

FocusNet: Imbalanced Large and Small Organ Segmentation with an End-to-End Deep Neural Network for Head and Neck CT Images

Yunhe Gao^{1*}, Rui Huang^{1*}, Ming Chen^{4,5,6}, Zhe Wang¹, Jincheng Deng³, Yuanyuan Chen^{4,5,6}, Yiwei Yang^{4,5,6}, Jie Zhang^{4,5,6}, Chanjuan Tao^{4,5,6}, and Hongsheng Li^{1,2(✉)}

¹ SenseTime Research, {claudegyh@gmail.com, huangrui@sensetime.com}

² The Chinese University of Hong Kong, {hsli@ee.cuhk.edu.hk}

³ Shenzhen Yino Intelligence Techonology,

⁴ Inst. of Cancer Research & Basic Medical Sciences of Chinese Academy of Sciences,

⁵ Cancer Hospital of University of Chinese Academy of Sciences,

⁶ Zhejiang Cancer Hospital

Abstract. In this paper, we propose an end-to-end deep neural network for solving the problem of imbalanced large and small organ segmentation in head and neck (HaN) CT images. To conduct radiotherapy planning for nasopharyngeal cancer, more than 10 organs-at-risk (normal organs) need to be precisely segmented in advance. However, the size ratio between large and small organs in the head could reach hundreds. Directly using such imbalanced organ annotations to train deep neural networks generally leads to inaccurate small-organ label maps. We propose a novel end-to-end deep neural network to solve this challenging problem by automatically locating, ROI-pooling, and segmenting small organs with specifically designed small-organ sub-networks while maintaining the accuracy of large organ segmentation. A strong main network with densely connected atrous spatial pyramid pooling and squeeze-and-excitation modules is used for segmenting large organs, where large organs' label maps are directly output. For small organs, their probabilistic locations instead of label maps are estimated by the main network. High-resolution and multi-scale feature volumes for each small organ are ROI-pooled according to their locations and are fed into small-organ networks for accurate segmenting small organs. Our proposed network is extensively tested on both collected real data and the *MICCAI Head and Neck Auto Segmentation Challenge 2015* dataset, and shows superior performance compared with state-of-the-art segmentation methods.

1 Introduction

In this paper, we attempt to solve the segmentation problem of imbalanced large and small organs in head and neck (HaN) CT images, where the key challenge is to precisely segment small organs whose volume is much smaller than

* Yunhe Gao and Rui Huang contributed equally to this work.

the average. In HaN radiotherapy planning task, it is of vital importance to accurately determine the locations and volumes of organs-at-risks (OARs). Oncologists would design radiotherapy plans such that the radiation can be concentrated on the lesion area without damaging normal organs. Currently, all OARs are manually annotated by oncologists, which is time consuming, tedious and prone to have high inter- and intra-observer variations. Therefore, a computer-aided head organ segmentation system could significantly lower the work load of doctors.

The main difficulty of the task is the severe imbalance between large and small organs (e.g., the smallest organ, *lens*, only occupy 0.0028% of the whole 3D volume, while the *parotid gland* is over 250 times larger than *lens*). State-of-the-art segmentation neural networks trained based on samples' natural frequencies would have poor performance on the small organs. In addition, due to the limitation of CT technology and the complex anatomical structure of the human head, the contrast between organs and their surroundings is often low. All these factors coupled together make it difficult to develop a method for segmenting both small and large organs simultaneously and accurately.

Over the past decade, many approaches were proposed to resolve the challenging problem of HaN organ segmentation. Early approaches include atlas-based methods, active contours, graph cut and etc. Atlas-based methods were commonly used where there is only a small number of annotated images available. However, atlas-based methods are based on image registration techniques and might generate incorrect organ maps if the organs are occupied by tumors.

Recently, convolutional neural networks (CNN), with its powerful feature representation capability, have made revolutionary progress in many tasks. 2D CNN models such as U-Net [8], and its 3D variants have achieved large performance gain in 2D and 3D segmentation than traditional methods. For OAR segmentation, Ibragimov et al. [4] proposed the first deep learning-based algorithm. Ren et al. [7] proposed a interleaved 3D-CNN for segmenting small organs in HaN, where the region of interest is obtained by registration. Zhu et al. [11] proposed a 3D Squeeze-and-Excitation U-Net for fast segmentation. However, existing segmentation CNNs are not optimized for imbalanced organ segmentation, these networks generally produce accurate segmentation maps for large organs, while the accuracy of small organs is often sacrificed.

To address the issue of imbalanced large and small organ segmentation, we observed how oncologists annotate OARs. For organs with large volume, they usually annotate them at the normal scale, while for the small organs, they first find the location and then zoom in to accurately mark them. According to this observation, we propose a novel end-to-end 3D convolutional neural network, FocusNet, which is delicately designed for accurate segmentation of both large organs and small organs.

The whole network has three parts: main segmentation network (S-Net), Small-Organ Localization branch (SOL-Net), Small Organ Segmentation branch (SOS-Net). S-Net is a strong backbone network, which is responsible for the segmentation of all large organs and also provides features for small-organ seg-

mentation. SOL-Net is trained for highlighting the center locations of the small organs. Based on the results of SOL-Net, high-resolution feature volumes concatenated with multi-scale features volumes are ROI-pooled and fed into to the SOS-Net for fine segmentation of small organs. All the three networks share feature volumes and are jointly optimized. In addition, a weighted focal loss [5] combined with generalized dice loss is also used to better deal with the severe sample imbalance problem. The proposed method was evaluated on two datasets: a self-collected HaN dataset with 50 CT scans, and MICCAI 2015 Head and Neck Auto Segmentation Challenge dataset. Our proposed algorithm outperforms state-of-the-art methods on HaN organ segmentation with a large margin.

2 Method

The overall structure of the proposed FocusNet is illustrated in Fig. 1. The FocusNet first segments large organs with main segmentation network (S-Net) and localizes the small-organ center locations with Small-Organ Localization branch (SOL-Net). Multi-scale features and high resolution feature are ROI-pooled for small organs to generate small-organ label maps. Therefore, the network decouples localization and segmentation of small organs, and solves the sample imbalance problem, which makes the segmentation of small organs much more easier. The results of large and small organs are then fused to generate the final segmentation label maps.

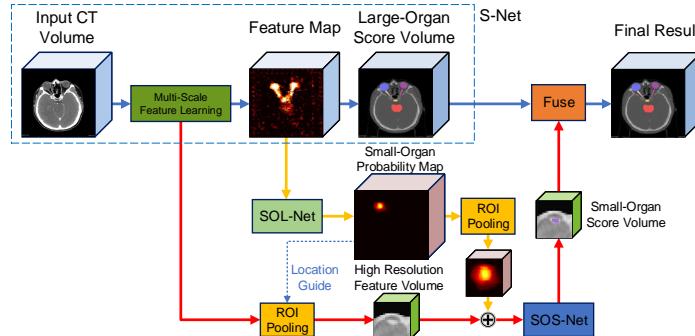


Fig. 1: Overall network structure of our proposed FocusNet.

Main Segmentation Network (S-Net). U-Net is a commonly used 2D CNN, many studies show that its 3D variants have better representation capability in 3D images as they can better capture volumetric contextual information. However, vanilla 3D U-Net has poor performance in OAR segmentation, we address this problem from two aspects. First, U-Net embeds high-resolution information into feature maps by four down-sampling operations, while the decoder reconstructs spatial resolution and obtain dense predictions by deconvolution

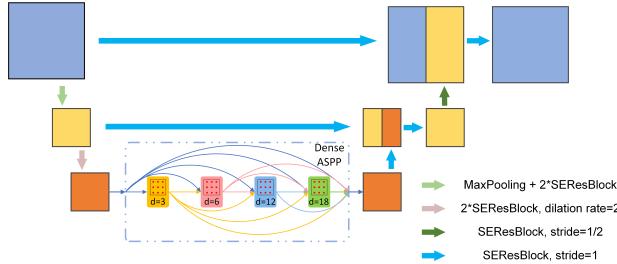


Fig. 2: Structure of multi-scale feature learning in S-Net. The squares denote feature volumes, and the side length indicates the feature volume size. The d in dense ASPP denotes the dilation rate of each convolution kernel.

operations. However, too many times of down-sampling leads to the loss of high-resolution information, which will have catastrophic effects on the small organs that only occupy a few voxels. The shortcut connection between encoder and decoder can only slightly mitigate this problem. Second, UNet can only capture features in fixed scales by downsampling, which limits its representation capability.

The S-Net is elaborately designed to solve the problem mentioned above. As shown in Fig. 2, S-Net has a strong backbone, which is a variant of 3D U-Net with residual connections. Squeeze-and-excitation modules [3] are used for channel-wise attention. To solve the first problem, the S-Net only performs down-sampling once. However, such structure has a disadvantage that the receptive field of convolution kernel is limited, which makes it difficult to integrate global image patterns to learn high-level features. Therefore, dilated convolution and densely connected atrous spatial pyramid pooling (DenseASPP) [10] are utilized in our S-Net. DenseASPP can be seen as the combination of the serial connected and parallel connected counterpart, which has the ability of combining arbitrary scales of features through adjusting dilation rate, and better feature reuse.

Small-Organ Localization Network (SOL-Net). We then mimic the way oncologists annotate the small organs, we propose to design an SOL-Net to first localize the center locations of small organs. As shown in Fig. 1, the feature volumes from the last layer of decoder of our S-Net is used as the input of SOL-Net. The training targets are the small-organ center location heat maps, which are created as 3D Gaussian distributions located at the center locations and each small organ has a separate map. The SOL-Net is trained to predict such location maps with a Mean-Square-Error loss. The SOL-Net consists of 2 Squeeze-and-Excitation Residual Blocks (SEResBlock) and a final $1 \times 1 \times 1$ convolution layer with sigmoid layer to output the small-organ location probability maps.

Small-Organ Segmentation Network (SOS-Net). Given the center locations of the small organs from the SOL-Net outputs, we further improve the segmentation accuracy by focusing on the surrounding regions of small organs. Specifically, we first identify the voxel with the highest location probability value

from SOL-Net as the small-organ center location, and ROI-pool a 3D feature volume around it. An SOS-Net, which contains 2 SEResBlock and a $1 \times 1 \times 1$ convolution layer, is created for each small organ for outputting the binary segmentation maps. The side-length of the ROI is determined as a fixed value, which is three times of the average diameter of the small organ. In this way, the unbalanced negative and positive sample problem is solved.

In order to make the best use of all available information, multi-scale feature volumes from the last layer of the S-Net decoder, raw image, and high resolution feature volumes from the first layer of the S-Net encoder are ROI-pooled from the small-organ ROI and concatenated together as the input of SOS-Net. Furthermore, the small-organ location probability heatmap is also concatenated as the spatial location prior. Intuitively, the multi-scale feature volumes from S-Net already encode small organs' segmentation results and the high-resolution feature volumes can help refine the segmentation results. Finally, we integrate the small-organ segmentation results and the large-organ segmentation results to obtain the final prediction for all organs.

Loss function. In our task, the ratio between the background and the smallest organ can reach nearly $10^5 : 1$, which makes the loss dominated by large numbers of easy background samples. We use a focal loss for multi-class classification to solve this problem, and further use weights to balance between organs,

$$L_{Focal} = \sum_{t=0}^C -\alpha_t(1 - p_t)^\gamma \log(p_t), \quad (1)$$

where C is the number of categories, p_t is the probability of class t , α_t is the weight of each organ, which is inversely proportional to each organ's average size. $(1 - p_t)^\gamma$ is the modulating factor which weights less on easy samples (voxels), whose prediction confidence p_t is close to 1. In our experiment, γ is set as 2.

Generalized dice loss is another loss function that directly optimize for the evaluation metrics. We take generalized dice loss as the following form:

$$L_{Dice} = \sum_{t=0}^C (1 - 2 \frac{\sum y_t p_t}{\sum y_t + \sum p_t}), \quad (2)$$

where the y_t and p_t are the label and probability of class t . In our experiment, the combination of focal loss and dice loss results in best segmentation accuracy, the total loss is as follow:

$$L_{total} = L_{Focal} + L_{Dice}, \quad (3)$$

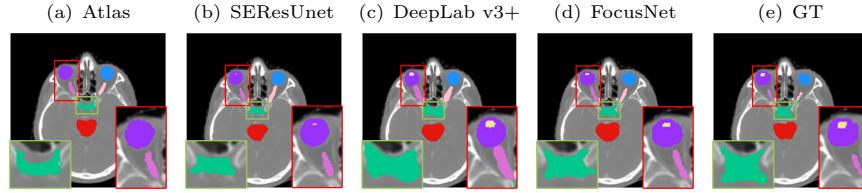


Fig. 3: Some segmentation examples of comparative experiment, GT is ground truth.

3 Experiments

The proposed FocusNet was evaluated on two datasets of HaN CT images. The first dataset, denoted as our dataset, consists of 50 collected samples from hospitals, and 18 organs were delineated manually by doctors. We randomly shuffled our dataset and selected 40 samples for training and 10 samples for testing. For fair comparison with state-of-art methods in OARs segmentation, we further evaluated FocusNet on a public dataset. The *MICCAI Head and Neck Auto Segmentation Challenge 2015* dataset, denoted as MICCAI'15 dataset, consists of 38 samples for training and 10 samples for testing, and has 9 organ annotations. Two evaluation metrics are used in this study: Dice score coefficient (DSC) and 95% Hausdorff Distance (95HD).

Training is performed in three stages. We first train the S-Net, and then train the SOL-Net while fixing the trained parameters of S-Net. The SOS-Net is trained at last because it needs the resulting feature volumes and the location probability maps from the SOL-Net. At last, we end-to-end finetune the whole network for joint optimization.

3.1 Experiments on our collected dataset

The average number of voxels of each organ is shown in supplementary material. The sample unbalance problem is severe, where large organs occupies over tens of thousands of voxels while small organs only occupies hundreds or even tens of voxels. Organs with voxels fewer than 1000 are considered as small organs, which results in 10 small organs in total.

Comparison with other methods. We compare our proposed method with 3D U-Net [8], a 3D variant of DeepLab-v3+ [2], and an atlas-based method [1]. U-Net [8] is popular in 2D medical image segmentation. We modify the U-Net (denoted as SERes U-Net) to replace original convolution with 3D Squeeze-and-Excitation Residual Blocks (SEResBlock) to increase its segmentation accuracy. DeepLab-v3+ [2] is successful for 2D semantic segmentation, we extend their network structure to 3D for volumetric segmentation. All the compared deep learning based methods were randomly initialized and trained using the same loss function as our proposed FocusNet. For the atlas-based method, Symmetric Normalization (SyN) [1], its implementation in ANTs software package is used to generate template image (atlas) and the template label from training set. When

Organs	Atlas-based	SERes U-Net	DeepLab	Ours	Method	AVG
Brain Stem	80.6±1.7	79.2±1.8	84.0±1.2	85.8±1.4	(All Organs)	
Eye L	78.7±3.5	86.1±2.6	86.7±2.2	87.6±0.8	SERes U-Net CE Loss	60.2
Eye R	83.9±3.4	88.2±1.6	90.5±0.6	91.2±0.9	SERes U-Net 1-ds	63.5
Oral Cavity	75.2±13.1	81.0±10.6	77.7±16.2	79.2±3.5	SERes U-Net 2-ds	63.3
Parotid L	80.6±4.3	75.1±6.4	78.8±4.0	77.2±4.2	SERes U-Net 3-ds	61.9
Parotid R	80.7±3.2	79.8±1.7	81.7±3.0	80.0±3.1	SERes U-Net+FL	64.4
Larynx	64.0±29.1	62.8±28.9	67.7±29.6	65.8±29.3	SERes U-Net+FL+DL	69.2
Spinal Cord	82.4±2.1	85.9±2.0	85.2±2.2	87.4±1.9	S-Net	70.6
Lens L	24.0±8.3	66.8±4.2	59.2±8.4	80.8±4.7	FocusNet	72.5
Lens R	26.9±1.5	57.1±6.2	64.3±6.6	79.0±6.4	Fat U-Net	70.5
Opt. Ner. L	47.7±10.6	56.0±3.7	52.7±9.9	63.9±3.9	(Small Organs Only)	
Opt. Ner. R	48.5±6.0	49.3±9.5	57.1±21.4	61.7±12.1	S-Net	61.5
Opt. Chiasm	54.8±9.0	54.0±7.6	55.6±11.3	63.8±11.4	FocusNet ROI Size 2x	63.9
Pituitary	44.6±12.0	67.6±12.7	78.1±10.8	76.9±7.2	FocusNet ROI Size 3x	65.0
Mid. Ear L	56.4±9.7	55.2±15.6	51.9±25.3	56.7±16.7	FocusNet ROI Size 5x	64.5
Mid. Ear R	56.2±14.5	47.4±13.4	46.6±21.8	52.2±20.9		
T.M.J. L	46.9±14.1	56.5±8.1	64.7±3.9	58.4±7.3		
T.M.J. R	50.3±18.8	55.1±12.1	66.1±8.4	57.2±5.6		
Average	62.3	66.5	69.4	72.5		

Table 1: Result of different approaches on the 18 HaN organs, measured by DSC. Below the dash line are small organs.

given a CT to be segmented, the optimal transformation between the atlas and target was obtained by registration, and then this transformation is applied to the template label.

Comparative results are shown in Table 1. For large organs, deep learning methods have slightly better results, but for small organs, deep learning methods have large advantages than atlas-based method SyN. It is because that small organs have more complex anatomical structures, atlas-based method has limited capability of dealing with complicated and diverse anatomy variations. Among deep learning based methods, our proposed FocusNet outperforms other methods by large margins. This is because it has specific mechanisms for handling small organs and could therefore generate more accurate resulting label maps.

For qualitative comparison, as seen in Fig. 3, for the optic nerves, the prediction given by DeepLab-v3+ is much larger than the ground truth, FocusNet gives a precise segmentation result. For the optic chiasm, the result of other methods are in a mess, it is difficult to see a clear shape, while FocusNet gives the best segmentation with X-like shape. SERes U-Net has poor results in lens, it is because lens only occupy tens of voxel, too much information is lost due to down-sampling.

Ablation Study of FocusNet. The ablation results are shown in Table 2. Our baseline is 3D SERes U-Net, which utilizes the SEResBlock but has 4 down-sampling operations, and adopts the cross-entropy loss. The baseline results in decent segmentation results for large organs while poor for small organs. We then test the baseline, SERes U-Net, with 1, 2 and 3 down-sampling operation respectively, and the one with 1 down-sampling results in highest performance. Utilization of the focal loss (FL) and then combine the dice loss (DL) improve the segmentation accuracy. Combining the DenseASPP module into the network, which is our S-Net, slightly increases the performance. Our proposed FocusNet boosts the final performance by adopting specifically designed small-organ networks.

We increase the number of channels of all layers of S-Net by a ratio, named Fat U-Net, so that it has the same parameters with FocusNet, result shows that without solving the organ imbalance problem, more parameters cannot boost the

Table 2: Ablation study of FocusNet. Above the midline are experiments about network structure and loss function, below are experiments about ROI size of SOS-Net.

performance. We also conducted experiments to find the best ROI size, which is obtained when the side length of the ROI is 3 times than that of the organ.

3.2 Experiments on MICCAI'15 dataset

Organs	MICCAI 2015 [6]	Ren et al. [7]	Wang et al. [9]	Zhu et al. [11]	S-Net	FocusNet
Extra Data	x	x	x	✓	x	x
Brain Stem	88.0	N/A	90.3±4	86.7±2	86.8±2.9	87.5±2.6
Chiasm	55.7	58±17	N/A	53.2±15	57.4±25.1	59.6±18.1
Mandible	93.0	N/A	94.4±1	92.5±2	92.5±1.5	93.5±1.9
Opt. Ner. L	64.4	72±8	N/A	72.1±6	71.8±6.9	73.5±9.6
Opt. Ner. R	63.9	70±9	N/A	70.6±10	71.9±9.9	74.4±7.2
Parotid L	82.7	N/A	82.3±6	88.1±2	86.1±2.6	86.3±3.6
Parotid R	81.4	N/A	82.9±6	87.4±4	87.8±4.6	87.9±3.1
Subman. L	72.3	N/A	N/A	81.4±4	79.4±9.8	79.8±8.1
Subman. R	72.3	N/A	N/A	81.3±4	79.7±4.5	80.1±6.1
Average	74.9	N/A	N/A	79.25	79.24	80.29

Table 3: Dice overlap coefficient of independent test set in the MICCAI 2015 dataset.

We also test our FocusNet on MICCAI 2015 Head and Neck dataset. All the settings of the FocusNet are the same as those used in experiments on our collected dataset, except the number of small organ branch is set as 3 since only 3 organs meet our previous definition of small organs: left and right optic nerve and optic chiasm. Visualization can be seen in supplementary material.

The evaluation results are shown in Table 3 and Table 5. We compared the highest score from the top four teams in MICCAI 2015 challenge [6]. For the result of Zhu et al. [11], it should be noted that they used 38 samples provided by the MICCAI 2015 Challenge combined with additional 216 samples for training.

Our backbone S-Net achieves state-of-the-art performance. It reaches comparable performance in Dice score to Zhu et al. [11] but with only 15% of training data, which shows that S-Net has stronger feature representation capability. Moreover, S-Net has much better result in terms of 95HD, because outliers are alleviated by enlarging the receptive field. After adding SOL-Net and SOS-Net, FocusNet achieves further improvement, especially for the three small organs.

Wang et al. [9] proposed a vertex regression-based method, which has good performance in brain stem and mandible, however, has relatively poorer performance in parotid glands, and they did not provide results of other organs. Compared with the registration-based region proposal and patch-based segmentation method used by Ren et al. [7], our approach integrates the localization and segmentation of small organs into a unified deep learning framework, which is much faster, has no redundant computation, and results in better performance.

4 Conclusion

We proposed an end-to-end deep neural network, FocusNet, which outperforms state-of-the-art methods on segmentation of imbalanced OARs in HaN CT images. By reducing the number of down-sampling and utilizing multi-scale features learned by DenseASPP, our S-Net can guarantee the accuracy of the segmentation of large organs. Trained for predicting small-organ center location maps,

Method	Avg 95HD
MICCAI 2015[6]	4.14
Zhu et al.[11]	6.72
FocusNet	2.62

Table 4: Average 95% Hausdorff distance of test set of MICCAI 2015 dataset. The complete table can be seen in supplementary material.

SOL-Net can generate accurate small-organ central locations. SOS-Net can solve the sample imbalance problem, and high-resolution feature volumes can be utilized to accurately segment small organs and thus can further boost the performance. A weighted focal loss combined with dice loss is introduced to mitigate the sample imbalance problem. Extensive experiments on real patients' data and the MICCAI 2015 dataset show the effectiveness of our proposed FocusNet.

Acknowledgements. This work has been supported in part by the General Research Fund through the Research Grants Council of Hong Kong under Grants CUHK14208417 and CUHK14239816, in part by the Hong Kong Innovation and Technology Support Programme (No. ITS/312/18FX), in part by National Key R&D Program of China (2017YFC0113201) and Zhejiang Key R&D Program (2019C03003).

References

1. Avants, B.B., et al.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis* 12(1), 26–41 (2008)
2. Chen, L.C., et al.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Computer Vision – ECCV 2018*. Springer (2018)
3. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks
4. Ibragimov, B., Xing, L.: Segmentation of organs-at-risks in head and neck ct images using convolutional neural networks. *Medical physics* 44(2), 547–557 (2017)
5. Lin, T.Y., et al.: Focal loss for dense object detection. In: *Proceedings of the IEEE Conference on CVPR*. pp. 2980–2988 (2017)
6. Raudaschl, P.F., et al.: Eval. of seg. methods on han ct: Auto-seg. challenge 2015. *Medical physics* 44(5), 2020–2036 (2017)
7. Ren, X., et al.: Interleaved 3D-CNN s for joint segmentation of small-volume structures in head and neck ct images. *Medical physics* 45(5), 2063–2075 (2018)
8. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image seg. In: *International Conf. on MICCAI*. pp. 234–241. Springer (2015)
9. Wang, Z., Wei, L., Wang, L., Gao, Y., Chen, W., Shen, D.: Hierarchical vertex regression-based segmentation of head and neck ct images for radiotherapy planning. *IEEE Transactions on Image Processing* 27(2), 923–937 (2018)
10. Yang, M., Yu, K., Zhang, C., Li, Z., Yang, K.: Denseaspp for semantic segmentation in street scenes. In: *Proceedings of the IEEE Conf. on CVPR*. pp. 3684–3692 (2018)
11. Zhu, W., et al.: Anatomynet: Deep 3d squeeze-and-excitation u-nets for fast and fully automated whole-volume anatomical segmentation. *Medical Physics* 46(2), 576–589 (2019)

Supplementary

In the supplementary material, we report a) the statistical results of the number of voxels of each organ on our collected dataset, b) some qualitative results of 95HD compared with state-of-the-art methods in OAR segmentation in MICCAI 2015 dataset, c) some visualization of segmentation results of MICCAI 2015 dataset.

Experiments on our collected dataset

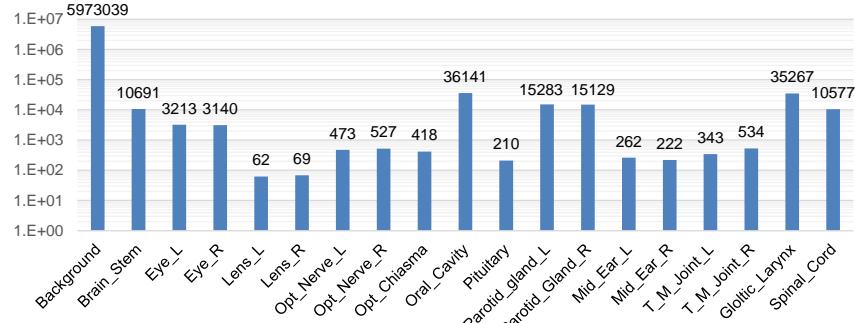


Fig. 4: Voxel number of 18 OARs in our collected dataset. The vertical axis is log scale. The sample imbalance problem is severe, where large organs occupy over tens of thousands of voxels while small organs only occupy tens of voxels.

Experiments on MICCAI'15 dataset

Organs	MICCAI 2015[6]	Ren et al. [7]	Zhu et al. [11]	FocusNet
Brain Stem	4.59	N/A	6.42±2.4	2.14±0.6
Chiasm	2.78	2.81±1.6	5.76±2.5	3.16±1.3
Mandible	1.97	N/A	6.28±2.2	1.18±0.3
Opt. Ner. L	2.76	2.33±0.8	4.85±2.3	3.76±2.9
Opt. Ner. R	3.15	2.13±1.0	4.77±4.3	2.65±1.5
Parotid L	5.11	N/A	9.31±3.3	2.52±1.0
Parotid R	6.13	N/A	10.08±5.1	2.07±0.8
Subman. L	5.35	N/A	7.01±4.4	2.67±1.3
Subman. R	5.42	N/A	6.02±1.8	3.41±1.4
Average	4.14	N/A	6.72	2.62

Table 5: 95% Hausdorff distance of independent test set in the MICCAI 2015 dataset. Our approach has better 95HD because the enlarged receptive field can alleviate outliers.

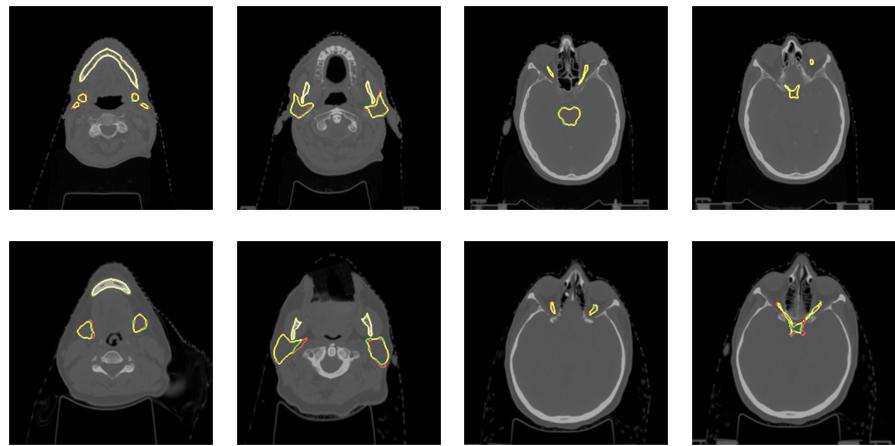


Fig. 5: Segmentation results on sample 0522c0576 (first row), and 0522c0661 (second row) by our FocusNet. The green lines are the contour of predicted organ boundaries, red lines are the contour of ground truth, and the yellow lines are the overlap.