Bone metastasis segmentation based on Improved U-NET algorithm

Jingyi Zhang\textsuperscript{1,2}, Mengge Huang\textsuperscript{1,2}, Tao Deng\textsuperscript{1,2,3,*}, Yongchun Cao\textsuperscript{2,3} and Qiang Lin\textsuperscript{1,2,3}

\textsuperscript{1} Key Laboratory of China’s Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou, 730030, P. R. China
\textsuperscript{2} Key Laboratory of Streaming Data Computing Technologies and Application, Northwest Minzu University, Lanzhou, 730030, P. R. China
\textsuperscript{3} School of Mathematics and Computer Science, Northwest Minzu University, Lanzhou, 730030, P. R. China
\* Corresponding author’s e-mail: dttom@lzu.edu.cn

Abstract. Whole body bone scan image analysis is widely used in nuclear medicine to assist nuclear medicine physicians in the detection of bone metastases. At present, the analysis of whole-body bone scan images mainly relies on the manual reading of nuclear medicine doctors. The doctors, based on personal knowledge and experience, look for abnormal lesion locations and diagnose them by examining the whole-body bone scan images. However, this method is prone to misdiagnosis and missed diagnosis. To solve the above problems, this study proposes an image segmentation method based on deep learning, which can automatically identify the location of bone metastases, so that doctors can make more accurate diagnosis. The Methods Attention mechanism was added to the jump connection of the original U-NET network to enhance the image feature selection. Experiments show that the algorithm in this study teaches traditional U-Net to show better results on the three indicators of MIoU Dice and MAP.

1. Introduction
Bone metastasis\cite{1} is one of the most serious hazards of malignant tumors and the main cause of clinical death in tumor patients\cite{2-3}. In the early stage, there are generally no symptoms, and bone isotope scans can find diseased bones throughout the body\cite{4-5}. SPECT bone scan\cite{6-7} is a nuclear medicine imaging examination of whole body bones. It has high sensitivity for the diagnosis of metastatic bone tumors and can detect bone metastases early\cite{8-11}. At present, the diagnosis of SPECT bone scan images is mainly read by doctors manually, but it is easily affected by the doctor’s subjective factors during the reading, and the lesion area of the SPECT image often appears as a small area with multiple lesions. In addition, the quality of the SPECT image is poor. There may be misdiagnosis or missed diagnosis. Therefore, the use of a computer diagnostic system to determine the size and location of the lesion area can greatly reduce the workload of doctors, reduce misdiagnosis and missed diagnosis, and improve the accuracy of diagnosis.

The algorithms used for SPECT bone scan image segmentation mainly include traditional methods and deep learning methods. Traditional algorithms include threshold segmentation, segmentation based on prior knowledge, cluster segmentation. For example, the image segmentation algorithm based on standard atlas adopted by Sadik et al.\cite{12} has high system sensitivity and specificity; the image
segmentation algorithm based on fuzzy set proposed by Huang et al.[13] is mainly used for image partitioning; the sensitivity and specificity of the feature point detection algorithm based on prior knowledge proposed by Šajn et al.[14] are easily affected by patient differences; xu et al.[15] proposed image segmentation based on kernel density estimation and K-means clustering algorithm; xu Lei et al.[16] proposed a segmentation algorithm based on gaussian mixture model and kernel density estimation.

With the development of deep learning technology, CNN has shown better and better results in image segmentation. Especially with the emergence of FCN[17] and U-Net[18], medical image segmentation has also achieved good results. Currently, the deep learning segmentation algorithms used in SPECT images include the BP neural network segmentation algorithm. For example, Zhu et al.[19] tried to use the BP neural network algorithm for image segmentation. The algorithm needs to build more network levels and has a large amount of calculation. The shortcomings such as slow convergence speed cannot reach the segmentation accuracy required for medical diagnosis. It can be seen that the SPECT bone scan image segmentation algorithm based on deep learning algorithm does not have a perfect algorithm.

In view of the shortcomings of the above deep learning algorithms, this study proposes a segmentation algorithm based on the improved U-NET model. The algorithm adds an attention mechanism between jump connections in the U-NET model to enhance the selection of data features, inhibit feature redundancy, and make the model better trained and learned.

2. Data and methods
The experimental data comes from the SPECT and medical clinical diagnosis records of the Nuclear Medicine Department of Gansu Provincial People’s Hospital from January to December 2017, two SPECT diagnoses of the front and back were collected 2~3 hours after the injection of Tc-99m methylene bisphosphate (MDP) radiopharmaceutical in each patient[20]. Each diagnosis contains SPECT imaging data and diagnosis report text. Among them, the SPECT imaging data is a data matrix composed of radiation values, and the matrix size is 256 (width) \times 1024 (height).

SPECT bone scan image data involves many diseases: bone metastasis, arthritis, myeloma. Among them, bone metastases often occur in the spine, thoracic cavity, pelvis and other parts, and are a frequent and frequent disease. Therefore, this article selects SPECT bone metastasis cases as the research content, and focuses the lesion area on the thoracic cavity. On the basis of 256 (width) \times 1024 (height) SPECT bone scan images of the whole province, the position coordinates of the thoracic cavity area are obtained through statistical methods, and the thoracic cavity is cut and saved as an experimental data set of 256 (width) \times 256 (height), and finally 125 sheets SPECT thoracic bone metastasis data matrix.

Due to the small number of data sets, sufficient features cannot be provided during model training. Therefore, the 128 original data sets are amplified by random horizontal and vertical displacement and random vertical tilt operations. The amplified data volume is 2280 sheets. According to the ratio of 8:2, the data is divided into training set (1830 sheets) and test set (450 sheets) to train and test the SPECT chest cavity bone scan bone transfer data.

2.1. Data annotation
The labeling process was completed by six members of the laboratory and three nuclear medicine experts from Gansu Provincial People’s Hospital. Among them, laboratory members used LabelMe (http://labelme.csail.mit.edu) labeling tools to mark the lesion area after being trained by doctors, the director of nuclear medicine experts is responsible for checking the marking results to ensure the correctness of the markings. If the three experts’ marks are consistent, the image is considered to be marked correctly. If the three experts’ marks are inconsistent, the image will be re-marked according to the principle of how much the minority should obey.

1) Sequence annotation.
In the process of medical image scanning, the position where the patient lies has a certain tilt angle relative to the machine direction. In order to compensate for the impact of this situation, the experiment performed a rotation operation on the data set. Taking into account the actual situation, the image rotation angle is \( \theta \in (0^\circ, 5^\circ) \). The mathematical formula corresponding to the rotation operation is shown
in formula (1), among them, \((x_1, x_2)\) and \((x, y)\) are the rotated coordinate point and the original coordinate point, \(\theta \in (\theta_1, \theta_2)\) is the rotation angle.

\[
\begin{bmatrix}
    x_r \\
    y_r
\end{bmatrix} = \begin{bmatrix}
    \cos \theta & \sin \theta \\
    -\sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
    x \\
    y
\end{bmatrix}
\]  
(1)

2) Translation.

In the process of medical image scanning, the position of the patient lying down is not absolute, but has a certain displacement relative to the machine. In order to compensate for the impact of this situation, the experiment performed a translation operation on the data set. The displacement distance should be set according to the actual situation. The geometric formula corresponding to the translation operation is shown in formula (2), among them \((x_1, x_2, x, y)\) \((\Delta x, \Delta y)\) respectively correspond to the coordinate point after translation, and \(a\) is the original coordinate point and translation amount.

\[
\begin{bmatrix}
    x_r \\
    y_r
\end{bmatrix} = \begin{bmatrix}
    1 & 0 \\
    0 & 1
\end{bmatrix} \begin{bmatrix}
    x \\
    y
\end{bmatrix} + \begin{bmatrix}
    \Delta x \\
    \Delta y
\end{bmatrix}
\]  
(2)

3) Horizontal mirroring.

The SPECT nuclear medicine test generates two images of the front bitmap and the back bitmap each time. In the process of storage and preservation, image loss often occurs. In order to compensate for the influence of this phenomenon on the experiment, a horizontal mirror operation is performed on the SPECT image. The geometric formula corresponding to horizontal mirroring is shown in formula (3), where \((x, y)\) and \((x_1, x_2)\) correspond to the original coordinate point and the coordinate point after the horizontal mirroring operation.

\[
\begin{bmatrix}
    x_r \\
    y_r
\end{bmatrix} = \begin{bmatrix}
    -1 & 0 & 0 & 1 \\
    0 & 1 & 0 & 0
\end{bmatrix} \begin{bmatrix}
    x \\
    y
\end{bmatrix}
\]  
(3)

The result of the data enhancement operation is shown in Figure 1.

---

Figure 1. Data enhancement operation ((a) is the original image, (b) is the image after rotation, (c) is the image after translation, (d) is the image after horizontal mirroring).

2.2. Model Structure

U-Net is a typical convolutional network, which has been widely used in medical image segmentation after being proposed, and has achieved good results. The method in this study is improved based on the U-Net model, introducing an attention mechanism between each jump connection of the network, allowing the network to focus its attention on the salient features useful for the task, and suppressing excessive extraction of irrelevant feature regions. The network model of this article is shown in Figure 2, which consists of an encoding part, a decoding part and a jump connection.
The model is a U-shaped symmetrical structure. First, the network is divided into left and right parts for analysis. The left side is the compression process. Reduce the image size through convolution and pooling, and extract some shallow features. The right part is the decoding process. Some deep features are obtained through convolution and upsampling. After each convolution, the size of the image will become larger. The jump connection part is used to connect the encoding part and the decoding part. Before splicing and fusing the features of the encoding part with the corresponding features of the decoding part, this study introduces an attention mechanism to suppress irrelevant feature responses and reduce the number of redundant features. Combining the feature map obtained in the encoding stage with the feature map obtained in the decoding stage is to combine deep and shallow features, refine the image, and then perform prediction segmentation based on the obtained features. It should be noted that the feature sizes of the two layers are different, so they need to be cropped. The last layer is classified by 1x1 convolution.

2.3. attention mechanism
In medical images, the detected SPECT thoracic image contains various parts such as normal bones, lesion areas, bladder. However, in segmentation training, only the lesion area is focused, and the rest are treated as background. Therefore, this study proposes to add an attention mechanism to the U-Net network. The core goal is to focus attention on the salient features useful for the task, such as a certain lesion area or organ tissue, and suppress excessive extraction of irrelevant feature areas. The structure of the attention mechanism is shown in Figure 3.
Among them, \( F_g \) is the feature map of the decoding part, and \( F_h \) is the feature map of the encoding part. \( C \) is the number of feature channels, \( H \) and \( W \) are the length and width of the feature, and \( D \) is the dimension in the feature map space.

1) Feature extraction

\[
F_h = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x(i,j)
\]

\[
F_g = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x(i,j)
\]

By formula (1) (2), \( F_h \), \( F_g \) will be changed from \([C,H,W]\) to \([C,1,1]\) through the average pooling method.

2) Attention calculation

\[
q_{\text{att}} = \psi^T(\sigma_2(W_f F_x + W_g F_h + b_g)) + b_p
\]

\[
\alpha = \sigma_2(q_{\text{att}}(s; \theta_{\text{att}}))
\]

Among them, \( \sigma_2 \), \( \sigma_1 \) are the Sigmoid activation function and ReLU function respectively. \( q_{\text{att}} \) is the attention calculation, and \( \alpha \) is the weight matrix after the feature weight is updated.

3) The feature weight is mapped to the feature map

\[
x' = x \cdot \alpha
\]

Multiply the updated weight \( \alpha \) with the feature map \( x \) to obtain a channel-attracting feature map \( x' \). This operation increases the weight of the target feature channel, reduces the weight of meaningless feature channels, and strengthens meaningful features.

3. Experiment

3.1. Experimental detail

The deep learning framework used in the experiment is Tensorflow 1.4, the computer operating system is Windows 8, and the GPU used for training and testing data. The batch size during training is 8, and the optimizer is Adam, where the initial learning rate is 0.00001, the exponential decay rate estimated for the first time is 0.9, and the exponential decay rate estimated for the second time is 0.999. When testing the model, use batch processing as 1 Perform the test. Both the output picture and the input picture are 256 (height) × 256 (width).

3.2. Result evaluation and analysis

In order to evaluate the performance of the method in this study, Dice similarity coefficients, MIoU and MAP are used for verification and evaluation. The calculation formula is shown in formula (9) – (11).

\[
\text{Dice} = \frac{2TP}{2TP + FP + FN}
\]

\[
\text{MIoU} = \frac{1}{(k+1)} \sum_{l=0}^{k} \frac{TP}{TP + FP + FN}
\]

\[
\text{MAP} = \frac{TP}{TP + FP + FN}
\]

Among them, \( TP \) (true example): predict a positive sample as a positive sample, \( FN \) (false negative example): predict a positive sample as a negative sample, \( FP \) (false positive example): predict a negative sample as a positive sample, \( TN \) (true negative example): Predict negative samples as negative samples.

3.3. Experimental results

In order to verify the optimal effect of the algorithm proposed in this study, the algorithm in this study is compared with U-Net and U-Net_Inception in MIoU, Dice and MAP. The results are shown in Table 1, and the segmentation results are shown in Figure 4.
Table 1. Index results of the three models for different categories.

| Model                      | MIoU  | Dice Coefficient | MAP  |
|----------------------------|-------|-----------------|------|
| U-Net                      | 0.614 | 0.552           | 0.723|
| The algorithm of this article | 0.633 | 0.571           | 0.761|
| U-Net_Inception            | 0.608 | 0.550           | 0.719|

From Table 1 and Figure 4, it can be seen that for the SPECT bone scan thoracic images with fewer lesions and more scattered, the algorithm in this study shows a more accurate segmentation result than the other two models. Because the algorithm in this study integrates the attention mechanism in the jump link, pays attention to the selection of relevant features, and suppresses irrelevant features. Experiments show that the algorithm in this study teaches the U-Net algorithm to improve the segmentation accuracy of bone metastases in SPECT bone scan thoracic images, and can further provide more auxiliary information for doctors' disease diagnosis and follow-up treatment.

4. Summary
Aiming at the automatic segmentation of bone metastases from SPECT bone scan thoracic images, this study studies and constructs an improved algorithm based on U-Net. First, the processing and expansion of SPECT bone scan data are explained; then, the structure and improvement of the algorithm are explained in detail Content; finally, based on real SPECT bone scan thoracic cavity image data, the constructed model is verified, and the experimental results verify the feasibility of this algorithm.

Acknowledgments
This research was funded by the National Natural Science Foundation of China [No. 32060234], the Gankeji ([2020] 11) Natural Science Foundation (20JR5RA511).

References
[1] Xu, W., Ma, W., Yang, Y., et al. (2020) Investigation and analysis of pain factors in patients with
bone metastases. Gansu Medical Journal, 39: 528–529.

[2] Dong, Q., Huang, L., Xiong, Z., et al. Clinical characteristics and risk factors of bone metastasis in patients with liver cancer. Tumor, 2020, 40: 775–783.

[3] Wang, D., Zhong, L., Tu, B., et al. (2020) Ultrasonic and pathological findings of bone metastases. Chinese Journal of Interventional Imaging and Therapy, 17: 636–638.

[4] Zheng, Q., Su, J., Li, X., et al. (2020) Differential Diagnostic Value of $^{18}$F-FDG PET/CT Imaging in ultiple Myeloma and Bone Metastases[J], Journal of Experimental Hematology, 28: 1267–1271.

[5] Yin, G., Pan, S. (2020) Advances in imaging studies of bone metastasis. Journal of Modern Oncology, 28: 1977–1980.

[6] Wang, Y. (2020) SPECT/CT Fusion Imaging in the Diagnosis of Bone Metastases. China Medical Device Information, 26: 155–156.

[7] Wang, C. (2020) Analysis of SPECT/CT Fusion Image and Whole Body Bone Imaging in the Diagnosis of Bone Metastasis. China Health Standard Management, 11: 84–86.

[8] Li, L. (2020) Diagnostic Value of SPECT/CT tomographic fusion imaging in bone metastasis of prostate cancer. China Modern Medicine, 27: 128–130.

[9] Tang, X., Zhu, G., Lai, C. (2018) Comparison and analysis of the diagnostic value of bone metastases with nuclide bone scan and magnet resonance general diffusion-weighted imaging. Chinese Youjiang Medical Journal, 46: 210–213.

[10] Šajn, L., Kononenko, I., Milčinski, M. (2007) Computerized segmentation and diagnostics of whole-body bone scintigrams. Comput Med Imag Grap, 31: 531–541.

[11] Guo, G., Xu, X. (2017) The Application of Image Fusion Technology of CT and SPECT. Medical Equipment, 30: 23–24.

[12] Sadik, M., Hamadeh, I., Nordblom P., et al. (2008) Computer-assisted interpretation of planar whole-body bone scans. J Nucl Med, 49: 1958–1965.

[13] Huang, J., Kao, P., Chen, Y. (2007) A set of image processing algorithms for computer-aided diagnosis in nuclear medicine whole body bone scan images. IEEE T Nucl Sci, 54: 514–522.

[14] Šajn, L., Kukar, M., Kononenko, I., et al. (2005) Computerized segmentation of whole-body bone scintigrams and its use in automated diagnostics. Comput Meth Prog Bio, 80: 47–55.

[15] Xu, L., Meng, Q., Yang, R., et al. (2015) Segmentation of bone scintigraphy image via the K-means clusters with kernel density estimation. Journal of Nanjing Medical University, 35: 585–589.

[16] Xu, L., Meng, Q., Yang, R., et al. (2016) A Research on the Segmentation Algorithm for the Whole Body SPECT Image via the Gaussian Mixture Model with Kernel Density Estimation. China Medical Devices, 31: 48–51+47.

[17] Ronneberger, O., Fischer, P., Brox, T. (2015) U-net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI). Cham. 234–241.

[18] Long, J., Shelhamer, E., Darrell, T. (2015) Fully convolutional networks for semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition. Boston 3431–3440.

[19] Zhu, C., Tian, L., Chen, P., et al. (2007) Segmentation of Whole Body Bone SPECT Image Based on BP Neural Network. Journal of Biomedical Engineering, 24: 1050–1053.

[20] Jia, P., Li, A., Sun, Y. (2015) ~(99m)Tc-MDP imaging in the diagnosis of sacrum tumor. Journal of Chinese Practical Diagnosis and Therapy, 29: 190–191.