Cycle-consistent adversarial network with polyphase U-Nets for liver lesion segmentation

Boah Kim
Dept. of Bio and Brain Engineering
KAIST, Daejeon, Republic of Korea
kba0111@kaist.ac.kr

Jong Chul Ye
Dept. of Bio and Brain Engineering
KAIST, Daejeon, Republic of Korea
jong.ye@kaist.ac.kr

Abstract

We proposed a novel deep learning algorithm for liver lesion segmentation using a cycle-consistent generative adversarial network (cycleGAN) architecture. In order to overcome the mode collapsing phenomenon from many-to-one mapping nature of segmentation, our method discovers relationships between the computed tomography (CT) images and segmentation-augmented CT images through a cyclical constraint. Moreover, to retain the accurate boundary information, we employ an improved U-Net architecture called the polyphase U-Net as a generator, inspired by the recent theory of deep convolutional framelets. The performance improvement by the proposed method was evaluated on the Liver Tumor Segmentation Challenge 2017 datasets.

1 Introduction

Liver cancer accounts for a high proportion of cancer types with high mortality rate. Radiologists typically diagnose liver cancer using computed tomography (CT) images, from which they estimate the extent of tumor from lesion segmentation to plan surgery or radiotherapy. Accordingly, lesion segmentation is an important part for tumor study. Recently, U-Net [1] has been one of the main workhorses for image segmentation area due to its capability to learn image hierarchy of increasingly complex features. However, U-Net has a limitation in that segmented boundaries tend to lose details. Moreover, the mapping from the original image to a segmentation is a many-to-one mapping, so it often exhibits instability during the training due to the mode-collapsing nature of the mapping. To address this problem, this paper proposes a systematic approach by synergistically combining two recent advances in this field. Specifically, based on the observation that radiologists often compare the segmented object data and the corresponding CT data side-by-side to overcome the ambiguity of the mapping, we propose a cycle-consistent loss to overcome the mode-collapsing behavior [2]. In addition, we propose a novel modification of U-Net, called *polyphase U-Net*, as the basic generator backbone of the network. We show that the polyphase U-Net satisfies the *frame condition* as required by the recent theory of deep convolutional framelets [3]; thus, it retains the accurate boundary details. By applying the proposed network in a successive manner (i.e. the liver segmentation followed by the lesion segmentation), the proposed method enhances segmentation accuracy for tiny tumor.

2 Main approach

The proposed method is based on cycleGAN [2], which has two generators and two discriminators as shown in Figure 1. One of generators, G_{AB}, generates a segmented mask on top of an input CT image, while the other, G_{BA}, generates a CT image from the segmented mask augmented CT image. The output of G_{AB} has three channels: two channels for cross entropy loss with segmentation ground-truth images and the other channel for L2 loss with object segmentation augmented CT

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

images. On the other hand, D_A discriminates images generated from G_{AB} and label images, X_B, while D_B distinguishes images generated from G_{BA} and CT images, X_A.

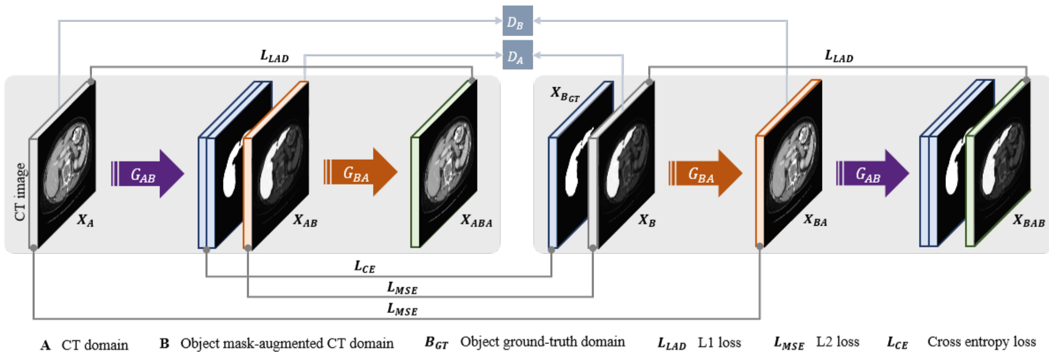

Figure 1: The flowchart of our modified cycleGAN model for segmentation.

As a generator for our cycle GAN, we propose the polyphase U-Net shown in Figure 2, which modifies the pooling and unpooling layers of the U-Net using the polyphase decomposition. Specifically, rather than using average or max pooling, the four neighbour pixels at the input images are decomposed into the four channel data with reduced size at the pooling layer, whereas the four reduced size channels are grouped together to an enlarged single channel at the unpooling layer. This polyphase decomposition satisfies the *frame condition* that is required to retain the high frequency components of the input signals [3]. Moreover, the polyphase decomposition does not change the polarity of the signal so that the ReLU nonlinearity can commute with the pooling layer. We found that these two properties are important to improve accuracy of the U-Net for segmentation applications.

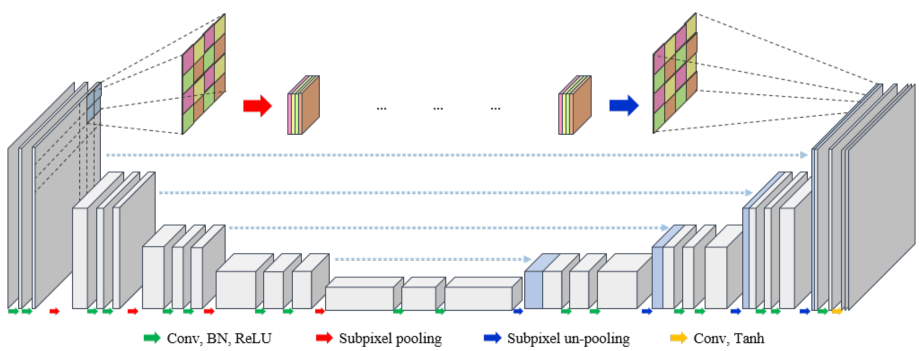

Figure 2: The architecture of subpixel U-Net.

3 Experiments and results

The proposed method was trained and evaluated on the Liver Tumor Segmentation Challenge (LiTS) 2017 datasets which provides 200 contrast-enhanced 3D abdominal CT scans. We trained two modified cycleGAN model with 73 scans and validated with 9 scans among the 130 CT scans provided with ground-truth of liver and liver lesions segmentation. One is to segment liver from corresponding CT scan, and the other is for liver lesion segmentation from the liver so that the second model segments liver lesion only in the liver. To train the model, we down-sampled each 512×512 2D slices to 256×256. We augmented datasets using random flipping, rotation, shift, and elastic deformation to improve segmentation performance without overfitting. Training the network was implemented in Python using pyTorch library.

We evaluated the network performance for validation cases using dice, precision, and recall score. As shown in Table 1, the proposed method outperformed with a dice score of 89% and 46% for each liver and lesion, which were higher than U-Net based cycleGAN. In addition, when we tested the proposed network on 70 test datasets from LiTS 2017, the proposed network scored a dice per case of

88.90% and 43.20% for each liver and lesion, while the conventional method scored a dice per case of 82.50% and 39.80% for each of them. Figure 3 shows results of liver and lesion segmentation and demonstrates that liver and lesion boundary are more accurately detected from the proposed method compared to the conventional method.

Table 1: Liver and liver lesion segmentation performance on 9 validation cases

Generator	Liver			Lesion		
	Dice	Precision	Recall	Dice	Precision	Recall
U-Net + Cycle GAN	0.8470	0.8356	0.8878	0.3924	0.3959	0.4737
polyphase U-Net +Cycle GAN	0.8907	0.8623	0.9414	0.4640	0.4803	0.5017

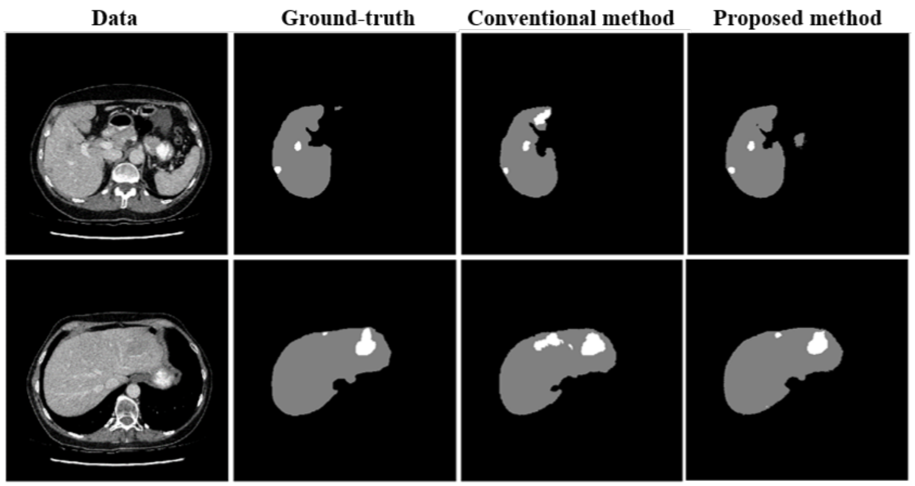

Figure 3: Segmentation results of the validation dataset; grey is liver and white is lesion.

4 Discussion and conclusion

The proposed method found the relationships between CT data and liver lesion segmentation-augmented CT data by comparing each domain as if radiologists conventionally performed. In addition, the polyphase U-Net satisfies the frame condition, so it does not lose high frequency information in contrast to the U-Net. By synergistically combining the two, our proposed method was able to perform liver and lesion segmentation in more detail.

References

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234-241, Springer, Cham, 2015.

[2] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2223-2232, 2017.

[3] JC. Ye, Y. Han, and E, Cha, "Deep convolutional framelets: a general deep learning framework for inverse problems," SIAM Journal on Imaging Sciences (in press), 2018.