



Liver segmentation in CT imaging with enhanced mask region-based convolutional neural networks

Xiaowen Chen^{1#^}, Xiaoqin Wei^{1#}, Mingyue Tang², Aimin Liu¹, Ce Lai¹, Yuanzhong Zhu¹, Wenjing He¹

¹School of Medical Imaging, North Sichuan Medical College, Nanchong, China; ²Department of Physics, School of Basic Medicine, North Sichuan Medical College, Nanchong, China

Contributions: (I) Conception and design: X Chen, X Wei; (II) Administrative support: Y Zhu; (III) Provision of study materials or patients: A Liu; (IV) Collection and assembly of data: M Tang, C Lai; (V) Data analysis and interpretation: X Chen, X Wei, W He; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Wenjing He, Master. School of Medical Imaging, North Sichuan Medical College, No. 234, Fujiang Road, Shunqing District, Nanchong 637000, China. Email: 18990715902@163.com.

Background: Liver segmentation in computed tomography (CT) imaging has been widely investigated as a crucial step for analyzing liver characteristics and diagnosing liver diseases. However, obtaining satisfactory liver segmentation performance is highly challenging because of the poor contrast between the liver and its surrounding organs and tissues, the high levels of CT image noise, and the wide variability in liver shapes among patients.

Methods: To overcome these challenges, we propose a novel method for liver segmentation in CT image sequences. This method uses an enhanced mask region-based convolutional neural network (Mask R-CNN) with graph-cut segmentation. Specifically, the k -nearest neighbor (k -NN) algorithm is employed to cluster the target liver pixels in order to get an appropriate aspect ratio. Then, anchors are adapted to the liver size using the ratio information. Thus, high-accuracy liver localization can be achieved using the anchors and rotation-invariant object recognition. Next, a fully convolutional network (FCN) is used to segment the foreground objects, and local fine-grained liver detection is realized by pixel prediction. Finally, a whole liver mask is obtained by Mask R-CNN proposed in this paper.

Results: We proposed a Mask R-CNN algorithm which achieved superior performance in comparison with the conventional Mask R-CNN algorithms in term of the dice similarity coefficient (DSC), and the Medical Image Computing and Computer-Assisted Intervention (MICCAI) metrics.

Conclusions: Our experimental results demonstrate that the improved Mask R-CNN architecture has good performance, accuracy, and robustness for liver segmentation in CT image sequences.

Keywords: Liver segmentation; computed tomography; mask region-based convolutional neural network (Mask R-CNN); k -nearest neighbor (k -NN); angle rotation

Submitted Oct 21, 2021. Accepted for publication Dec 06, 2021.

doi: 10.21037/atm-21-5822

View this article at: <https://dx.doi.org/10.21037/atm-21-5822>

Introduction

Liver segmentation in medical imaging data is of great clinical significance for lesion resections and liver transplantations

(1-3). In addition, accurate liver segmentation offers a substantial aid to physicians in diagnosing and treating liver diseases (4-6). However, manual liver segmentation

[^] ORCID: 0000-0002-0313-6363.

takes a lot of time and effort. Automatic liver segmentation algorithms have been proposed to reduce these costs, but the performance outcomes of such algorithms are still limited due to several challenges in computed tomography (CT) imaging. First, the liver grayscale patterns in CT images are very similar to those of the surrounding organs (such as the stomach, pancreas, kidneys, and muscles) (7-9). Secondly, liver shapes and sizes vary widely across patients (10,11). Thirdly, differences in CT scanning equipment lead to large variations in CT image appearance and liver location (12,13).

State-of-the-art liver segmentation performance has progressively improved with the emergence of deep network architectures (14,15). However, segmentation accuracy and speed are still unsatisfactory for complex scenes (16-18). In addition, the use of fine-grained classification in liver segmentation is complicated by the small inter-class differences that arise from the close similarities between subcategories and by the large intra-class variabilities in position, scale, and orientation (19,20). Indeed, fine-grained recognition is popular in computer vision and pattern recognition applications (21,22); it facilitates the learning of object parts and helps differentiate between object subclasses, and hence can be employed to learn liver patterns more accurately (23).

In 2016, Girshick *et al.* (24) designed a region-based convolutional neural network (R-CNN), which used candidate region proposals and a CNN-based classification algorithm for detection. This R-CNN architecture boosted the performance of target detection and recognition systems and inspired the creation of more powerful deep-learning algorithms for such systems. In particular, the R-CNN algorithm adopted 4 steps, comprised of conventional target detection framework, feature extraction, image classification, and non-maximum suppression. Nevertheless, the R-CNN algorithm exploited CNN-based features instead of traditional hand-crafted features such as those of the scale-invariant feature transform (SIFT) (25,26) or the histogram of oriented gradients (HOGs) (27-29). The fast R-CNN (Fast R-CNN) (30) and mask R-CNN (Mask R-CNN) (31) were both built as variants of the R-CNN algorithm.

This study sought to improve liver segmentation performance through an enhanced variant of the Mask R-CNN algorithm (32-35). A novel idea in the study is that the k -NN was employed for data clustering and obtaining an appropriate aspect ratio during the training phase. Compared with other models, our proposed model can effectively reduce computational resources and improve computational accuracy. Enhanced Mask R-CNN algorithm

is based on original algorithm and combined with k -NN, we applied this model to liver segmentation for the first time. At present, this model has not been widely used in clinical trials, and we will study this model as soon as possible and apply it to clinical treatment.

The experimental results demonstrate significant performance improvements based on the proposed method.

The key aspects of our study were as follows:

- (I) The study highlighted the importance of random variations in liver images. With these variations, the proposed Mask R-CNN method provided remarkable improvements in segmentation compared to the conventional Mask R-CNN method. Labeled data was augmented through image rotation operations during the data preparation phase to improve generalization and reduce overfitting;
- (II) During the training phase, the k -nearest neighbor (k -NN) algorithm was employed for data clustering and obtaining an appropriate aspect ratio. Moreover, a fully convolutional network (FCN) was adopted to realize the segmentation algorithm after accounting for width-height reversal and noise;
- (III) The performance of the proposed Mask R-CNN algorithm was evaluated in comparison with the conventional Mask R-CNN algorithm according to 3 metrics, namely, detection accuracy (DA), detection speed (DS), and false-detection rate (FD). The effectiveness and feasibility of the proposed algorithm were verified in comparison to state-of-the-art segmentation algorithms through the measurement of the volumetric overlap error (VOE), the relative volume difference (RVD), the average symmetric surface distance (ASD), the root-mean-square symmetric surface distance (RMSD), and the maximum symmetric surface distance (MSD).

We present the following article in accordance with the MDAR reporting checklist (available at <https://dx.doi.org/10.21037/atm-21-5822>).

Methods

Conventional Mask R-CNN

Figure 1 shows the general framework and the associated network structure for the conventional Mask R-CNN algorithm. First, feature extraction is performed through convolutional layers on input liver images of arbitrary sizes

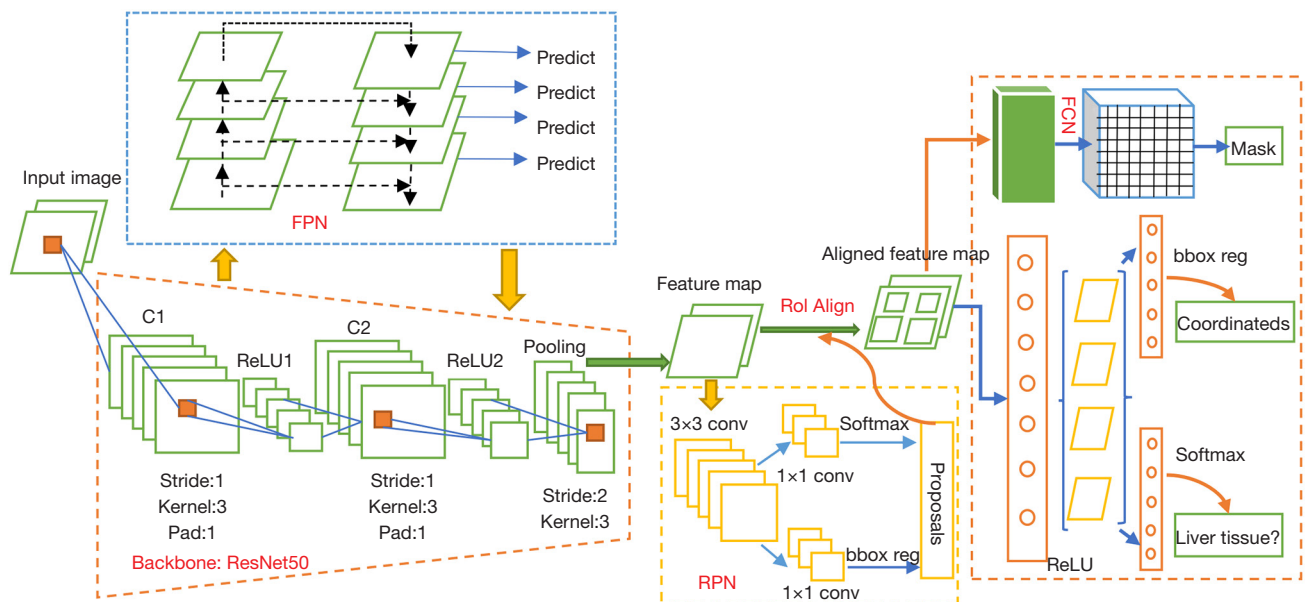


Figure 1 Framework of the conventional Mask R-CNN algorithm. FPN, feature pyramid network; RPN, region proposal network; FCN, fully convolutional network; Mask R-CNN, mask region-based convolutional neural network.

to form feature maps. Then, a region proposal network (RPN) exploits the convolutional layer outputs for domain (or proposal) generation, as well as category and bounding-box regression, in order to speed up the computations. Pixel-wise segmentation of the target liver and liver localization in the image are carried out by a parallel feature pyramid network (FPN).

The conventional Mask R-CNN algorithm follows a two-stage object detection method. Firstly, a candidate target region is generated, and the boundary box of the candidate object is proposed in line with the Faster R-CNN method. Secondly, binary masks, predicted classes, and bounding-box offsets are returned by the conventional Mask R-CNN algorithm for each region of interest (RoI), where the classification outcomes are dependent on mask predictions (36,37). In the training phase, a multi-objective loss function is defined for each RoI sample as (38):

$$L = L_{cls} + L_{box} + L_{mask} \quad [1]$$

where L_{cls} , L_{box} , and L_{mask} denote the classification loss, the bounding-box loss, and the segmentation loss, respectively.

The conventional Mask R-CNN algorithm proposes a RoI Align layer that obtains image values at pixel points with floating-point coordinates through bilinear interpolation. This approach avoids any quantization of the RoI boundaries or intervals and thus ensures the

continuity of the entire process of feature aggregation. Specifically, the RoI Align layer does not carry out pooling by merely supplementing the coordinate points on the candidate region boundary. Instead, this layer first traverses each candidate region and keeps the floating-point boundary unquantized. Each region is then subdivided into $K \times K$ units with unquantized boundaries, and 4 fixed coordinate positions are found in each unit through bilinear interpolation. Hence max pooling is applied. The RoI Align layer clusters local features in the Mask R-CNN algorithm reduces misalignment caused by the two quantization operations in RoI pooling and thus improves the DA (39,40).

Enhanced Mask R-CNN algorithm

In this study, novel improvements were made to the conventional Mask R-CNN detection framework. The proposed network framework is outlined in *Figure 2*. This framework can be divided into two steps. The first step was the candidate region identification, and in the second step, the global and local features of image blocks were learned, mainly by a part-based segmentation algorithm and the FCN architecture. The details of the two steps are as follows.

Step 1: the RPN was exploited to extract features and generate a feature map at the last layer. The whole image was scanned by a sliding window to get target anchors.

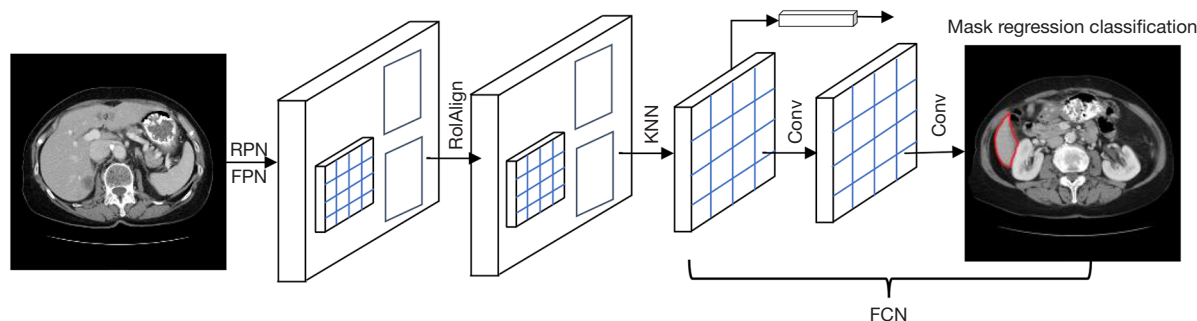


Figure 2 The enhanced Mask R-CNN framework. RPN, region proposal network; FPN, feature pyramid network; k-NN, k-nearest neighbor; FCN, fully convolutional network; Mask R-CNN, mask region-based convolutional neural network.

For each image position, it was possible to identify tens of thousands of overlapping candidate target regions with different dimensions and aspect ratios. These regions covered the whole image. We used the k -NN algorithm to optimize the number of target regions, largely reduce the computational overhead, and improve computational accuracy (Figure 2). In addition, the RPN generated two outputs for each anchor, namely, the anchor category and the border tuning parameter. For multiple overlapping anchors, we adopted non-maxima suppression to obtain rough target results, where the anchor of the highest foreground score was retained. Therefore, it was possible to use the RPN prediction outcomes to select the best anchor containing the target, and fine border adjustment was applied.

Step 2: image segmentation was performed using an FCN architecture that could take an input image of arbitrary resolution and produce an output of the same size. This architecture located targets in fine-grained images and treated segmentation prediction as target masking. For effective mask learning, all fine-grained training and testing images retained their original resolutions. The FCN-based mask learning process is shown in Figure 2. First, FCN-based prediction was applied to obtain a local target mask in a given input image. If a pixel was predicted to be a local target position, the actual mask value was retained. Thus, fine-grained liver detection was carried out. Otherwise, if the pixel-wise prediction indicated a background region, the corresponding mask values were reset to zero. The global and local liver features in each image were learned, and the FCN algorithm returned a more accurate target mask. In addition, the obtained target masks could locate the target positions by finding the bounding rectangles. In this study, the FCN was used for target mask learning and prediction.

In addition, in Figure 3, the 3 streams shown correspond

to the angle rotations of 3 image blocks. The characteristics of each of these blocks were learned through a series of operations such as convolution, activation, pooling, and discriminator selection. Indeed, combining image features of different scales in this study enhanced the robustness of liver detection.

Statistical analysis

In our experiments, we use 6 metrics to evaluate the segmentation performance of different algorithms. Firstly, we use the dice similarity coefficient (DSC) (41,42), which reflects the degree of spatial coincidence between a segmentation output region U_1 and the corresponding ground-truth region U_2 . This coefficient can be defined mathematically as

$$DSC(U_1, U_2) = \frac{2|U_1 \cap U_2|}{|U_1| + |U_2|} \quad [2]$$

where $|U_i|$ denotes the cardinality of the set U_i , and \cap denotes set intersection. The higher the DSC value, the better the segmentation performance. The best segmentation performance was achieved when the DSC had a value of 1 (i.e., when the U_1 and U_2 sets were identical), while the worst performance was indicated by a DSC value of 0 (when the U_1 and U_2 sets were mutually disjoint).

In addition to the DSC metric, we also employed 5 metrics suggested by the Medical Image Computing and Computer-Assisted Intervention (MICCAI) society. These metrics were the VOE, the RVD, the ASD, the RMSD, and the MSD (5,43).

$$VOE(U_1, U_2) = 100 \times \left(1 - \frac{|U_1 \cap U_2|}{|U_1 \cup U_2|} \right) \quad [3]$$

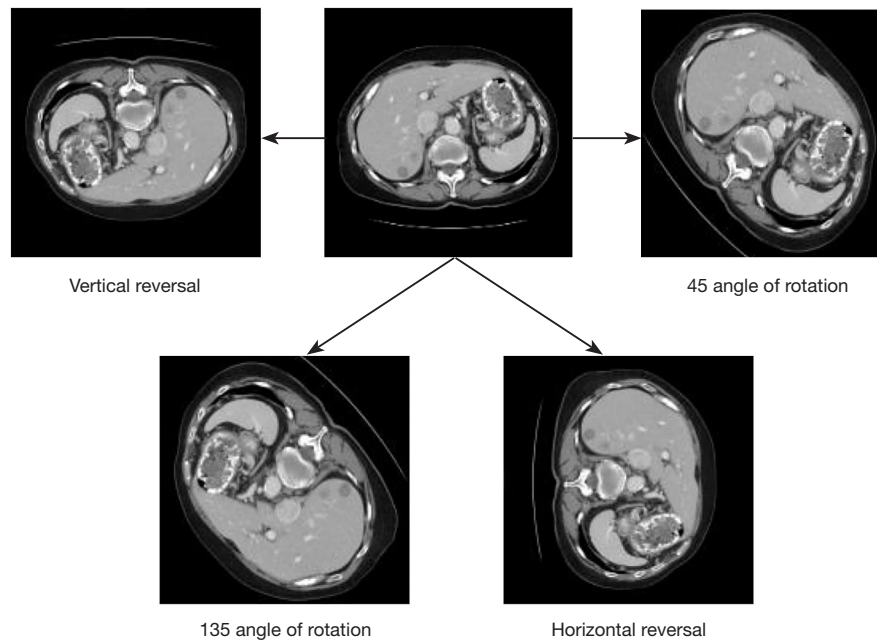


Figure 3 A demonstration of liver image enhancement.

$$RVD(U_1, U_2) = 100 \times \left(\frac{|U_1| - |U_2|}{|U_2|} \right) \quad [4]$$

$$ASD(U_1, U_2) = \frac{1}{|S(U_1)| + |S(U_2)|} \left(\sum_{S_{U_1} \in S(U_1)} d(S_{U_1}, S(U_2)) + \sum_{S_{U_2} \in S(U_2)} d(S_{U_2}, S(U_1)) \right) \quad [5]$$

$$RMSD(U_1, U_2) = \sqrt{\frac{1}{|S(U_1)| + |S(U_2)|} \times \left(\sum_{S_{U_1} \in S(U_1)} d^2(S_{U_1}, S(U_2)) + \sum_{S_{U_2} \in S(U_2)} d^2(S_{U_2}, S(U_1)) \right)} \quad [6]$$

$$MSD(U_1, U_2) = \max \left\{ \max_{S_{U_1} \in S(U_1)} d(S_{U_1}, S(U_2)), \max_{S_{U_2} \in S(U_2)} d(S_{U_2}, S(U_1)) \right\} \quad [7]$$

where $S(U_1)$ and $S(U_2)$ denote the boundaries of the two regions, respectively. Also, for any pixel v , $d(v, S(U_1))$ denotes the shortest distance between v and $S(U_1)$, i.e., $d(v, S(U_1)) = \min_{s_{U_1} \in S(U_1)} \|v - s_{U_1}\|$, where $\|\cdot\|$ is the Euclidean distance.

The lower the value of each of the VOE, ASD, RMSD, and MSD, the better the segmentation performance. As the RVD metric could be negative (in the case of under-segmentation), the absolute value of this metric was used for evaluating the segmentation performance. The smaller the absolute RVD value, the better the segmentation

performance. In general, each of these 5 metrics had a 0 value for perfect segmentation.

Results

Experimental data and environment

In our experiments, we used a CentOS7 with an Intel Core i7 processor and a 48-GB GPU NVIDIA RTX8000 graphics card with 128GB of memory. We carried out the algorithm implementation and simulation in Python

3.6. For system training and testing, we exploited the liver segmentation dataset of the Codalab competition (<https://competitions.codalab.org/>). This dataset contains 723 sequences of enhanced CT with a resolution of 512×512 (39,40) and has 3-phase ground-truth data of liver CT images. The total CT image count was 3,034 after accounting for liver image rotations. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Table 1 Learning parameters for the enhanced Mask R-CNN architecture

Description	Value
Momentum	0.9
Decay	0.0002
Batch size	32
Learning rate	0.001
Iteration	50,000

Mask R-CNN, mask region-based convolutional neural network.

Parameter settings

In this study, the enhanced Mask R-CNN architecture was trained for liver segmentation. Reasonable parameters values were selected to speed up training and prevent overfitting. The specific parameter settings employed herein are shown in *Table 1*.

Experimental analysis

Experiments were conducted to compare the performance of the conventional Mask R-CNN against the enhanced Mask R-CNN, which employed sequence information and a graph-cut function and technical repeat 12 times. Six slices of liver images (with normal and pathological cases) were selected to investigate the effect of using sequence information and the graph-cut function in improving the Mask R-CNN segmentation outputs. *Figure 4* gives a comparison of the results, which are shown in red contours.

As shown in *Figure 4*, the conventional Mask R-CNN method obviously missed lesion areas of adjacent greyscale liver regions during liver slice processing. Also, this

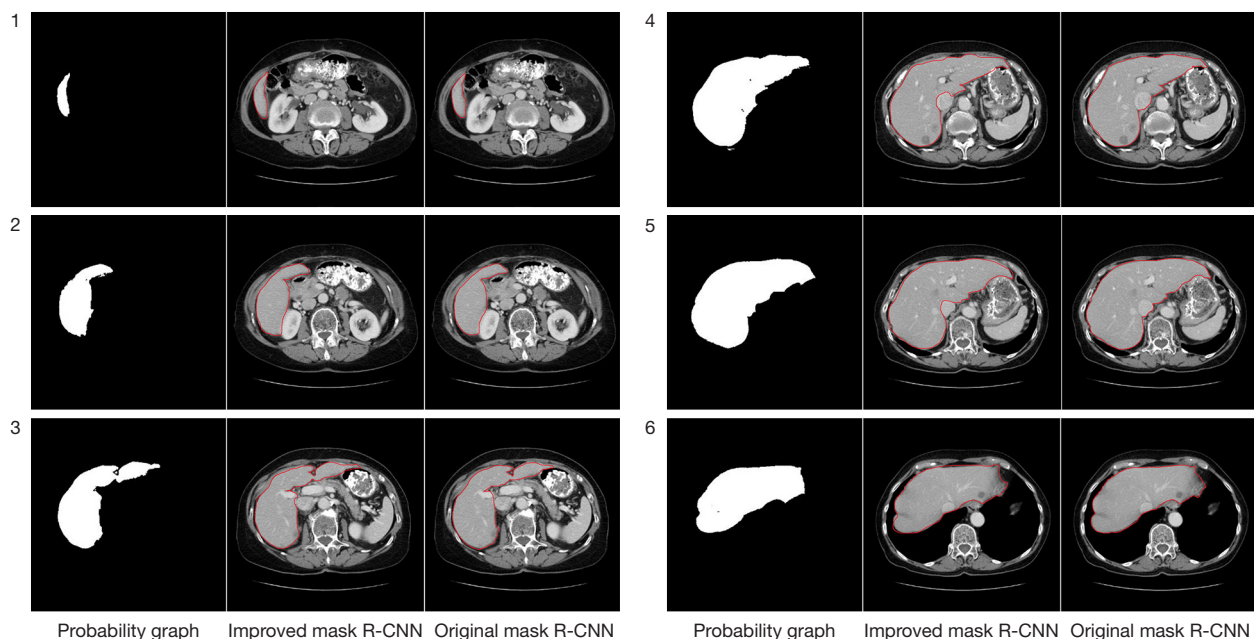


Figure 4 A comparison of segmentation results between the enhanced and conventional Mask R-CNN algorithms. Mask R-CNN, mask region-based convolutional neural network.

Table 2 Comparison of segmentation algorithms before and after enhancements

Algorithm type	AP (%)	FPR (%)	DR (s)
Original Mask R-CNN	79.78	4.34	0.332
Improved Mask R-CNN	83.23	2.34	0.112

Mask R-CNN, mask region-based convolutional neural network; AP, average precision; FPR, false-positive rate; DR, detection rate.

method misjudged small liver regions, resulting in mis-segmentation. Our proposed method managed to correct these mis-segmentation problems by incorporating the sequence aspect ratio information. Our method could, to a certain extent, improve segmentation accuracy and robustness for each slice in the sequence. This is an obvious advantage of our enhanced Mask R-CNN over the conventional Mask R-CNN.

As shown in *Table 2*, the experimental results of the enhanced Mask R-CNN were evaluated in terms of the following indicators: the average precision (AP), the false-positive rate (FPR) (also called the complemented sensitivity), and the detection rate (DR). The enhanced algorithm had relatively high AP and FPR values, as well as very low temporal complexity. This performance can be ascribed to the training adequacy and the relative roughness of the output boundary, though there were *mis-segmentation* or *over-segmentation* errors.

Further experiments were carried out to assess the impact of sequence information and the graph-cut function on improving the Mask R-CNN method in comparison to the FCN-8s algorithm (44), the 2D-dense-FCN algorithm (45), and the U-Net algorithm (46). Parameter settings for these networks were seen in *Table 1* and the performance was evaluated using the DSC, VOE, RVD, ASD, RMSD, and MSD metrics. A comparison of the segmentation results of the 4 algorithms is shown in *Figure 5*. The test data for this comparison included slices with large, medium, and small liver regions.

The results in *Figure 5* show that for the 6 liver slices, some under-segmentation or over-segmentation errors were made by the FCN-8s, U-Net, and 2D-dense-FCN algorithms. Also, some of the segmented liver slices did not

exhibit complete boundaries, while others had extraneous parts that did not belong to the original liver slices. However, our enhanced Mask R-CNN method was able to produce more solid boundaries and return segmented liver slices with no extra holes.

Table 3 indicates that the enhanced Mask R-CNN method significantly outperformed the other algorithms, except for the U-Net method, which shows a better VOE, and the 2D-dense-FCN method, which shows a better RVD (but less stability) compared to our method. However, our algorithm clearly outperformed the U-Net algorithm for all other indicators. In addition, our algorithm showed superior performance for all 5 indicators in comparison to all the other algorithms. The relatively large ASD, RMSD, and MSD values for both the U-Net and FCN-8s algorithms indicate large differences between the liver segmentation results and the corresponding ground-truth regions.

Conclusions

Automatic algorithms for liver segmentation in CT images seek to handle peripheral organs and the large inter-personal differences in liver characteristics (47). Filtering a certain number of anchors can greatly improve accuracy and reduce time consumption. To address the weaknesses of the conventional Mask R-CNN algorithm, we proposed a novel enhanced Mask R-CNN algorithm. Specifically, we augmented the conventional method with rotation angle adjustment and filtered out a certain anchor ratio. We also used the enhanced Mask R-CNN for liver slice segmentation as well as the creation of a probability map. Our proposed solution enhanced the Mask R-CNN algorithm by incorporating the advantages of the *k*-NN methodology. In addition, our solution improved the segmentation accuracy and robustness using the rotation information obtained from liver image sequences. While we focused on enhancing and employing the Mask R-CNN algorithm in liver segmentation in this study, the methodology can be extended to enhancing the segmentation outcomes for other organs (48,49). Otherwise, for the same CT images, we can apply this model to segment any type of bodily tissue, but adjusting the parameters appropriately was required.

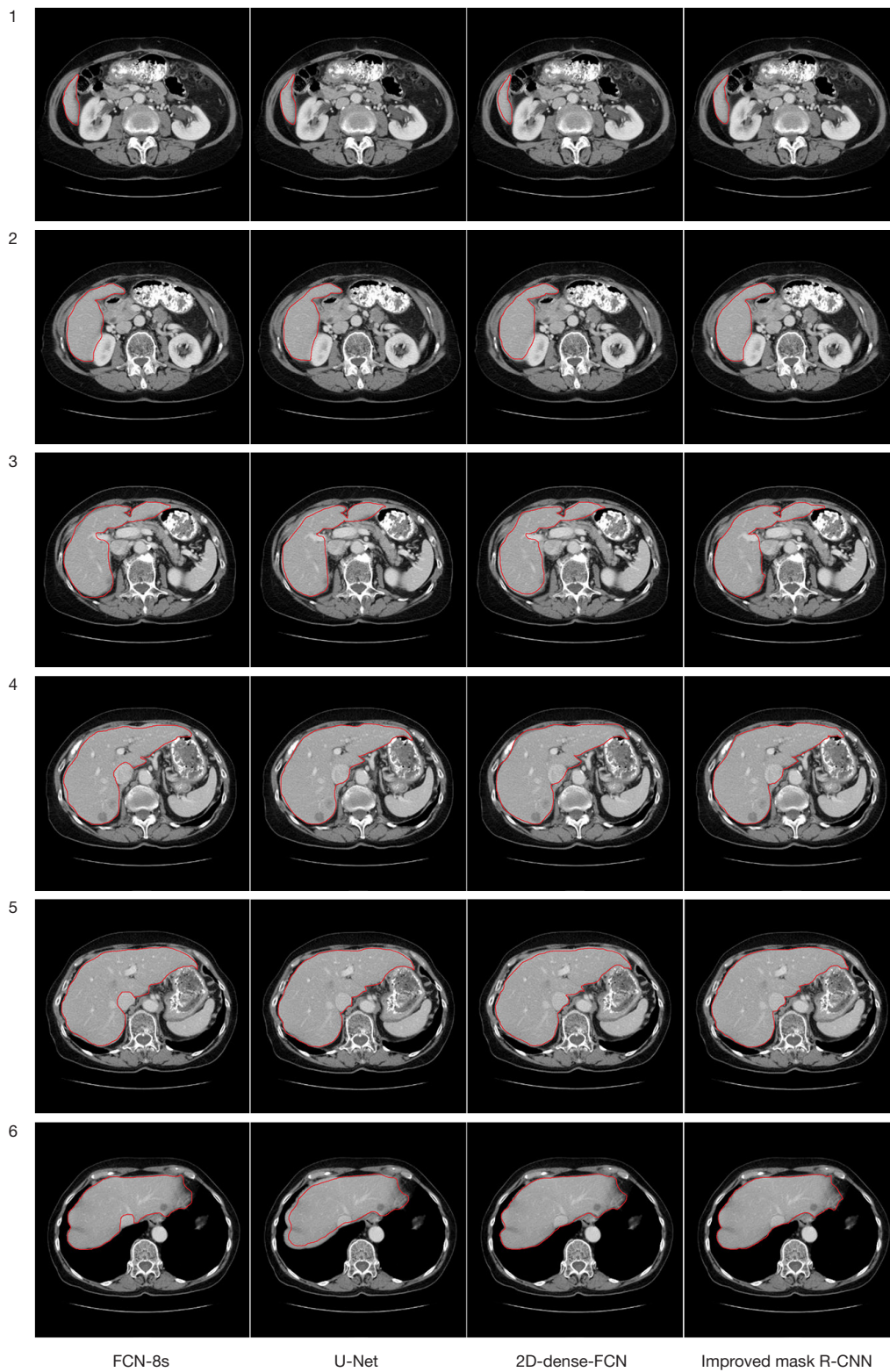


Figure 5 A comparison of the segmentation results between the algorithms of FCN-8s, U-Net, 2D-dense-FCN, and enhanced Mask R-CNN. FCN, fully convolutional network; Mask R-CNN, mask region-based convolutional neural network.

Table 3 A comparison of 4 liver segmentation algorithms based on 6 metrics

Algorithm type	DSC (%)	VOE (%)	RVD (%)	ASD (mm)	RMSD (mm)	MSD (mm)
FCN-8s	82.34	13.02	0.23	5.32	12.42	44.44
U-Net	89.89	8.21	1.21	6.32	6.23	32.45
2D-dense-FCN	92.11	21.24	-2.12	3.34	5.23	21.23
Improved Mask R-CNN	94.34	9.34	0.64	2.23	3.32	19.32

FCN, fully convolutional network; Mask R-CNN, mask region-based convolutional neural network; DSC, dice similarity coefficient; VOE, volumetric overlap error; RVD, relative volume difference; ASD, average symmetric surface distance; RMSD, root-mean-square symmetric surface distance; MSD, maximum symmetric surface distance.

Acknowledgments

Funding: This research was partly supported by the Project of Sichuan Provincial Department of Education (grant No. 17ZA0165) and the City-School Science and Technology Strategic Cooperation Project in Nanchong city (grant No. 19SXHZ0239, 18SXHZ0393).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at <https://dx.doi.org/10.21037/atm-21-5822>

Data Sharing Statement: Available at <https://dx.doi.org/10.21037/atm-21-5822>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/atm-21-5822>). The authors report that this research was partly supported by the Project of Sichuan Provincial Department of Education (grant No. 17ZA0165) and the City-School Science and Technology Strategic Cooperation Project in Nanchong city (grant No. 19SXHZ0239, 18SXHZ0393). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons

Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Guo X, Schwartz LH, Zhao B. Automatic liver segmentation by integrating fully convolutional networks into active contour models. *Med Phys* 2019;46:4455-69.
- Spinczyk D, Krasoń A. Automatic liver segmentation in computed tomography using general-purpose shape modeling methods. *Biomed Eng Online* 2018;17:65.
- Wang Z, Zhang C, Jiao T, et al. Fully Automatic Segmentation and Three-Dimensional Reconstruction of the Liver in CT Images. *J Healthc Eng* 2018;2018:6797102.
- Borhani AA. Editorial Comment on "Liver Steatosis Categorization on Contrast-Enhanced CT Using a Fully Automated Deep Learning Volumetric Segmentation Tool: Evaluation in 1204 Healthy Adults Using Unenhanced CT as a Reference Standard". *AJR Am J Roentgenol* 2021;217:367.
- Liao M, Zhao YQ, Wang W, et al. Efficient liver segmentation in CT images based on graph cuts and bottleneck detection. *Phys Med* 2016;32:1383-96.
- Ota T, Kamiyama T, Kato T, et al. A rare case of cavernous hemangioma accompanied with diffuse hepatic hemangiomatosis. *Surg Case Rep* 2020;6:251.
- Hashemi MS, Baniassadi M, Baghani M, et al. A novel machine learning based computational framework for homogenization of heterogeneous soft materials: application to liver tissue. *Biomech Model Mechanobiol*

- 2020;19:1131-42.
8. Qin W, Wu J, Han F, et al. Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation. *Phys Med Biol* 2018;63:095017.
 9. Jung Won Cha, Farhangi MM, Dunlap N, et al. Volumetric analysis of respiratory gated whole lung and liver CT data with motion-constrained graph cuts segmentation. *Annu Int Conf IEEE Eng Med Biol Soc* 2017;2017:3405-8.
 10. Kloenne M, Niehaus S, Lampe L, et al. Domain-specific cues improve robustness of deep learning-based segmentation of CT volumes. *Sci Rep* 2020;10:10712.
 11. Shimizu Y, Takamatsu S, Yamamoto K, et al. Segmental analysis of respiratory liver motion in patients with and without a history of abdominal surgery. *Jpn J Radiol* 2018;36:511-8.
 12. Bharti P, Mittal D, Ananthasivan R. Characterization of chronic liver disease based on ultrasound images using the variants of grey-level difference matrix. *Proc Inst Mech Eng H* 2018;232:884-900.
 13. Park J, Kim JH, Kim JE, et al. Prediction of liver regeneration in recipients after living-donor liver transplantation in using preoperative CT texture analysis and clinical features. *Abdom Radiol (NY)* 2020;45:3763-74.
 14. Salvi M, Molinaro L, Metovic J, et al. Fully automated quantitative assessment of hepatic steatosis in liver transplants. *Comput Biol Med* 2020;123:103836.
 15. Actor JA, Fuentes DT, Rivière B. Identification of Kernels in a Convolutional Neural Network: Connections Between Level Set Equation and Deep Learning for Image Segmentation. *Proc SPIE Int Soc Opt Eng* 2020;11313:1131317.
 16. Zhang R, Zhou Z, Wu W, et al. An Improved Fuzzy Connectedness Method for Automatic Three-Dimensional Liver Vessel Segmentation in CT Images. *J Healthc Eng* 2018;2018:2376317.
 17. Almotairi S, Kareem G, Aouf M, et al. Liver Tumor Segmentation in CT Scans Using Modified SegNet. *Sensors (Basel)* 2020;20:1516.
 18. Guo X, Huang S, Fu X, et al. Vascular segmentation in hepatic CT images using adaptive threshold fuzzy connectedness method. *Biomed Eng Online* 2015;14:57.
 19. Tsurusaki M, Numoto I, Oda T, et al. Assessment of Liver Metastases Using CT and MRI Scans in Patients with Pancreatic Ductal Adenocarcinoma: Effects of Observer Experience on Diagnostic Accuracy. *Cancers (Basel)* 2020;12:1455.
 20. Worm ES, Høyer M, Fledelius W, et al. On-line use of three-dimensional marker trajectory estimation from cone-beam computed tomography projections for precise setup in radiotherapy for targets with respiratory motion. *Int J Radiat Oncol Biol Phys* 2012;83:e145-51.
 21. Zaleska-Freljan KI, Kosicka B, Zbiegieni B. The histological changes in some organs of the laboratory mice after intragastrically given bromfenvinphos and mixture of bromfenvinphos with methoxychlor. *Pol J Pharmacol Pharm* 1983;35:185-93.
 22. Hunt C, Ropella GE. Using an in silico liver to evaluate a hepatic enzyme induction mechanism. *Annu Int Conf IEEE Eng Med Biol Soc* 2008;2008:2415-8.
 23. Mühlberg A, Holch JW, Heinemann V, et al. The relevance of CT-based geometric and radiomics analysis of whole liver tumor burden to predict survival of patients with metastatic colorectal cancer. *Eur Radiol* 2021;31:834-46.
 24. Girshick R, Donahue J, Darrell T, et al. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Trans Pattern Anal Mach Intell* 2016;38:142-58.
 25. Huang D, Zhu C, Wang Y, et al. HSOG: a novel local image descriptor based on histograms of the second-order gradients. *IEEE Trans Image Process* 2014;23:4680-95.
 26. Zheng J, Li B, Xin M, et al. Structured fragment-based object tracking using discrimination, uniqueness, and validity selection. *Multimed Syst* 2019;25:487-511.
 27. Dalal N, Triggs B. editors. Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005;1:866-93.
 28. Wang S, Yan J, Yan Z. Improved moving object detection algorithm based on local united feature. *Chinese Journal of Scientific Instrument* 2015;36:2241-8.
 29. Viola P, Jones M. editors. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001;1:I-I.
 30. Girshick R. editor. Fast r-cnn. 2015 IEEE International Conference on Computer Vision (ICCV), 2015;1440-8.
 31. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal Mach Intell* 2017;39:1137-49.
 32. Shi J, Li Z, Zhu T, et al. Defect Detection of Industry Wood Veneer Based on NAS and Multi-Channel Mask R-CNN. *Sensors (Basel)* 2020;20:4398.

33. Xi D, Qin Y, Wang Y. Vision Measurement of Gear Pitting Under Different Scenes by Deep Mask R-CNN. *Sensors (Basel)* 2020;20:4298.
 34. Zhang L, Wu J, Fan Y, et al. An Efficient Building Extraction Method from High Spatial Resolution Remote Sensing Images Based on Improved Mask R-CNN. *Sensors (Basel)* 2020;20:1465.
 35. Zhang Y, Chu J, Leng L, et al. Mask-Refined R-CNN: A Network for Refining Object Details in Instance Segmentation. *Sensors (Basel)* 2020;20:1010.
 36. Dai J, He K, Sun J. editors. Instance-aware semantic segmentation via multi-task network cascades. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:3150-8.
 37. Li Y, Qi H, Dai J, et al. editors. Fully convolutional instance-aware semantic segmentation. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017:4438-46.
 38. Wang T, Zhong L, Yuan J, et al. Quantitative analysis of functional filtering bleb size using Mask R-CNN. *Ann Transl Med* 2020;8:709.
 39. Armato SG 3rd, Huisman H, Drukker K, et al. PROSTATEx Challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images. *J Med Imaging (Bellingham)* 2018;5:044501.
 40. Chen L, Song H, Wang C, et al. Liver tumor segmentation in CT volumes using an adversarial densely connected network. *BMC Bioinformatics* 2019;20:587.
 41. Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26:297-302.
 42. Heimann T, van Ginneken B, Styner MA, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans Med Imaging* 2009;28:1251-65.
 43. Liao M, Zhao Y, Zeng Y, et al. Liver segmentation from abdominal ct volumes based on graph cuts and border marching. *J Electron Inf Technol* 2016;38:1552-6.
 44. Long J, Shelhamer E, Darrell T, editors. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015:3431-40.
 45. Kaluva KC, Khened M, Kori A, et al. 2d-densely connected convolution neural networks for automatic liver and tumor segmentation. (2018). Available online: <https://arxiv.org/abs/1802.02182>
 46. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, et al. editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer, 2015:234-41.
 47. Godoy P, Hewitt NJ, Albrecht U, et al. Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. *Arch Toxicol* 2013;87:1315-530.
 48. Chiao JY, Chen KY, Liao KY, et al. Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine (Baltimore)* 2019;98:e15200.
 49. Yang Z, Yuan Y, Zhang M, et al. Safety Distance Identification for Crane Drivers Based on Mask R-CNN. *Sensors (Basel)* 2019;19:2789.
- (English Language Editor: L. Roberts)

Cite this article as: Chen X, Wei X, Tang M, Liu A, Lai C, Zhu Y, He W. Liver segmentation in CT imaging with enhanced mask region-based convolutional neural networks. *Ann Transl Med* 2021;9(24):1768. doi: 10.21037/atm-21-5822