Wnet++: A Nested W-shaped network with multiscale input and adaptive deep supervision for osteosarcoma segmentation

Limei Shuai[1], Xin Gao[2], Jiajun Wang[1]

[1]School of Electronic and Information Engineering, Soochow University, Suzhou, China.

[2]Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Science, Suzhou, China

*Email: 20194228021@stu.suda.edu.cn

Introduction

- **Background**: Osteosarcoma is one of the most common primary malignant bone tumors which most commonly occurs in adolescents and children. Therefore, an accurate and reliable automatic segmentation method is urgently needed in clinic.
- **Aim**: In order to solve the demand for more accurate segmentation in medical images, we further investigate effective ways to improve the segmentation results of the networks.
- **Contribution**: We hypothesize that the cursory predicted maps from the first network could be sequentially improved by being fed into the second network, meanwhile, the dense skip connections combine the characteristic of the same scale in different levels. Therefore, W-net++ combined with multi-scale input, adaptive deep supervision and CAM was proposed to achieve a more accurate segmentation in osteosarcoma CT images.

Method

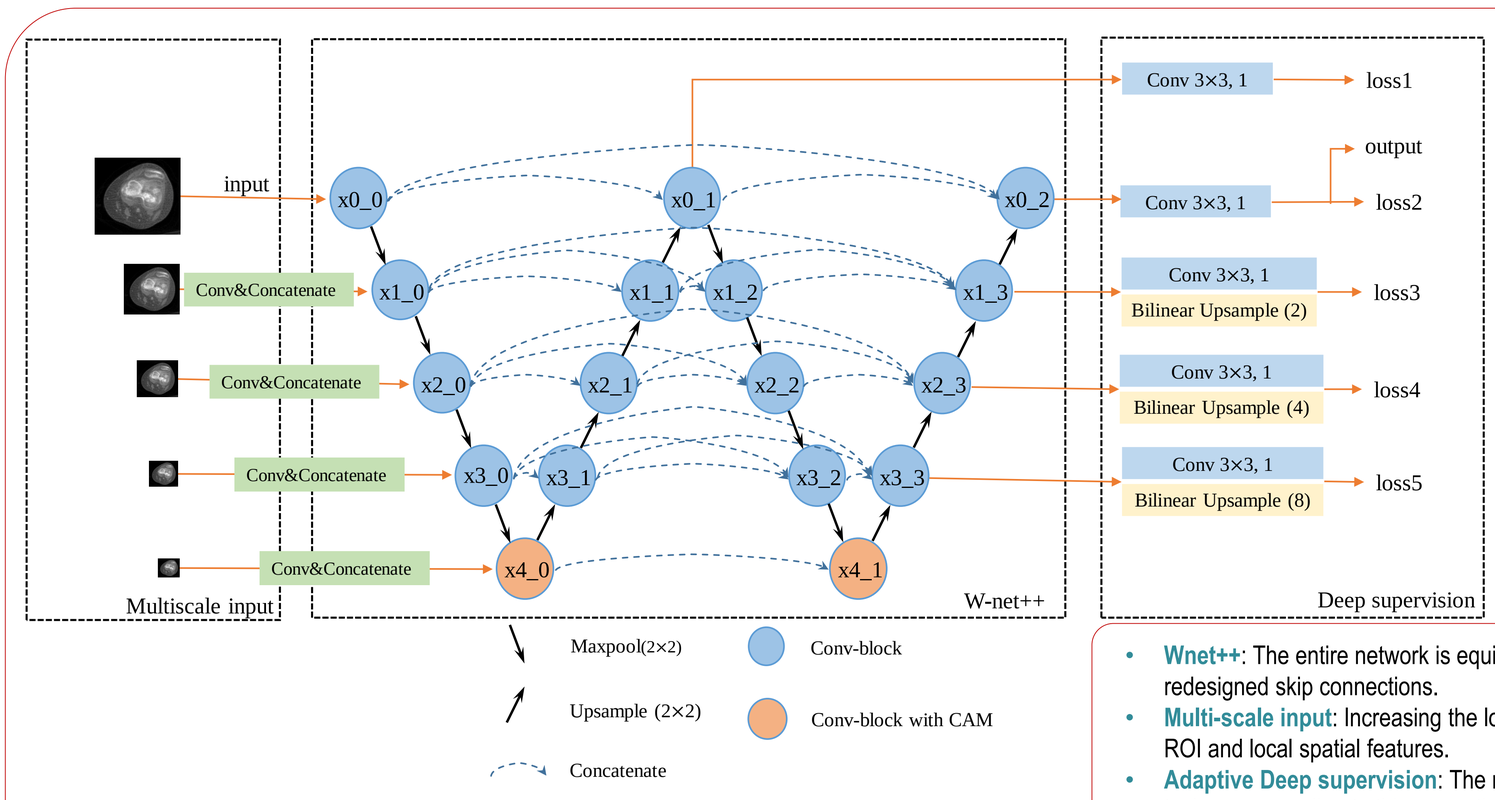

Fig.1 The overall architecture of the network

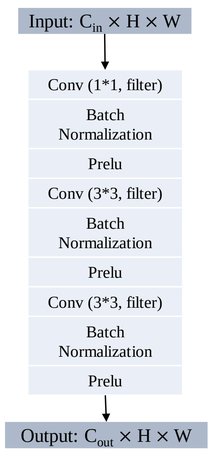

Fig.2 The Conv-block

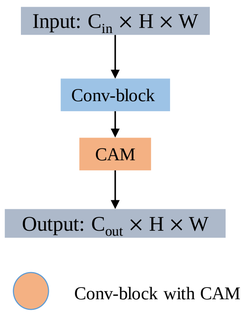

Fig.3 The Conv-block with CAM

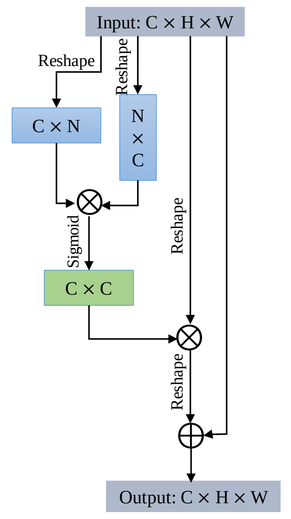

Fig.4 The CAM

- **Wnet++**: The entire network is equivalent to two cascaded U-Net forming a W-shape network, and a nested network through redesigned skip connections.
- **Multi-scale input**: Increasing the low-level semantic feature in the high-level maps facilitates the network learning the features of small ROI and local spatial features.
- **Adaptive Deep supervision**: The method is to calculate the proportion of each loss in the total loss, and then assign weight to each task loss according to this proportion, and multiply each loss by the new weight and sum again to get the total loss after reweighting.
- **Conv-block with CAM**: Add Channel Attention Module(CAM) to the backbone of Conv-block, and embed it into the bottom blocks of the model which contain the maximum channels, so as to learn more about the correlation characteristics between channels.

Experiments & Results

- **Materials**: The dataset we built contains 2303 annotated osteosarcoma CT images of 23 osteosarcoma patients ranged in ages from 8 to 30 years old.
- **Method**: Experiments were performed by a five-fold cross-validation method, and the average results of five experiments were taken to evaluate the segmentation performance of the model.
- **Preprocessing**: (a) Images were cropped to 320 × 320; (b) Normalization; (c) Histogram equalization
- **Evaluation Metrics**: Dice Similarity Coefficient (DSC), Jaccard Similarity Coefficient, Precision, Sensitivity, Specificity were used to evaluate the prediction quantitatively.
- **Results**: Fig.5 shows segmentation results of the tumor lesions for four osteosarcoma CT images where the ground truth delineated by radiologist, results from the U-Net, U-Net++, U-Net3+, MSFCN, MSRN and proposed model are presented, respectively. it demonstrated that our method achieved an average DSC gain of 6.17 points, 1.91 points, 1.55 points over U-Net, U-Net++, MSRN respectively.

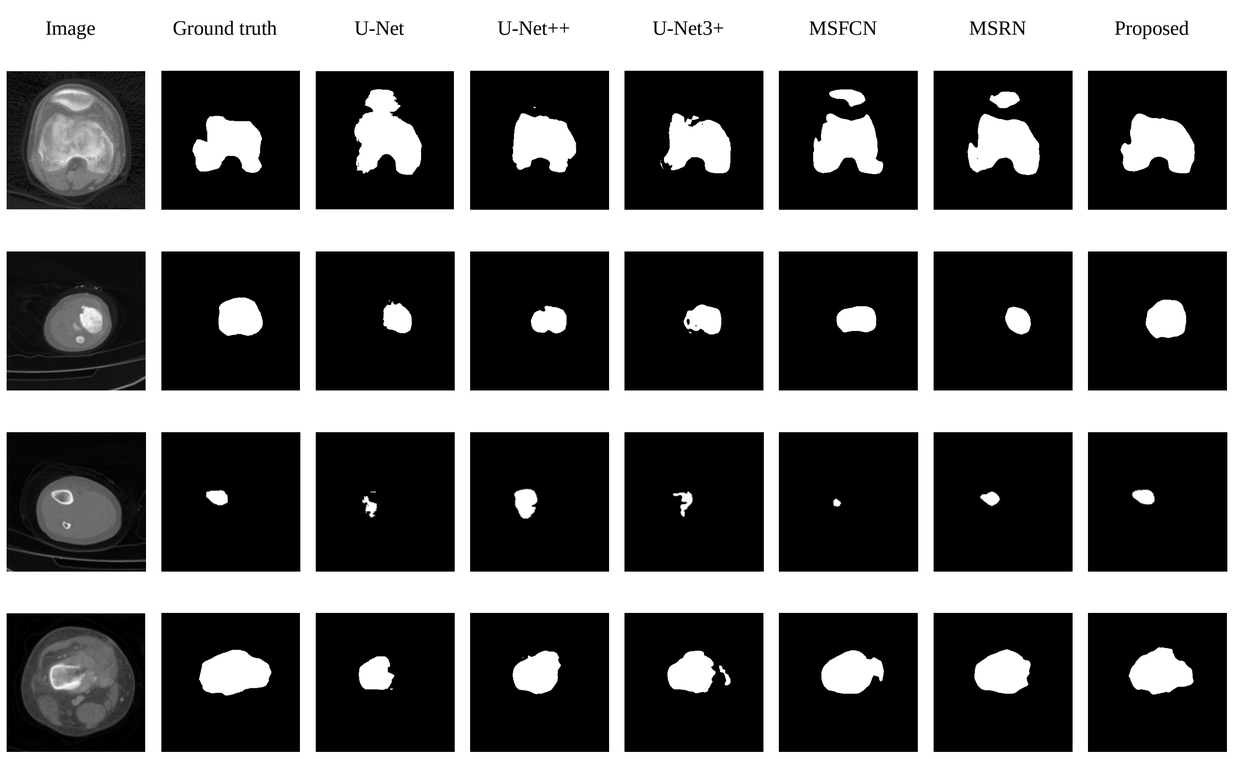

Fig.5 Results on four subjects on the osteosarcoma dataset. The proposed network achieves qualitatively better results than other state-of-the-art networks.

Conclusions

The W-net++ with multiscale input and adaptive deep supervision is proposed and applied in osteosarcoma CT image segmentation. We have evaluated our method on osteosarcoma dataset, and it demonstrated that our method achieved an average DSC gain of 6.17 points, 1.91 points, 1.55 points over U-Net, U-Net++, MSRN respectively.

References

[1] L. Huang, W. Xia, B. Zhang, B. Qiu, and X. Gao, “MSFCN-multiple supervised fully convolutional networks for the osteosarcoma segmentation of CT images,” Comput. Methods Programs Biomed., vol. 143, pp. 67–74, 2017, doi: 10.1016/j.cmpb.2017.02.013.

[2] D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen, and H. D. Johansen, “Double U-Net: A deep convolutional neural network for medical image segmentation,” Proc. - IEEE Symp. Comput. Med. Syst., vol. 2020-July, no. 1, pp. 558–564, 2020, doi: 10.1109/CBMS49503.2020.00111.

Acknowledgement

Natural Science Foundation of China No.61473243

Natural Science Foundation of Jiangsu Province No.BK20171249.