

Short Paper of Ocean for FLARE21 Challenges

Yongze Guo
Sun Yat-sen University
Guangzhou, China
guoyongze1012@163.com

Yue Li
Sun Yat-sen University
Guangzhou, China
liyue6768@qq.com

Abstract

In the challenge of FLARE21, we need to segment four kinds of abdominal organs in CT images from different medical centers. To satisfy the clinical requirements, the segmentation method should be accurate and efficient. nnUNet is the SOTA method in many medical imaging segmentation tasks. In our experiments, it has achieved the efficiency requirements. Therefore, we focus on generalizing a model on data from different medical centers and we hope its accuracy will not drop a lot. In this study, we proposed a pre-processing method that could transfer CT data from different centers into a common image domain, where body anatomic structure was kept but individual noise was removed. The data released in training stage were divided into two groups according to their source. The 281 data from center2 were used as training group to train our segmentation model and the remaining 80 data from center1 were used as test data to evaluate the adaptation ability of our model in different data center. However, the results did not meet our expectations. Compared with the baseline nnUNet trained in the same training data, the Dice and Jaccard of our model were lower for most of organ. Thus, this idea is not applicable to this problem and needs further improvement.

1. Introduction

One of the main difficulties of organ segmentation task is the model’s generalization ability in data from different medical centers. We always found significant difference when we used a model trained on single-center in the inference of data from another center. To address this problem, we hope to design a scheme which can extract images’ common texture feature and project them into a same sub-space to summarize the common characteristics of these data so that the model’s generalization ability may be improved.

Since nnUNet is popular in medical image segmentation [1], we added our idea on nnUNet framework and took it as our baseline in experiment. The details of our method

is explained as following.

2. Method

2.1. Preprocessing

Pre-processing is the most important part in our study.

The patients from different centers actually have similar anatomic structures. The main differences between these data are from acquisition equipment, manufacturer’s brand and technician’s operation habits. As a result, CT images from different center have different noise. CT images from different phase have different enhanced region. However, no matter for which center or which phase, the patients’ anatomic structures are not changed. This gives us the idea that if we can extract the anatomic structure, data from different center can find a common representation.

We use a sliding window to implement whitening operation in local region in each slice.

$$\hat{p}_i = \frac{p_i - \bar{p}_i}{\sigma_i}$$

where $\bar{p}_i = \sum_{j \in U(i)} w_{ij} p_j$ is the weighted intensity average value of pixel p_i in its neighbor window $U(i)$ and w_{ij} is its corresponding weight. $\sigma_i^2 = \sum_{j \in U(i)} w_{ij} (p_j - \bar{p}_i)^2$ is the weighted variance in window $U(i)$.

In the code, w_{ij} is the ij -th element of a Gaussian kernel with the same size of sliding window, and $\hat{\sigma}_i = \sigma_i + \epsilon$ is used to replace the σ_i to avoid too large value showed in final image ($\epsilon = 0.1$ in our experiment).

2.2. Proposed Method

After preprocessing, the standard nnUNet was used. The default loss function, learning rate, total number of epochs, and other setting of nnUNet were used in this study. We choose the 3D version of nnUNet and the full resolution images were as model inputs. The trainer was nnUNetTrainerV2.

In the training procedure, only the 281 data from center 1 were used. In the inference step, the remaining 80 data

were used as test data. Several metrics including Dice, Jaccard, Precision, and Recall rate were calculated. The results would shown in the next part.

2.3. Post-processing

Since the model is trained on the 281 data for only one time. No any ensembling process need to be implemented. This can save the inference time. As a result, there is no post-processing should be done.

3. Dataset and Evaluation Metrics

3.1. Dataset

- Training data: 281 training cases from data center 1 with segmentation annotations provided by FLARE21.
- Test data: 80 training cases from data center 2.

3.2. Evaluation Metrics

The nnUNet has satisfy the requirement of model efficiency. So we do not evaluate the model efficiency and focus on the metrics of Dice, Jaccard.

4. Implementation Details

4.1. Environments and requirements

This model is implemented on pytorch 1.9 in python 3.9. The cuda version is 11.0. The server has a Intel(R) Xeon(R) Silver 4210 CPU and RTX8000 GPU.

5. Results

5.1. Preprocessing results

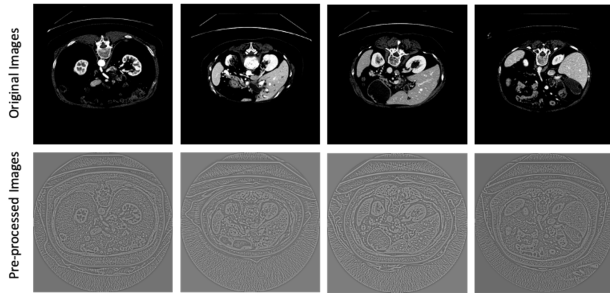


Figure 1. Several examples of our preprocessing result. It can be seen that after local weighted whitening, the intensity information of original images is removed and only the anatomic structure is kept for each case

Figure 1 shows the local weighted whitening method. This step kept the anatomic structure, which is very important for tissue segmentation task.

5.2. Inference results compared with baseline

We calculated the Dice and Jaccard in the 80 test set for both our model and standard nnUNet, which was the baseline. The results are shown in Table 1 and Table 2.

Table 1. The evaluation metric values for our proposed model

Organ	Dice	Jaccard
Liver	97.84 ± 0.46	96.81 ± 1.24
Kidney	96.59 ± 1.56	95.02 ± 2.77
Spleen	97.40 ± 2.22	96.21 ± 3.30
Pancreas	80.12 ± 8.73	73.85 ± 11.45

Table 2. The evaluation metric values for baseline model nnUNet

Organ	Dice	Jaccard
Liver	97.78 ± 0.50	96.71 ± 1.31
Kidney	97.10 ± 1.74	95.75 ± 2.88
Spleen	97.82 ± 0.81	96.78 ± 1.60
Pancreas	81.17 ± 5.88	74.93 ± 9.29

It can be seen that our proposed model was worse than nnUNet. This meant that the idea of whitening was not appropriate for the task. As a result, we submitted the nnUNet trained on 281 data from center 1 as our final model in FLARE21.

6. Discussion and Conclusion

Local weighted whitening method did not show improvement for domain adaptation in organ segmentation tasks. More modifications will be studied in the future.

Acknowledgment

We did not use any pre-trained model and additional datasets.

References

- [1] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, “nnu-net: a self-configuring method for deep learning-based biomedical image segmentation,” *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021. 1