

Research Article

Tooth CT Image Segmentation Method Based on the U-Net Network and Attention Module

Sha Tao  and Zhenfeng Wang 

School of Electrical Engineering, Tongling University, Tongling 244000, China

Correspondence should be addressed to Zhenfeng Wang; wzf860826@163.com

Received 20 August 2021; Revised 5 July 2022; Accepted 19 July 2022; Published 19 August 2022

Academic Editor: Huiling Chen

Copyright © 2022 Sha Tao and Zhenfeng Wang. 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.

Traditional image segmentation methods often encounter problems of low segmentation accuracy and being time-consuming when processing complex tooth Computed Tomography (CT) images. This paper proposes an improved segmentation method for tooth CT images. Firstly, the U-Net network is used to construct a tooth image segmentation model. A large number of feature maps in downsampling are supplemented to downsampling to reduce information loss. At the same time, the problem of inaccurate image segmentation and positioning is solved. Then, the attention module is introduced into the U-Net network to increase the weight of important information and improve the accuracy of network segmentation. Among them, subregion average pooling is used instead of global average pooling to obtain spatial features. Finally, the U-Net network combined with the improved attention module is used to realize the segmentation of tooth CT images. And based on the image collection provided by West China Hospital for experimental demonstration, compared with other algorithms, our method has better segmentation performance and efficiency. The contours of the teeth obtained are clearer, which is helpful to assist the doctor in the diagnosis.

1. Introduction

In recent years, deep learning has developed rapidly and has been applied to many fields of image processing [1]. Among them, image segmentation is one of the key topics in modern computer vision research [2]. The effect of image segmentation in the medical field is getting more and more attention. Whether it is pathological analysis, diagnosis and treatment, or clinical surgical navigation, image segmentation can play a better auxiliary treatment effect. It can quickly identify the lesion and reduce the cost of medical care [3]. Traditional image segmentation is done manually layer by layer by experienced doctors, which takes a long time. The segmentation results of these regions of interest (RoI) are quite different among different doctors [4]. The automatic segmentation of CT images can significantly reduce the workload of doctors and improve the accuracy and consistency of RoI segmentation [5]. At present, modern medical CT mainly includes multiple spiral CT and cone beam CT (CBCT). CBCT is more commonly used in dentistry, which has the

characteristics of short scanning time and low radiation dose [6]. Because the bones of the teeth are closely intertwined with some soft tissues, the CT scan image of the tooth will become very blurred. And there are differences in the grayscale value distribution between teeth. There may also be differences in the grayscale value of the same tooth, the topological structure is more complicated, and it is not easy to distinguish [7]. Therefore, accurate and efficient medical image segmentation algorithms are necessary.

Image segmentation currently does not have a unified general-purpose algorithm. Most algorithms are aimed at specific types of objects, looking for their characteristic points to construct specific energy functions to achieve segmentation [8, 9]. In medical CT image segmentation, the traditional template-based segmentation method has gradually become one of the common methods of RoI delineation in radiotherapy [10]. Kaur et al. [11] proposed a CT image segmentation technique based on a mixture of spatial intuitionistic fuzzy C-means clustering and spatial regularization level set methods. It can integrate the details of the spatial

image and accurately and effectively extract the kidney disease image in the CT image. Li et al. [12] proposed a new validity index after evaluating the tightness and separation of the dataset twice. The effectiveness of the proposed metrics is verified using a large number of synthetic datasets and some typical CT images. This metric is suitable for datasets with arbitrarily shaped clusters that are susceptible to noise corruption. It is beneficial to cluster analysis and computer-aided detection systems. Shi et al. [13] proposed an improved potential well function combined with a level set model for tooth segmentation. The segmentation speed and accuracy are improved. Ji et al. [14] proposed a level set method for tooth shape extraction from head CBCT images, which achieved high accuracy.

However, due to the frequent changes in the anatomy of human organs, it is difficult for traditional methods to establish a universal template based on fixed images. Moreover, even the same RoI may have large differences due to differences in patient size and age [15]. In recent years, deep learning methods have been widely used. For example, stacked autoencoder, deep belief networks (DBN), recurrent neural network (RNN), and convolutional neural network (CNN) have been extensively studied in related fields [16, 17]. Sun [18] proposed the Darknet-53 network for extracting tooth image features and multiscale prediction. The SE module is added to the network to improve the network performance. Li et al. [19] proposed a CNN-based stereo image segmentation framework. The use of coherent parallax propagation solves the inevitable inaccuracy problem of the parallax calculated from the stereo image pair. Then, the pixels propagated through the coherent parallax and the high-confidence pixels of the target probability map generated by the CNN architecture will be used to generate the initial reliable pixels to perform the segmentation based on the energy minimization framework. CNN automatically extracts multilevel visual features by learning a large number of mapping relationships between inputs and outputs [20]. Zhou et al. [21] proposed a three-level network based on 3D U-Net to segment a single tooth. But the relationship between alveolar bone and teeth is very important when orthodontists make orthodontic plans for patients. Kadry et al. [22] used the pretrained VGG-UNet system to extract ISL fragments from brain MRI slice test images and extract COVID-19 infection from lung CT images. Tong et al. [23] propose an extended version of U-Net, which uses the concept of triple attention mechanism to segment skin lesions. The segmentation results are more accurate for complex skin lesions such as shape, size, color, and texture. But the algorithm is more complicated and time-consuming. Zhang et al. [24] proposed a triple crossover U-Nets (TIU-Nets) for glioma segmentation. The proposed TIU-Nets are composed of binary class segmentation U-Net (BU-Net) and multiclass segmentation U-Net (MU-Net) and have achieved better performance. Tooth CT images are relatively simple for skin lesion images and brain MRI images. Miki et al. [25] studied the application of CNN to classify tooth types on dental CT images, which is beneficial to obtain high classification accuracy without

the need for accurate tooth segmentation. Chung et al. [26] proposed a neural network for per-pixel labeling to segment frames using instances that are robust to metal artifacts.

Compared with CNN, U-Net adds upsampling and downsampling processes, generates images of different resolutions, and convolves images of various resolutions [27]. To this end, a tooth CT image segmentation method using the U-Net network is proposed. For this paper, the main contributions are as follows:

- (1) Since the segmentation effect of fully convolutional networks (FCN) is relatively rough, the proposed method uses the U-Net network for image segmentation. And through data enhancement, a more ideal tooth CT image segmentation result can be obtained
- (2) In order to improve the segmentation speed of the network and reduce the loss of precision, the attention module is integrated in the U-Net network and improved. The Subregional Average Pooling (SAP) module is used to obtain the initial channel weight information, and the CE module is used to expand the low-level channel attention

This paper is divided into four sections in total. Section 1 introduces the research significance and technical difficulties of topic selection and classifies the existing image semantic segmentation methods. On this basis, summarize the innovative points of the proposed method. Section 2 elaborates on the proposed method, using the U-Net network combined with the improved attention module to segment the tooth CT image. Section 3 discusses the experimental results and uses relevant standards to evaluate the performance of the proposed method. Section 4 is the conclusion.

2. Materials and Methods

2.1. U-Net. U-Net is a U-Net network structure proposed by Olaf Ronneberger in 2015. It is a semantic segmentation network based on FCN, which matches the form of downsampling during the upsampling process to keep it consistent [28, 29]. On this basis, a large number of feature maps in the downsampling stage are added to the downsampling to fill in the information lost during the operation. Because its structure is symmetrical like U-shaped, it is named the U-Net network.

U-Net has special advantages in processing medical images. In solving the problem of the lack of sample images in medical images, elastic deformation is used to complete data enhancement [30]. Elastic deformation is a relatively common type of deformation in actual cells, so it is very suitable for medical image processing [31]. The algorithm of data enhancement is used to make the neural network model learn the invariance of elastic deformation. In the case of a small dataset, the network has good elastic deformation adaptability. The segmentation can be completed correctly when meeting the elastic deformation of the medical image [32]. The U-Net network has excellent effect on cell segmentation, and its segmentation effect is shown in Figure 1.

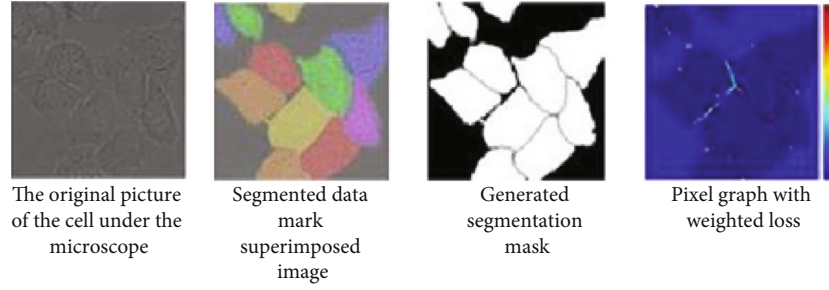


FIGURE 1: U-Net network segmentation cell diagram.

The U-Net network has attracted attention from the medical field since its inception. On the basis of U-Net, many scholars have improved this structure. For example, Milletari's 3D deformation structure V-Net and Drozdal's changes to the skip structure have added many variants to the U-Net network.

2.2. Attention Module. In order to increase the weight of effective features and improve the accuracy of segmentation, we have added an attention mechanism module to the network [33]. The human visual attention mechanism is divided into data-driven and task-driven [34]. This paper uses a data-driven attention mechanism module. The level of attention weight is positively correlated with the importance of the corresponding location information. The input unit with high weight has a decisive effect on the output result. Firstly, global average pooling (GAP) is performed on each channel to obtain the vector of $1 \times 1 \times K$. More than two fully convolutional (FC) layer transformations are done. Sigmoid and ReLU activation functions are used. The extraction network is shown in Figure 2.

In the attention module, first use GAP to shrink the data output from the downsampling layer from $W \times H \times K$ to $1 \times 1 \times K$, which is expressed as follows:

$$z_c = f_{sq}(u_c) \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), \quad (1)$$

where f_{sq} represents the GAP function.

After that, the network performs two FC operations. C represents the dimensionality reduction coefficient, which can be adjusted according to the specific network. The best performance is obtained when $C = 16$ is demonstrated experimentally; the formula is as follows:

$$s_c = f_{ex}(z_c, W) = \varepsilon(g(z_c, W)) = \varepsilon(W_2 \delta(W_1 z_c)),$$

$$W_1, W_2 \in R^{(K/C) \times K}, \quad (2)$$

where ε and δ represent sigmoid and ReLU activation functions, respectively. f_{ex} is the FC function.

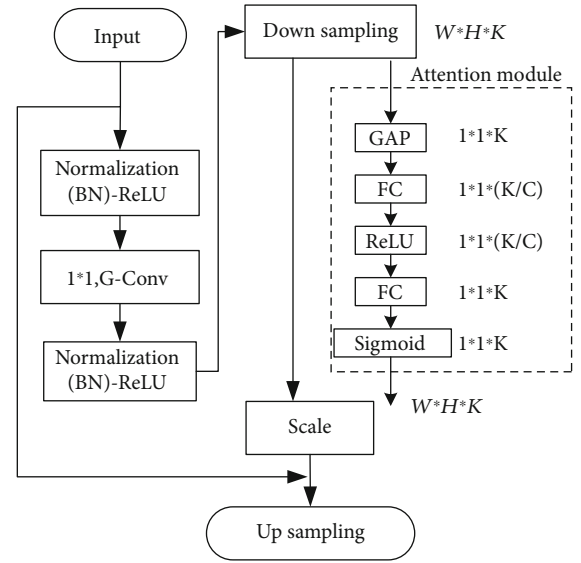


FIGURE 2: Feature extraction network with the attention module.

Then, perform the scale operation. Based on the number of channels unchanged, the data is changed to $W \times H \times K$; the calculation is as follows:

$$X_c = f_{scale}(u_c, s_c) = s_c \cdot u_c, \quad (3)$$

where f_{scale} is the scale operation function.

Before the network joins the attention module, for an input image with a batch of 256 and a size of 224×224 , a forward propagation takes 42 ms. After adding the attention module, it takes 47 ms. After reducing the complexity, there is still time to increase, but it can be ignored compared with the improvement of segmentation accuracy [35].

Because segmentation task scenes are often complex and changeable, using GAP is too simple. The features needed for segmentation cannot be extracted effectively [36]. In response to this problem, the SAP structure was proposed. The structures of GAP and SAP are shown in Figure 3.

As shown in Figure 3, GAP turns a channel feature with a dimension of $H \times W \times K$ into a feature with the size of $1 \times 1 \times K$. In SAP, a feature with a size of $H \times W \times K$ is transformed into a new feature with a size of $m \times n \times k$, so that some spatial information is retained. Then, through a convolution kernel of size $m \times n$, the obtained feature

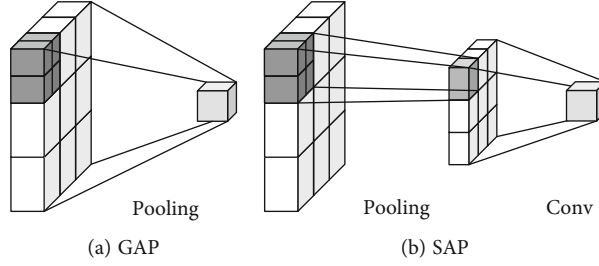


FIGURE 3: Structure comparison between GAP and SAP.

map is transformed into a feature of size $H \times W \times K$. The feature extraction network with the attention module is shown in Figure 4.

First increase the number of channels and use a 1×1 convolution. Because 1×1 convolutional layers are usually used to adjust the number of channels between network layers and control the model complexity, it enables interaction and information integration across channels. At the same time, the nonlinear characteristics can be greatly increased while keeping the size of the feature map unchanged, making the network very deep. Then, add the high-level and low-level channel attention weights. Finally, multiply the new weights with higher-order features [37, 38]. The size range of the updated weight is $(0, 2)$, which can not only reduce but also expand the feature value [39].

The obtained channel feature \hat{F} connects a 1×1 convolution φ_r and sigmoid activation function ε , $r \in [1, K]$. The output ε_{SAP} of the SAP module is

$$\varepsilon_{\text{SAP}} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_r, \dots, \varepsilon_K],$$

$$\varepsilon_r = \varepsilon \left(\sum_{i=1}^K \varphi_r * \hat{F}_i \right). \quad (4)$$

Finally, define the high-order attention distribution obtained by the SAP module as ε_h . The low-level attention distribution after passing the CE module is $\varepsilon_l(\cdot)$. Then, it can be used for the updated attention distribution ε_F :

$$\varepsilon_F = \varepsilon_l(\cdot) + \varepsilon_h(\cdot). \quad (5)$$

The recalibrated characteristic map \hat{U} can be obtained as

$$\hat{U} = U\varepsilon_F. \quad (6)$$

2.3. Network Structure. The attention module is integrated into the U-Net network to obtain a network model for tooth CT image segmentation. The structure is shown in Figure 5.

Using 3 layers of 3×3 convolutional layers, change the ordinary 3×3 convolution to dilated convolution with expansion rate 2, and the resolution of the output feature map is 1/16 of the input. Atrous spatial pyramid pooling (ASPP) is used to provide multiscale information in the U-Net network. That is, on the basis of spatial pyramid pooling, a hole convolution with different expansion rates is added, and global information is added through SAP

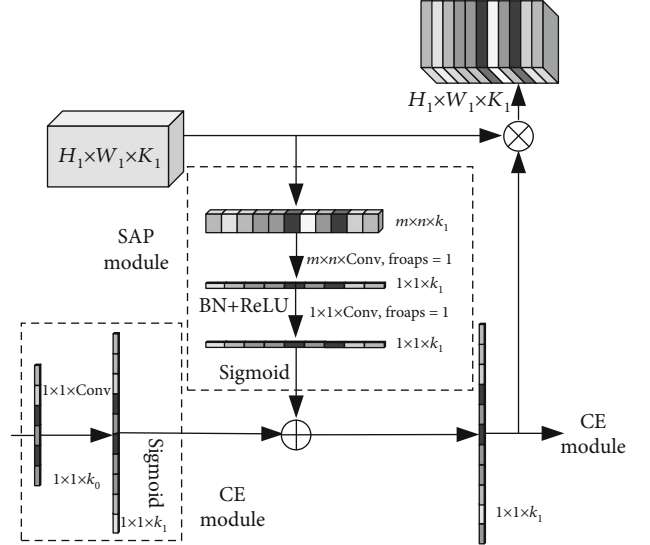


FIGURE 4: Feature extraction network with the attention module.

and image features. The features output by ASPP are upsampled, added to a low-level feature that has undergone a 3×3 convolution, and then merged with a 3×3 convolution. Finally, it is upsampled and restored to the original image size [40].

3. Results and Discussion

The experiment is based on the Keras framework, which is developed by Google. It uses TensorFlow and Theano packages to integrate many basic neural network structures and some mature algorithms. The experimental dataset used is dental CT data provided by West China Hospital. The operating system is Windows 10, the CPU is an 8 GB Intel Core i7-6700, and the GPU is NVIDIA Ge Force GTX 1070.

When training, the optimization method uses random gradient descent (SGD). Its momentum parameter is set to 0.9, and the weight decay rate is 1×10^{-4} . The learning rate uses the initial learning rate learning rate = 8×10^{-3} multiplied by $(1 - \text{current_iter}/\text{max_iter})^{\text{power}}$ to reduce the strategy. Where power = 0.9, current_iter is the current number of iterations, and max_iter is the maximum number of iterations in the training process. The number of experimental training iterations is 800.

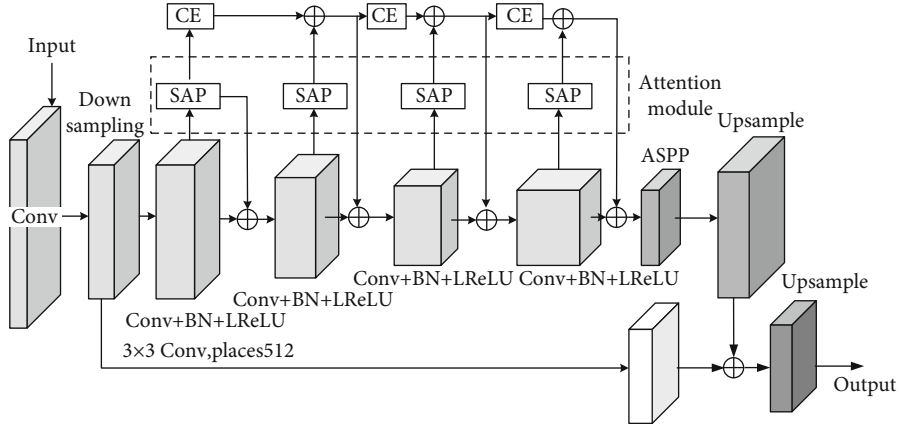


FIGURE 5: U-Net network structure combined with the improved attention module.

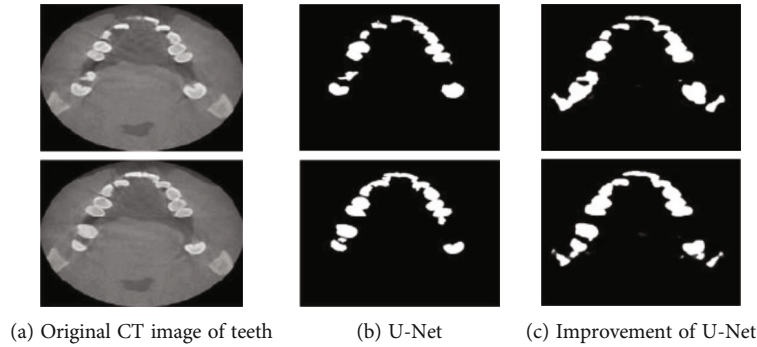


FIGURE 6: Comparison of segmentation results before and after the improvement of the U-Net network.

3.1. Data Preprocessing. Based on the segmentation task of the U-Net network, the final output is closely related to the quality of the input image. In view of the fact that some of the tooth CT images have similar grayscale value, shapes, or textures, image enhancement methods are used to preprocess the input image data to improve image contrast. Histogram equalization (HE) is a basic method of image enhancement. And adaptive histogram equalization (AHE) is improved through HE. AHE improves local contrast, enhances the edge sharpness of each area of the image, and improves the shape, texture, and boundary information of the teeth. Before entering the network, the normalized image grayscale value is (0,1), and the image size is 512×512 .

In order to overcome the overfitting phenomenon of neural networks, the proposed method adopts enhancement techniques such as random cropping, flipping, grayscale perturbation, and shape perturbation on the data image to overcome this problem. The gray level disturbance in the training set can improve the stability of the network, thereby improving the performance of the prediction set network.

The CT image and contour image are deformed by affine transformation to form shape disturbance. In the deformation, the coordinates of the 3 vertices (upper left, upper right, and lower left) are first obtained. Then, each point moves randomly, and the range of random movement is the image length. Finally, affine transformation is performed on the entire image.

TABLE 1: Comparison of segmentation performance before and after the improvement of the U-Net network.

Methods	MPA (%)	MIoU (%)	Execution time (ms)
U-Net	82.37	80.19	98
Improved U-Net	85.91	83.73	126

After image enhancement, the training set (400 images) and the validation set (100 images) are both enlarged 3 times. A total of 1500 image data were obtained for experimental training and parameter adjustment.

3.2. Evaluation Index. In the experiment, mean pixel accuracy (MPA) and mean intersection over union (MIoU) are used as indicators to evaluate the performance of the proposed method.

Assume that n_{ii} represents the number of correct segmentation. n_{ij} represents the number of pixels that originally belonged to class i but were divided into class j . n_{ji} represents the number of pixels that originally belonged to class j but were divided into class i . There are $k + 1$ categories (including k classes and an empty class or background class).

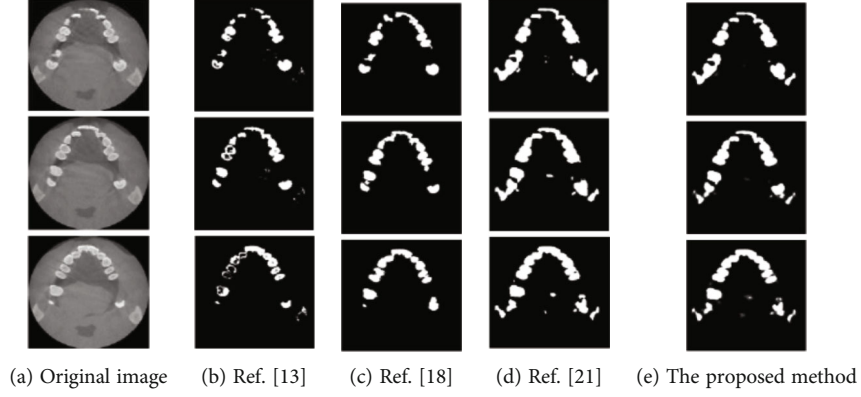


FIGURE 7: Segmentation results of different methods.

MPA calculates the proportion of pixels that are correctly segmented in each class and then finds the average of all classes. The calculation is as follows:

$$\text{MPA} = \frac{1}{k+1} \sum_{i=0}^k \frac{n_{ii}}{\sum_{j=0}^k n_{ij}}. \quad (7)$$

MIOU calculates the ratio of intersection and union of two sets. Calculate the pixel intersection ratio within each pixel category, and then, calculate the average:

$$\text{MIOU} = \frac{1}{k+1} \sum_{i=0}^k \frac{n_{ii}}{\sum_{j=0}^k n_{ij} + \sum_{j=0}^k n_{ji} - n_{ii}}. \quad (8)$$

Due to its strong representativeness, high efficiency, and simplicity, MIOU has become the current general image segmentation evaluation index. Therefore, MIOU is used as the main evaluation index of the experiment.

3.3. Comparison of the Proposed Algorithms. According to the proposed method, the final tooth CT image segmentation result is obtained. The segmentation comparison result of the proposed method and the traditional U-Net network is shown in Figure 6.

It can be seen from Figure 6 that the proposed method can segment the tooth CT image more effectively than the traditional U-Net network. The effect of tooth edge extraction is significantly better than that of U-Net. However, affected by some noise, there is a certain shadow in the center of the segmented image. Therefore, the performance of the U-Net network before and after the improvement is quantitatively analyzed. The results of MPA and MIOU and the time to complete the segmentation are shown in Table 1.

As can be seen from Table 1, the proposed method improves the U-Net network by using the attention mechanism. MPA is 85.91%, and MIOU is 83.73%, which is superior to the U-Net network in terms of segmentation accuracy. However, due to the complexity of the segmentation model, the time to complete the segmentation has also increased. Compared with the U-Net network, the time has increased by 28 ms. However, from an overall point of view,

TABLE 2: Comparison of segmentation performance of different methods.

Methods	MPA (%)	MIOU (%)	Execution time (ms)
Ref. [13]	69.11	68.05	74
Ref. [18]	80.03	79.24	167
Ref. [21]	84.76	83.19	183
The proposed method	85.91	83.73	126

the proposed method still has a large application advantage in tooth CT image segmentation.

In order to further demonstrate the performance of the proposed method, compare it with References [13, 18] and [21]. The segmentation results of tooth CT images obtained by each method are shown in Figure 7.

It can be seen from Figure 7 that the tooth profile obtained by the proposed method is clearer and is closest to the original image. Figure 7(b) adopts the traditional segmentation method combined with the level set model [13], the segmentation effect is not good, and the contour is blurred. Figures 7(c) and 7(d) use learning algorithms for segmentation [18, 21], but a single algorithm needs to be improved in processing complex tooth images. Therefore, the effectiveness of the proposed method for segmentation of tooth CT images can be demonstrated.

The quantitative analysis results of different methods are shown in Table 2.

It can be seen from Table 2 that the segmentation performance of the proposed method is the best, with MPA of 85.91% and MIOU of 83.73%. It completes the detection speed of 126 ms, but the model is more complicated than that of Reference [13], so the time is slightly longer. The traditional method used in Reference [13] has low segmentation accuracy for complex teeth CT images, with MIOU of only 68.05%. But it takes the shortest time to complete the detection, which is 74 ms. Reference [18] chooses the Darknet-53 network for extracting tooth image features and multiscale prediction. Overall, it has good segmentation performance, but the algorithm is not effective in learning complex features, and the segmentation accuracy needs to

be improved. Reference [21] segmented teeth based on a three-level network of 3D U-Net. But the model is more complicated, so it takes a long time, 183 ms. From a comprehensive point of view, the proposed method has obvious application advantages in tooth CT image segmentation.

4. Conclusions

With the development of today's society, medical standards have put forward higher requirements in people's lives. There are many representative computational intelligence algorithms, such as monarch butterfly optimization (MBO) [41], earthworm optimization algorithm (EWA) [42], elephant herding optimization (EHO) [43], and moth search (MS) algorithm [44, 45]. With the continuous increase in the number of CT images and MRI images, the segmentation of medical images by manpower alone can no longer meet the demand. Therefore, the use of deep learning for automatic segmentation has attracted wide attention from scholars. We propose a tooth CT image segmentation method using the U-Net network. The attention module is integrated into the U-Net network, and the SAP and CE modules are used to improve the performance of the network. Based on the tooth CT image dataset provided by West China Hospital, the improved network was demonstrated experimentally. With comprehensive comparison, the method in this paper has better segmentation performance and segmentation efficiency. The contours of the teeth obtained are clearer, which is helpful to assist the doctor in the diagnosis. The proposed method only performs segmentation for tooth images and has not been used for segmentation of other medical images. In the following research, the proposed method will be applied to other organizations to improve the generalization ability of the proposed method. And in-depth pruning of the network is carried out to improve network efficiency.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This work was supported by the Anhui Provincial Natural Science Key Foundation (No. KJ2018A0484).

References

- [1] Y. Chen, L. Liu, J. Tao et al., "The improved image inpainting algorithm via encoder and similarity constraint," *The Visual Computer*, vol. 37, no. 7, pp. 1691–1705, 2021.
- [2] Y. Chen, H. Zhang, L. Liu et al., "Research on image inpainting algorithm of improved total variation minimization method," *Journal of Ambient Intelligence and Humanized Computing*, vol. 6, no. 8, pp. 271–280, 2021.
- [3] F. Guo, C. Shi, X. Li, X. Wu, J. Zhou, and J. Lv, "Image segmentation of nasopharyngeal carcinoma using 3D CNN with long-range skip connection and multi-scale feature pyramid," *Soft Computing*, vol. 24, no. 16, pp. 12671–12680, 2020.
- [4] E. Cuevas, H. Becerra, A. Luque, and M. A. Elaziz, "Fast multi-feature image segmentation," *Applied Mathematical Modelling*, vol. 90, no. 9, pp. 742–757, 2021.
- [5] X. Jiang, H. Yu, and S. Lv, "An image segmentation algorithm based on a local region conditional random field model," *International Journal of Communications, Network and System Sciences*, vol. 13, no. 9, pp. 139–159, 2020.
- [6] H. Tonbul and T. Kavzoglu, "A spectral band based comparison of unsupervised segmentation evaluation methods for image segmentation parameter optimization," *International Journal of Environment and Geoinformatics*, vol. 7, no. 2, pp. 132–139, 2020.
- [7] T. Song, F. Meng, A. Rodriguez-Paton, P. Li, P. Zheng, and X. Wang, "U-next: a novel convolution neural network with an aggregation U-net architecture for gallstone segmentation in CT images," *Access*, vol. 7, no. 6, pp. 166823–166832, 2019.
- [8] Y. Wang, H. Wang, K. Shen, J. Chang, and J. Cui, "Brain CT image segmentation based on 3D slicer," *Journal of Complexity in Health Sciences*, vol. 3, no. 1, pp. 34–42, 2020.
- [9] H. Su, D. Zhao, H. Elmannai et al., "Multilevel threshold image segmentation for COVID-19 chest radiography: a framework using horizontal and vertical multiverse optimization," *Computers in Biology and Medicine*, vol. 146, article 105618, 2022.
- [10] L. Zhou and Z. Zhang, "An image segmentation method based on Mumford-Shah model with mask factor and neighborhood factor," *Pattern Analysis and Applications*, vol. 23, no. 1, pp. 85–94, 2020.
- [11] R. Kaur, M. Juneja, and A. K. Mandal, "A hybrid edge-based technique for segmentation of renal lesions in CT images," *Multimedia Tools and Applications*, vol. 78, no. 10, pp. 12917–12937, 2019.
- [12] Q. Li, S. Yue, M. Ding, J. Li, and Z. Wang, "Determining the correct number of clusters in the CT image segmentation," *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 11, pp. 2675–2680, 2020.
- [13] Q. Shi, F. Yan, Y. Yang, Y. Chen, X. Lin, and Y. Wang, "Image segmentation of tooth and alveolar bone with the level set model," *Chinese Journal of Magnetic Resonance*, vol. 38, no. 2, pp. 182–193, 2021.
- [14] D. X. Ji, S. H. Ong, and K. Foong, "A level-set based approach for anterior teeth segmentation in cone beam computed tomography images," *Computers in Biology & Medicine*, vol. 50, pp. 116–128, 2014.
- [15] S. Zhou and Z. Xu, "Automatic grayscale image segmentation based on affinity propagation clustering," *Pattern Analysis and Applications*, vol. 23, no. 1, pp. 331–348, 2020.
- [16] S. Wang, Y. Cong, H. Zhu et al., "Multi-scale context-guided deep network for automated lesion segmentation with endoscopy images of gastrointestinal tract," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 514–525, 2021.
- [17] B. Lu, "Algorithm improvement of neural network in endoscopic image recognition of upper digestive tract system," *Expert Systems*, vol. 2, no. 6, article e12912, pp. 311–320, 2021.
- [18] S. Z. H. Tooth, *Segmentation Method Based on Deep Learning*, Xi'an University of Science and Technology, 2020.

- [19] X. Li, H. Huang, H. Zhao, Y. Wang, and M. Hu, "Learning a convolutional neural network for propagation-based stereo image segmentation," *The Visual Computer*, vol. 36, no. 1, pp. 39–52, 2020.
- [20] A. P. Bajcsy, S. Feldman, M. Brady, and K. Snyder, "Semantic SEM image segmentation of concrete with contextual labels," *Microscopy and Microanalysis*, vol. 25, no. S2, pp. 144–145, 2019.
- [21] X. R. Zhou, W. Y. Jin, and L. Chai, "Dental CBCT image segmentation," *Modern Computer*, vol. 4, pp. 122–126, 2021.
- [22] S. Kadry, R. Damaševičius, D. Taniar, V. Rajinikanth, and I. A. Lawal, "U-Net supported segmentation of ischemic-stroke-lesion from brain MRI slices," in *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, pp. 1–5, Chennai, India, 2021.
- [23] X. Tong, J. Wei, B. Sun, S. Su, Z. Zuo, and P. Wu, "ASCU-Net: attention gate, spatial and channel attention U-Net for skin lesion segmentation," *Diagnostics*, vol. 11, no. 3, p. 501, 2021.
- [24] J. Zhang, J. Zeng, P. Qin, and L. Zhao, "Brain tumor segmentation of multi-modality MR images via triple intersecting U-Nets," *Neurocomputing*, vol. 421, pp. 195–209, 2021.
- [25] Y. Miki, C. Muramatsu, T. Hayashi et al., "Classification of teeth in cone-beam CT using deep convolutional neural network," *Computers in Biology & Medicine*, vol. 80, pp. 24–29, 2017.
- [26] M. Chung, M. Lee, J. Hong et al., "Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation," *Computers in Biology and Medicine*, vol. 120, no. 11, article 103720, 2020.
- [27] C. Ye, W. Wang, S. Zhang, and K. Wang, "Multi-depth fusion network for whole-heart CT image segmentation," *IEEE Access*, vol. 7, no. 5, pp. 23421–23429, 2019.
- [28] E. A. Audenaert, J. Van Houcke, D. F. Almeida et al., "Cascaded statistical shape model based segmentation of the full lower limb in CT," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 22, no. 6, pp. 644–657, 2019.
- [29] A. Kozei, N. Nikolov, O. Haluzynskyi, and S. Burburska, "Method of threshold CT image segmentation of skeletal bones," *Innovative Biosystems and Bioengineering*, vol. 3, no. 1, pp. 4–11, 2019.
- [30] R. Kulkarni and K. Bhavani, "Analysis of CT DICOM image segmentation for abnormality detection," *International Journal of Engineering and Manufacturing*, vol. 9, no. 5, pp. 46–55, 2019.
- [31] Z. Liu, Y. Q. Song, V. S. Sheng et al., "Liver CT sequence segmentation based with improved U-Net and graph cut," *Expert Systems with Application*, vol. 126, no. 7, pp. 54–63, 2019.
- [32] M. A. Koenrades, E. M. Struijs, A. Klein et al., "Quantitative stent graft motion in ECG gated CT by image registration and segmentation: in vitro validation and preliminary clinical results," *European Journal of Vascular and Endovascular Surgery*, vol. 58, no. 5, pp. 746–755, 2019.
- [33] P. K. Barik and R. C. Barik, "Feed forwarded CT image registration for tumour and cyst detection using rigid transformation with HSV colour segmentation," *International Journal of Computational Systems Engineering*, vol. 5, no. 5/6, pp. 277–286, 2019.
- [34] C. Fan, Z. Wang, G. Li et al., "A novel 3D U-Net deep network with paralleling structure for stroke lesion image segmentation," *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 3, pp. 724–730, 2020.
- [35] B. Park, H. Park, S. M. Lee, J. B. Seo, and N. Kim, "Lung segmentation on HRCT and volumetric CT for diffuse interstitial lung disease using deep convolutional neural networks," *Journal of Digital Imaging*, vol. 32, no. 6, pp. 1019–1026, 2019.
- [36] R. Hemke, C. G. Buckless, A. Tsao, B. Wang, and M. Torriani, "Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment," *Skeletal Radiology*, vol. 49, no. 3, pp. 387–395, 2020.
- [37] 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.
- [38] X. Dong, Y. Lei, S. Tian et al., "Synthetic MRI-aided multi-organ segmentation on male pelvic CT using cycle consistent deep attention network," *Radiotherapy and Oncology*, vol. 141, no. 1, pp. 192–199, 2019.
- [39] M. Sakamoto, Y. Hiasa, Y. Otake et al., "Bayesian segmentation of hip and thigh muscles in metal artifact-contaminated CT using convolutional neural network-enhanced normalized metal artifact reduction," *Journal of Signal Processing Systems*, vol. 92, no. 3, pp. 335–344, 2020.
- [40] J. Fang, Q. B. Zhou, and S. Wang, "Segmentation technology of nucleus image based on U-Net network," *Scientific Programming*, vol. 2021, no. 7, 10 pages, 2021.
- [41] Y. Feng, S. Deb, G. G. Wang, and A. H. Alavi, "Monarch butterfly optimization: a comprehensive review," *Expert Systems with Applications*, vol. 168, article 114418, 2021.
- [42] A. Skrk, B. Ks, and A. NI, "Development of deer hunting linked earthworm optimization algorithm for solving large scale traveling salesman problem," *Knowledge-Based Systems*, vol. 227, article 107199, 2021.
- [43] W. Li and G. G. Wang, "Elephant herding optimization using dynamic topology and biogeography-based optimization based on learning for numerical optimization," *Engineering with Computers*, vol. 1, pp. 1–29, 2021.
- [44] G. G. Wang, "Moth search algorithm: a bio-inspired meta-heuristic algorithm for global optimization problems," *Mematic Computing*, vol. 10, no. 2, pp. 151–164, 2018.
- [45] W. Li, G. Wang, and A. Gandomi, "A survey of learning-based intelligent optimization algorithms," *Archives of Computational Methods in Engineering*, vol. 28, no. 5, pp. 3781–3799, 2021.