

# Progressively Training an Enhanced U-Net Model for Segmentation of Kidney Tumors

XueJian HE<sup>1</sup>, Ping Shun Leung, Lu Wang

The Hong Kong Applied Science and Technology Research Institute Company Limited

**Abstract:** An enhanced U-Net model with multi-scale inputs and deep supervision are adopted for Kidney tumor segmentation. Focal Tversky Loss is used to train the model, in order to improve the model performance of detecting small tumors. Progressive training is proposed for facilitating model converge. A simple post-processing method is used to remove segmentation noises. The preliminary results indicate that the proposed model can segment the normal kidney with a satisfactory result; for the tumors with small sizes in low contrast or extreme sizes, there is still a room for improvement.

**Keywords:** segmentation, kidney tumor, deep learning, CT images.

## 1. Methods

An enhanced U-Net model is used to segment kidney CT images. The U-Net model was enhanced by multi-scale inputs and deep supervision as shown in Figure 1, inspired by the work in [1]. The original input images are down-sampled by average pooling with  $2 \times 2$  kernels, so that four inputs with different scales are generated. The rightest side of the network architecture is the deep supervision, which performs convolution operations to four-level intermediate feature maps. It generates four segmentation outputs at different scales. The difference to the model in [1] is that attention gates are also added to the most upper layer.

Focal Tversky Loss (FTL) [1] was adopted to train the model, in order to improve the model performance of detecting small objects, such as tumors. Since there are four outputs, the final

---

<sup>1</sup>jimhe@astri.org

loss is composed by weighted FTL at four levels. The weights are  $\{0.2, 0.2, 0.2, 0.4\}$  from bottom to top levels.

From experiments it was found that the model training with FTL could be instable, which means it is hard for the model to converge. Hence, progressive training as shown in Figure 2 was proposed. In the training loop, the model was progressively fine-tuned with the previously trained weights using randomly selected training dataset. After each training loop was finished, the model was evaluated with testing dataset. The training loop would continue, till there was no performance improvement on the testing dataset.

A simple post-processing method is used to remove noises in the segmented label images. In the process, the kidney and tumor are considered as one tissue, so that it could be assumed that the segmented volumetric image contains two major volumes with respect to the two kidneys. In the post-processing, first it finds the two kidneys including tumors with measuring the isolated object volumes. Then it checks other small objects. If the distances between the small objects and the two kidneys are larger than a threshold, then these small objects are considered as noises. Consequently, they will be removed from the segmentation results.

## **2. Results**

The model was evaluated with 40 subjects which are randomly selected from 90 testing dataset as provided by KiTS19 [2]. The preliminary results are summarized in Table 1. It can be seen the tumor DICE [3] value is low as compared with that of kidney. From the visual observation, it is found that the model misses some small tumors, which causes small even zero DICE values. These small tumors are usually presented with low contrast in the images, which makes the model difficult to distinguish it from normal kidneys. Another possible reason is that it seems

that the testing dataset contains more cases with extremely large tumors as compared with the training dataset.

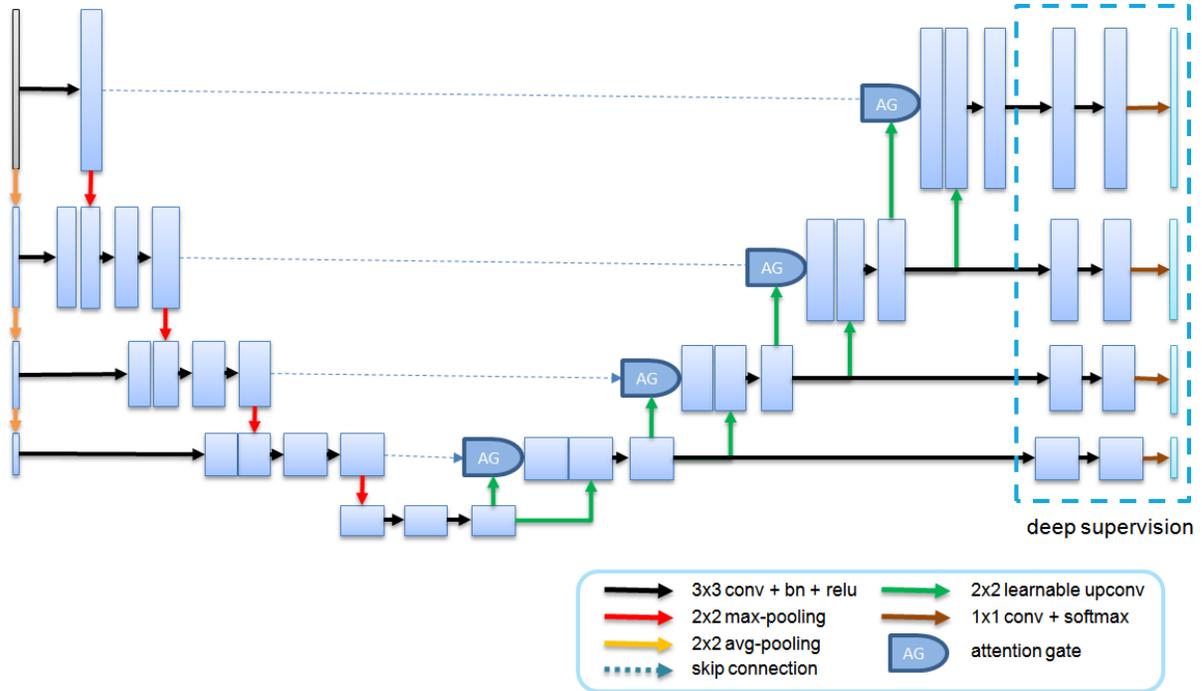


Figure 1. The enhanced U-Net.

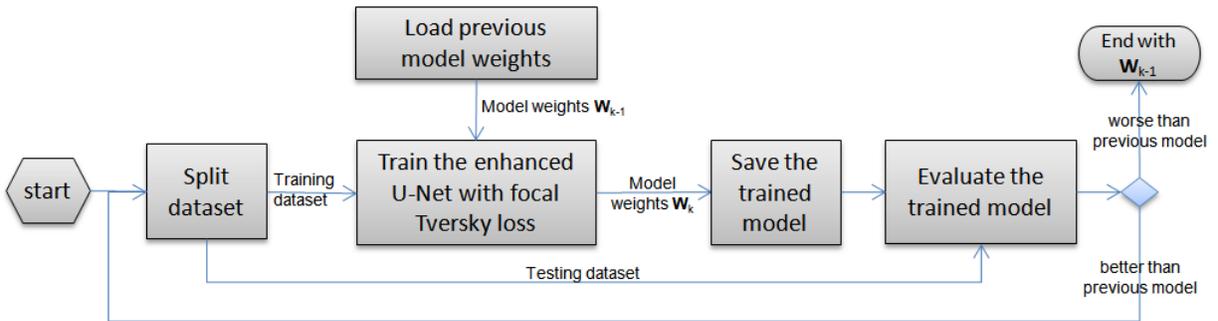


Figure 2. The workflow of progressive training.

Table 1. The preliminary results measured by DICE values.

	DICE
<b>Kidney</b>	0.953
<b>Tumor</b>	0.572
<b>Average</b>	0.763

## References

1. Nabila Abraham, Naimul Mefraz Khan, A Novel Focal Tversky loss function with improved Attention U-Net for lesion segmentation, submitted to 2019 IEEE International Symposium on Biomedical Imaging (ISBI).
2. <https://kits19.grand-challenge.org/home/>
3. Lee R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.