

A BOUNDARY OPTIMIZATION SCHEME FOR LIVER TUMORS FROM CT IMAGES

Ming Gong¹, John Soraghan¹, Gaetano Di Caterina¹, Xiaoquan Li¹, Derek Grose²

¹Dept. of Electronic and Electrical Engineering, University of Strathclyde, Glasgow

²Beatson West of Scotland Cancer Center, Glasgow, UK

ABSTRACT

Liver CT scan image analysis plays an important role in clinical diagnosis and treatment. Accurate segmentation of liver tumor from CT images is the prerequisite for targeted therapy and liver resection. Existing semi-automatic segmentation based on graph cuts or fully automatic segmentation methods based on deep learning have reached the level close to that of radiologists. To improve tumor segmentation on liver CT images, we propose a fully automatic post-processing scheme to optimize tumor boundaries. This method improves boundary prediction performance by optimizing a sequence of patches extracted along the initial predicted boundary. The proposed boundary refinement segmentation network obtains strong semantic information and precise location information through the information interaction between different branches, to achieve precise segmentation. The Liver Tumor Segmentation (LiTS) dataset is used to evaluate the relative segmentation performance obtaining an average dice score of 0.805 using the new method.

Index Terms— CT data, Liver tumour, Boundary Refinement, High resolution segmentation

1. INTRODUCTION

Liver cancer is the third most deadly malignant tumor in the world. Globally around 905,700 people were diagnosed with liver cancer in 2020, and around 830,200 patients have already died from the disease [3]. Medical imaging examination has become an important examination method and diagnostic basis in the diagnosis of liver diseases. Computed tomography (CT)-based medical images can provide rich information including the location, structure and function of organs and lesions with a relative lower cost. In clinical diagnosis, CT image segmentation of liver tumor is highly relevant for the diagnosis and treatment of liver cancer, such as preoperative preparation for tumor resection and subsequent radiation therapy. However, liver tumor segmentation is still a challenging task due to variable shape, diverse pathologies, and proximity to other vital organs. In clinical application, liver tumor segmentation of CT scans is usually accomplished by expert radiologists, and it is a time-consuming task, with the average scans taking up to 90

minutes to complete per patient. In addition, it is a highly subjective job which relies on vision and work experience, even some experts have difficulty in giving the same segmentation results for one CT image.

Conventional image segmentation including region growing, have achieved some results but still require additional human intervention such as feature selection and initial point determination. Recently, Deep Learning has become the main research direction in computer vision problems. Mature models successfully applied to a range of image processing challenges include Convolutional neural network, Recurrent neural network, Generative adversarial networks, and Transformer model. Specifically, the U-Net is a popular semantic segmentation network based on fully convolutional neural network that is widely used in biomedical image segmentation applications.

Recently researchers used different methods to perform liver tumour segmentation from CT scans. The modified U-Net proposed by Seo et al.[7] adds a residual path with deconvolution and activation structure in the skip path to reduce the repetition of low-resolution information. It also uses an extra convolution operation in skip connection to get achieve better feature extraction. Li et al.[6] proposed a hybrid densely connected U-Net, which merged a 2D Dense U-Net and a 3D counterpart that efficiently extracts intra-slice and inter-slice features. Amer et al.[11] proposed cascaded U-Net to determine the region of interest firstly according to the liver segmentation prediction through the first U-Net, then the tumour segmentation results are obtained with a smaller size input with the second U-Net. Chen et al.[12] use more spatial features in coronal slices to improve the performance of liver segmentation. In addition, they also designed a new deep residual module and attention module to fuse feature channels and spatial correlation.

These approaches have achieved significant results in liver tumor segmentation. Subsequent optimization on the segmentation results can improve the accuracy. Christ et al.[13] used a 3D conditional random field (CRF) to refine the segmentation results from a cascaded-fully CNN model. Lu et al. [14] used graph cut and previously learned probability maps from 3D CNN to perform accuracy improvements on the initial segmentation. However, these optimization schemes need to optimize each scan individually, which takes a lot of time.

This paper proposes a fully automatic optimized algorithm for CT tumor segmentation based on human behavior and experience in segmentation. It aims to improve the global segmentation result by optimizing a series of boundary patches obtained from a coarse segmentation result. In addition, a new optimized network that integrates information from multiple resolutions is proposed.

The remainder of the paper is organized as follows. Section 2 introduces the tumour segmentation refinement framework including boundary patch extraction and a modified segmentation network. Section 3 shows the performance of this method on public LiTS data sets. Concluding remarks are provided in Section 4.

2. METHOD

2.1. Proposed Optimization Workflow

The proposed segmentation optimization workflow shown in Figure 1, consists of four major steps. The first step is to obtain all coarse segmentation results for all images in training set. The second step is to extract all image patches and mask patches according to the coarse prediction result for training neural networks. The third step is an optimized refined segmentation network for boundary patch segmentation. The final step is to restore optimized boundary to the corresponding position of the image.

2.2. Boundary Patch Extraction

Expert radiologists complete segmentation tasks, by first finding the target area and then carefully depicting the boundary. Segmentation results provide location information to determine the distribution of target regions. Boundary patches are then obtained on these target areas for subsequent optimization. Initially all images are used in the training data set through a prediction network such as cascaded U-Net[8] to obtain a coarse segmentation result. This segmentation is expected to have a typical average dice value greater than 0.6 for obtaining boundary patches. In this coarse prediction results, some images cannot detect the presence of tumour because the size of tumor is too small, so we manually mark the predicted value of this part of the image with a 5×5 pixel block in order to ensure the richness of the training set. Consequently, the dataset has been expanded from a pair of image and ground truth, to an image set, a ground truth set, and a coarse segmentation set as in Figure 1(a) and (b).

All images in the coarse segmentation prediction set as shown in Figure 1 (d) are processed as binary images, since all boundary values can be easily obtained. Based on the coarse segmented prediction, several boundary patches are extracted by moving along the tumor border with each tumor border pixel as the centre point. At the same time, the boundary patches on the corresponding image and ground truth are acquired at the same position, which are indicated

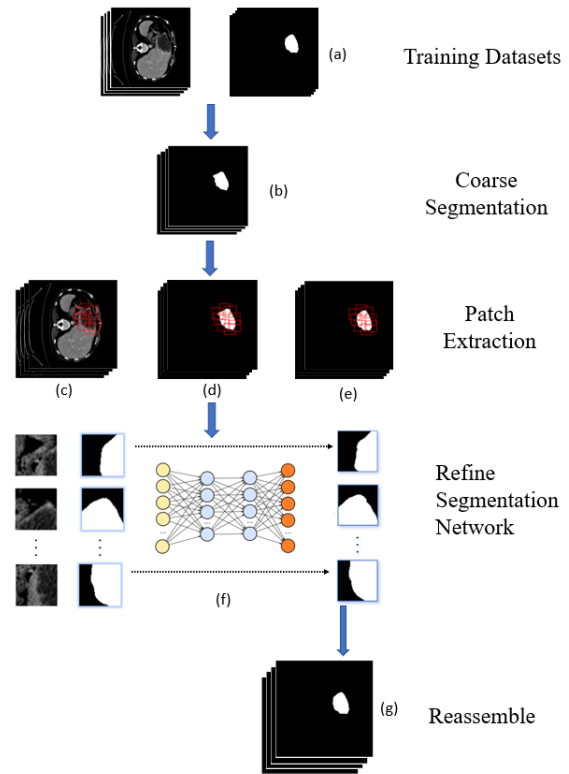


Figure 1 Proposed optimization workflow, (a) liver CT image and corresponding ground truth, (b) extracted coarse segmentation (c) image and (d) mask boundary patches from the coarse segmentation (e) mask boundary patches from the ground truth (f) all patches pass the boundary refinement network (g) final refined segmentation result.

in Figure 1(c) and (e). The size of patches is designed as $1/8$ of the original image. However, this will inevitably lead to a large amount of information redundancy, accompanied by a large amount of computational cost. We design the sliding strides to be $2/3$ of the patch's length to avoid this problem. So far, three expanded datasets containing rich boundary information have been produced and used for the subsequent refinement stage.

2.3. Boundary Refinement Segmentation

We employ a multi-level information fusion network for accurate segmentation. As illustrated in Figure 2, this is a two-input single output segmentation network. The concatenation of original image and coarse segmentation image is fed into the network. The purpose of concatenating of the input image and the coarse segmentation result is to attach a strong constraint to the input image to make the network pay more attention to the information near the boundary. It can accelerate the convergence rate and converge in the right direction.

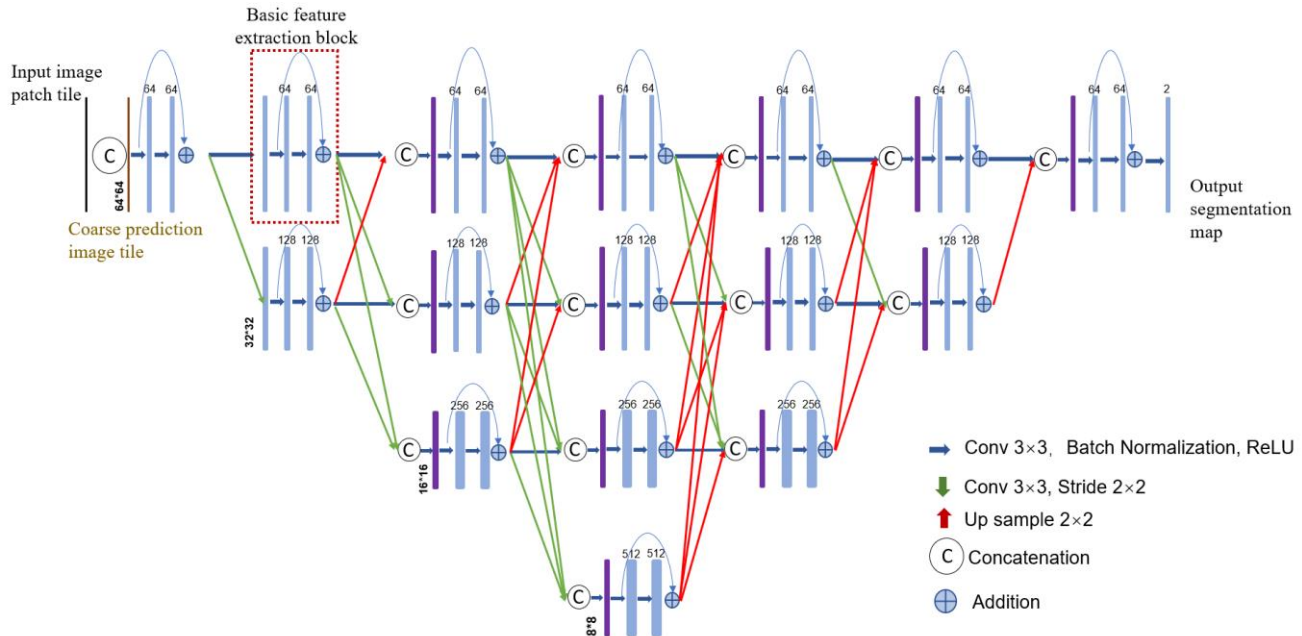


Figure 2 The architecture of proposed boundary refinement segmentation network.

The network shown in Figure 1(f) and in more detail in Figure 2 comprises a feature extraction module, several down-sampling, up-sampling, and concatenation. As shown in the red dotted box in Figure 2, the basic feature extraction is processed with four 3×3 convolutions with batch normalization and ReLU activation function. The size of convolution kernels increases from 64 to 512 in a double increment with the number of down-sampling. As the size of input images has significantly decreased compared to the original images, a convolution operation with a stride of 2 is used to accomplish the down-sampling rather than pooling. Although it will inevitably increase some computational cost, more information can be retained in the case of less input information.

Each feature extraction block is passed with residual structure to avoid gradient vanishing caused by the increase of network depth. We set 7,5,3,1 feature extraction block in four different resolutions respectively. As the number of down sampling increases, more abstract features and information are extracted with a smaller feature size as shown in Figure 2. Fully fusing information in different layers enables the model to learn more complex patterns and better capture the correlation of data. When the features received by each basic feature extractor come from different dimensions, the information is fused together through a cascade operation as shown in the purple bar in Figure 2 to extract new features. The previous feature map is adjusted to an appropriate size using a 1×1 convolution, down sampling (convolution with a stride size of 2) or up sampling (bilinear interpolation) to satisfy the concatenation requirements.

More fine-grained information is obtained in the low resolution by using multiple feature extractors which is used

in multi-level information fusion. The fine-grained information in the shallow layer reduces the resolution by down sampling, while abstracted information in deep layer expands the feature size by interpolation. Features in different dimensions are fused together by concatenation for information interaction to obtain richer semantic information and precise location information.

2.4. Reassemble Process

The image patches and corresponding mask patches are simultaneously fed into a trained optimized segmentation

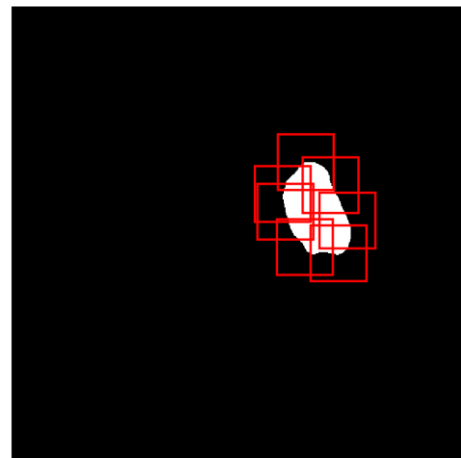


Figure 3 Overlapping refined patches.

TABLE I. COMPARISON OF OUR MODEL WITH OTHER EIGHT TUMOUR SEGMENTATION METHODS BASED ON SIX MEASURE METRICS. THE SYMBOL "\ " REPRESENTS UNREPORTED RESULT. BOLD FONT REPRESENTS THE HIGHEST SCORE ON EACH MEASUREMENT.

Model	Tumor					
	Global dice	Dice per case	VOE(%)	RVD(%)	ASSD(mm)	MSD(mm)
Shape-parameter[1]	\	0.754	0.709	0.124	1.6	\
U-Net+Level set[2]	0.700	\	\	\	\	\
AHCnet[4]	0.591	0.574	1.507	0.329	1.462	7.538
2D-dense[5]	\	0.725	0.589	\	\	\
H-dense U-Net[6]	0.824	0.722	\	\	\	\
CU-Net[8]	\	0.595	0.460	\	\	\
MCG-FRN[9]	0.764	0.674	0.324	0.194	4.408	7.113
Hybrid attention[10]	0.798	0.762	0.395	0.327	0.887	7.302
Our model	0.805	\	0.325	0.191	0.841	7.359

network to obtain accurate patch segmentation results as illustrated in Figure 3, which shows 7 overlapping patches. The segmentation results of these patches are binarized into 0 and 1 and need to be restored to the original position according to the position of the boundary pixels. The original prediction is directly replaced with the refined boundary patch and form a new prediction map. Those overlapping pixels, where the sum of their pixel values is greater than or equal to 1 are considered to be a tumor pixel.

3. EXPERIMENTS AND DISCUSSION

The data set used for training and testing is the Liver Tumour Segmentation Challenge (LiTS), which comes from clinical sites around the world [15]. The data set contains CT scans of 130 patients. All scans have been provided in *nii* format with an axial size of 512×512 (400mm×400mm). For each patient data includes, hundreds of cross-sectional scans of the abdominal cavity and annotated images provided by professional radiologists. We select all CT scan images containing tumors, for a total of 70 patients for training and testing. The training set and test set contained 50 patients and 20 patients, respectively. The training set of patches contains 30183 patches to train the refinement segmentation network. The size of patch is set to 64×64 (50mm×50mm), which reduce the amount of calculation cost while retaining sufficient information around the tumor area. This work is implemented using Keras based on the TensorFlow backend. The proposed boundary refinement network is trained for 300 epochs using Adam optimizer with learning rate 1e-4 to 1e-5.

To evaluate our tumor segmentation results, we measure the network's performance in terms of five indexes, dice score, volume overlap error (VOE), relative volume difference (RVD), average symmetric surface distance (ASSD) and the maximum symmetric surface distance

(MSD). Table 1 shows the results of eight different segmentation methods on LiTS datasets. As seen in Table 1, a 0.325 VOE and 0.841 ASSD is obtained based on our method, which is the highest score among all segmentation methods. We also get a relatively high score on global dice score as the most important evaluation metric. In addition, the prediction time is much faster than traditional optimization algorithms, and it only takes ten seconds to optimize 100 pictures on average.

Figure 4 shows some visualizations of tumor segmentation for four patients. The green line represents the optimized tumor boundary, the yellow line represents the

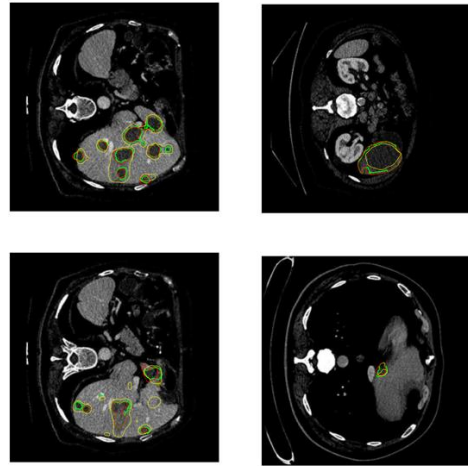


Figure 4 Visualization results of four different patients. Red line presents the ground truth of tumor boundary, yellow line present segmentation results need to be refined, green line presents refined segmentation boundary.

coarsely segmented tumor boundary, and the red line represents the ground truth boundary. It is seen that our

method performs well in optimization for large, medium,

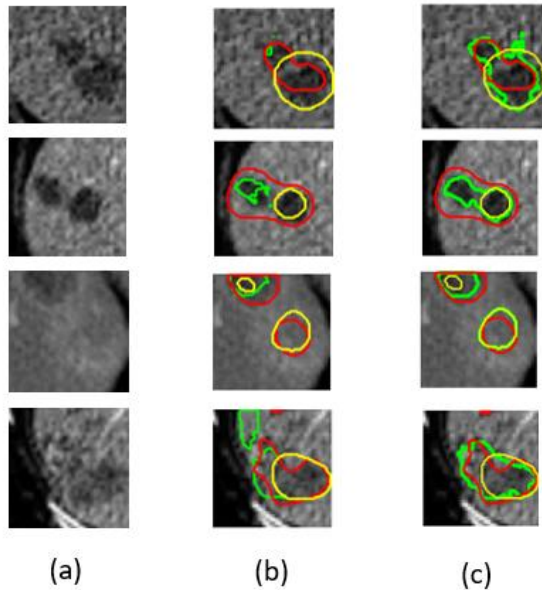


Figure 5 Comparison of boundary optimization segmentation network with and without mask patch. Red: Ground Truth; Yellow: Coarse Segmentation; Green: Refined Segmentation

and small tumors. The optimization algorithm can effectively shrink or expand the rough segmentation results to approach the real results. Figure 5 shows the effect of the mask patch on the optimization algorithm. Column (a) are initial CT images, column (b) are segmentation results without the mask patch, and column (c) are segmentation results with the mask patch. Red, yellow, and green line present ground truth, coarse segmentation, and refined segmentation boundary respectively. When the model remains unchanged, the prediction results without the participation of the mask patch become very unreliable. For some small tumors, the model cannot obtain accurate prediction results or even completely wrong predictions, which are not as good as the results before optimization. The use of the mask patch allows the network to focus more on the pixels near the area to be optimized, so that the optimization result is closer to the real value which can be clearly seen from Figure 5.

4. CONCLUSION

In this paper, we presented a new fully automatic method to optimize liver tumors segmentation results from CT scans, which can achieve end-to-end refinement for liver tumor segmentation. We extract the image patches and mask patches according to the coarse segmentation results. We also designed a specialized boundary refinement network for patch size images segmentation. We use mask patch to

strengthen the network's attention to the boundary area to improve segmentation performance. Compared to traditional U-shape deep neural networks, our proposed network uses more feature extraction block to replace the skip connection to obtain more information in the case of limited input information. We obtained an average dice score of 0.805 and volume overlap error of 0.325 on the liver tumour segmentation challenge. Due to the reduced input image size, the problem of class imbalance is improved.

5. REFERENCES

- [1] M. G. Linguraru *et al.*, "Tumor burden analysis on computed tomography by automated liver and tumor segmentation," *IEEE transactions on medical imaging*, vol. 31, no. 10, pp. 1965-1976, 2012.
- [2] O. I. Alirri, "Deep learning and level set approach for liver and tumor segmentation from CT scans," *Journal of Applied Clinical Medical Physics*, vol. 21, no. 10, pp. 200-209, 2020.
- [3] H. Runggay *et al.*, "Global burden of primary liver cancer in 2020 and predictions to 2040," *Journal of Hepatology*, vol. 77, no. 6, pp. 1598-1606, 2022.
- [4] H. Jiang, T. Shi, Z. Bai, and L. Huang, "Ahnctnet: An application of attention mechanism and hybrid connection for liver tumor segmentation in ct volumes," *IEEE Access*, vol. 7, pp. 24898-24909, 2019.
- [5] K. C. Kaluva, M. Khened, A. Kori, and G. Krishnamurthi, "2D-densely connected convolution neural networks for automatic liver and tumor segmentation," *arXiv preprint arXiv:1802.02182*, 2018.
- [6] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663-2674, 2018.
- [7] H. Seo, C. Huang, M. Bassenne, R. Xiao, and L. Xing, "Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images," *IEEE transactions on medical imaging*, vol. 39, no. 5, pp. 1316-1325, 2019.
- [8] A. A. Albishri, S. J. H. Shah, and Y. Lee, "CU-Net: Cascaded U-Net model for automated liver and lesion segmentation and summarization," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2019: IEEE, pp. 1416-1423.
- [9] Z. Bai, H. Jiang, S. Li, and Y.-D. Yao, "Liver tumor segmentation based on multi-scale candidate generation and fractal residual network," *IEEE Access*, vol. 7, pp. 82122-82133, 2019.
- [10] M. Gong, B. Zhao, J. Soraghan, G. Di Caterina, and D. Grose, "Hybrid attention mechanism for liver tumor segmentation in CT images," in *2022 10th European Workshop on Visual Information Processing (EUVIP)*, 2022: IEEE, pp. 1-6.
- [11] A. Amer, T. Lambrou, and X. Ye, "MDA-unet: a multi-scale dilated attention U-net for medical image segmentation," *Applied Sciences*, vol. 12, no. 7, p. 3676, 2022.
- [12] Y. Chen *et al.*, "A deep residual attention-based U-Net with a biplane joint method for liver segmentation from CT scans," *Computers in Biology and Medicine*, vol. 152, p. 106421, 2023.
- [13] P. F. Christ *et al.*, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," in *International conference on medical image computing and computer-assisted intervention*, 2016: Springer, pp. 415-423.
- [14] F. Lu, F. Wu, P. Hu, Z. Peng, and D. Kong, "Automatic 3D liver location and segmentation via convolutional neural network and graph cut," *International journal of computer assisted radiology and surgery*, vol. 12, pp. 171-182, 2017.
- [15] P. Bilic *et al.*, "The liver tumor segmentation benchmark (lits)," *arXiv preprint arXiv:1901.04056*, 2019.