABLE 3: The effect of different improved 3D U-net models on brain tumor segmentation.

Index	Segmented organization	[31]	[32]	[33]	Proposed
Dice	WT	0.7432	0.8328	0.8024	0.8787
	TC	0.6995	0.8006	0.7895	0.8456
	ET	0.6006	0.7670	0.7162	0.8004
	Mean	0.6811	0.8011	0.7964	0.8416
Hausdorff	WT	3.4675	2.8254	3.0052	2.5002
	TC	2.3006	1.8145	2.1723	1.5101
	ET	3.4783	2.8380	3.1016	2.6443
	Mean	3.0821	2.4926	2.7597	2.2182

suitable for medical IS. In the three indicators of Dice, PPV, and Sensitivity, 3D U-net has a larger index value than U-net. On the Hausdorff index, 3D U-net has a smaller value than U-net segmentation. This shows that 4 indicators all prove that 3D U-net has better segmentation effect than U-Net. From the data in the table, it can be seen that the method proposed in this paper has greatly improved the scores of Dice, PPV, and Sensitivity, and the value of Hausdorff scores has been significantly reduced. The 3D U-net that introduces the SECDC module can automatically learn the importance of different layers, thereby improving the segmentation accuracy and robustness of the model so that the segmentation effect of the model used in this article is optimal. Specifically, in terms of the Dice indicator, compared with CNN, U-Net, and 3D U-net, the model used has increased by 10.46%, 5.68%, and 2.50%, respectively. The Dice index of tumor segmentation of the whole tumor area exceeded 0.87, and the segmentation Dice index of tumor enhancement area reached 0.80. In the PPV index, compared with CNN, U-Net, and 3D U-net, the model used has increased by 10.50%, 5.95%, and 3.54%, respectively. In the Sensitivity index, compared with CNN, U-Net, and 3D U-net, the models used have increased by 6.94%, 2.46%, and 2.07%, respectively. In the Hausdorff index, compared with

CNN, U-Net, and 3D U-net, the models used are reduced by 16.93%, 7.05%, and 6.70%, respectively. From the analysis of the data, it can be seen that the model proposed in this paper is superior to other methods in terms of performance, accuracy, and sensitivity.

Since there are many improvement strategies for 3D U-net, many studies have applied their improved 3D U-net models to IS. In order to compare the effects of different improvement strategies on IS performance, and this paper compares the model used with other improved 3D U-Net models. Table 3 shows the performance comparison between the proposed model and other types of improved 3D U-Net models. In order to compare the experimental data more vividly, this study uses a histogram to show the difference between each model in each indicator, as shown in Figure 7.

Analyzing the segmentation performance of the three parts, the model segmentation in [31] has the worst effect, followed by [33], and then [32]. The IS performance based on the proposed model is relatively good. On the indicator Dice, for [31–33], the proposed model increased by 23.56%, 5.06%, and 5.68%, respectively. On the indicator Hausdorff, for [31–33], the proposed model increases by 28.03%, 11.01%, and 19.62%, respectively. According to the experimental

FIGURE 7: Segmentation comparison of various studies on different indicators.

results, it can be seen that the proposed model is superior to other methods in different segmentation areas of the tumor, and the overall segmentation performance is better.

5. Conclusion

Aiming at the unsatisfactory effect of traditional CNN on the segmentation of 3D medical images, this paper selects the U-Net model with higher performance for brain IS. Considering that the 3D U-Net model has problems such as complex network structure and large amount of calculation, this paper introduces the SECD module into 3D U-Net, thereby constructing a high-precision lightweight segmentation model. The improved 3D U-Net network uses $1 \times 1 \times 1$ convolution to reduce the amount of parameters. Based on the normal convolution and the dilated convolution with the expansion rate of 2, the image features under different fields of view are effectively explored. At the same time, the 3D-SE module is introduced to effectively automatically learn the importance of different layers, thereby improving the robustness of the model. The experimental results on the BraTS2019 dataset prove the superiority of this method. However, in practical applications, there are still problems such as large sample labeling workload and long model segmentation time. According to these problems, the algorithm in this paper can be further optimized to achieve rapid network segmentation and efficient diagnosis by doctors. In addition, future work will continue to study the relationship between encoding and decoding and make full use of low-level features and semantic information to optimize the results.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Research Project of Jiangsu Province Colleges and Universities, under Grant 18KJD510011, and Jiangsu Province High-Level Key Professional Construction Project funding under Grant Su Jiaogao (2017) no. 17.

References

- [1] M. Gridach, "PyDiNet: pyramid dilated network for medical image segmentation," *Neural Networks*, vol. 140, pp. 274–281, 2021.
- [2] O. Avalos, E. Ayala, F. Wario, and M. Pérez-Cisneros, "An accurate Cluster chaotic optimization approach for digital medical image segmentation," *Neural Computing & Applications*, vol. 33, no. 16, pp. 10057–10091, 2021.
- [3] S. Rakesh and S. Mahesh, "Nodule segmentation of lung CT image for medical applications," *Global Transitions Proceedings*, vol. 2, no. 1, pp. 80–83, 2021.
- [4] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019.
- [5] L. Lu, M. Sun, Q. Lu, T. Wu, and B. Huang, "High energy X-ray radiation sensitive scintillating materials for medical imaging, cancer diagnosis and therapy," *Nanomaterials and Energy*, vol. 79, Article ID 105437, 2021.
- [6] D. Avola, L. Cinque, A. Fagioli, G. Foresti, and A. Mecca, "Ultrasound medical imaging techniques," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–38, 2021.
- [7] P. Li, T. Shi, Y. Zhao, and A. Lu, "Design of threshold segmentation method for quantum image," *International Journal of Theoretical Physics*, vol. 59, no. 2, pp. 514–538, 2020.

- [8] Y. Song, Y. L. Xu, and W. Gao, "A threshold segmentation method of chromosome microscope image," *Acta Microscopica*, vol. 28, no. 2, pp. 188–194, 2019.
- [9] Y. Yin, H. Luo, J. Sa, and Q. Zhang, "Study and application of improved level set method with prior graph cut in PCB image segmentation," *Circuit World*, vol. 46, no. 1, pp. 55–64, 2020.
- [10] M. Bernier, P.-M. Jodoin, O. Humbert, and A. Lalande, "Graph cut-based method for segmenting the left ventricle from MRI or echocardiographic images," *Computerized Medical Imaging and Graphics*, vol. 58, pp. 1–12, 2017.
- [11] S. Liu, Z. Liu, and K. Jiang, "Image target segmentation method based on fuzzy Renyi entropy and region growing," *Systems Engineering and Electronics*, vol. 40, no. 8, pp. 1693–1701, 2018.
- [12] J. Ma, K. Du, L. Zhang, F. Zheng, J. Chu, and Z. Sun, "A segmentation method for greenhouse vegetable foliar disease spots images using color information and region growing," *Computers and Electronics in Agriculture*, vol. 142, pp. 110–117, 2017.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Neural Information Processing Systems*, vol. 25, no. 2, pp. 1097–1105, 2012.
- [14] P. Matlani and M. Shrivastava, "Hybrid deep VGG-NET convolutional classifier for video smoke detection," *Computer Modeling in Engineering and Sciences*, vol. 119, no. 3, pp. 427–458, 2019.
- [15] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 3431–3440, 2015.
- [16] O. Ronneberger, P. Fischer, T. Brox et al., "U-net: convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer Assisted Intervention*, Springer, New York, NY, USA, pp. 234–241, 2015.
- [17] V. Iglovikov and A. Shvets, "TernausNet: U-net with VGG11 encoder pre-trained on ImageNet for image segmentation," 2018, <https://arxiv.org/abs/1801.05746>.
- [18] O. Oktay, J. Schlemper, L. L. Folgoc et al., "Attention U-net: learning where to look for the pancreas," 2018, <https://arxiv.org/abs/1804.03999>.
- [19] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Transactions on Medical Imaging*, vol. 39, no. 6, pp. 1856–1867, 2020.
- [20] H. Sun, C. Li, B. Liu et al., "AUNet: attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms," *Physics in Medicine and Biology*, vol. 65, no. 5, Article ID 055005, 2020.
- [21] F. Milletari, N. Navab, and S. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the International Conference on 3D Vision*, pp. 565–571, Stanford, CA, USA, October 2016.
- [22] T. G. W. Boers, Y. Hu, E. Gibson et al., "Interactive 3D U-net for the segmentation of the pancreas in computed tomography scans," *Physics in Medicine and Biology*, vol. 65, no. 6, Article ID 065002, 2020.
- [23] X. Li, H. Chen, X. Qi, Q. Du, C. W. Fu, and P. A. Heng, "H-DenseUNet: hybrid densely connected UNet for ltumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
- [24] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: learning dense volumetric segmentation from sparse annotation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Springer, Athens, Greece, October 2016.
- [25] Y. Wang, L. Zhao, M. Wang, and Z. Song, "Organ at risk segmentation in head and neck CT images using a two-stage segmentation framework based on 3D U-net," *IEEE Access*, vol. 7, pp. 144591–144602, 2019.
- [26] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 2011–2023, 2020.
- [27] B. Li, Z.-W. Tan, P. P. Shum, C. Wang, Y. Zheng, and L. j. Wong, "Dilated convolutional neural networks for fiber Bragg grating signal demodulation," *Optics Express*, vol. 29, no. 5, pp. 7110–7123, 2021.
- [28] Y. Wang, S. Hu, G. Wang, C. Chen, and Z. Pan, "Multi-scale dilated convolution of convolutional neural network for crowd counting," *Multimedia Tools and Applications*, vol. 79, pp. 1057–1073, 2020.
- [29] M. Havaei, A. Davy, D. Warde-Farley et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017.
- [30] A. Barakat and P. Bianchi, "Convergence and dynamical behavior of the ADAM algorithm for nonconvex stochastic optimization," *SIAM Journal on Optimization*, vol. 31, no. 1, pp. 244–274, 2021.
- [31] M. Bhalerao and S. Thakur, "Brain tumor segmentation based on 3D residual U-Net," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11993, pp. 218–225, 2020.
- [32] N. Boutry, J. Chazalon, E. Puybureau, G. Tochon, H. Talbot, and T. Géraud, "Using separated inputs for multimodal brain tumor segmentation with 3d u-net-like architectures," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11992, pp. 187–199, 2020.
- [33] W. Jun, X. Haoxiang, and Z. Wang, "Brain tumor segmentation using dual-path attention U-net in 3D MRI images," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 12658, pp. 183–193, 2021.