BABT
BRAZILIAN ARCHIVES OF
BIOLOGY AND TECHNOLOGY

Article - Engineering, Technology and Techniques

Segmentation for Athlete's Ankle Injury Image Using Residual Double Attention U-Net Model

Jing Zhang[1]
https://orcid.org/0009-0007-6015-8220

Jian Zhou[1]*
https://orcid.org/0000-0002-5489-7744

Ming Huang[2]
https://orcid.org/ 0009-0000-5501-0407

Raja Soosaimarian Peter Raj[3]
https://orcid.org/0000-0002-7216-2207

[1]China Three Gorges University, College of Physical Education, Hubei, China; [2]Harbin No. 5 Middle School, Harbin, China; [3]Vellore Institute of Technology, School of Computer Science and Engineering, Vellore, Tamilnadu, India.

*Correspondence: zhoujian@ctgu.edu.cn (J.Z.).

HIGHLIGHTS

- We propose a segmentation using the Residual Double Attention U-Net model.

- Adjusting the gradient propagation of the segmentation framework using the residual structure.

- Solved the problem of low Correspondence Ratio and F1 values in traditional algorithms.

- Using multiple data sets to test the application effect of proposed algorithm.

Abstract: The image of an athlete's ankle joint injury can help to check whether the athlete's ankle joint is damaged, and plays a very important role in clinical diagnosis. To address the problem of poor segmentation effect of traditional athletes' ankle injury image segmentation algorithm, an ankle injury image segmentation algorithm based on residual double attention U-Net model is proposed. First, the region of interest is extracted from the original ankle injury image. After translation, rotation and turnover, the image data is expanded. Second, the residual structure is used to adjust the gradient propagation and residual feedback of the segmentation framework, extract the attribute information in the region of interest, and combine the two to retain more image features. Finally, combined with the double attention module to improve the weight ratio of image features, the athlete ankle injury image segmentation is realized in the image segmentation framework based on residual double attention U-Net model. The results demonstrate that the maximum values of DSC, ASSD, PM, and CR for the proposed algorithm are 0.93, 0.1, 0.96, and 0.95, respectively, and the F1 score is 95.7%, indicating that the segmentation effect of this algorithm is closer to the theoretical segmentation effect, and higher precision in segmentation, and the segmented image has a high degree of similarity to the original image, resulting in excellent segmentation performance.

INTRODUCTION

Professional athletes should receive corresponding training during their adolescence and youth. However, long-term overload training causes the ankle joint to be affected by different forces in different directions and of different sizes, accompanied by certain wear and impact, resulting in serious ankle injury. It has been reported that ankle joint is one of the most vulnerable parts of professional athletes. The bones of athletes in adolescence and youth have not yet fully developed. Thus, long-term training will have a certain impact on the development and structure of ankle joints. Medical images can help to check whether the ankle joint structure is damaged, and thus they play a very important role in clinical diagnosis. In the field of image segmentation, there are several technical challenges that need to be addressed. Firstly, there is the issue of accurately extracting relevant features for segmentation, especially when dealing with complex images. Secondly, there is the problem of handling image noise and variability. Thirdly, there is a need to improve the speed and efficiency of segmentation algorithms. Fourthly, there is a need to develop more robust and accurate evaluation metrics to assess the segmentation performance of various algorithms. In recent years, with the rapid development of medical image technology, more and more experts and scholars have turned their research focus to the field of medical image segmentation whose main purpose of medical image segmentation is to select the region of interest in the medical image with the help of automatic or semi-automatic segmentation algorithm and segment the image completely [1]. The segmented images can help doctors quickly diagnose the loss of ankle joints of athletes and formulate corresponding treatment plans, which is of great significance for guiding athletes, coaches and team doctors to carry out follow-up rehabilitation treatment and recovery training [2].

Aiming at the important research topic of image segmentation of athletes' ankle joint injury, Huang W and coauthors [3] realized image segmentation by training neural networks. The algorithm determined the prior knowledge of the topological structure of the segmented object, which was then introduced into the training network, and the differentiable property was analyzed through the topological data; the required number of topologies was determined according to the Betti number of divided objects; finally, the segmentation included the features of topological structure to realize image segmentation. After testing, it was found that the value of the algorithm is relatively low, indicating that there is a problem of mis-segmentation in the image segmentation process of ankle joint injury of athletes using the algorithm, and the actual application effect is not good. Chen Z and coauthors [4] used task-driven generation of confrontation networks to realize image segmentation of retinal blood vessels. In the generation model, the U network was used to segment retinal blood vessels. In the discrimination model, multi-scale discriminators with different receptive fields were used to help generate more segmentation details; The task-driven model based on perceptual loss completed feature matching and finally realized image segmentation. However, it was found in practical application that the overlap between the theoretical segmentation effect and the actual segmentation effect of the algorithm is relatively low, indicating that the segmentation effect of the algorithm is poor and difficult to be applied in practice. The difference between the theoretical segmentation result and the actual segmentation result of the algorithm is relatively large. Wang B and coauthors [5] realized automatic segmentation of complex lung tumour images with an effective deep network. The encoder-decoder model was used to connect the global attention units in the image, and the region of interest was extracted by multi-scale semantic information. Finally, the segmentation ability of the algorithm was improved by Tversky loss and boundary loss. However, after testing, it was found that the performance of the algorithm is relatively low, indicating that the degree of under-segmentation in ankle joint injury images by the algorithm is relatively high, and the segmentation effect is poor, resulting in low applicability. Karani N and coauthors [6] proposed an image segmentation method based on an adaptive neural network. The segmented convolutional neural networks (CNN) was designed as a series of two subnetworks: a relatively shallow image normalized CNN, and then a deep CNN, which was used to segment the normalized image. In this process, an independently trained de-noising automatic encoder was used to de-noising the data, and the adaptive neural network was used to realize image segmentation. However, the difference between the theoretical segmentation result of the algorithm and the actual segmentation result of the algorithm is large. Hua L and coauthors [7] proposed a medical image segmentation algorithm based on local edge regions. The local ACM of gradient information was constructed based on the probability score of fuzzy k-nearest neighbour classifier, and the gradient information was then detected. The local feature function was introduced, and the edge information based on the probability score was used to construct the energy of the local region, so that the evolution curve stopped at the precise boundary of the region of interest, and the image segmentation was

realized by combining the boundary localization results. However, this method has the problem of missing segmentation, and the actual application effect is poor.

It was found that traditional segmentation algorithms for athlete ankle injury images have lower values of Dice Similarity Coefficient (DSC), Average Symmetric Surface Distance (ASSD), Prevent Match (PM), Correspondence Ratio (CR), and F1, resulting in lower segmentation accuracy and poorer segmentation performance. To address the issues in traditional segmentation algorithms, a new algorithm is proposed which utilizes a residual double attention U-Net model for athlete ankle injury image segmentation. The main contributions of this paper are as follows: (1) The traditional algorithm does not translate, rotate and flip the original image, resulting in the inability to accurately segment small cavities and capillaries. To address this problem, this paper expands the image data after the image translation, rotation and inversion, and realizes the image preprocessing, to lay a solid foundation for the subsequent accurate image segmentation. (2) After pre-processing the images, the residual double attention U-Net model was utilized to highlight the features of the regions of interest through the use of residual recurrence units and dual attention modules. This coupled with the feature-based segmentation method was used to segment athlete ankle injury images, resulting in a solution for the issue of lower CR and F1 values present in traditional algorithms, thereby improving the overall segmentation quality. (3) The application effectiveness of the proposed algorithm was tested using multiple datasets, with DSC, ASSD, PM, CR, and F1 values used as evaluation criteria. Through experimentation, it was demonstrated that the proposed algorithm has an outstanding effect on reducing the segmentation of athlete ankle injury images, while effectively preserving important information and ideal edge processing. The segmented images have no jagged edges and the overall segmentation performance is good.

METHODOLOGY

In the original U-Net model, a residual structure that can extract attribute information in the region of interest was established, which effectively addresses issues such as gradient disappearance and explosion, and improves the convergence speed and accuracy of the model.
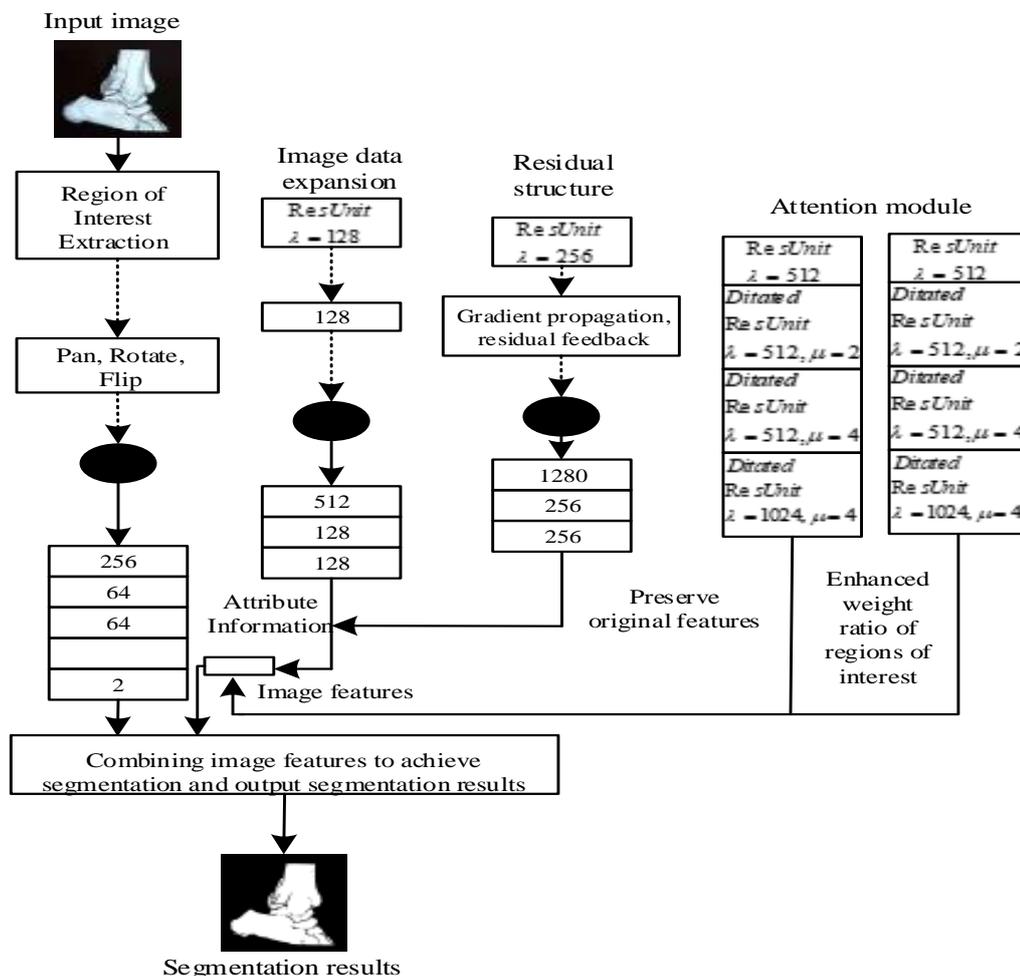
Figure 1. The framework of the proposed algorithm

To address the one-way transmission of attention information in the traditional U-Net model, the residual double attention U-Net model designed a bidirectional attention module, which can better capture local and global features, increase the weight of image features, and cope with complex backgrounds and uneven distribution of objects, thereby improving the segmentation accuracy of athlete ankle injury images. The athlete ankle injury image segmentation algorithm framework using the residual double attention U-Net model is shown in Figure 1.

According to Figure 1, the region of interest is extracted from the input athlete ankle injury image, and the image data is expanded by performing translations, rotations, and flips. The residual structure is utilized to adjust the gradient propagation and residual feedback of the segmentation framework, extract attribute information from the region of interest, and combine the two to preserve more image features. The double attention module is applied to enhance the weight proportion of the region of interest image and retain more original features, ultimately achieving athlete ankle injury image segmentation based on the combined image features.

Data preprocessing

After the Computed Tomography (CT) image samples of ankle joint loss were acquired, the region of interest needs to be extracted first [8], and then preprocessed. Because a CT image may contain some regions unrelated to diagnosis, this part of the region was removed and only the region of interest was left, as shown in Figure 2.
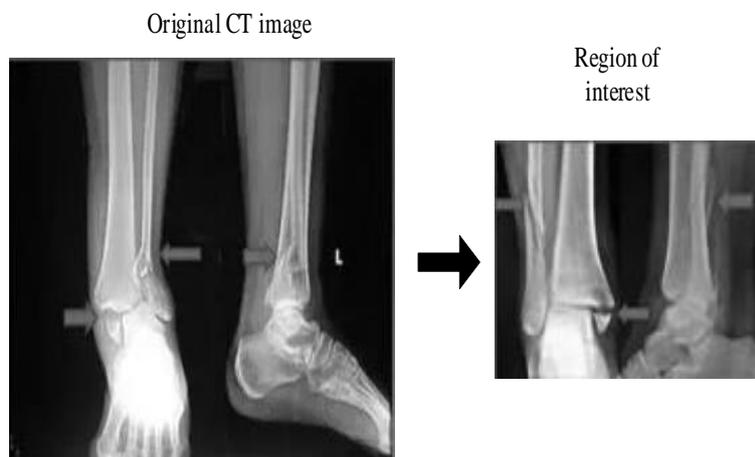
Figure 2. Region of interest extraction in CT images

This algorithm calculated the mean and standard deviation of the sample image. After subtracting from the mean value and then dividing from the standard deviation, the grey-scale regularization of the region of interest was realized, which is convenient for more accurate segmentation [9].

Data expansion

Based on data preprocessing, the data was expanded by translation, rotation and flipping. The number of ankle injury images allowed to be used is limited, and it is difficult to form a complete U-Net model. Therefore, this paper used a series of operations such as flip, translation, rotation and image deformation to expand the existing data. In which, flipping, translation and rotation are only simple deformations of the image, and there is no significant difference [10-11]. Image deformation can generate image data of various shapes for U-Net model training.

The process of data expansion in this paper is as follows: first, the original image was deformed twice, and then it was translated, rotated and flipped to complete the expansion of the original CT image.

Image segmentation using residual double attention U-Net model

Based on data preprocessing and data expansion, the residual double attention U-Net model was used to segment the ankle injury image to ensure the segmentation quality and speed. It is composed of a loop residual unit and a double attention module. The in-depth diagram of image segmentation is shown in Figure 3.
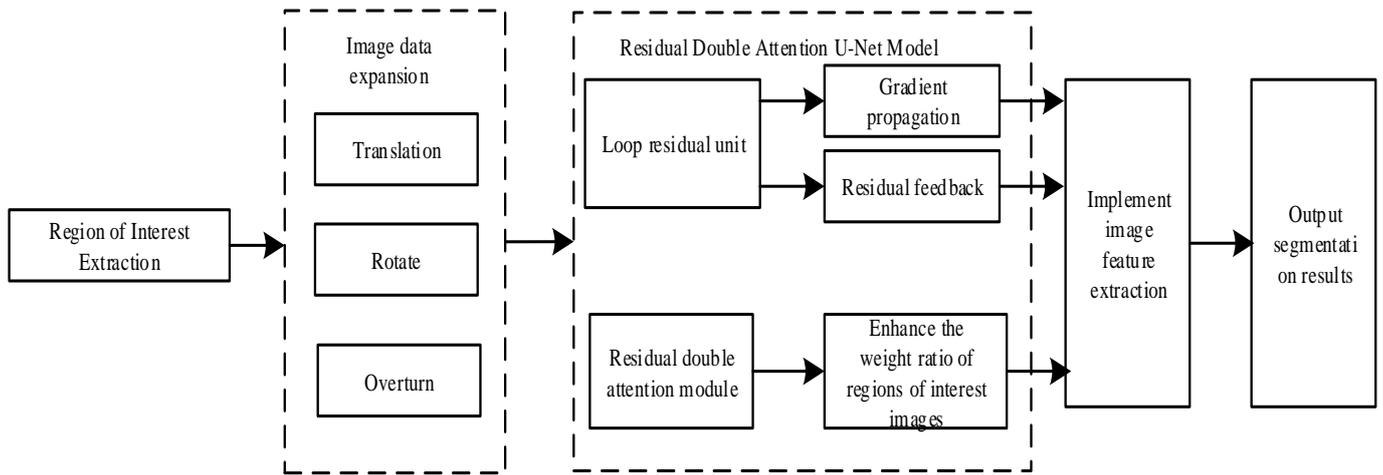
Figure 3. In-depth diagram of ankle joint injury image segmentation

Figure 3 is an end-to-end deep network model. The original CT image is input into the model, and the output result is a binary segmentation map. The white area of the image is the segmentation target. The whole framework completes feature extraction through the training of convolution layer[12-13]. Each training process can generate high-level context information, and to obtain more accurate edge information, an integration process is added to the segmentation framework [14], the residual double attention U-Net model selects a 1 × 1 convolutional enhancement of the low-level feature depth before cascading. At the end of the frame, the training result was converted into a binary classification problem using the Softmax layer, and the $soft-\max$ value of each pixel in the image was calculated through the energy function. The definition formula of $soft-\max$ is

$$m_k(x) = \frac{\exp(\alpha_k(x))}{\sum_{i=1}^{T} \exp(\alpha_k(x))} \quad （1）$$

where $\alpha_k(x)$ refers to the activation value[15] of the pixel characteristic channel at the point x, k is image category, $m_k(x)$ refers to the approximate maximum function, T is the linear conversion parameters, and i refers to energy function coefficient.

In this paper, the binary cross entropy of pixels was used as the target to train the U-Net model [16-17]. The Gaussian distribution $Q(0, 0.4)$ was used to initialize the convolution kernel. The gradient descent method was used to control the loss function to the lowest value. The formula is

$$\delta_j = \delta_j - \beta \frac{\partial}{\delta_j} J(\delta) \quad （2）$$

where β refers to the learning rate, j refers to the loss function coefficient, and J refers to the constraint coefficient of gradient descent method. ∂ is represents the cost function.

Loop residual unit

Ankle injury image segmentation is very difficult to work. Only when the depth of the training model reaches a certain level can accurate feature extraction be achieved. Therefore, the residual structure was used to achieve ideal gradient propagation in the segmentation framework [18]. The definition formula of residual structure is

$$u = F(w) + w \quad （3）$$

where w refers to the input content of the framework, u is the input result, $F(w) + w$ means that after experiencing w the two convolution layers, the obtained $F(w)$ is integrated with w through jumping.

The residual structure combines the pre-convolution content and the post-convolution content using jump connection [19], so that the error information in the image is directly transmitted to the bottom layer of the frame, effectively avoiding the disappearance of the gradient in the calculation process. In addition, the residual structure is similar to the recall mechanism of human brain. When people come into contact with new content, they will probably forget the content they came into contact with before. At this time, the recall mechanism is needed to help people remember these fuzzy memories. The residual structure strengthens the original feature information in the output result using jump connection, thus effectively avoiding the problem of network degradation. The residual structure in this paper consists of two convolution layers and one jump connection.

The residual feedback [20] can automatically extract the attribute information in the image. Compared with the recall mechanism of the residual structure, the residual feedback is similar to the consolidation mechanism of the human brain, which deepens the impression of the known things through consolidation review. The extracted features of the region of interest are taken as the input content, and the feature extraction is performed again. The feature information is enhanced to enhance the impression of the known things.

In this paper, the residual structure and residual feedback are combined to obtain the loop residual unit. The definition formula is

$$u_s = G\left(u_f\right) + w \qquad (4)$$

where u_f refers to the output vector of the first residual propagation, and u_s refers to the enhancement vector of the second residual feedback. $G(\cdot)$ represents the residual feedback function.

In the process of image segmentation of ankle joint injury, the looped residual element completely preserves the injury characteristics of ankle joint using jump connection; with the loopback connection, the overall feature extraction ability of the algorithm is improved, and the basic preparation for accurate segmentation is made.

The double attention module

The double attention module is composed of two parts [21]: trunk branch and soft mask branch. In which, the role of the trunk branch is to preserve the original features of the CT image, and the role of the soft mask branch is to preserve the features of the region of interest and enhance the weight ratio of the region of interest in the trunk branch. The definition formula of double attention module is

$$q_{att}^l = \psi^T \left(\sigma_1 \left(W_w^T w_i^l + W_g^T g_i + b_g \right) \right) + b_\psi \qquad (5)$$

where w_i^l is the content input to the attention module. g_i represents the gate signal provided by higher-level contextual information. σ_2 is the activation function. W_w^T, W_g^T represents different discrete coefficients. ψ^T represents the linear conversion parameter. b_g and b_ψ represent different bias parameters.

The formula for calculating batch normalization parameters is as follows:

$$\delta_i^l = \sigma_2 \left(q_{att}^l \left(w_i^l \cdot g_i; \Theta_{att} \right) \right) \qquad (6)$$

where q_{att}^l represents the normalization parameter. Θ_{att} is the dual attention parameter. The calculation formula of the activation function is as follows:

$$\sigma_2 = \frac{1}{1 + e^{(-w_i, c)}} \qquad (7)$$

where $e^{(-w_i, c)}$ stands for exponential function.

The calculation formula for output result $y_{i,c}^l$ of the attention module is as follows:

$$y_{i,c}^l = w_{i,c}^l \cdot \delta_i^l \qquad (8)$$

where w_i^l represents the content input to the attention module. $y_{i,c}^l$ is the output result of the attention module. g_i represents the gate signal provided by the higher-level context information. Θ_{att} indicates double attention

parameter. ψ , W and c are linear conversion parameters. b_g , $e^{(-w_i,c)}$ and b_{ψ} are offset parameters.

In summary, the double attention module is to convolute high-level features and low-level features to reduce the number of channels in the segmentation framework. Then the high-level features and low-level features are integrated, and the weighted vector is obtained after a series of operations such as convolution layer, batch normalization processing and up-sampling.

The proposed algorithm

To improve the segmentation effect of athletes' ankle injury images, the segmentation algorithm of athletes' ankle injury images based on residual double attention U-Net model is improved.
Input: athlete ankle joint injury image
Output: segmentation result of ankle joint injury image of athletes
The region of interest is extracted from the original ankle injury image of athletes. After translation, rotation and turning, the image data is expanded. The residual structure is used to adjust the gradient propagation and residual feedback of the segmentation framework, extract the attribute information in the region of interest, and combine the two to retain more image features. It is combined with the double attention module to improve the weight ratio of image features, the image segmentation of athletes' ankle joint injury is realized in the image segmentation framework based on the residual dual attention U-Net model. It is shown in Figure4.
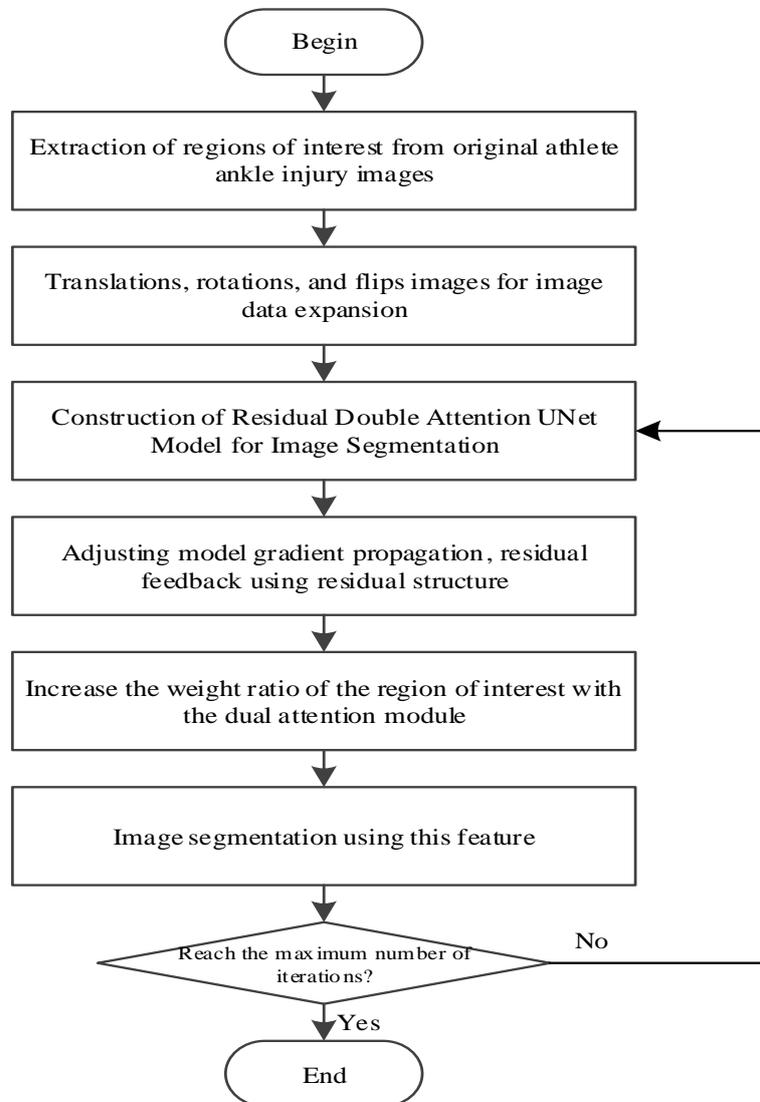
Figure 4. Process of the proposed algorithm

Figure 4 reveals that the region of interest is extracted from the original athlete ankle injury image, and the image data is expanded by performing translations, rotations, and flips. The residual structure is utilized to adjust the gradient propagation and residual feedback of the segmentation framework, extract attribute

information from the region of interest, and combine the two to preserve more image features. The double attention module is applied to improve the weight proportion of the image region of interest, and the image segmentation of athlete ankle injury is achieved by combining the feature extraction results.

EXPERIMENTAL RESULTS AND ANALYSIS

Experimental environment and Data sets

The experimental environment of this paper is based on windows10 64-bit system, with 16GB memory, and the GPU is NVIDIA GeForce Titan X. The deep learning PyTorch framework is built to realize the training of the model

Two data sets were selected in the experiment: (1) MPII human pose data set (http://human-pose.mpi-inf.mpg.de/) includes 165 images of the wrist, elbow, knee and ankle, including 775 × 522 pixels of ankle injury image. 85 of them are randomly selected as training images and the remaining 80 are test images. (2) UCI machine learning data set (http://archive.ics.uci.edu/ml/datasets.php). The data set used in this experiment is from the localization data for person activity data set in the life science class of UCI machine learning data set. There are 164860 samples and 8 features in this dataset, and the number of samples × the number of features > 500,000, including the position coordinates of the left and right ankles, waist and chest of five people at different time points, including 58 pieces of 896 × 768-pixel ankle joint image. In which, there were 26 ankle injury images and 32 normal ankle images, all of which were in 24-bit RGB format.

The experimental steps are as follows:

(1) Because the two data sets selected in the experiment are small, there are only 223 images in total. (2) After translation, rotation and flipping by using the data expansion method, 26,000 images are finally obtained as experimental data. After the experimental data set is set, 20000 ankle injury images are used as the training set, and 6000 ankle injury images are used as the experimental set. (3) After the above operations are completed, this part of data is taken as the data set of the experiment, and the experimental operation process is completed based on the two data sets. (4) NVIDIA GeForce Titan X GPU training includes 75 stages, and each stage trains 20 images. Set $\beta = 0.01$, and after every 1,000 times of training, β is multiplied by 0.1.

Evaluation criteria

The algorithm in MDAN [3], the algorithm in RVSTGAN [4], the algorithm in DNAS [5], the algorithm in TANN [6] and the algorithm in LEAC [7] as well as the algorithm experimental comparison method in this paper are compared, and five performance evaluation indexes are selected to verify the segmentation performance of different methods.

DSC: DSC is used to evaluate the coincidence degree between the actual segmentation effect and the theoretical segmentation effect of the algorithm. The larger the value of DSC, the better the segmentation effect of the algorithm.

$$DSC = 2 \times \frac{V(M \cap N)}{V(M) + V(N)} \qquad （9）$$

where $V(\cdot)$ refers to the size of the segmentation region in the image, M and N refer to the theoretical segmentation effect and the actual segmentation effect of the algorithm.

ASSD: ASSD is used to calculate the difference between the theoretical segmentation result and the actual segmentation result of the algorithm. The smaller the value of $ASSD$, the closer the segmentation result of the algorithm is to the theoretical segmentation result.

$$ASSD = \frac{X + Y}{M + N} \qquad （10）$$

where

$$X = \sum_{m \in M} \left[\min_{n \in N} \left\{ dist(m,n) \right\} \right] \qquad （11）$$

$$Y = \sum_{n \in N} \left[\min_{m \in M} \left\{ dist(n,m) \right\} \right] \qquad （12）$$

where $dist(n,m)$ and $dist(m,n)$ are the distance between the pixel point m and pixel point n in the images.

PM: PM is used to measure the degree of missing segmentation of ankle joint injury image by the algorithm. The larger the value of PM is, the less the algorithm misses segmentation.

$$PM = \frac{TPs}{GT} \times 100\% \qquad (13)$$

where TPs refers to the size of the region correctly segmented by the algorithm, and GT refers to the theoretical segmentation region.

CR: CR is used to measure the degree of error segmentation of the algorithm to the ankle injury image. The larger the value of CR is, the less the false segmentation in the actual algorithm.

$$CR = \frac{TPs - 0.5 \times FPs}{GT} \times 100\% \qquad (14)$$

where FPs represents the size of the algorithm's erroneous segmentation region.

F1 value: F1 value is used to judge the segmentation accuracy of the algorithm. The higher the value, the higher the segmentation accuracy of the algorithm.

$$F_1 = \frac{2 \cdot \Pr ecision \cdot \mathrm{Re} call}{\Pr ecision + \mathrm{Re} call} \qquad (15)$$

where $\Pr ecision$ represents the segmentation accuracy of the algorithm. $\mathrm{Re} call$ represents the recall rate of the algorithm.

RESULTS AND DISCUSSION

The comparison results of DSC values of different algorithms are shown in Figure 5.
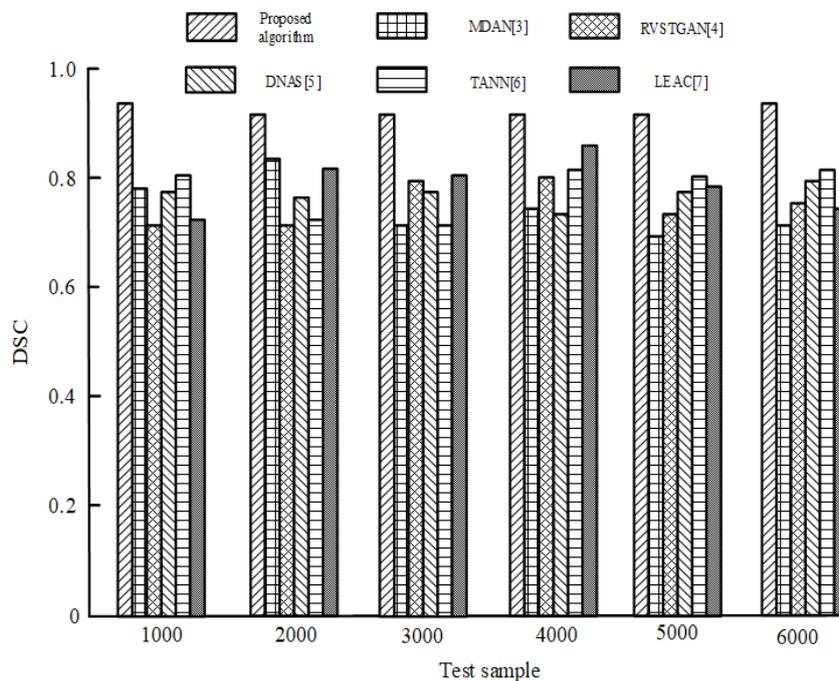
Figure 5. Comparison results of DSC value

According to the data in Figure 5, the maximum DSC value of the proposed algorithm is 0.93, which is 0.01, 0.13, 0.13, 0.1 and 0.06 higher than the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], respectively; The minimum DSC value of the proposed algorithm is 0.93, which is 0.14, 0.19, 0.18, 0.19 and 0.18 higher than the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], respectively. It shows that the DSC value of the proposed algorithm is higher than the algorithm in MDAN [3], the algorithm in RVSTGAN [4], the algorithm in DNAS [5], the algorithm in TANN [6] and the algorithm in

LEAC [7], which indicates that the actual segmentation effect of the proposed algorithm is higher than the theoretical segmentation effect, and the actual application effect is better.

The comparison results of ASSD values of different algorithms are shown in Figure 6.
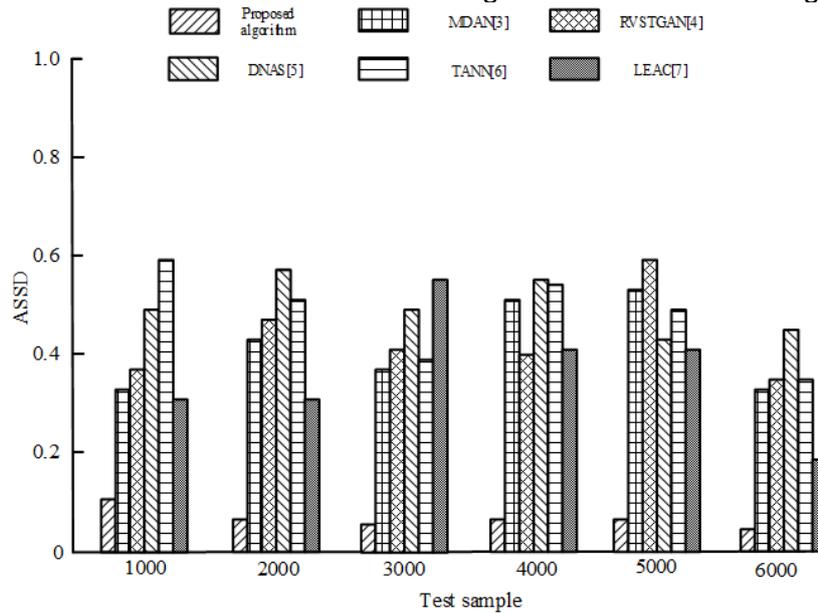
Figure 6. Comparison results of ASSD value

According to the data in Figure 6, the maximum ASSD of the proposed algorithm is 0.1, which is lower than the algorithm in MDAN [3], the algorithm in RVSTGAN [4], the algorithm in DNAS [5], the algorithm in TANN [6] and the algorithm in LEAC [7] by 0.33, 0.38, 0.5, 0.49 and 0.46, respectively; The minimum value of ASSD of the proposed algorithm is 0.1, which is lower than 0.26, 0.31, 0.38, 0.26 and 0.12 of the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], respectively. It shows that the ASSD value of the proposed algorithm is lower than that in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], which indicates that the closer the segmentation result of the proposed algorithm is to the theoretical segmentation result, the better the actual application effect.

The comparison results of PM values of different algorithms are shown in Figure 7.
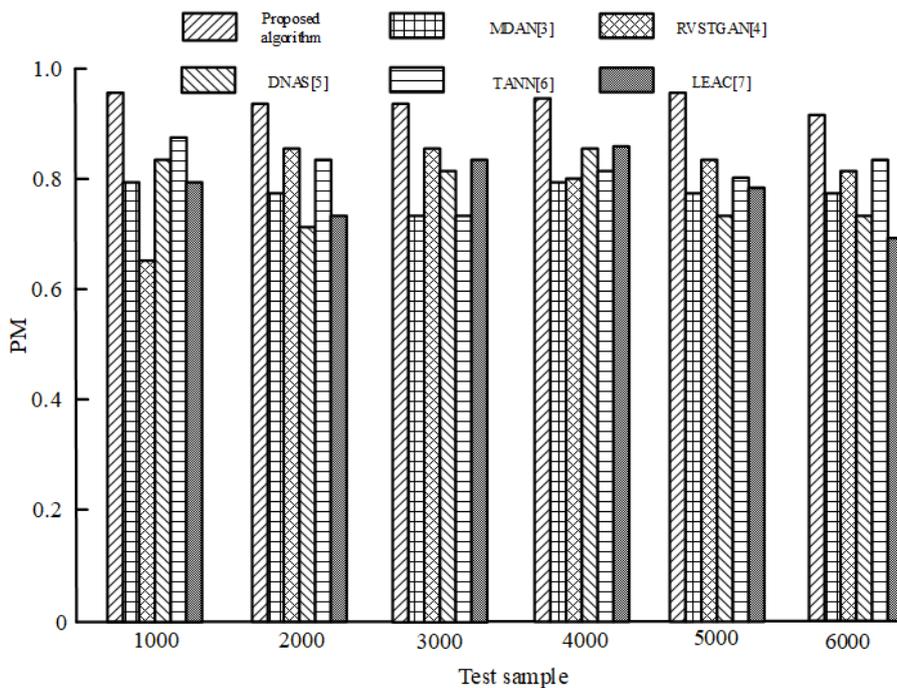
Figure 7. Comparison results of PM value

The comparison results of CR values of different algorithms are shown in Figure 8.
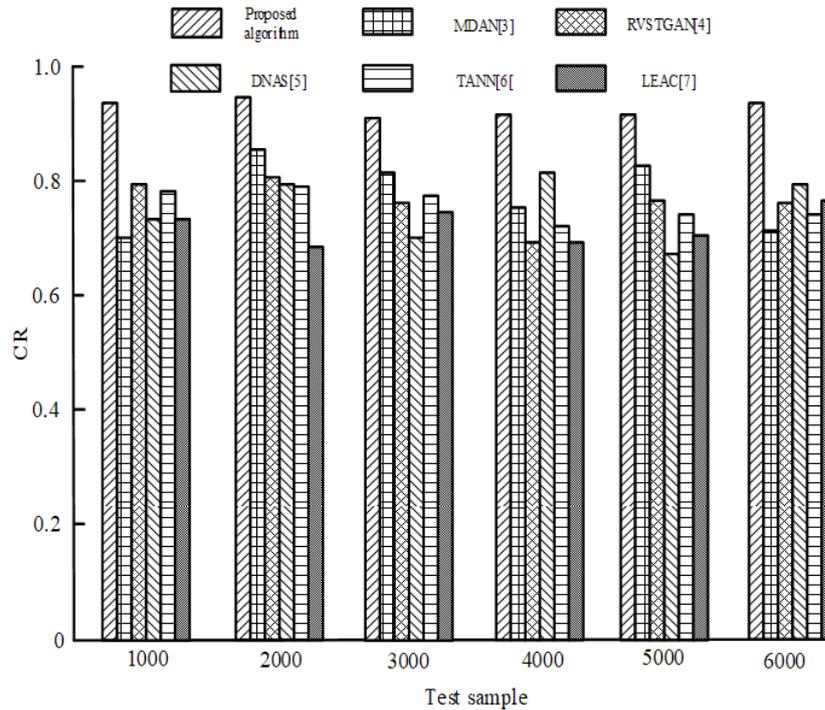
Figure 8. Comparison results of CR value

According to the data in Figure 8.The maximum CR of the proposed algorithm is 0.95, which is 0.09, 0.14, 0.16, 0.17 and 0.08 higher than the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7]; The minimum CR value of the proposed algorithm is 0.93, which is 0.23, 0.14, 0.25, 0.21 and 0.2 higher than the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], respectively. It shows that compared with the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], the CR value of the proposed algorithm is higher, which indicates that the proposed algorithm has fewer false segmentation and better practical application effect.

The F1 values of different algorithms are compared, and the results are shown in Table 1.

Table 1. Comparison results of F1 value (%)

Test sample	Proposed algorithm	MDAN[3]	RVSTGAN[4]	DNAS[5]	TANN[6]	LEAC[7]
1000	96.8	82.6	81.3	79.8	86.3	71.5
2000	97.8	84.7	82.6	78.5	84.7	84.6
3000	95.5	82.3	84.7	85.6	85.2	72.3
4000	92.3	85.6	85.6	82.5	86.3	74.6
5000	94.7	87.4	84.7	84.7	84.7	80.3
6000	96.8	86.3	80.4	83.1	85.2	82.6
Average value	95.7	84.8	83.2	82.4	85.4	77.7

According to the data in Table 1. The average F1 of the proposed algorithm is 95.7%, which is 10.9%, 12.5%, 13.3%, 10.3% and 8% higher than the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], respectively. Compared with the algorithms in MDAN [3], RVSTGAN [4], DNAS [5], TANN [6] and LEAC [7], the F1of the proposed algorithm is higher, which indicates that the segmentation accuracy of this method is higher and the actual application effect is better.
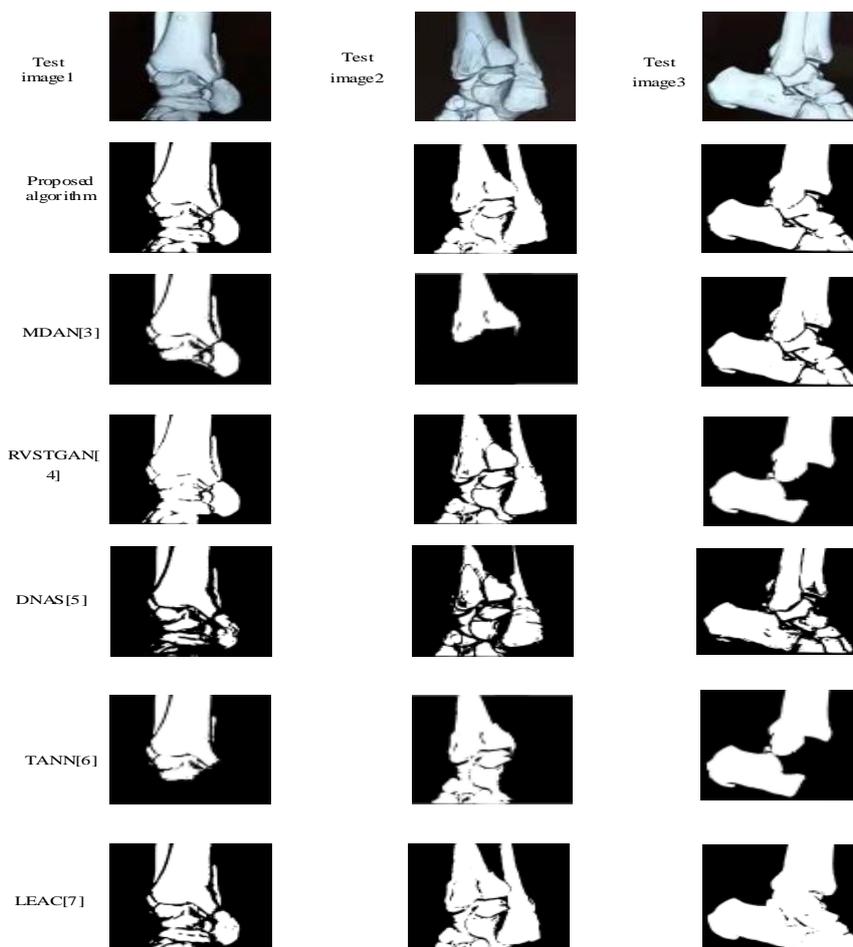
Figure 9. Comparison of segmentation results of ankle joint images

According to the data in Figure 9. Three ankle injury images were randomly selected from the test set, and the proposed algorithm was compared with the algorithm in MDAN [3], the algorithm in RVSTGAN [4], the algorithm in DNAS [5], the algorithm in TANN [6] and the algorithm in LEAC [7] for image segmentation, as shown in Figure7. There are many noises and impurities in the segmentation results of the algorithm in MDAN [3], and important details of test image 1 and test image 2 are lost; The segmentation results of the algorithm in RVSTGAN [4] are not well processed, and there are many burrs; The algorithm in DNAS [5] has some missing segmentation; The algorithm in TANN [6] has the problem of losing important details; The segmentation result details of the algorithm in LEAC [7] are lost, and there are many impurities in the image. However, the proposed algorithm has no missed segmentation and false segmentation, and the edge processing is also ideal. There is no burr, and the similarity between the segmented image and the original image is very high, which indicates that the proposed algorithm has high segmentation accuracy.

The calculation results of PSNR, MSE, AD, LMSE, NAE for the proposed algorithm and other algorithms are shown in Table 2.

Table 2. Comparison Results of PSNR, MSE, AD, LMSE, NAE

Algorithms	PSNR	MSE	AD	LMSE	NAE
MDAN[3]	32.1	0.45	0.06	0.39	0.05
RVSTGAN[4]	33.4	0.69	0.05	0.45	0.09
DNAS[5]	25.6	0.78	0.06	0.44	0.08
TANN[6]	24.8	0.77	0.04	0.46	0.07
LEAC[7]	30.7	0.89	0.07	0.37	0.05
Proposed algorithm	41.2	0.12	0.01	0.11	0.01

According to the comparison results of different methods in Table 2, we can see that the PSNR of the proposed algorithm is 41.2, which is much higher than other methods, and the MSE of the proposed algorithm is only 0.12. Among the other methods, the MSE of LEAC [7] algorithm is as high as 0.89, with the largest mean square error. For the index AD, the proposed algorithm has the lowest value of 0.01, and among other algorithms, the AD of LEAC [7] algorithm has the highest value of 0.07. The comparison results of LMSE index show that the LMSE value of the proposed

algorithm is only 0.11, and the LMSE of other methods is above 0.3. In the comparison of NAE index, the proposed algorithm still occupies a significant advantage, and the NAE is the lowest, which is only 0.01. Through the above comparison, it can be seen that the proposed algorithm has obvious advantages.

CONCLUSIONS

In this paper, a segmentation algorithm based on residual double attention U-Net model is proposed for ankle injury image, which highlights the characteristics of the region of interest in the image and reduces the probability of segmentation for other regions. Through the comparative simulation experiments with other methods, the DSC value of the proposed algorithm is between 0.90 and 0.93, the ASSD value is between 0.07 and 0.10, the PM value is between 0.92 and 0.96, the CR value is between 0.93 and 0.95, and the average value of F1 is 95.7%. It shows that the proposed algorithm has no missed segmentation or false segmentation. At the same tithe edge processing is also ideal, and there is no burr. The similarity between the segmented image and the original image is very high, and the segmentation effect is good. However, it is found that the proposed algorithm is difficult to achieve accurate segmentation for the slices with small shapes. Future research can focus on applying the proposed residual dual attention U-Net model for segmenting images of other body parts beyond athlete ankle injuries. Additionally, further exploration can be made to improve the pre-processing techniques for extracting regions of interest.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

1. Aljabri M, AlGhamdi M. A review on the use of deep learning for medical images segmentation. Neurocomputing. 2022; 506:311-335. doi:10.1016/j.neucom.2022.07.070
2. Pang S, Du A, Orgun M, Wang Y, Sheng Q, Wang S, et al. Beyond CNNs: exploiting further inherent symmetries in medical image segmentation. Ieee T Cybernetics. 2022, doi: 10.1109/TCYB.2022.3195447.
3. Huang W, Shao Z, Luo M, Zhang P, Zha Y. A novel multi-loss-based deep adversarial network for handling challenging cases in semi-supervised image semantic segmentation. Pattern Recongn Lett. 2021; 146(11):208-14. doi:10.1016/j.patrec.2021.03.017
4. Chen Z, Jin W, Zeng X, Zeng X, Xu L. Retinal vessel segmentation based on task-driven generative adversarial network. Let Image Process. 2021;14(7): 4599-605.doi:10.1049/iet-ipr.2020.1032
5. Wang B, Chen K, Tian X, Yang Y, Zhang X. An effective deep network for automatic segmentation of complex lung tumors in CT images. Med Phys, 2021; 48(9):5004-16. doi:10.1002/mp.15074
6. Karani N, Erdil E, Chaitanya K, Konukoglu E. Test-time adaptable neural networks for robust medical image segmentation. Med Image Anal. 2021;68(5):101907. doi:10.1016/j.media.2020.101907
7. Liu H, Fang J, Zhang Z, Lin Y. Localised edge‐region‐based active contour for medical image segmentation. Let Image Process. 2021; 15(7):1567-82.doi:10.1049/ipr2.12126
8. Quan F, Lang B, Liu Y. ARRPNGAN: Text-to-image GAN with attention regularization and region proposal networks. SIGNAL PROCESS-IMAGE.2022;106:116728. doi.org/10.1016/j.image
9. Shao M, Zhang G, Zuo W, Meng D. Target attack on biomedical image segmentation model based on multi-scale gradients. Inform Sciences. 2021;554(1)33-46. doi:10.1016/j.ins.2020.12.013
10. Law H, Choi G P, Lam K C, Lui L M. Quasiconformal model with CNN features for large deformation image registration. INVERSE PROBL IMAG. 2022; 16(4):1019-46. doi: 10.3934/ipi.2022010
11. Gao C, Ye H, Cao F, Wen C, Zhang Q, Zhang F. Multiscale fused network with additive channel-spatial attention for image segmentation. Know-based Syst. 2021; 214(28):106754. doi:10.1016/j.knosys.2021.106754
12. Jiang J, Hu Y C, Tyagi N, Rimner A, Lee N, Deasy J, et al. PSIGAN: Joint probabilistic segmentation and image distribution matching for unpaired cross-modality adaptation based MRI segmentation. Ieee T Med Imaging. 2020; 39(12):4071-84. doi: 10.1109/TMI.2020.3011626.
13. Jin B, Liu P, Wang P, Shi L, Zhao J. Optic disc segmentation using attention-based U-Net and the improved cross-entropy convolutional neural network. Entropy-switz. 2020;22(8):844. doi:10.3390/e22080844
14. Tian F, Gao Y, Fang Z, Gu J. Automatic coronary artery segmentation algorithm based on deep learning and digital image processing. Appl Intell. 2021; 51(1):8881–95. doi:10.1007/s10489-021-02197-6
15. Duan W, Chen Y, Zhang Q, Lin X, Yang X. Refined tooth and pulp segmentation using U-Net in CBCT image. Dentomaxillofac Rad. 2021; 12(1):20200251. doi:10.1259/dmfr.20200251
16. Li W, Zhu XY, Wang XC, Wang F, Liu JY, Chen MY, et al. Segmentation and accurate identification of large carious lesions on high quality x-ray images based on attentional U-Net model. A proof of concept study. J APPL PHYS. 2022;132(3):033103. doi:10.1063/5.0084593
17. Lu Y, Qin X, Fan H, Lai T, Li Z. WBC-Net: A white blood cell segmentation network based on UNet++ and ResNet. Appl Soft Comput. 2021;101(12):107006. doi:10.1016/j.asoc.2020.107006

18. Chen G, Dai Y, Li R, Zhao Y, Cui L, Yin X. SDFNet: automatic segmentation of Kidney ultrasound images using multi-scale low-level structural feature. Expert Syst Appl. 2021;185(12):115619. doi:10.1016/j.eswa.2021.115619
19. Li X F, Wang Y W, Cai Y J. Automatic annotation algorithm of medical radiological images using convolutional neural network. Pattern Recogn Lett. 2021;152:158-65. doi:10.1016/j.patrec.2021.09.011
20. Gupta S, Patil A T, Purohit M, Parmar M, Patel M, Patil H, et al. Residual neural network precisely quantifies dysarthria severity-level based on short-duration speech segments. Neural Networks. 2021;139(6):105-17. doi:10.1016/j.neunet.2021.02.008
21. Shu X, Gu Y, Zhang X, Hu C, Cheng K. FCRB U-Net: A novel fully connected residual block U-Net for fetal cerebellum ultrasound image segmentation. Comput Biol Med. 2022, 148:105693. doi:10.1016/j.compbiomed.2022.105693