Accurate and Efficient Segmentation for Gross Tumor Volume in Head and Neck CT

华中科技大学

Kaixiang Yang, Xin Fang, Xin Li, Qiang Li, Zhiwei Wang

Wuhan National Laboratory for Optoelectronics and with MoE Key Laboratory for Biomedical Photonics

School of Engineering Sciences, Huazhong University of Science and Technology, Wuhan, China

{kxyang, zwwang}@hust.edu.cn

MICCAI2023 Vancouver CANADA

Background

- Nasopharyngeal Cancer : malignant tumor with rapid increase and high mortality; Screening tool: no-contrast and contrast-enhanced CT;
- Treatment: Radiotherapy with external beam radiation; Radiotherapy planning: Delineating gross tumor volume (GTVs) by radiologists

Key Challenge

- Inaccurate segmention caused by ①Individual differences, ②Low contrast of soft tissue in CT, ③Unbalance of bg and fg, ④Multi-regions

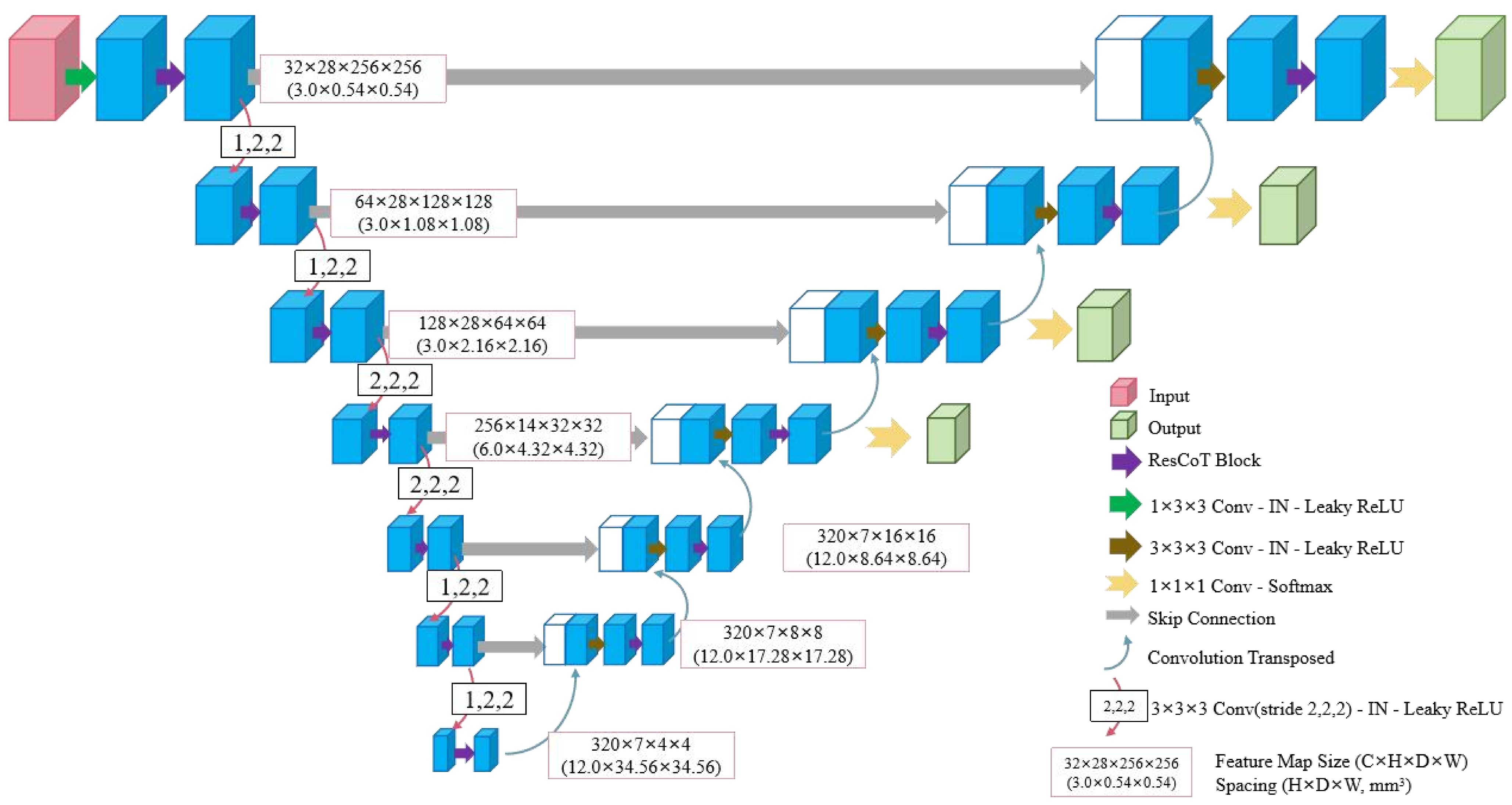
Our Contributions

- We combine the Focal loss to make the model focus on the GTVs which are not prone to distinguish.
- Train a single model without pre-training, ensemble or additional data to obtain accurate delineation, has certain clinical significance.

Motivations and Method

A pure 3D U-Net and the deep supervision strategy are utilized to decrease model parameters and enhance learning ablility in different stages, respectively. Two CT scans (no-contrast and contrast-enhanced CT) are concatenated on the channel dimension as network input.

Thus, the framework is as follows:



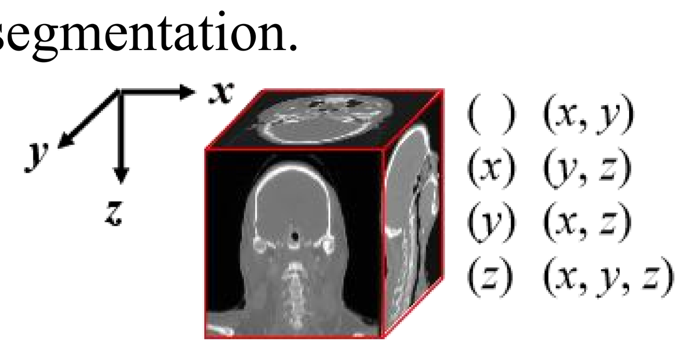
Optimization

Dice Loss plus Focal Loss

$$
\begin{aligned}
L &= L_{Dice} + L_{Focal} \\
&= \sum_{c=0}^{C-1}\left(1-\frac{TP_p(c)}{TP_p(c)+\alpha FP_p(c)+\beta FN_p(c)}\right)/C \\
&\quad -\sum_{n=1}^{N}\sum_{c=0}^{C-1}\left[g_n(c)(1-p_n(c))^2\log(p_n(c))\right]/N
\end{aligned}
$$

Inference

Flipping Testing Time Augmentation (TTA) is used in inference to performance more accurately GTVs segmentation.



Experiments

1. Training and Validation:

Target	Dice and CE loss		Dice and Focal loss	
	DSC(%)	NSD(%)	DSC(%)	NSD(%)
GTV_p	79.86±6.80	72.28±12.26	80.44±5.92	72.61±11.71
GTV_{nd}	77.37±8.04	73.12±7.80	77.26±9.44	73.09±9.09
Average	78.61±7.55	72.70±10.28	78.85±8.04	72.85±10.49
AVG Score	75.66±8.38		**75.85±8.70**	

✦Dice-Focal Loss performs higher than Dice-CE Loss, making model focus on those hard-to-segment.

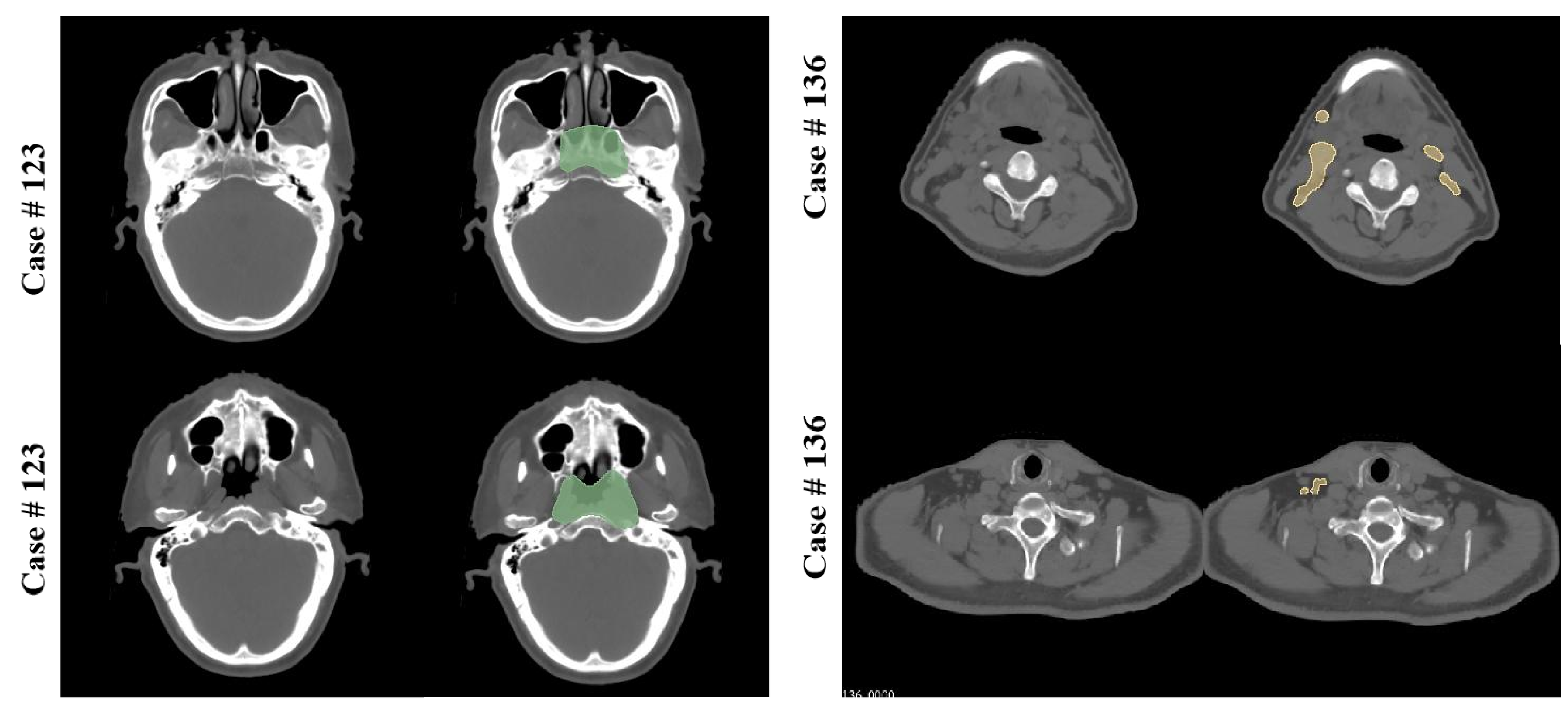
✦ Validation: ***75.85***±***8.70*** ranking ***4th*** on leaderboard

2. Test:

Target	Dice and Focal loss	
	DSC(%)	NSD(%)
GTV_p	78.76±6.60	35.92±11.05
GTV_{nd}	67.41±13.78	63.08±15.37
Average	73.09±12.21	49.50±19.07
AVG Score	61.29±12.18	

✦ Final Result: ***61.29***±***12.18*** on 60 testing cases

3. Visulization:



References

[1] Isensee, Fabian, et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." Nature methods. 2021.

[2] Ronneberger, et al. "U-net: Convolutional networks for biomedical image segmentation." Int. Conf. on MICCAI. 2015.

[3] Lin, Tsung-Yi, et al. "Focal loss for dense object detection." ICCV. 2017.