Uncertainty-aware Mean Teacher Framework with Inception and Squeeze-and-Excitation Block for MICCAI FLARE22 Challenge

Hui Meng[1], Haochen Zhao[1], Ziniu Yu[1], Qingfeng Li[1], and Jianwei Niu[1,2]

[1] Research center of Big Data and Computational Intelligence, Hangzhou Innovation Institute of Beihang University, Hangzhou 310051, China.

[2] State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, also with the Beijing Advanced Innovation Center for Big Data and Brain Computing (BDBC), Beihang University, Beijing 100191, China.

Abstract. Semi-supervised learning has attracted extensive attention in the field of medical image analysis. However, as a fundamental task, semi-supervised segmentation has not been investigated sufficiently in the field of multi-organ segmentation from abdominal CT. Therefore, we propose a novel uncertainty-aware mean teacher framework with inception and squeeze-and-excitation block (UMT-ISE). Specifically, the UMT-ISE consists of a teacher model and a student model, in which the student model learns from the teacher model by minimizing segmentation loss and consistency loss. Additionaly, we adopt an uncertainty-aware algorithm to make the student model learn accurate and reliable targets by making full use of uncertainty information. To capture multi-scale features, the inception and squeeze-and-excitation block are incoporated into the UMT-ISE. It is worth noting that abdominal CT of test cases are first extracted before multi-organ segmentation in the inference phase, which significantly improves segmentation accuracy.

Keywords: semi-supervised learning, multi-organ segmentation, uncertainty estimation, multi-scale features

1 Introduction

Accurate segmentation of medical images is essential for many clinical applications, such as disease diagnosis and tumor localization [3]. Nowadays, manual segmentaion results given by radiologists are widely regarded as gold standards. However, manual segmentation is tedious and time consuming. Additionally, manual segmentation heavily depends on radiologists' experience and suffers from intra- and inter-observer variabilities. Therefore, many researchers have developed different automatic segmentation methods [12], which are supposed to assist radiologists to make accurate diagnosis.

For abdominal organ segmentation, most research work focus on single organ segmentation, such as kidney [6] or blood vessels [9]. Compared with single-organ segmentation, multi-organ segmentation faces two major challenges. The first one is that large morphological differences between multiple organs limits accurate segmentation of all organs. The second one is that it's difficult to obtain large dataset with accurate annotations for multi-organ segmentation. Therefore, it is necessary to make full use of unlabeled medical images to improve the multi-organ segmentation accuracy [4].

To utilize unlabeled medical images effectively, we propose a novel uncertainty-aware mean teacher framework with inception and squeeze-and-excitation block (UMT-ISE) for segmenting multiple organs from 3D abdominal CT. The UMT-ISE is constructed based on conventional teacher-student model [2], which consists of a teacher model and a student model. For the same unlabeled data under different perturbations, the segmentation predictions of the teacher model and the student model are constrained to be consistent [13]. Different from the conventional teacher-student model, the UMT-ISE adopts framework of uncertainty-aware mean teacher. The teacher model in the UMT-ISE generates multiple predictions for each target under Monte Carlo sampling and gives uncertainty evaluation. The predictions with high uncertainty are filtered out and the predictions with low uncertainty are retained to compute consistency loss. Based on the design of the uncertainty evaluation, the teacher model tends to generate high-quality predictions and the student model can be constantly optimized. Considering multiple organs have different sizes, the inception and squeeze-and-excitation (ISE) block are incoporated into the UMT-ISE to capture multi-scale features.

2 Method

2.1 Preprocessing

The following preprocessing techniques were utilized in this work:

- Cropping strategy:
  The range of CT scans varies depending on the situation. For example, some patients may have CT scans not only of the abdominal area, but of the entire chest, lower abdomen and even the legs. In some cases, only the abdominal region containing the target organs is present. Therefore, it is necessary to filter out some irrelevant and redundant slices.We train a UA-MT network to perform a rough abdominal segmentation to extract abdominal regions
  For training data with labels, we tailor them according to the range of target organs in annotations. For training data without labels, validation data and test data, we first implement rough segmentation of target organs and then crop CT scans according to the scope of target organs.
- Clipping in the x and y directions:
  Because different images occupy different fields of vision in the x and y directions, some of which occupy very small proportion, part of the redundant parts in the x and y directions should be deleted to make the abdominal area occupy a larger proportion in the whole picture. For images with a

large proportion, organs are often attached to the edge of the image, resulting in missegmentation. For such images, some redundancy should be added in the x and y directions, so as to make the abdominal region in an appropriate range of image proportion.

- Adjusting window level and window width:
  In order to have a good contrast between the organ and the background area, it is of great significance to adjust the window width and window level of the original image. According to the doctor's observation, the window width and window level of the image are respectively adjusted to 40 and 255 in this paper.
- Resampling method for anisotropic data:
  In this paper, the model uses the whole CT image as the network input. In order to match the input size of the model, the size of all and images is uniformly adjusted to 192×192×96.
- Intensity normalization method:
  A z-score normalization is applied based on the mean and standard deviation of the intensity values.
- Data augmentation method:
  Data enhancement operations such as random clipping to 112×112×80 size and horizontal flipping are performed on the training set

2.2 Proposed Method

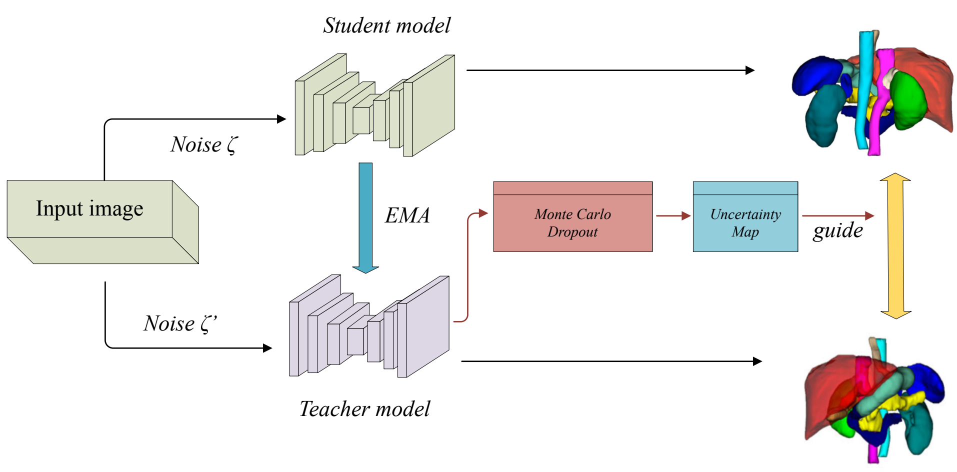

Fig. 1. Network architecture

Strategies to use the unlabelled cases:

The inputs of teacher model and student model are added to the same image with different noises, and the two outputs are constrained by unsupervised loss function calculation

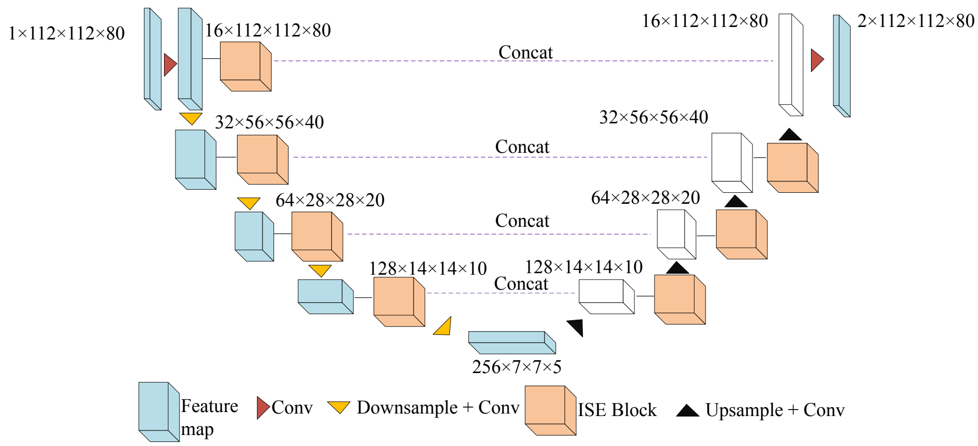

Fig. 2. A V-Net with ISE block

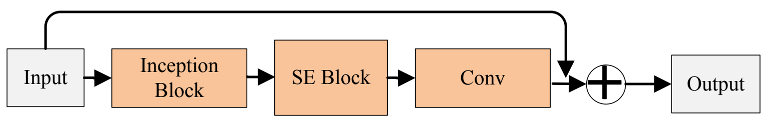

Fig. 3. The detailed architecture of the ISE block.

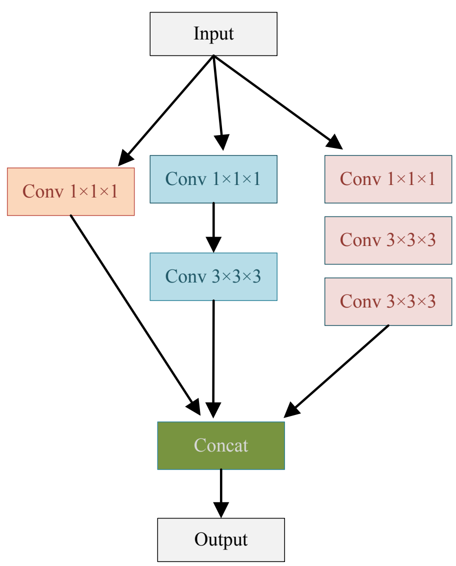

Fig. 4. The detailed architecture of the Inception block.

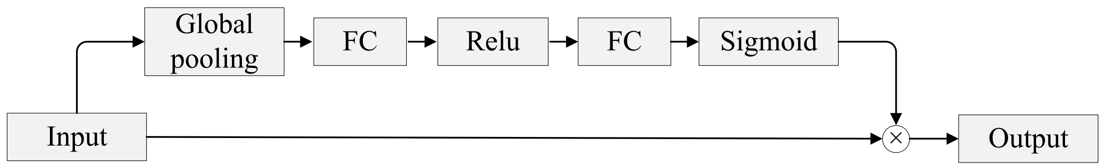

Fig. 5. The detailed architecture of the SE block.

Network architecture details :

The network architecture is shown in Fig.1.

The whole network model is composed of two V-Net models, teacher model and student model, and the two models share the weight. In order to ensure the accuracy of multi-organ segmentation and considering the problem of small organ segmentation, we design to add ISE block in V-Net network. Not only multi-scale problem is considered, but also channel attention is added to improve segmentation accuracy. The network structure is shown in Fig.2.

The structure of the proposed ISE block in Fig.3., which integrates the residual block, Inception block, and a squeeze-and-excitation block (SE block) block.The convolution kernels of different sizes are used in the Inception block to obtaint the receptive fields of different sizes. Then, features from different kernels are fused to get multi scale features to reduce the impact of different resolutions of anisotropic CT images on organ segmentation. The Inception block for this work is shown in Fig.4.

Although abundant features are obtained, a deal of redundant features are collected, which weaken the important and target-related features and can reduce the discriminability of the network. Thus, a squeeze-and-excitation block (SE block) is employed here to recalibrate the importance of the multi scale features obtained by the Inception block. The specific components and structure of the SE block is illustrated in Fig.5. In the SE block, a global average pooling layer is used to aggregate the global information, which is followed by two full connection layers to capture the channel-wise relationships. Then, the features obtained by the Inception block is recalibrated by the channel-wise relationships through point-wise multiply operation.

Loss function:

We use the summation between Dice loss and cross entropy loss because compound loss functions have been proved to be robust in various medical image segmentation tasks.

2.3 Post-processing

The post-processing operations used in the work include removing small connected areas and filling the holes to reduce false positive islands. For the CT whose z-axis direction is restored to spacing 1 and greater than 800, we believe that it may contain the whole sequence from head to foot. Even the crude segmentation model is difficult to effectively segment the abdominal region. Therefore, we divide the whole sequence into three equal parts, predict the three parts separately, and finally put the three parts together. Reserving the maximum number of z-direction pieces containing the foreground region can effectively reduce missegmentation

3 Experiments

3.1 Dataset and evaluation measures

The FLARE2022 dataset is curated from more than 20 medical groups under the license permission, including MSD [11], KiTS [7,8], AbdomenCT-1K [10], and TCIA [1]. The training set includes 50 labelled CT scans with pancreas disease and 2000 unlabelled CT scans with liver, kidney, spleen, or pancreas diseases. The validation set includes 50 CT scans with liver, kidney, spleen, or pancreas diseases. The testing set includes 200 CT scans where 100 cases has liver, kidney, spleen, or pancreas diseases and the other 100 cases has uterine corpus endometrial, urothelial bladder, stomach, sarcomas, or ovarian diseases. All the CT scans only have image information and the center information is not available.

The evaluation measures consist of two accuracy measures: Dice Similarity Coefficient (DSC) and Normalized Surface Dice (NSD), and three running efficiency measures: running time, area under GPU memory-time curve, and area under CPU utilization-time curve. All measures will be used to compute the ranking. Moreover, the GPU memory consumption has a 2 GB tolerance.

3.2 Implementation details

Environment settings The environments and requirements are presented in Table 1.

Table 1. Environments and requirements.

Windows/Ubuntu version	Windows 10
CPU	Intel(R) Core(TM) i9-9900K CPU @ 3.60GHz
RAM	16×2GB;
GPU (number and type)	NVIDIA Tesla V100 GPU
CUDA version	11.1
Programming language	Python 3.6
Deep learning framework	Pytorch (Torch 1.7.0, torchvision 0.8.0)
Specification of dependencies	None
(Optional) Link to code	

Training protocols

Table 2. Training protocols.

Network initialization	"he" normal initialization
Batch size	16
Patch size	192×192×96
Total epochs	2000
Optimizer	SGD with nesterov momentum ($\mu = 0.99$)
Initial learning rate (lr)	0.05
Lr decay schedule	halved by 1000 epochs
Training time	48 hours
Number of model parameters	9.44M[3]
Number of flops	41.40G[4]
CO_2eq	1 Kg[5]

4 Results and discussion

4.1 Quantitative results on validation set

Table 3 and Table 4 illustrates the results on the validation cases.

Table 3. Quantitative results on validation set(with unlabel data).

Organ	DSC (%)
Mean DSC	0.7458
Liver	0.9547
RK	0.8500
Spleen	0.8885
Pancreas	0.7143
Aorta	0.8497
IVC	0.7650
RAG	0.6116
LAG	0.5108
Gallbladder	0.6429
Esophagus	0.6825
Stomach	0.8169
Duodenum	0.5471
LK	0.8608

4.2 Segmentation efficiency results

Fig.6. and Fig.7. show two examples with good segmentation results and two examples with bad segmentation examples. We think there are two reasons for

Table 4. Quantitative results on validation set(only label data).

Organ	DSC (%)
Mean DSC	0.7193
Liver	0.9221
RK	0.8280
Spleen	0.8118
Pancreas	0.7148
Aorta	0.8020
IVC	0.7331
RAG	0.6158
LAG	0.5473.
Gallbladder	0.5677
Esophagus	0.6548
Stomach	0.7629
Duodenum	0.5748
LK	0.8164

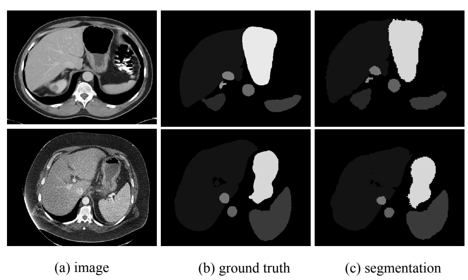

Fig. 6. Well-segmented examples from validation sets

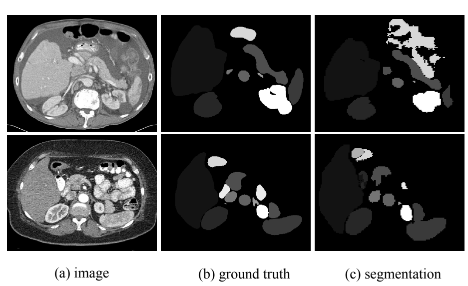

Fig. 7. Challenging examples from validation sets

this problem. One is that the image has a large void, which leads to poor segmentation accuracy, the other is that the model has a poor segmentation effect on small organs.

5 Conclusion

This method has better results for large organ segmentation, smaller number of model parameters and faster operation. In addition, our method can be tested in CPU, which is more convenient to complete some clinical tasks. However, the method in this paper still has some limitations. For some small organs, their shapes and positions are easily affected by tumors and edema, so the segmentation results are not good. Similarly, in the case of CT with too many sections, such as head to foot, the segmentation effect is poor and it is difficult to extract abdominal organs. How to make the segmentation method more robust is still a problem worthy of further discussion.

6 Acknowledgment

The authors of this paper declare that the segmentation method they implemented for participation in the FLARE22 challenge didn't use any pre-trained models or additional datasets other than those provided by the organizers.

References

1. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., et al.: The cancer imaging archive (tcia): maintaining and operating a public information repository. Journal of Digital Imaging **26**(6), 1045–1057 (2013) 6
2. Cui, W., Liu, Y., Li, Y., Guo, M., Li, Y., Li, X., Wang, T., Zeng, X., Ye, C.: Semi-supervised brain lesion segmentation with an adapted mean teacher model. In: International Conference on Information Processing in Medical Imaging. pp. 554–565. Springer (2019) 2
3. Fu, Y., Mazur, T.R., Wu, X., Liu, S., Chang, X., Lu, Y., Li, H.H., Kim, H., Roach, M.C., Henke, L., et al.: A novel mri segmentation method using cnn-based correction network for mri-guided adaptive radiotherapy. Medical physics **45**(11), 5129–5137 (2018) 1
4. Ganaye, P.A., Sdika, M., Benoit-Cattin, H.: Semi-supervised learning for segmentation under semantic constraint. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 595–602. Springer (2018) 2
5. Hall, E.J.: Intensity-modulated radiation therapy, protons, and the risk of second cancers. International Journal of Radiation Oncology* Biology* Physics **65**(1), 1–7 (2006)
6. Heimann, T., Van Ginneken, B., Styner, M.A., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., et al.: Comparison and evaluation of methods for liver segmentation from ct datasets. IEEE transactions on medical imaging **28**(8), 1251–1265 (2009) 1

7. Heller, N., Isensee, F., Maier-Hein, K.H., Hou, X., Xie, C., Li, F., Nan, Y., Mu, G., Lin, Z., Han, M., et al.: The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. Medical Image Analysis **67**, 101821 (2021) 6
8. Heller, N., McSweeney, S., Peterson, M.T., Peterson, S., Rickman, J., Stai, B., Tejpaul, R., Oestreich, M., Blake, P., Rosenberg, J., et al.: An international challenge to use artificial intelligence to define the state-of-the-art in kidney and kidney tumor segmentation in ct imaging. American Society of Clinical Oncology **38**(6), 626–626 (2020) 6
9. Kirbas, C., Quek, F.: A review of vessel extraction techniques and algorithms. ACM Computing Surveys (CSUR) **36**(2), 81–121 (2004) 1
10. Ma, J., Zhang, Y., Gu, S., Zhu, C., Ge, C., Zhang, Y., An, X., Wang, C., Wang, Q., Liu, X., Cao, S., Zhang, Q., Liu, S., Wang, Y., Li, Y., He, J., Yang, X.: Abdomenct-1k: Is abdominal organ segmentation a solved problem? IEEE Transactions on Pattern Analysis and Machine Intelligence (2021). https://doi.org/10.1109/TPAMI.2021.3100536 6
11. Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., Van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., et al.: A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063 (2019) 6
12. Wang, Y., Zhou, Y., Shen, W., Park, S., Fishman, E.K., Yuille, A.L.: Abdominal multi-organ segmentation with organ-attention networks and statistical fusion. Medical image analysis **55**, 88–102 (2019) 1
13. Yu, L., Wang, S., Li, X., Fu, C.W., Heng, P.A.: Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 605–613. Springer (2019) 2