Automatic Couinaud Segmentation from CT Volumes on Liver Using GLC-UNet

Jiang Tian(✉), Li Liu, Zhongchao Shi, and Feiyu Xu

AI Lab, Lenovo Research, Beijing, China
{tianjiang1,liuli16,shizc2,fxu}@lenovo.com

Abstract. Automatically generating Couinaud segments on liver, a prerequisite for modern surgery of the liver, from computed tomography (CT) volumes is a challenge for the computer-aided diagnosis (CAD). In this paper, we propose a novel global and local contexts UNet (GLC-UNet) for Couinaud segmentation. In this framework, intra-slice features and 3D contexts are effectively probed and jointly optimized for accurate liver and Couinaud segmentation using attention mechanism. We comprehensively evaluate our system performance (98.51% in terms of Dice per case on liver segmentation, and 92.46% on Couinaud segmentation) on the Medical Segmentation Decathlon dataset (task 8, hepatic vessels and tumor) from MICCAI 2018 with our annotated 43,205 CT slices on liver and Couinaud segmentation. (https://github.com/GLCUnet/dataset).

1 Introduction

Liver cancer is one of the most common cancer diseases and the fourth leading cause of cancer death in the world.[1] "A good knowledge of the anatomy is a prerequisite for modern surgery of the liver." [5] The Couinaud segmentation (detailed definition in Sect. 1 in the Supplementary) is currently the most widely used system to describe functional liver anatomy [8]. It is the preferred anatomy classification system since it divides the liver into eight independent functional units allowing resection of segments without damaging other segments. The system uses the vascular supply in the liver to separate the functional units [9].

In the clinical diagnosis, partitioning sub-regions in Couinaud segmentation is with 3D surfaces. As shown in Fig. 4, partitioning of green and red regions is obtained by a line containing the inferior vena cava (IVC) and the major branch

[1] http://www.who.int/news-room/fact-sheets/detail/cancer.

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-32692-0_32) contains supplementary material, which is available to authorized users.

© Springer Nature Switzerland AG 2019
H.-I. Suk et al. (Eds.): MLMI 2019, LNCS 11861, pp. 274–282, 2019.
https://doi.org/10.1007/978-3-030-32692-0_32

of the right hepatic vein (RHV) (shifts from slice to slice). Couinaud segmentation of a CT volume is consequently on a slice-by-slice basis. It is worth noting that the caudate lobe of liver, a very important sub-region (light blue region in Fig. 4, missing in traditional method [10]), is partitioned using a curve. The hepatic artery and portal vein inflow to and hepatic vein drainage of the liver have many anatomic variations. More specifically, there may be an early bifurcation, early trifurcation or even multiple hepatic veins running in the hepatic fissure. Meanwhile, the liver is a common site of primary or secondary tumor development. Their heterogeneous and diffusive shape will make the situation even worse. Consequently, it is difficult to deduce segmental anatomy of the liver. Since the treatment planning enters "segment era", accurate measurement of each segment is a prerequisite for modern surgery of the liver.

In the clinical diagnosis, it is necessary to probe the spatial information far along the z-axis to make Couinaud segmentation, which is time-consuming. Therefore, fully automatic methods are highly demanded. Recently, fully convolutional neural networks (FCNs) [1,3,7] have achieved great success on a broad array of computer vision problems. Many researchers advance this stream using deep learning methods in the liver and tumor segmentation problem [2]. However, 2D FCN based methods ignore the contexts along the z-axis, which would lead to limited segmentation accuracy. Meanwhile, a 3D FCN, where convolutions in regular 2D FCN are replaced by 3D convolutions with volumetric data input, may probe the 3D contexts. However, the high memory consumption limits the depth of the network to capture long-range context.

To address the aforementioned problem, in this paper, we propose a framework, GLC-UNet, where 3D contexts are effectively probed using attention mechanism for accurate Couinaud segmentation. Meanwhile, experienced doctors usually observe the hepatic hilum and the second hepatic hilum first to obtain a holistic overview of hepatic vessels, which serves as a "skeleton" to constrain the Couinaud segmentation of each slice. Inspired by this fact, in this paper, the proposed framework utilizes this "skeleton" as the global context in Couinaud segmentation. The global context is a sampling of the CT volume[2] around the hepatic hilum and the second hepatic hilum, which are highly correlated with the slices which contain the right and left branch of portal veins (the RPV and the LPV). Incorporating such clinical domain knowledge into network design boosts the performance. Furthermore, we apply this system to a publicly available liver CT dataset. We make annotation on both liver and Couinaud segments of this dataset. Finally, we conduct comprehensive experiments to analyze the effectiveness of our proposed GLC-UNet.

2 Couinaud Segmentation Using GLC-UNet

Given a CT volume, a GLC-UNet is first utilized to obtain segmentation of the liver, which serves as the region of interest for Couinaud segmentation. Next, the RPV and the LPV are detected. The slices contain these two veins are baseline

[2] A set of CT slices selected from the CT volume.

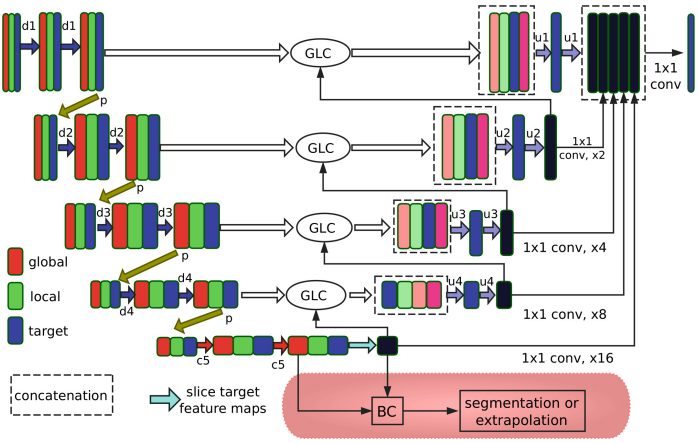

Fig. 1. Overall illustration of the proposed GLC-UNet. In the contracting path, except the bottom level, the global, local, and target feature maps are transferred into two directions, one to GLC and the other to the following convolutional layers. On the bottom level, only the feature maps of the target slice is transferred into the expanding path. In the expanding path, feature maps from GLC are transferred to the following upsampling layer. Meanwhile, feature maps from different layers are fused (1×1 convolution followed by bilinear resize) before the final prediction. In the end, the features, from all the branches, are combined into a single tensor via concatenation. The segmentation or extrapolation classification in the shaded area is for Couinaud segmentation only. Operation d1 denotes concatenation of 3 different convolutional layers: "$6 \times 6, 12, a2$", "$6 \times 6, 12, a2$", and "$3 \times 3, 24, c$. Note that "$6 \times 6, 12, a2$" corresponds to atrous convolutional layer (rate $= 2$) with kernel size of 6×6 and 12 features, and "$3 \times 3, 24, c$" stands for regular convolutional layer with kernel size of 3×3 and 24 features. Detailed operations parameters are listed in Sect. 2 in the Supplementary. GLC model is shown in Fig. 2, and Fig. 3 illustrates the corresponding binary classification used in Couinaud segmentation.

for global context for Couinaud segmentation. Finally, another multi-task GLC-UNet is used to obtain Couinaud segmentation of liver.

GLC-UNet with Multi-scale Information. As shown in Fig. 1, the proposed GLC-UNet efficiently probes intra-slice and inter-slice features. The inter-slice feature extraction consists of two parts. One is the global context, and the other is the local context. A context at the input level is defined as a set of CT slices selected from a CT volume. Denote by $\mathbf{G}$ the input global context. We have the corresponding global context feature maps at level t for the i-th slice in $\mathbf{G}$ as $\mathbf{G}_{ti}$ (similarly for the input local context $\mathbf{L}$ and the corresponding $\mathbf{L}_{ti}$). It is worth noting that convolutional and pooling operations in the contracting path independently applied on each input in this set. For intra-slice feature, the idea is to provide the model with multi-scale information. To do that, we add a series of atrous convolutions with different dilation rates. These rates are designed to capture long-range in-plane context.

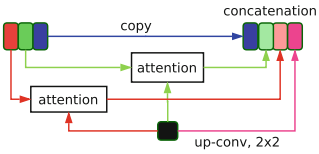

Fig. 2. GLC architecture.

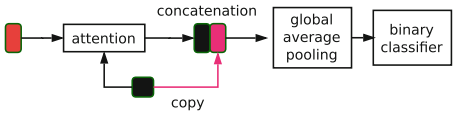

Fig. 3. Segmentation or extrapolation.

In the contracting path, the feature maps are transferred into two directions, one to GLC and the other to the following convolutional layers. The global context, local context, and the target slice are combined using attention mechanisms as shown in Fig. 2, which distill the visual features by concatenating the global and the local contexts, the target slice's feature maps, and the upsampled feature maps. On the bottom level, only the feature map of the target slice is transferred into the expanding path to generate intra-slice features. In the expanding path, feature maps from GLC are transferred to the following upsampling layer. Meanwhile, feature maps from different levels are fused before the final prediction. In the end, the features, from all the branches, are combined into a single tensor via concatenation.

Global and Local Contexts Composition. Let $(S_1, \ldots, S_\Psi)$ be the set of slices in a CT volume, and tar be the index along z-axis of the target slice.

For a specific CT volume, in Couinaud segmentation, the global context (a set of slices) is the same for all slices to be segmented in this CT volume. It is driven by liver anatomy specifically. Throughout the paper, we will denote by Π the total number of the global context slices, and by zt the top index of the slice with liver region. We have $\mathbf{G}$, the input global context, as follows.

$$ze = (zl + zt)/2, \tag{1}$$

$$zb = min(ze - (\Pi - 1), zr), \tag{2}$$

$$step = (ze - zb)/(\Pi - 1), \tag{3}$$

$$\mathbf{G} = (S_{zb}, S_{zb+step}, \ldots, S_{ze-step}, S_{ze}), \tag{4}$$

where zl and zr are the corresponding coordinates along z-axis of the LPV and the RPV, respectively. These coordinates are obtained via detection of the RPV and the LPV via a faster R-CNN network.

In clinical, a radiologist usually observes a couple of adjacent slices to segment liver. Therefore, for liver segmentation, $\mathbf{G}$ is obtained as follows.

$$\mathbf{G} = (S_{tar-4}, S_{tar-2}, S_{tar+2}, S_{tar+4}), \tag{5}$$

where the range along z-axis is empirically set to ± 4.

Adjacent slices, $\mathbf{L} = (S_{tar-1}, S_{tar+1})$, are employed to provide local context, which serves as a smoothing factor both for liver and Couinaud segmentation.

GLC Model. As shown in Fig. 2, information extracted from coarse level is used in GLC to disambiguate irrelevant and noisy responses in skip connections,

and to guide the global and local contexts composition. It is performed right before the concatenation operation to merge only relevant activations.

The global context feature maps $\mathbf{G}_t = \sum_{i=1}^{\Pi} \alpha_{ti}\mathbf{G}_{ti}$ is dynamic maps that represent the relevant part of features at level t, where α_{ti} is a scalar weighting of feature maps $\mathbf{G}_{ti}$ at level t, defined as follows.

$$\alpha_{ti} = exp(e_{ti}) / \sum_{k=1}^{\Pi} exp(e_{tk}), \qquad e_{ti} = f_{att}(\mathbf{a}_{ti}, \mathbf{h}_{t+1}). \tag{6}$$

f_{att} is a function that determines the amount of attention allocated to feature maps $\mathbf{G}_{ti}$, conditioned on the target feature maps $\mathbf{H}_{t+1}$ from coarser level. This function is implemented as a multilayer perceptron as $f_{att} = \mathbf{w}^T tanh(U\mathbf{h}_{t+1} + W\mathbf{a}_{ti} + \mathbf{b})$. Note that by construction $\sum_{i=1}^{\Pi} \alpha_{ti} = 1$. Feature vectors $\mathbf{a}_{ti}$ and $\mathbf{h}_{t+1}$, reduced from three dimensional tensors, are obtained using global average pooling (GAP) as $\mathbf{a}_{ti} = GAP(\mathbf{G}_{ti})$ and $\mathbf{h}_{t+1} = GAP(\mathbf{H}_{t+1})$.

Similarly, the local context feature maps are obtained as $\mathbf{L}_t = \sum_{i=1}^{2} \beta_{ti}\mathbf{L}_{ti}$, where β_{ti} is a scalar weighting of feature maps $\mathbf{L}_{ti}$ at level t. Skip-connection brings feature maps $\mathbf{H}_t$ in the encoder to the decoder of the same level. Feature maps $\mathbf{D}_t$ generated on the same level of the expanding path are merged through concatenation to combine global and local contexts as $\mathbf{Z}_t = [\mathbf{H}_t, \mathbf{L}_t, \mathbf{G}_t, \mathbf{D}_t]$.

Segmentation or Extrapolation. There are a couple of slices with no visible vascular information in the upper and the lower part of a liver. In this situation, an experienced radiologist extrapolates the Couinaud segments from the nearest slice with acceptable visual feature. Consequently, for each slice to be processed, GLC-UNet has to first decide whether segmentation should proceed or come to an extrapolation.

Feature vector $\mathbf{v}$ is obtained by concatenating feature vectors (in Fig. 3) from the global context feature maps and the target one as $\mathbf{v} = [GAP(\mathbf{G}_t), GAP(\mathbf{H}_t)]$, where $\mathbf{G}_t = \sum_{i=1}^{\Pi} \alpha_{ti}\mathbf{G}_{ti}$, α_{ti} is a scalar weighting of feature maps $\mathbf{G}_{ti}$ at level t defined in Section **GLC Model**. However, the amount of attention allocated to feature maps $\mathbf{G}_{ti}$ is conditioned on the feature maps $\mathbf{H}_t$ from the same level. A linear projection from $\mathbf{v}$ and a logistic classifier produce a distribution over $[SEG = 0, EXT = 1]$ as $p(ext|\mathbf{v}) \propto exp(G_{ext}\mathbf{v} + B_{ext})$, where G_{ext} and B_{ext} are parameters to be learned. If $p(ext|\mathbf{v})$ is greater than predefined threshold (e.g. 0.5), then GLC-UNet will stop producing segmentation and make extrapolation.

3 Experiments

We use a publicly available dataset of MICCAI 2018 Medical Segmentation Decathlon (task 8) [4]. Both training and testing datasets have no annotated ground truth segmentation masks for liver and Couinaud segments. For liver, we annotate 443 CT volumes in total (31,668 slices). For Couinaud segments, we annotate 193 sampled CT volumes in total (11,537 slices). To be consistent, we use the same 193 CT volumes for both liver and Couinaud segmentation.

We randomly select 50 cases for testing. Training is conducted on the remaining 143 ones. By default, the number of input global context slices Π for Couinaud segmentation is set to 10 in the experiments considering both performance and model complexity.

Comparison with Other Methods. We compare the proposed GLC-UNet with UNet [3], UNet 2.5D, UNet 3D, and UNet with convolutional LSTM [6] both on liver and Couinaud segmentation. For the latter one, in order to make fair comparison, we only use slices with Couinaud segments annotation. The input for UNet 2.5D, UNet 3D, and UNet with convolutional LSTM are 7 consecutive slices. UNet and UNet 2.5D are in their canonical configurations. Both methods are inherently not designed to capture long-range 3D context effectively. Due to GPU memory limit, in Couinaud segmentation, it is not possible to feed 13 slices (input for GLC-UNet, NVIDIA Tesla P100) into UNet 3D, and UNet with convolutional LSTM. Meanwhile, two different experiments, one with ground truth liver mask and another based on predicted mask, are conducted.

We evaluate the performance on the Dice per case score over all of the CT volumes. The results are given in Table 1. The comparison on Couinaud Dice per case score indicates not only that model using inter-slice context improves over the baseline model (UNet) with only intra-slice features, but also that the richer the context information, the bigger the performance difference compared with the baseline model. It suggests that our framework provides better alignment from Couinaud segments to provided vascular features, and global context becomes particular necessary. In order to assess how the results of GLC-UNet will generalize to the whole dataset (193 CT volumes in total), we perform a 4-fold cross validation. The dataset is divided into four parts randomly. Three-fourths are utilized for training and one-fourth for validation. The last row in Table 1 demonstrates that GLC-UNet achieves good performance consistently.

Table 1. Dice per case scores on liver segmentation and Couinaud segmentation. All values (mean ± std) are reported as percentage (%).

Method	Couinaud segmentation		Liver segmentation
	Ground truth liver	Predicted liver	
GLC-UNet	95.10 ± 3.06	92.80 ± 3.08	98.18 ± 0.85
UNet	91.31 ± 3.67	83.86 ± 5.44	95.33 ± 2.78
UNet 2.5D	92.41 ± 3.30	89.04 ± 4.30	97.61 ± 1.59
UNet 3D	92.92 ± 3.46	88.93 ± 4.44	96.97 ± 1.65
UNet with convolutional LSTM	93.26 ± 3.12	90.06 ± 4.01	97.73 ± 1.73
GLC-UNet (cross validation)	94.51 ± 3.62	92.46 ± 3.84	98.51 ± 0.74

Effectiveness of Multi-scale Intra-slice Information One advantage in the proposed method is that we provide the model with multi-scale intra-slice information, which consists of atrous convolutions with different dilation rates

and a 3×3 convolution. Here, we analyze the behaviors of GLC-UNet with and without atrous convolutions. Both two experiments are conducted under the same experimental settings. As shown in Table 2 (Methods 1 and 2), it is clearly observed that with the atrous convolution model, GLC-UNet achieves better Dice per case score, which shows the importance of utilizing the atrous convolution to capture long-range in-plane context.

Table 2. Dice per case scores for Couinaud segmentation by ablation study of our methods on the dataset. All values (mean ± std) are reported as percentage (%).

#	Method	Dice per case	#	Method	Dice per case
1	GLC-UNet	95.10 ± 3.06	2	without atrous convolution	94.49 ± 3.09
3	$\Pi = 3$	94.25 ± 3.10	4	without 3×3 convolution	94.53 ± 3.11
5	$\Pi = 5$	94.58 ± 3.64	6	without resized features	94.53 ± 3.20
7	$\Pi = 7$	94.97 ± 2.98	8	3×3 convolution in decoder	94.75 ± 2.87

Method 4 in Table 2 demonstrates that the 3×3 convolution in in the contracting path can help the network achieve better performance by extracting abundant vascular information, which plays an important role in achieving the promising results. Combining feature maps from different levels just before the final prediction makes local predictions respect global structure, which is confirmed by the result in Method 6 in Table 2. Meanwhile, as shown in Method 8 in Table 2, Couinaud segments are global information, small convolutional kernel (3×3) in the expanding path will degrade the performance.

Effectiveness of Inter-slice Information. In the clinical diagnosis, an experienced radiologist usually observes and makes Couinaud segmentation according to many slices along the z-axis. Here, we analyze the behaviors of GLC-UNet with different number of global context slices. Other experimental settings are the same for all experiments. As shown in Table 2 (Methods 3, 5 and 7), increasing the number of global context slices Π can generally lead to performance improvements. This is due to the fact that more slices increase the possibility of sampling important landmark veins, which in turn leads to better performance.

Couinaud Segmentation Results. Among $1,272$ slices (from the testing dataset) with liver region for segmentation, 67 ones are predicted as extrapolation. In comparison, among 877 slices with liver region for extrapolation, 63 ones are predicted as segmentation. Misclassification usually occurs in the transition area (from extrapolation to segmentation or vice versa). The Couinaud segments on the slices in Fig. 4 marked with red circles are extrapolated from the nearest slice with acceptable visual feature.

As far as performance of locating the slices which contain the LPV and the RPV is concerned, the recall rate is 79%. If we set the threshold to ± 1 slice, which means that the LPV and the RPV are located at adjacent slices of the

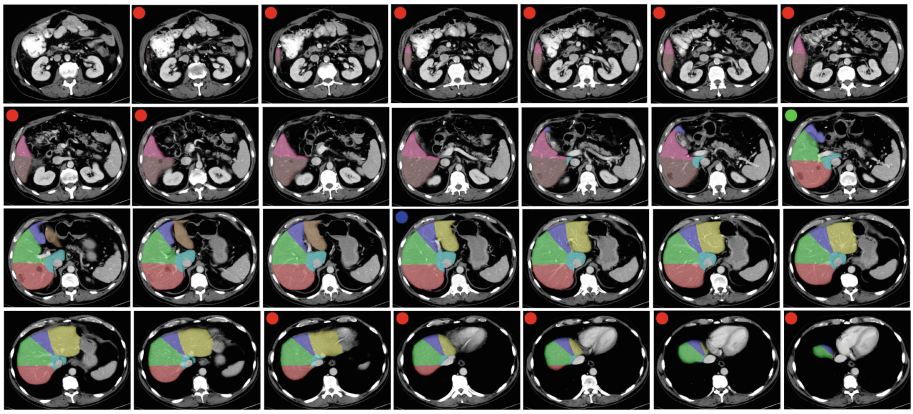

Fig. 4. An example of generating Couinaud segmentation (more results are available in Sect. 4 in the Supplementary). Segments I to VIII are marked with different colors. Slices marked with red circle are from extrapolation. RPV is on the slice with green circle, and LPV is on the one with blue circle. (Color figure online)

target, the rate will be 96%. As shown in Fig. 4, the last slice in the second row and the first one in the third row are almost indistinguishable to be marked as the slice which contains the RPV.

Experiment on Other Dataset. We perform an 8-fold cross validation on the training dataset of the task 09 (spleen segmentation) in Medical Segmentation Decathlon [4] to validate the effectiveness of GLC-UNet. The same GLC-UNet as the one used in the liver segmentation in the paper is utilized here. The Dice score is 95.4%. It shows that the design is generalizable for other organ segmentation tasks, and has been validated by spleen segmentation.

4 Conclusion

This paper investigates automatically generating Couinaud segmentation, which is the preferred anatomy classification system in clinical, on liver from CT volumes. We propose a novel convolutional encoder-decoder network for 2D slice segmentation where 3D contexts are efficiently probed using attention mechanism. Finally, extensive experiments on the Medical Segmentation Decathlon dataset from MICCAI 2018 with our annotated 43,205 CT slices on liver and Couinaud segmentation demonstrate the effectiveness of our proposed GLC-UNet.

References

1. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: CVPR (2015)
2. Sun, C., et al.: Automatic segmentation of liver tumors from multiphase contrast-enhanced ct images based on FCNs. AI Med. **83**, 58–66 (2017)
3. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
4. MICCAI 2018 Medical Segmentation Decathlon. http://medicaldecathlon.com/
5. Rutkauskas, S., Gedrimas, V., Pundzius, J., Barauskas, G., Basevicius, A.: Clinical and anatomical basis for the classification of the structural parts of liver. Medicina-Lithuania **42**(2), 98–106 (2006)
6. Zhang, Y., et al.: SequentialSegNet: combination with sequential feature for multi-organ segmentation. In: ICPR (2018)
7. Milletari, F., Navab, N., Ahmadi, S.-A.: V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: 3DV (2016)
8. Lowe, M.C., D'Angelica, M.I.: Anatomy of hepatic resectional surgery. Surg. Clin. North Am. **96**(2), 183–195 (2016)
9. Jarnagin, W.R., Belghiti, J., Blumgart, L.H.: Blumgart's Surgery of the Liver, Biliary Tract, and Pancreas. Elsevier, Amsterdam (2012)
10. Oliveira, D.A.B., Feitosa, R.Q., Correia, M.M.: Automatic couinaud liver and veins segmentation from CT images. In: BIOSIGNALS (2008)